CrossMark

J. Fluid Mech. (2018), *vol.* 839, *pp.* 95–134. © Cambridge University Press 2018
doi:10.1017/jfm.2017.913

Geometric generalised Lagrangian-mean theories

Andrew D. Gilbert[1] **and Jacques Vanneste**[2,†]

[1]Department of Mathematics, College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter EX4 4QF, UK

[2]School of Mathematics and Maxwell Institute for Mathematical Sciences, University of Edinburgh, Edinburgh EH9 3FD, UK

(Received 21 December 2016; revised 23 August 2017; accepted 11 December 2017; first published online 25 January 2018)

Many fluctuation-driven phenomena in fluids can be analysed effectively using the generalised Lagrangian-mean (GLM) theory of Andrews & McIntyre (*J. Fluid Mech.*, vol. 89, 1978, pp. 609–646) This finite-amplitude theory relies on particle-following averaging to incorporate the constraints imposed by the material conservation of certain quantities in inviscid regimes. Its original formulation, in terms of Cartesian coordinates, relies implicitly on an assumed Euclidean structure; as a result, it does not have a geometrically intrinsic, coordinate-free interpretation on curved manifolds, and suffers from undesirable features. Motivated by this, we develop a geometric generalisation of GLM that we formulate intrinsically using coordinate-free notation. One benefit is that the theory applies to arbitrary Riemannian manifolds; another is that it establishes a clear distinction between results that stem directly from geometric consistency and those that depend on particular choices. Starting from a decomposition of an ensemble of flow maps into mean and perturbation, we define the Lagrangian-mean momentum as the average of the pull-back of the momentum one-form by the perturbation flow maps. We show that it obeys a simple equation which guarantees the conservation of Kelvin's circulation, irrespective of the specific definition of the mean flow map. The Lagrangian-mean momentum is the integrand in Kelvin's circulation and distinct from the mean velocity (the time derivative of the mean flow map) which advects the contour of integration. A pseudomomentum consistent with that in GLM can then be defined by subtracting the Lagrangian-mean momentum from the one-form obtained from the mean velocity using the manifold's metric. The definition of the mean flow map is based on choices made for reasons of convenience or aesthetics. We discuss four possible definitions: a direct extension of standard GLM, a definition based on optimal transportation, a definition based on a geodesic distance in the group of volume-preserving diffeomorphisms, and the 'glm' definition proposed by Soward & Roberts (*J. Fluid Mech.*, vol. 661, 2010, pp. 45–72). Assuming small-amplitude perturbations, we carry out order-by-order calculations to obtain explicit expressions for the mean velocity and Lagrangian-mean momentum at leading order. We also show how the wave-action conservation of GLM extends to the geometric setting. To make the paper self-contained, we introduce in some detail the tools of differential geometry and main ideas of geometric fluid dynamics on which we rely. These include variational formulations which we use for alternative derivations of some key results. We mostly focus on the Euler equations

† Email address for correspondence: j.vanneste@ed.ac.uk

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

for incompressible inviscid fluids but sketch out extensions to the rotating–stratified Boussinesq, compressible Euler, and magnetohydrodynamic equations. We illustrate our results with an application to the interaction of inertia-gravity waves with balanced mean flows in rotating–stratified fluids.

Key words: mathematical foundations, wave–turbulence interactions, waves/free-surface flows

1. Introduction

The interactions between waves, or more broadly fluctuations, and mean flows are at the heart of a broad range of fluid-dynamical phenomena, from acoustic streaming to the formation of jets in geo- and astrophysical fluids. These phenomena involve the nonlinear rectification of fluctuations, leading to the emergence of persistent mean-flow structures. It has long been realised that they are best described using some form of Lagrangian representation of fluid motion, especially when dissipative processes play a secondary role. In essence, this is because persistent flows can often be identified with changes in vorticity which, in turn, is controlled by material transport in the absence of dissipation. The theoretical articulation of this observation, pioneered by Dewar (1970), Bretherton (1971), Soward (1972), Grimshaw (1975) and others, culminated in the development by Andrews & McIntyre (1978*a*) of the generalised Lagrangian-mean (GLM) theory which has since become an organising centre for research on wave–mean-flow interactions. We refer the reader to the book by Bühler (2014) for a detailed exposition of the GLM formalism and of some of its more recent applications.

One of the advantages of GLM is its generality: it considers an arbitrary ensemble of flows and an arbitrary averaging operator, and deduces from the Navier–Stokes equations or their generalisations averaged equations governing the fully nonlinear dynamics of a suitably defined Lagrangian mean flow. These can be specialised to specific settings, typically involving assumptions of multiple time scales, small amplitudes and spatial scalings, but in their generality they exhibit fundamental properties such as an averaged form of Kelvin's circulation theorem. They also capture, at once, multiple asymptotic regimes, something that standard asymptotic expansion methods do not achieve (compare, for instance, the GLM approach of Bühler & McIntyre (1998) and Holmes–Cerfon Bühler & Ferrari (2011) with the multiscale approach of Wagner & Young (2015)).

The GLM formalism has however one important deficiency, especially for incompressible fluids, which are our main focus. Because it has been developed in Cartesian coordinates and makes use of Euclidean parallel transport in its foundations, GLM is constructed outside the natural configuration space of incompressible fluid dynamics, namely the space SDiff(M) of all volume-preserving diffeomorphisms from the fluid domain manifold M to itself. In this sense GLM does not respect the intrinsic geometry of fluid dynamics – neither the finite-dimensional geometry of M, nor the infinite-dimensional geometry of SDiff(M).

This has two consequences. First, GLM does not apply straightforwardly to arbitrary domain manifolds M. This is not only an academic matter: non-trivial manifolds such as the sphere or spherical shells arise in many applications. Attempts to use GLM for these, using either the Euclidean structure of the embedding space or treating local coordinates as if they were Cartesian, lead to unsatisfactory results, with mean

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

velocity fields that fail to be tangent to M or that depend on the choice of coordinates (McIntyre 1980; Bühler 2014, § 10.2.1).

Second, even in simple Euclidean geometry, the GLM Lagrangian-mean velocity in an incompressible fluid is generically divergent, even though it is an average of divergence-free velocity fields. This phenomenon, which has been called the 'divergence effect' (McIntyre 1988; Bühler 2014, § 10.2.5), implies that the GLM Lagrangian-mean velocity does not belong to the tangent space of SDiff(M).

To understand this state of affairs, let us recall how the Lagrangian mean flow is defined. Consider a fluid particle that is mapped by an ensemble of flows to the positions $\boldsymbol{x}' = \boldsymbol{x}'^{\alpha}$, where α indexes the ensemble member. (In discussing Andrews & McIntyre's GLM, we will follow closely their notation, in particular in omitting the label α, thus making implicit the dependence on ensemble member.) The Lagrangian-mean position $\boldsymbol{x}$ of the particle and the (ensemble-dependent) displacement $\boldsymbol{\xi}$ are then defined by

$$\boldsymbol{x}' = \boldsymbol{x} + \boldsymbol{\xi}(\boldsymbol{x}, t) \quad \text{and} \quad \langle \boldsymbol{\xi}(\boldsymbol{x}, t) \rangle = 0, \tag{1.1a,b}$$

where the brackets denote an ensemble average. The difficulty in interpreting (1.1) on non-Euclidean manifolds is clear: particle positions are then points, not vectors, and linear operations including the addition and averaging in (1.1) do not apply to them. The definition

$$\overline{\boldsymbol{u}}^{L}(\boldsymbol{x}, t) = \langle \boldsymbol{u}(\boldsymbol{x} + \boldsymbol{\xi}(\boldsymbol{x}, t), t) \rangle \tag{1.2}$$

of the Lagrangian-mean velocity $\overline{\boldsymbol{u}}^{L}$ is similarly problematic: the vectors $\boldsymbol{u}(\boldsymbol{x} + \boldsymbol{\xi}(\boldsymbol{x}, t), t)$ on the right-hand side belong to different tangent planes for different ensemble members α and hence cannot be averaged in a geometrically intrinsic way. As noted above, a consequence is that $\overline{\boldsymbol{u}}^{L}$ is not tangent to non-Euclidean manifolds even though the vector fields $\boldsymbol{u}$ are, and another is that $\boldsymbol{\nabla} \cdot \overline{\boldsymbol{u}}^{L}$ can be non-zero even though $\boldsymbol{\nabla} \cdot \boldsymbol{u}$ is zero for each ensemble member. Whether these consequences are a price worth paying for the simplicity of (1.1)–(1.2) and of the resulting averaged equations of motion is a matter of opinion that will depend on the application considered.

The present paper aims at overcoming the deficiency of GLM associated with its lack of geometrically intrinsic meaning. To do so, we develop a generalisation that rests on solid geometric foundations. The main benefits are that the theory developed (i) applies to any real (Riemannian) manifold, (ii) is independent of the choice of coordinate system, (iii) leads to a generalised mean velocity that is divergence free for incompressible fluids, and (iv) provides clear interpretations for the computations leading to averaged equations, distinguishing steps imposed by geometric consistency from those requiring choices to be made.

The motivations for our paper and some of our results are similar to those of Roberts & Soward (2006*a*,*b*) and Soward & Roberts (2010, 2014), building on the pioneering work of Soward (1972) and on more recent developments by Gjaja & Holm (1996) and Holm (2002*a*,*b*). One difference is that these authors restrict their attention to Euclidean space. A more significant difference is that we formulate our theory in the language of modern coordinate-free differential geometry rather than classical tensor calculus. The advantages of a coordinate-free treatment are well established in many branches of physics (e.g. Frankel 1997): it enables compact derivations, guarantees the coordinate independence of the results, and often offers clear geometric interpretations. A drawback in the present context is that it may not be familiar to the majority of fluid dynamicists. To remedy this, the paper starts in § 2 with a self-contained introduction to the geometric formulation of fluid dynamics,

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

including the necessary differential geometry. This formulation, initiated by Arnold (1966), interprets the Euler equations for incompressible perfect fluid motion as geodesic equations in SDiff(M) and defines the geometric structure of incompressible fluid dynamics. Our approach is to work as much as possible within this geometric structure or later its counterpart for compressible fluids.

One of the insights of Roberts & Soward (2006*a*,*b*), reflected in our work, is that the form of the averaged equations governing the Lagrangian-mean dynamics is to a large extent independent of the specific definition chosen for the mean flow (which, in standard GLM, is set uniquely by the condition $\langle \boldsymbol{\xi} \rangle = 0$ in (1.1)). A key idea here is that the mean flow is defined at the level of the flow maps – the diffeomorphisms that send fluid positions from their initial to their current configurations – rather than velocity fields. Thus, given an ensemble of flow maps ϕ^α associated with perturbed fluid configurations, a single α-independent mean flow map $\bar{\phi}$ is chosen to represent the ensemble. Thinking of the ϕ^α as forming a cluster of points in SDiff(M), as is the case for wave-like motion, this only requires $\bar{\phi}$ to be located well within the cluster, in a sense that specific definitions of the mean flow map make precise. We emphasise that $\bar{\phi}$ is in general not constructed by application of an averaging operator since the perturbed maps ϕ^α are nonlinear objects that cannot be averaged in any straightforward sense. The mean velocity field associated with the mean flow map $\bar{\phi}$ is naturally defined as its time derivative and is generally not a straightforward average of the perturbed velocities.

We start our analysis in § 3 with a derivation of an averaged momentum equation for an arbitrary choice of the mean flow. The derivation shows that the key element of the averaging procedure, equivalent to that of Soward (1972) and Andrews & McIntyre (1978*a*), is the 'pull-back' of the dynamical equations from the perturbed configuration to the mean configuration. The pull-back operation, to be defined in § 2.1, brings various fluctuating quantities into a single vector space, which allows an averaging operator to be applied to construct mean quantities, including the Lagrangian-mean momentum one-form (or covariant vector). The Lagrangian-mean momentum appears as the integrand in the averaged Kelvin circulation theorem that follows and generalises the GLM quantity $\overline{\boldsymbol{u}}^L - \boldsymbol{p}$, with $\boldsymbol{p}$ the GLM pseudomomentum. A more general notion of pseudomomentum emerges: it is interpreted as a one-form that encapsulates the averaged effect of the fluctuations on the Kelvin circulation, and measures the lack of commutation between averaging on the one hand, and the transformation of the mean velocity vector into a one-form using the manifold's metric, on the other.

We consider the definition of the mean flow in § 4 where we discuss four possible choices. The first extends standard GLM to arbitrary manifolds M without rectifying the problem of divergent mean velocity. The other three do rectify that problem, the second using ideas from optimal transportation, and the third – in our view the most natural – relying on a geodesic distance between flow maps in SDiff(M) to define the mean flow map as a Riemannian centre of mass in SDiff(M). The fourth is the 'glm' definition recently proposed by Soward & Roberts (2010, 2014) which extends work by Holm (2002*a*,*b*). This definition relates the perturbation flow maps to vector fields and imposes a zero-average condition on the latter.

Although our theoretical developments hold formally for finite-amplitude perturbations, practical applications rest on perturbative expansions in powers of a small-amplitude parameter, which are often truncated to the first non-trivial order. In § 5, we derive coordinate expressions for such leading-order expansions for both mean velocity and Lagrangian-mean momentum for each of the four mean-flow

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

definitions. We illustrate the use of these expressions in § 6 with an application to the effect of waves on the balanced dynamics of a rapidly rotating, strongly stratified Boussinesq fluid, giving some insight into the manner in which the divergence effect is resolved. While we do not pursue perturbative computations beyond the first non-trivial order, we note that our formulation is well suited to the systematic computation of higher-order corrections.

In GLM, the equations governing the mean dynamics need to be complemented by a model for the dynamics of the fluctuations. An exact model is of course provided by subtracting the mean equations from the equations describing the full dynamics. This is rarely useful, however: in practice, simplifying assumptions are made on the perturbation dynamics, in order to obtain tractable, closed models of wave–mean flow interactions. Typical simplifications include linearisation, short-wave approximation (see Bühler 2014, and references therein) and heuristic closures (e.g. Holm 2002*a*,*b*). Fundamental constraints on the perturbation dynamics need however to be respected, the most important of which is the conservation of wave action. Andrews & McIntyre (1978*b*) derived an exact, general GLM version of wave-action conservation (see also Grimshaw 1984; Bühler 2014, § 10.3). We obtain a geometric counterpart of their derivation in § 7.

We conclude this introduction with three remarks. First, our paper focusses chiefly on the most basic fluid model, that is, the Euler equations for incompressible perfect fluids enclosed by rigid boundaries. Admittedly, this is not the most interesting model for wave–mean-flow interactions, but we think it best to establish the main ideas and techniques in as simple a context as possible. We illustrate in § 3.3 the relevance of the theory developed to more complicated models by sketching how key results extend to Boussinesq and compressible perfect fluids, and to ideal magnetohydrodynamics (MHD); see also § 6. We leave a complete treatment of these and other relevant models for future work. Second, we note that the developments of GLM have followed two parallel strands, with some authors basing their derivations on the equations of motion (e.g. Soward 1972; Andrews & McIntyre 1978*a*,*b*; Bühler 2014), and others relying on variational formulations (e.g. Dewar 1970; Bretherton 1971; Grimshaw 1984; Gjaja & Holm 1996; Holm 2002*a*,*b*; Salmon 2013). We acknowledge the importance of variational arguments by providing alternative variational derivations of a few essential results. Finally, a word of caution about notation is in order: throughout the paper we have adopted the 'light' notation of differential geometers (not using bold for vectors, for instance), reserving the usual notation of fluid dynamics for references to classical GLM expressions, as in (1.1)–(1.2). This enables a clear distinction between objects that have related but different meanings in the old and new theories. Our notation also avoids some of the shorthands traditionally used in GLM. Here we identify explicitly ensemble-dependent quantities by a superscript α, and make a careful distinction between points in the fluid domain M, and maps between such points, i.e. elements of $\mathrm{SDiff}(M)$.

2. Background

In this section, we provide a brief, informal introduction to the geometric formulation of fluid dynamics pioneered by Arnold (1966). More recent treatments can be found, for instance, in Marsden & Hughes (1983), Arnold & Khesin (1998), Holm, Marsden & Ratiu (1998), Boyland (2001), Brenier (2003), Holm, Schmah & Stoica (2009) and Besse & Frisch (2017). To make our introduction as self-contained as possible, we start by reviewing the main tools of differential geometry needed; these are standard and described, for example, in the books by Hawking & Ellis (1973), Schutz (1980) or Frankel (1997).

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

2.1. *Elements of differential geometry*

In this subsection, we focus on fundamentals of differential geometry. We consider fluid motion on an n-dimensional manifold M which at this point does not have to be equipped with a metric. The manifold M is taken to be fixed, that is, time independent, and to have fixed boundary ∂M (which could be empty). Free-surface effects are therefore ignored, though they could be incorporated in essentially the same formalism. Obvious examples of manifolds M of interest are the Euclidean space $\mathbb{R}^n$, with $n = 2, 3$, the torus $\mathbb{T}^n$, the 2-sphere $\mathbb{S}^2 \subset \mathbb{R}^3$ given by $r = 1$ and the spherical shells $r_0 \leqslant r \leqslant r_1$ in $\mathbb{R}^3$, where r is the radial component of spherical polar coordinates. We treat M in a coordinate-independent way but, since we deal with non-relativistic fluids, retain the special role of the time coordinate t.

The basic object in fluid mechanics, and more generally in continuum mechanics, is the flow map, ϕ_t, which maps the positions of fluid particles at the initial time $t = 0$ to their positions at time t. Thus, if x denotes the position at time t of the particle initially at a,

$$x = \phi_t a. \tag{2.1}$$

The flow map is a diffeomorphism, that is, a smooth invertible map from M to itself. We write $\phi_t : M \to M$ and $\phi_t \in \mathrm{Diff}(M)$, where $\mathrm{Diff}(M)$ is the set of all diffeomorphisms from M to M. We shall be mostly be concerned with the subset $\mathrm{SDiff}(M)$ of $\mathrm{Diff}(M)$ consisting of all the volume-preserving diffeomorphisms from M to M. For convenience, we generally omit the subscript t from the flow map in what follows, simply remembering that ϕ depends on time, as essentially do most of the objects introduced. Being fixed, the boundary ∂M is preserved under the action of any flow map ϕ; if it has more than one connected component, ϕ must map each component to itself, otherwise ϕ would be not be continuously connected to the identity map, the flow map at $t = 0$.

In the standard Eulerian description of fluids, ϕ does not appear explicitly; instead, it is the velocity field $u = u(\cdot)$ that appears as the fundamental dynamical variable. It is defined by

$$\dot{\phi} a = u(x) = u(\phi a), \tag{2.2}$$

where the overdot denotes time derivative and $a \in M$ is arbitrary. Equivalently,

$$u = \dot{\phi} \circ \phi^{-1}, \tag{2.3}$$

where $\circ$ denotes composition. In writing (2.2) we have adhered to our convention of omitting to make explicit the time dependence of u on t. When the velocity field is known, the flow map is deduced by integrating the ordinary differential equations (2.2) with the initial condition $\phi|_{t=0} = \mathrm{id}$, that is, the identity map.

The velocity field is an example of a vector field on M. In general, a vector at x such as $v(x)$ can be identified as an equivalence class of curves through $x \in M$ sharing the same tangent. If $\psi_s x$ is one such curve parameterised by s and with $\psi_0 x = x$, we write that

$$v(x) = \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \psi_s x \in T_x M, \tag{2.4}$$

where $T_x M$ is the tangent space to M at x (the set of all vectors at x). This formalises the interpretation of vectors as displacements – arrows – between two infinitesimally close points x and $y = \psi_s x$ with $s \to 0$, as illustrated in figure 1. (Note that (2.2) is an instance of (2.4) with $\psi_s = \phi_{t+s}$.) A vector field, then, is a map v from M to

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

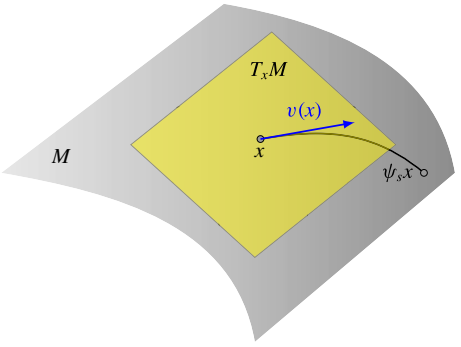


FIGURE 1. (Colour online) A vector $v(x)$ in the tangent space T_xM at $x \in M$ is defined by a curve $\psi_s x$ according to (2.4).

the tangent bundle TM (the union of the tangent spaces T_xM at all points $x \in M$) with the property that $v(x) \in T_xM$. We emphasise this property because traditional GLM formulations involve objects such as $u \circ \phi$ which are not vector fields (since $(u \circ \phi)(x) = u(\phi x) \in T_{\phi x}M \neq T_xM$) and, in our opinion, are best avoided for clarity in a general setting. Note that if M has a fixed boundary ∂M, the velocity field is tangential to it, $u \parallel \partial M$ (meaning that $u(x) \in T_x\partial M$ for boundary points $x \in \partial M$).

Fluid dynamics on arbitrary manifolds is formulated most naturally using fields of differential one-forms in addition to vector fields. A one-form at x is defined as a linear map from vectors in T_xM to the real numbers. One-forms at x make up a vector space, the cotangent space T_x^*M, and their application to vectors of T_xM is interpreted as a pairing (or interior product) which produces a real number. If v is a vector field on M and ν is a field of one-forms, the pairing at x is denoted as

$$\nu(x)(v(x)) = \nu(v)(x) \in \mathbb{R} \tag{2.5}$$

and defines a scalar field, a real function over the whole manifold M. In coordinates, with $v^i(x)$ and $\nu_i(x)$ denoting the components of the vector and one-form, the pairing (2.5) is simply given by $\nu(v) = v^i\nu_i$ at each point x. Note that the coordinates x^i need not be orthogonal; indeed 'orthogonality' is meaningless since no metric has yet been introduced. Pictorially, one-forms can be represented as stacks of infinitesimally close planes (in dimension $n = 3$; hyperplanes of dimension $n - 1$ in general) and their pairing with a vector interpreted as measuring the number of planes pierced by the vector. (Note that, although we will not adopt this terminology, vectors such as $v(x)$ are traditionally referred to as 'contravariant vectors' and one-forms such as $\nu(x)$ as 'covariant vectors', and their pairing as 'contraction'.)

Higher-order forms (and fields thereof) are defined similarly as linear maps on sets of vectors, with added skew-symmetry properties: a k-form α acts linearly on k vectors $v_1, \ldots, v_k$ to produce the real number $\alpha(v_1, \ldots, v_k)$ which is invariant under even permutations of its arguments and flips signs under odd permutations. For example, a two-form α acts linearly on pairs of vectors u, v, with $\alpha(u, v) = -\alpha(v, u) \in \mathbb{R}$; in coordinates, it has components α_{ij} that are anti-symmetric, $\alpha_{ij} = -\alpha_{ji}$, and satisfy $\alpha(u, v) = \alpha_{ij}u^iv^j$. Zero-forms are simply scalars. For an n-dimensional M, the set of k-forms makes up a vector space of dimension $\binom{n}{k}$, so that all n-forms are scalar multiples of one another and therefore proportional to a reference n-form, termed the 'volume-form' and denoted by ω. Differential forms may be naturally integrated, along

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

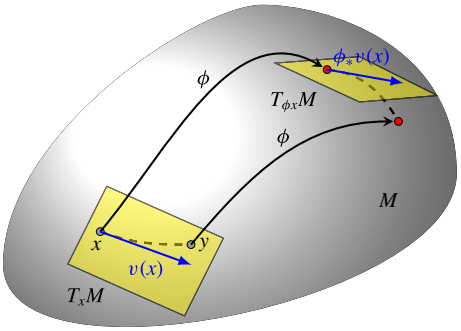


FIGURE 2. (Colour online) Push-forward of a vector $v(x)$ at x: a curve $\psi_s x$ tangent to $v(x)$ at x (dashed line) which includes a point y is mapped to a curve passing through ϕx and ϕy. The tangent of the latter curve at ϕx defines the push-forward $\phi_* v(x)$.

curves for one-forms, surfaces for two-forms, etc. (e.g. Schutz 1980; Frankel 1997). The wedge (or exterior) product $\wedge$ combines a k-form α with an l-form β to produce a $(k+l)$-form $\alpha \wedge \beta$ and is defined as an anti-symmetrised tensor product.

The geometric tools used in this paper include the push-forward and pull-back of vectors and differential forms (and indeed any tensor) by diffeomorphisms such as the flow map. The definition of the push-forward ϕ_* by a map $\phi : M \to M$ recognises that the action of diffeomorphisms on points naturally induces an action on vectors thought of as infinitesimal displacements. The push-forward $\phi_* v(x) \in T_{\phi x}M$ of $v(x)$ is a vector at the image point ϕx defined as the tangent to the curve $\phi(\psi_s x)$, with $\psi_s x$ a curve tangent to v at x to satisfy (2.4); explicitly, we have

$$\phi_* v(x) = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \phi(\psi_s x); \tag{2.6}$$

see figure 2. The push-forward ϕ_*, often termed the differential of ϕ, is linear. It provides a way to map entire vector fields v according to

$$(\phi_* v)(\phi x) = \phi_* v(x), \tag{2.7}$$

and arises naturally in applications of the chain rule, for example in (3.3) below. In coordinates x^i, the push-forward takes the form

$$(\phi_* v)^i(\phi x) = v^j(x)\, \partial_j \phi^i(x). \tag{2.8}$$

A well-known example of the push-forward of a vector field arises in MHD with the observation that, for incompressible flow in the absence of diffusion, the magnetic field is a 'frozen' vector field, transported by the flow according to the Cauchy solution. The magnetic field b at any t is then the push-forward of the initial magnetic field b_0 by the volume-preserving flow map: $b = \phi_* b_0$.

Once the push-forward of a vector is defined, the definitions of the push-forward and pull-back of various tensors follow naturally. The pull-back ϕ^* of a vector $v(x) \in T_x M$ by ϕ is simply the push-forward under the inverse map ϕ^{-1}, that is,

$$\phi^* v(x) = (\phi^{-1})_* v(x) \in T_{\phi^{-1} x}M, \tag{2.9}$$

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

using the invertibility of the diffeomorphism. The pull-back of one-forms is defined using their duality with vectors: if $\nu(x) \in T^*_x M$ is a one-form, its pull-back $\phi^*\nu(x) \in T^*_{\phi^{-1}x} M$ satisfies

$$(\phi^*\nu)(v) = \nu(\phi_* v) \quad \text{for any } v \in T_{\phi^{-1}x} M. \tag{2.10}$$

With this definition, the coordinate expression of the pull-back of a one-form field $\nu = \nu_i \,\mathrm{d}x^i$ is readily obtained from (2.8) as

$$(\phi^*\nu)_i(\phi^{-1}x) = \nu_j(x)\partial_i\phi^j(\phi^{-1}x). \tag{2.11}$$

The pull-back of other differential forms is defined similarly via pairing with vectors; the push-forward is defined as the pull-back by the inverse map in a way that mirrors (2.9). Note that the pull-back and push-forward of scalar functions (zero-forms) are simple compositions: $\phi^* f = f \circ \phi$ and $\phi_* f = f \circ \phi^{-1}$.

Another important differential-geometric tool is the Lie derivative with respect to a vector field v. For any tensor field (vector or form) τ, the Lie derivative is defined by

$$\mathcal{L}_v \tau = \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \psi_s^* \tau, \tag{2.12}$$

where the map ψ_s is the flow of v in the sense that

$$\frac{\mathrm{d}}{\mathrm{d}s}\psi_s x = v(\psi_s x), \quad \psi_0 = \mathrm{id}. \tag{2.13}$$

An application to the fluid flow map ϕ and associated (time-dependent) vector field u gives the useful formula

$$\partial_t(\phi^*\tau) = \phi^*(\partial_t\tau + \mathcal{L}_u\tau), \tag{2.14}$$

holding when τ depends explicitly on t. Another useful property concerns the action of pull-back on Lie derivatives: for any vector field v,

$$\phi^*(\mathcal{L}_v\tau) = \mathcal{L}_{\phi^* v}(\phi^*\tau). \tag{2.15}$$

Both properties (2.14) and (2.15) are derived in § A.1, and in Soward & Roberts (2010, 2014) using coordinate methods. Coordinate expressions for Lie derivatives are obtained by writing the map ψ_s in (2.12) as $(\psi_s x)^i = x^i + s v^i(x) + o(s)$ and using the coordinate expression of the pull-back. This gives

$$\mathcal{L}_u f = v^j\partial_j f, \quad (\mathcal{L}_v u)^i = v^j\partial_j u^i - u^j\partial_j v^i, \quad (\mathcal{L}_v \nu)_i = v^j\partial_j\nu_i + \nu_j\partial_i v^j \tag{2.16a–c}$$

for the Lie derivatives of scalar fields f (zero-forms), vector fields v and one-form fields ν, respectively. An interpretation of the Lie derivative of a vector field is used repeatedly in what follows: if a map $\psi_{s,t}$ depends on the two parameters s and t, we can define the two s- and t-dependent vector fields

$$u = \frac{\mathrm{d}}{\mathrm{d}t}\psi_{s,t} \circ \psi_{s,t}^{-1} \quad \text{and} \quad v = \frac{\mathrm{d}}{\mathrm{d}s}\psi_{s,t} \circ \psi_{s,t}^{-1}, \tag{2.17a,b}$$

in the manner of (2.2). The Lie derivative of v with respect to u then satisfies

$$\mathcal{L}_u v = -\mathcal{L}_v u = \partial_s u - \partial_t v. \tag{2.18}$$

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

The final operation needed in what follows is the exterior derivative d which acts on a k-form field to produce a $(k+1)$-form field. It is defined by its action on a scalar field f, giving a one-form field $\mathrm{d}f$ with the property that $(\mathrm{d}f)(v) = \mathcal{L}_v f$ for any vector field v, by the property that $\mathrm{d}^2\alpha = 0$ for any form α, and by its action on the exterior (wedge) products of forms, $\mathrm{d}(\alpha \wedge \beta) = \mathrm{d}\alpha \wedge \beta + (-1)^k \alpha \wedge \mathrm{d}\beta$ where α is a k-form. For our purpose, an important property of the exterior derivative is that it commutes with pull-backs/push-forwards by diffeomorphisms and with Lie derivatives:

$$\phi_*(\mathrm{d}\alpha) = \mathrm{d}(\phi_*\alpha), \quad \phi^*(\mathrm{d}\alpha) = \mathrm{d}(\phi^*\alpha), \quad \mathrm{d}(\mathcal{L}_v\alpha) = \mathcal{L}_v(\mathrm{d}\alpha). \tag{2.19a–c}$$

Another is Cartan's formula

$$\mathcal{L}_v\alpha = \iota_v\,\mathrm{d}\alpha + \mathrm{d}(\iota_v\alpha), \tag{2.20}$$

relating the Lie derivative of a k-form α to its exterior derivative by means of the interior product $\iota_v\alpha$ defined as the $(k-1)$-form such that $(\iota_v\alpha)(v_1, \ldots, v_{k-1}) = \alpha(v, v_1, \ldots, v_{k-1})$ for all vectors $v_1, \ldots, v_{k-1}$.

2.2. *Fluid dynamics on a Riemannian manifold*

We now review the geometric formulation of fluid dynamics which naturally emerges from Arnold's (1966) interpretation of the motion of a perfect incompressible fluid on an arbitrary manifold as geodesic motion (see also Ebin & Marsden 1970; Arnold & Khesin 1998; Holm, Marsden & Ratiu 1998). This requires the manifold M to have a Riemannian structure, that is, to be equipped with a metric $g(\cdot, \cdot)$, a positive-definite bilinear form which, at each point $x \in M$, maps pairs of vectors to non-negative real numbers, thus generalising the scalar product of Euclidean space. The metric induces a preferred n-form, which we take as the reference volume form ω. This volume form defines the volume element on M and appears naturally in integrals over the domain such as $\int_M f\omega$ where f is a scalar function. A fluid motion is incompressible provided the associated flow map ϕ preserves volume in the sense that

$$\phi^*\omega = \omega. \tag{2.21}$$

This is the explicit form of $\phi \in \mathrm{SDiff}(M) \subset \mathrm{Diff}(M)$.

The metric g induces a distance $d(x, y)$ between points x and y on M. This is the length of the shortest path from x to y, which is defined by

$$d(x, y) = \inf_{\gamma_s:[0,S]\to M} \int_0^S [g(\gamma_s', \gamma_s')]^{1/2}\,\mathrm{d}s, \quad \gamma_0 = x, \gamma_S = y, \tag{2.22}$$

where γ_s denotes paths joining x and y, and the prime denotes an s-derivative. The distance here does not depend on the parameterisation of the path, including the choice of interval $[0, S]$. It is then a standard result that obtaining $d(x, y)$ by the above variational principle is equivalent to a more convenient formula for the distance squared,

$$d^2(x, y) = \inf_{\gamma_s:[0,1]\to M} \int_0^1 g(\gamma_s', \gamma_s')\,\mathrm{d}s, \quad \gamma_0 = x, \gamma_1 = y. \tag{2.23}$$

Here the parameterisation is no longer arbitrary, and the minimising path γ_s is characterised by a constant kinetic energy $g(\gamma_s', \gamma_s')/2$. We refer the reader to Arnold

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

& Khesin (1998) for more discussion; we write other types of distances using formulae of the type (2.23) in what follows without further comment.

The distance $d(x, y)$ in (2.22) or (2.23) between points in M does not play an immediate role in the dynamics of fluids on M. The key role is rather played by the distance between two points in SDiff(M), that is, between pairs of volume-preserving diffeomorphisms ϕ, ψ. This distance is defined by

$$D^2(\phi, \psi) = \inf_{\gamma_s:[0,1]\to \mathrm{SDiff}(M)} \int_0^1 \int_M g(\gamma_s', \gamma_s')\omega \,\mathrm{d}s, \quad \gamma_0 = \phi, \gamma_1 = \psi, \tag{2.24}$$

where γ_s now denotes paths joining ϕ to ψ in SDiff(M). Introducing the s-dependent vector field $q = \gamma_s' \circ \gamma_s^{-1}$ and using volume preservation, this can be rewritten as

$$D^2(\phi, \psi) = \inf_{\gamma_s:[0,1]\to \mathrm{SDiff}(M)} \int_0^1 \int_M g(q, q)\omega \,\mathrm{d}s, \quad \gamma_0 = \phi, \gamma_1 = \psi. \tag{2.25}$$

This makes explicit the right invariance of the metric associated with D^2, that is, the invariance of the integral over M in (2.24) with respect to the right translations $\gamma_s \mapsto \gamma_s \circ \chi$ for an arbitrary s-independent $\chi \in \mathrm{SDiff}(M)$. This invariance implies that $D(\phi, \psi) = D(\phi \circ \psi^{-1}, \mathrm{id})$. In physical terms, D^2 is twice the s-integrated kinetic energy for a flow that joins the two configurations ψ and ϕ of the fluid.

The geometric view of the Euler equations for an incompressible perfect fluid on M is that they arise as geodesic equations on SDiff(M) equipped with the distance (2.25). Specifically, the Euler equations are obtained by minimising the action

$$\mathscr{A}[\phi] = \frac{1}{2} \int_0^T \int_M g(u, u)\omega \,\mathrm{d}t \tag{2.26}$$

over all time-dependent flow maps $\phi \in \mathrm{SDiff}(M)$ (with corresponding velocity $u = \dot{\phi} \circ \phi^{-1}$) which join two fixed flow maps between $t = 0$ and $t = T$. The motion which achieves the minimum is an incompressible flow with constant kinetic energy. When the Euler equations are derived in this fashion, the natural dynamical variable is not the velocity field $u = \dot{\phi} \circ \phi^{-1}$ but the momentum conjugate to it, that is, the one-form ν obtained from u using the metric g according to

$$\nu = g(u, \cdot) = u_\flat, \tag{2.27}$$

say, meaning that $\nu(v) = g(u, v)$ for all $v \in T_x M$. Equation (2.27) defines the so-called inertia operator which associates a one-form such as ν to a vector such as u; we say that ν is dual to u. In coordinates (2.27) amounts to the familiar lowering of an index: $\nu_i = g_{ij} u^j$. This observation motivates the musical notation $u_\flat$ (and its inverse $\nu_\sharp$). The one-form (2.27) dual to the velocity field will be called the 'momentum one-form'.

Minimising the action $\mathscr{A}$ in (2.26) leads to the Euler equations in the form

$$\partial_t \nu + \mathcal{L}_u \nu = -\mathrm{d}\pi, \quad \mathrm{div}\, u = 0, \tag{2.28a,b}$$

where π is a Lagrange multiplier enforcing volume preservation and $\mathrm{div}\, u$ is defined by $\mathcal{L}_u \omega = (\mathrm{div}\, u)\omega$ (see appendix B for a derivation, and Holm, Marsden & Ratiu (1998) for a general formulation applicable to a broad class of fluid models). In this derivation, ν appears as the variational derivative of the Lagrangian $\int_M g(u, u)\omega$

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

with respect to u, hence its interpretation as the conjugate momentum, noted above. Applying the metric, (2.28) can be rewritten in terms of the vector field $u = \nu_\sharp$ as

$$\partial_t u + \nabla_u u = -\nabla p, \quad \operatorname{div} u = 0, \tag{2.29}$$

where $p = \pi + g(u, u)/2$ is the usual pressure, $\nabla p = (\mathrm{d}p)_\sharp$ and ∇_u is the covariant derivative based on g (see (2.36) below for a coordinate expression). The derivation of (2.29) uses the identity

$$\mathcal{L}_u \nu = (\nabla_u u)_\flat + \tfrac{1}{2}\,\mathrm{d}g(u, u) \tag{2.30}$$

(e.g. Arnold & Khesin 1998, Theorem IV.1.17). While (2.29) may have the more familiar form, (2.28) turns out to be superior in the context of GLM theories. Its advantages are hinted at by the ease with which Kelvin's circulation theorem follows, as we now show. Using (2.14) and (2.19), (2.28) can be rewritten as

$$\partial_t(\phi^*\nu) = -\mathrm{d}(\phi^*\pi). \tag{2.31}$$

Now, let $C_0 \subset M$ be a closed curve in the initial fluid configuration and ϕC_0 its image as deformed under the flow map ϕ, then

$$\frac{\mathrm{d}}{\mathrm{d}t}\oint_{\phi C_0} \nu = \frac{\mathrm{d}}{\mathrm{d}t}\oint_{C_0} \phi^*\nu = \oint_{C_0} \partial_t(\phi^*\nu) = -\oint_{C_0} \mathrm{d}(\phi^*\pi) = 0. \tag{2.32}$$

This shows that the circulation around any material curve

$$\oint_{\phi C_0} \nu = \text{const.} \tag{2.33}$$

The vorticity is the two-form $\mathrm{d}\nu$. It is transported by the flow: taking the exterior derivative d of (2.28) and using that $\mathrm{d}^2 = 0$ gives the vorticity equation

$$\partial_t\,\mathrm{d}\nu + \mathcal{L}_u\,\mathrm{d}\nu = 0. \tag{2.34}$$

This means that vorticity is pushed forward by the flow map from the initial condition, $\mathrm{d}\nu = \phi_*\,\mathrm{d}\nu|_{t=0}$; see Besse & Frisch (2017) for discussion of this and applications to helicity.

For completeness, we record the coordinate version of (2.28) and (2.29). Using (2.16), (2.28) is written explicitly as

$$\partial_t \nu_i + u^j\partial_j\nu_i + \nu_j\partial_i u^j = -\partial_i\pi, \quad |g|^{-1/2}\partial_i(|g|^{1/2}u^i) = 0, \tag{2.35a,b}$$

where the expression for the divergence is obtained using that $\omega = |g|^{1/2}\,\mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \cdots \wedge \mathrm{d}x^n$, with $|g|$ the determinant of g_{ij}. Introducing $\nu_i = g_{ij}u^j$ into the first equation of (2.35) leads after a short computation to

$$\partial_t u^i + u^j\left(\partial_j u^i + \Gamma^i_{jk}u^k\right) = -g^{ij}\partial_j p, \tag{2.36}$$

where $\Gamma^i_{jk} = g^{il}(\partial_j g_{lk} + \partial_k g_{lj} - \partial_l g_{jk})/2$ is the Christoffel symbol (Hawking & Ellis 1973; Schutz 1980). This is the coordinate version of (2.29).

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

2.3. *Other fluid models*

To conclude this section we note that more complicated fluids than the perfect incompressible fluids considered so far are governed by equations that have a structure similar to that of the Euler equations. The three-dimensional ($n = 3$) Boussinesq equations are widely used to model rotating–stratified geophysical fluids. In coordinate-free notation, they take the form

$$\partial_t \nu_a + \mathcal{L}_u \nu_a = -\mathrm{d}\pi + \theta\, \mathrm{d}\Phi, \tag{2.37a}$$

$$\partial_t \theta + \mathcal{L}_u \theta = 0, \tag{2.37b}$$

$$\operatorname{div} u = 0, \tag{2.37c}$$

where θ is the buoyancy acceleration, Φ the geopotential divided by g, given by z in the Euclidean case, and ν_a is the absolute momentum one-form

$$\nu_a = \nu + A. \tag{2.38}$$

Here A is a one-form that encodes planetary rotation; on the f-plane, with f the constant Coriolis parameter and in Cartesian coordinates, it is given by

$$A = \tfrac{1}{2} f(x\,\mathrm{d}y - y\,\mathrm{d}x), \tag{2.39}$$

up to an irrelevant differential. Like the Euler equations, the Boussinesq equations have a variational formulation: (2.37) may be obtained by minimising the action

$$\mathscr{A}[\phi] = \int_0^T \int_M \left(\frac{1}{2} g(u, u) + A(u) - \theta \Phi \right) \omega\, \mathrm{d}t \tag{2.40}$$

over time-dependent flow maps ϕ and with $\theta = \phi_* \theta_0$ transported from its initial value (e.g. Holm, Marsden & Ratiu 1998).

Taking the exterior derivatives of (2.37*a*) and (2.37*b*) gives

$$(\partial_t + \mathcal{L}_u)\, \mathrm{d}\nu_a = \mathrm{d}\theta \wedge \mathrm{d}\Phi \quad \text{and} \quad (\partial_t + \mathcal{L}_u)\, \mathrm{d}\theta = 0. \tag{2.41a,b}$$

The two-form $\mathrm{d}\nu_a$ corresponds to the absolute vorticity. The conservation

$$(\partial_t + \mathcal{L}_u)\mathcal{Q} = 0 \tag{2.42}$$

of the potential-vorticity three-form

$$\mathcal{Q} = \mathrm{d}\nu_a \wedge \mathrm{d}\theta \tag{2.43}$$

follows immediately. The more familiar scalar Rossby–Ertel potential vorticity Q is defined by $Q\omega = \mathcal{Q}$ and is also conserved: $(\partial_t + \mathcal{L}_u)Q = 0$. We note that $\mathcal{Q}$ can be interpreted as the mass of Haynes & McIntyre's (1987) potential-vorticity substance and its impermeability property related to the alternative form $\mathcal{Q} = \mathrm{d}(\nu_a \wedge \mathrm{d}\theta)$ (see also Haynes & McIntyre 1990; Bretherton & Schär 1993).

Similarly, a perfect compressible fluid, with equation of state $p = P(\rho, \theta)$, where ρ and θ are the density and entropy, is governed by the equations

$$\partial_t \nu + \mathcal{L}_u \nu = -\rho^{-1} \mathrm{d}p + \tfrac{1}{2}\, \mathrm{d}g(u, u), \tag{2.44a}$$

$$\partial_t m + \mathcal{L}_u m = 0, \tag{2.44b}$$

$$\partial_t \theta + \mathcal{L}_u \theta = 0. \tag{2.44c}$$

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

Here the mass m is defined as the n-form $m = \rho\omega$, so that (2.44*b*) is equivalent to the continuity equation $\partial_t\rho + \mathrm{div}(\rho u) = 0$. The conservation

$$(\partial_t + \mathcal{L}_u)\,\mathrm{d}\nu \wedge \mathrm{d}\theta = 0 \tag{2.45}$$

of the potential-vorticity three-form $\mathrm{d}\nu \wedge \mathrm{d}\theta$ follows from observing that $\mathrm{d}p \wedge \mathrm{d}\rho \wedge \mathrm{d}\theta = 0$ by virtue of the dependence $p = p(\rho, \theta)$. In three dimensions, $n = 3$, the more usual scalar Rossby–Ertel potential vorticity Q is defined by the relation $Qm = \mathrm{d}\nu \wedge \mathrm{d}\theta$ (Schutz 1980). Its conservation (as a scalar) is an immediate consequence of the conservation of the two three-forms $\mathrm{d}\nu \wedge \mathrm{d}\theta$ and m.

Finally, we formulate the equations governing the dynamics of ideal MHD flows in a coordinate-free way, in $n = 3$ dimensions (e.g. Frankel 1997). We let b be the magnetic (vector) field, $\beta = g(b, \cdot) = b_\flat$ its dual one-form, and the flux two-form B be defined by the interior product $\iota_b\omega = \omega(b, \cdot, \cdot) = B$. The condition that the magnetic field be divergence free, $\mathrm{div}\, b = 0$, amounts to requiring that $\mathrm{d}B = 0$ (the two-form B is said to be closed). This is preserved under the induction equation, which is simply transport of B by the flow u,

$$\partial_t B + \mathcal{L}_u B = 0. \tag{2.46}$$

The corresponding equation for magnetic field is then

$$\partial_t b + \mathcal{L}_u b + (\mathrm{div}\, u)b = 0. \tag{2.47}$$

This extra term ensures that b remains divergence free and means that b is transported as a 'tensor of weight -1', as is density ρ (see below (2.44)). We will not use the machinery of tensor weights in this paper; see the parallel development in Soward & Roberts (2014) and note that if the flow is incompressible, $\mathrm{div}\, u = 0$ and b is transported as a vector field.

The momentum equation may be written, in general, as

$$\begin{aligned}\rho(\partial_t v + \mathcal{L}_u v) &= \iota_b\,\mathrm{d}\beta - \mathrm{d}p + \tfrac{1}{2}\rho\,\mathrm{d}g(u, u)\\ &= \mathcal{L}_b\beta - \mathrm{d}g(b, b) - \mathrm{d}p + \tfrac{1}{2}\rho\,\mathrm{d}g(u, u).\end{aligned} \tag{2.48}$$

The two versions of the Lorentz force term are linked by the Cartan formula, giving here $\mathcal{L}_b\beta = \iota_b\,\mathrm{d}\beta + \mathrm{d}(\iota_b\beta)$. Compressible MHD is governed then by (2.46) or (2.47), (2.48), conservation of mass (2.44*b*) and an equation of state such as $p = P(\rho)$. Note that the Lorentz force term involves both the magnetic (vector) field b and the corresponding one-form β as discussed by Holm (2002*b*) and Roberts & Soward (2006*a*).

3. Lagrangian-mean dynamics

3.1. *Generalising GLM to arbitrary manifolds*

GLM and other Lagrangian-mean theories can be thought of as extensions to SDiff(M) of the classical (Krylov–Bogoliubov–Mitropolsky) averaging of ordinary differential equations (e.g. Nayfeh 1973; Sanders & Verhulst 1985). The central idea is to seek a variable transformation that represents the state of the system as the sum of a mean position and a perturbation. Crucially the mean is defined, if possible, in such a way that the perturbation remains bounded and small for long times, so that the fluctuating states stay close to the mean state. The mean dynamics is of course affected by the

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

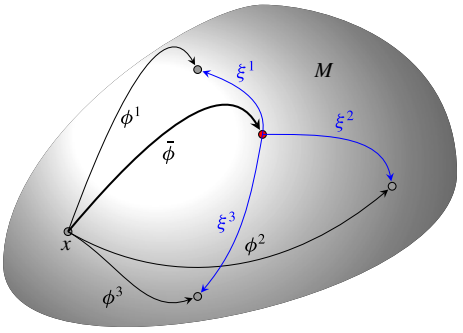


FIGURE 3. (Colour online) Decomposition $\phi^\alpha = \xi^\alpha \circ \bar{\phi}$ of the flow maps shown acting on a single point $x \in M$.

perturbation; as a result, the equations governing the mean dynamics contain terms representing the average effect of the perturbation. Often these terms depend only on a few features of the perturbation so there is no need to work out all the details of its evolution.

In a similar manner, in GLM one wishes to define a mean configuration of the fluid – a mean flow map $\bar{\phi}$ – that stays close to the exact fluctuating flow maps but evolves in a simpler manner because it filters out fluctuations. The theory applies to different types of averaging: the fast-time averaging that appears in asymptotic treatments of two-time-scale models, averaging along a particular direction, or a probabilistic average when fluctuations model uncertainty. For this reason it is customary since Andrews & McIntyre (1978*a*) to develop GLM theories for an arbitrary ensemble average, only stipulating the ensemble when applications are made specific. We follow this approach and consider an ensemble of flow maps ϕ^α, where the parameter α indexes the ensemble. It is useful to think of α taking a few integer values, with the ensemble consisting of a few different flows, or real values, so that the ensemble consists of a one-parameter family of flows. The theory is not restricted to these cases however and applies to the most general situation when the ensemble is determined by specifying a measure on SDiff(M).

It is natural to write each fluctuating flow map ϕ^α in the ensemble as the composition

$$\phi^\alpha = \xi^\alpha \circ \bar{\phi} \tag{3.1}$$

of a perturbation map ξ^α with a mean flow map $\bar{\phi}$. The maps ϕ^α then stay close to $\bar{\phi}$ when the maps ξ^α stay close to the identity. Expression (3.1) was used, for example, by Holm (2002*a*,*b*) and Marsden & Shkoller (2003), and naturally extends to arbitrary manifolds the sum $\boldsymbol{x} + \boldsymbol{\xi}(\boldsymbol{x}, t)$ of mean position $\boldsymbol{x}$ and perturbation $\boldsymbol{\xi}(\boldsymbol{x}, t)$ of standard GLM. It can be visualised either by thinking of the decomposition of $\phi^\alpha x$ for a particular point $x \in M$ (see figure 3), or by thinking of ϕ^α as a point on the infinite-dimensional manifold SDiff(M) (see figure 4). Note that we do not insist that $\phi^\alpha = \mathrm{id}$ at $t = 0$. The ensemble of maps can accommodate an ensemble of different initial conditions.

Of course the foregoing cannot uniquely define the mean map. Nor is the problem of defining it straightforward, since we cannot use linear averaging. The definition has to be geometrically intrinsic, thus coordinate independent, and preferably ensure that $\bar{\phi} \in \mathrm{SDiff}(M)$. We discuss possible definitions in § 4, and in § 5 examine their form

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

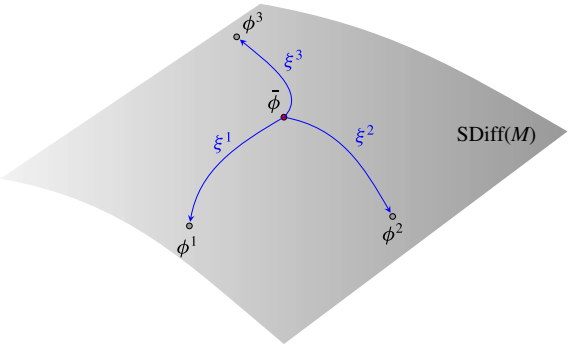


FIGURE 4. (Colour online) Decomposition $\phi^\alpha = \xi^\alpha \circ \bar{\phi}$ of the flow maps, with each map thought of as a point in the infinite-dimensional group of diffeomorphisms SDiff(M).

under the simplifying assumption of small-amplitude perturbations. In the remainder of the current section, we derive averaged dynamical equations that hold for arbitrary $\bar{\phi}$. This highlights properties that depend only on the geometric structure of the Euler and related equations and on the mean–perturbation decomposition (3.1), independently of any particular definition of $\bar{\phi}$.

3.2. *Geometric GLM*

We start by defining the mean velocity field $\bar{u}$ as the time derivative of the mean flow map:

$$\bar{u} = \dot{\bar{\phi}} \circ \bar{\phi}^{-1}. \tag{3.2}$$

We emphasise that this is not the Eulerian-mean velocity: $\bar{u} \neq \langle u^\alpha \rangle$. While the notation $\bar{u}$ is often employed for the Eulerian-mean velocity, we will reserve the explicit $\langle u^\alpha \rangle$ for the few instances when this velocity will be needed. While our choice of notation is unconventional, it is completely natural in view of the right-hand side of (3.2) and we trust that it will not lead to confusion. Our $\bar{u}$ is a Lagrangian-mean velocity in the sense that the flow map it generates takes fluid particles a to positions $\bar{\phi}a$ that are as close as possible, in some sense, to the ensemble of positions $\phi^\alpha a$. We nonetheless refrain from using the notation $\overline{u}^L$ or terming it 'Lagrangian-mean velocity' for reasons that will become apparent shortly.

Taking the time derivative of (3.1) and evaluating it at $x = \bar{\phi}a$ leads to

$$\dot{\xi}^\alpha x + (\xi^\alpha_* \bar{u})(\xi^\alpha x) = u^\alpha(\xi^\alpha x), \tag{3.3}$$

with $u^\alpha = \dot{\phi}^\alpha \circ (\phi^\alpha)^{-1}$ as in (2.3). Defining the ensemble of perturbation velocity fields

$$w^\alpha = \dot{\xi}^\alpha \circ (\xi^\alpha)^{-1}, \tag{3.4}$$

equation (3.3) can be written as an equality between vector fields,

$$w^\alpha + \xi^\alpha_* \bar{u} = u^\alpha. \tag{3.5}$$

Pulling back by ξ^α and averaging leads to the equality

$$\bar{u} = \xi^{\alpha *}(u^\alpha - w^\alpha) = \langle \xi^{\alpha *}(u^\alpha - w^\alpha) \rangle, \tag{3.6}$$

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

with the second equality following from $\langle\bar{u}\rangle = \bar{u}$. This can be thought of as a geometrically consistent version of the GLM equation $\bar{\boldsymbol{u}}^L(\boldsymbol{x}) = \langle \boldsymbol{u}(\boldsymbol{x} + \boldsymbol{\xi}(\boldsymbol{x}))\rangle$. On a general manifold M the latter equality would involve averaging vectors lying in different tangent planes and so is simply not defined. Similarly, (3.5) indicates that w^α and $\xi^{\alpha *} w^\alpha$ are geometrically consistent relatives of GLM's $\boldsymbol{u}^\ell(\boldsymbol{x}) = \boldsymbol{u}(\boldsymbol{x}+\boldsymbol{\xi}(\boldsymbol{x})) - \bar{\boldsymbol{u}}^L(\boldsymbol{x})$.

The main aim of Lagrangian-mean theories is the derivation of equations governing the mean flow that take as simple a form as possible. For a perfect incompressible fluid, this is best achieved starting with the Euler equation (2.28) in terms of the one-form ν^α for each member α of the ensemble, pulling back with ξ^α then averaging. For stratified Boussinesq fluids, compressible fluids and MHD flows, a similar procedure can be applied to (2.37), (2.44) and (2.46), (2.48), respectively. A simple property, derived in Soward & Roberts (2010, 2014) and in § A.2, is key to this procedure: for any ensemble- and time-dependent tensor τ^α,

$$\xi^{\alpha *}(\partial_t + \mathcal{L}_{u^\alpha})\tau^\alpha = (\partial_t + \mathcal{L}_{\bar{u}})(\xi^{\alpha *}\tau^\alpha). \tag{3.7}$$

Note how the Lie derivative on the right-hand side is taken with respect to the mean velocity $\bar{u}$ which, however the mean flow map is chosen, is fluctuation free, i.e. the same for each realisation α. Applying this to the Euler equation (2.28), using (2.19) and averaging yields

$$(\partial_t + \mathcal{L}_{\bar{u}})\langle\xi^{\alpha *}\nu^\alpha\rangle = -\mathrm{d}\langle\xi^{\alpha *}\pi^\alpha\rangle. \tag{3.8}$$

This suggests defining the Lagrangian mean of any tensorial object τ^α as

$$\overline{\tau}^L = \langle\xi^{\alpha *}\tau^\alpha\rangle. \tag{3.9}$$

We emphasise the following: with this definition, the mean velocity $\bar{u}$ is not the Lagrangian mean of the velocity in the sense of the Lagrangian-mean operator (3.9),

$$\bar{u} \neq \overline{u}^L. \tag{3.10}$$

For the equality to hold, that is to say for $\bar{u} = \langle\xi^{\alpha *}u^\alpha\rangle$ to hold, we would have to impose that $\overline{w}^L = \langle\xi^{\alpha *}w\rangle = 0$ (see (3.6)). However, as detailed in § 4, while this apparently provides a possible definition for the mean flow map, it is unacceptable because it leads to a mean flow that drifts away from the fluctuating flows, even if the latter are wavelike and form a cluster in SDiff(M). Note that standard GLM manages to maintain the equality in (3.10) but only by sidestepping intrinsic geometry (by transforming vector fields as $\boldsymbol{u}(\boldsymbol{x}+\boldsymbol{\xi}(\boldsymbol{x}))$, i.e. using parallel transport, rather than following proper pull-back transformation rules involving $\boldsymbol{u}(\boldsymbol{x}+\boldsymbol{\xi}(\boldsymbol{x}))\cdot(\boldsymbol{\nabla}(\boldsymbol{x}+\boldsymbol{\xi}(\boldsymbol{x})))^{-1}$, with the superscript -1 denoting matrix inverse, as appropriate even in Euclidean spaces; see Soward & Roberts (2010)).

With the definition (3.9), $\overline{\nu}^L = \langle\xi^{\alpha *}\nu^\alpha\rangle$ and (3.8) takes the simple form

$$\partial_t\overline{\nu}^L + \mathcal{L}_{\bar{u}}\overline{\nu}^L = -\mathrm{d}\overline{\pi}^L. \tag{3.11}$$

This is a central result for any GLM-like theory: irrespective of the specific definition of the mean flow map $\bar{\phi}$, the Lagrangian-mean momentum, defined as a one-form, satisfies the averaged Euler equations (3.11). An Euclidean-space version was obtained by Soward & Roberts (2010); the manipulations of Andrews & McIntyre (1978*a*), following Soward (1972) and leading to their Theorem I, can similarly be interpreted

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

as the pull-back of the momentum equation leading to (3.11). Note that (3.11) can be derived from an averaged variational principle, in a manner that generalises to arbitrary manifolds the derivation of Salmon (2013). This is described in appendix C.

A consequence of (3.11) is that the Lagrangian-mean Kelvin circulation is conserved,

$$\oint_{\bar{\phi}C_0} \overline{v}^L = \text{const.}, \tag{3.12}$$

when the contour of integration is transported by the mean velocity $\bar{u}$. Equivalently, the Lagrangian-mean vorticity

$$\mathrm{d}\overline{v}^L = \overline{\mathrm{d}v}^L \tag{3.13}$$

is conserved as a two-form:

$$\partial_t \,\mathrm{d}\overline{v}^L + \mathcal{L}_{\bar{u}}\,\mathrm{d}\overline{v}^L = 0. \tag{3.14}$$

Naturally, these also follow by applying $\xi^{\alpha *}$ to (2.33) and (2.34), and averaging. Key to these calculations are the properties (2.19) that the operator d commutes with pull-backs and Lie derivatives.

The physics of the wave–mean-flow interactions is encoded in the relation between the mean velocity $\bar{u}$ and Lagrangian-mean momentum $\overline{v}^L$, that is, between the advecting velocity and advected momentum in (3.11). Unlike in the original Euler equations, these are not related through the inertia operator $\flat$ associated with the metric g. Rather, they are related in a more complicated way, which depends on the precise definition of the mean flow map and on the transport of the metric by the ξ^α, arising when the pull-back is applied to v^α as defined in (2.27).

It is natural to consider the difference

$$-p = \overline{v}^L - (\bar{u})_\flat \tag{3.15}$$

between the Lagrangian-mean momentum and the one-form dual to the mean velocity vector since this characterises the manner in which waves or perturbations modify the mean Kelvin circulation, hence the mean vortical dynamics, compared with that of the Euler equations. The one-form p defined in (3.15) generalises the pseudomomentum $\boldsymbol{p}$ of Andrews & McIntyre (1978*a*), and we will use the same terminology. (Note that the sign of p is chosen according to a GLM standard convention ensuring that p reduces to the familiar product of wave action with wavevector in a small-amplitude, Wentzel–Kramers–Brillouin (WKB) regime; see Andrews & McIntyre 1978*a*). Parameterising the effect of perturbations on the mean flow can then be interpreted as providing an expression for p, for example by means of a closure in terms of $\bar{u}$.

Expressions for p making it clear that it is quadratic to leading order in the perturbation amplitude ε, say, (and is thus an $O(\varepsilon^2)$ 'wave property') are readily found from (3.15). Defining the perturbation velocity $\tilde{w}^\alpha$ alternative to w^α by

$$\tilde{w}^\alpha = u^\alpha - \bar{u} = w^\alpha + (\xi^\alpha_* - \mathrm{id})\bar{u}, \tag{3.16}$$

we rewrite (3.15) as

$$-p = \langle (\xi^{\alpha *} - \mathrm{id})\bar{u}_\flat \rangle + \langle \xi^{\alpha *} \tilde{w}^\alpha_\flat \rangle \tag{3.17a}$$

$$= \langle \xi^{\alpha *} (\xi^\alpha_* \bar{u})_\flat - \bar{u}_\flat \rangle + \langle \xi^{\alpha *} w^\alpha_\flat \rangle. \tag{3.17b}$$

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

The right-hand sides are quadratic since, to leading order, ξ^α is the identity at $O(1)$ and the velocities $\tilde{w}^\alpha$ and w^α have vanishing averages at $O(\varepsilon)$. The second term in each of these equivalent expressions is reminiscent of the standard GLM definition of pseudomomentum. Equation (3.17*b*) can also be written as

$$-p = \overline{g}^L(\bar{u}, \cdot) - \bar{u}_\flat + \langle \xi^{\alpha *} w^\alpha_\flat \rangle, \tag{3.17c}$$

where $\overline{g}^L = \langle \xi^{\alpha *} g \rangle$, consistent with our general definition (3.9) of Lagrangian averaging. The first two terms represent the effect on the mean flow of the wave-induced change in the average geometry; the third term represents the effect of the wave velocities, as in standard GLM.

3.3. *Other fluid models*

We now sketch the construction of mean equations for the more complex fluid models of § 2.3. The Lagrangian-mean version of the three-dimensional Boussinesq equations (2.37) is obtained by applying $\xi^{\alpha *}$ to (2.37*a*)–(2.37*b*) to obtain the momentum and buoyancy equations

$$\partial_t \overline{v}^L_a + \mathcal{L}_{\bar{u}} \overline{v}^L_a = -\mathrm{d}\overline{\pi}^L + \overline{\theta \, \mathrm{d}\Phi}^L, \tag{3.18a}$$

$$\partial_t \overline{\theta}^L + \mathcal{L}_{\bar{u}} \overline{\theta}^L = 0, \tag{3.18b}$$

and to (2.42) to obtain the conservation of the Lagrangian-mean potential vorticity

$$(\partial_t + \mathcal{L}_{\bar{u}}) \, \overline{\mathrm{d}v_a \wedge \mathrm{d}\theta}^L = 0. \tag{3.19}$$

The last term in (3.18*a*), explicitly $\langle (\xi^{\alpha *} \theta^\alpha) \, \mathrm{d}(\xi^{\alpha *} \Phi^\alpha) \rangle$, is inconvenient but can be simplified by assuming that the buoyancy θ^α in each flow realisation α results from the transport of the same initial buoyancy distribution, θ_0 say. With this assumption,

$$\xi^{\alpha *} \theta^\alpha = \bar{\phi}_* \theta_0 = \overline{\theta}^L \tag{3.20}$$

is independent of α and hence equal to its average. The momentum and potential-vorticity equations then simplify as

$$\partial_t \overline{v}^L_a + \mathcal{L}_{\bar{u}} \overline{v}^L_a = -\mathrm{d}\overline{\pi}^L + \overline{\theta}^L \, \mathrm{d}\overline{\Phi}^L \tag{3.21}$$

and

$$(\partial_t + \mathcal{L}_{\bar{u}}) \overline{Q}^L = 0, \quad \text{where } \overline{Q}^L = \mathrm{d}\overline{v}^L_a \wedge \mathrm{d}\overline{\theta}^L. \tag{3.22}$$

The Lagrangian-mean Rossby–Ertel potential vorticity $\overline{Q}^L = \langle \xi^{\alpha *} Q^\alpha \rangle$ is also transported by the flow. When $\bar{u}$ is non-divergent (i.e. $\bar{\phi}$ preserves volume), it is related straightforwardly to its three-form counterpart by $\overline{Q}^L \omega = \overline{\mathcal{Q}}^L$. When $\bar{u}$ is divergent, this relationship is more complicated: introducing a mean density $\bar{\rho}$ such that $\bar{\phi}_* \omega = \bar{\rho} \omega$, it reads as $\overline{Q}^L \bar{\rho} \omega = \overline{\mathcal{Q}}^L$, noting that $\xi^{\alpha *} \omega = \bar{\rho} \omega$ is independent of α. The potential-vorticity equation is most useful in the limit of rapid rotation and strong stratification, when the mean flow can be assumed to be near geostrophic and in hydrostatic balance, and entirely slaved to the potential vorticity. We discuss this further in § 6.

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

For a compressible fluid governed by (2.44*a*), applying $\xi^{\alpha *}$ and averaging leads to

$$\partial_t \overline{v}^L + \mathcal{L}_{\bar{u}} \overline{v}^L = -\overline{\rho^{-1}\,\mathrm{d}p}^L + \tfrac{1}{2}\,\mathrm{d}\overline{g(u,u)}^L, \tag{3.23a}$$

$$\partial_t \overline{m}^L + \mathcal{L}_{\bar{u}} \overline{m}^L = 0, \tag{3.23b}$$

$$\partial_t \overline{\theta}^L + \mathcal{L}_{\bar{u}} \overline{\theta}^L = 0, \tag{3.23c}$$

on using (3.7). The momentum equation (3.23*a*) is complicated by a right-hand side that is not immediately expressed in terms of the Lagrangian mean of p^α, ρ^α and ν^α. Furthermore, averaging the equation of state gives $\overline{p}^L = \langle P(\xi^{\alpha *}\rho^\alpha, \xi^{\alpha *}\theta^\alpha)\rangle$, with a right-hand side that is, again, not expressed in terms of Lagrangian means. However, the conservation of the potential vorticity $\mathrm{d}\nu^\alpha \wedge \mathrm{d}\theta^\alpha$ in (2.45) implies that the Lagrangian-mean potential vorticity

$$(\partial_t + \mathcal{L}_{\bar{u}})\,\overline{\mathrm{d}\nu \wedge \mathrm{d}\theta}^L = 0 \tag{3.24}$$

is conserved. A conserved scalar in three dimensions is obtained as the quotient of the conserved three-forms $\overline{\mathrm{d}\nu \wedge \mathrm{d}\theta}^L$ and $\overline{m}^L$.

As in the Boussinesq case, simplifications arise if we assume that the mass m^α and entropy θ^α in each of the flow realisations α are obtained by transporting single initial distributions, m_0 and θ_0, say. In this case,

$$\xi^{\alpha *} m^\alpha = \bar{\phi}_* m_0 = \overline{m}^L \quad \text{and} \quad \xi^{\alpha *}\theta^\alpha = \bar{\phi}_*\theta_0 = \overline{\theta}^L \tag{3.25a,b}$$

are both independent of α and thus equal to their average. (The first assumption is also made by Bhat *et al.* (2005) in their Lagrangian averaging of compressible flows.) This simplifies the averaged equation of state into $\overline{p}^L = \langle P(\xi^{\alpha *}\rho^\alpha, \overline{\theta}^L)\rangle$ and, more importantly, the form of the conserved Lagrangian potential-vorticity volume form,

$$\overline{\mathrm{d}\nu \wedge \mathrm{d}\theta}^L = \mathrm{d}\overline{\nu}^L \wedge \mathrm{d}\overline{\theta}^L, \tag{3.26}$$

and of its scalar equivalent.

The averaging of the MHD equations parallels that of the Euler equations; see Holm (2002*b*) for GLM and related developments. Applying $\xi^{\alpha *}$ to the induction equation (2.46) and averaging gives

$$(\partial_t + \mathcal{L}_{\bar{u}})\overline{B}^L = 0. \tag{3.27}$$

The mean magnetic field $\bar{b}$ defined by $\iota_{\bar{b}}\omega = \overline{B}^L$ remains divergence free since $\mathrm{d}\overline{B}^L = 0$, even though the mean velocity $\bar{u}$ may be divergent (this is equivalent to the pull-back of b under ξ^α with tensor weight -1 in Soward & Roberts (2014)). To make some progress with the momentum equation we again assume that the mass distribution and magnetic flux are transported in each realisation from the same initial mass m_0 and flux B_0, so that

$$\xi^{\alpha *} m^\alpha = \bar{\phi}_* m_0 = \overline{m}^L \quad \text{and} \quad \xi^{\alpha *} B^\alpha = \bar{\phi}_* B_0 = \overline{B}^L, \tag{3.28a,b}$$

independent of α. With $\bar{\rho}\omega = \overline{m}^L$ defining a mean density $\bar{\rho}$, we find that $\iota_{\bar{b}/\bar{\rho}}\,\overline{m}^L = \overline{B}^L$; since $\overline{m}^L$ and $\overline{B}^L$ are both transported by the flow, so is $\bar{b}/\bar{\rho}$ (as a vector field). We note that

$$\xi^{\alpha *}(b/\rho) = \overline{b/\rho}^L = \bar{b}/\bar{\rho} \tag{3.29}$$

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

follows from pulling back $\iota_{b^\alpha/\rho^\alpha} m^\alpha = B^\alpha$ by ξ^α to obtain

$$\iota_{\xi^{\alpha*}(b^\alpha/\rho^\alpha)} \xi^{\alpha*} m^\alpha = \iota_{\xi^{\alpha*}(b^\alpha/\rho^\alpha)} \overline{m}^L = \xi^{\alpha*} B^\alpha = \overline{B}^L = \iota_{\bar{b}/\bar{\rho}} \overline{m}^L \tag{3.30}$$

and averaging.

Applying $\xi^{\alpha*}$ to (2.48) written as

$$(\partial_t + \mathcal{L}_{u^\alpha}) v^\alpha = \iota_{b^\alpha/\rho^\alpha} \,\mathrm{d}\beta^\alpha - (\rho^\alpha)^{-1} \,\mathrm{d}p^\alpha + \tfrac{1}{2} \,\mathrm{d}g(u^\alpha, u^\alpha), \tag{3.31}$$

averaging and using (3.29) yields

$$(\partial_t + \mathcal{L}_{\bar{u}}) \overline{v}^L = \iota_{\bar{b}/\bar{\rho}} \,\mathrm{d}\overline{\beta}^L - \overline{(\rho^\alpha)^{-1} \,\mathrm{d}p^\alpha}^L + \tfrac{1}{2} \,\mathrm{d}\overline{g(u^\alpha, u^\alpha)}^L. \tag{3.32}$$

It is then natural to define the one-form field h by, say,

$$-h = \overline{\beta}^L - (\bar{b})_\flat, \tag{3.33}$$

which, like the pseudomomentum (3.15), encodes the effect of the waves on the mean field in (3.32) and measures the lack of commutation between averaging and application of the inertia operator, in this instance to the magnetic field. This field is termed (minus) the 'magnetisation induced by the Lagrangian mapping' by Holm (2002*b*), but could also be called the 'pseudomagnetising force' (cf. Roberts & Soward 2006*a*) or perhaps just the 'pseudofield' for short.

Finally suppose that the original MHD flow is incompressible, the density ρ is a constant, and all the maps ξ^α are volume preserving (which excludes standard GLM). Then $\bar{\rho}$ also remains constant throughout and, taking this as unity, we can write

$$(\partial_t + \mathcal{L}_{\bar{u}}) \bar{b} = 0, \tag{3.34a}$$

$$(\partial_t + \mathcal{L}_{\bar{u}}) \overline{v}^L = \iota_{\bar{b}} \,\mathrm{d}\overline{\beta}^L - \mathrm{d}\overline{\pi}^L = \mathcal{L}_{\bar{b}} \,\overline{\beta}^L - \mathrm{d}(\overline{\pi}^L + \iota_{\bar{b}} \overline{\beta}^L), \tag{3.34b}$$

$$\operatorname{div} \bar{u} = 0 \quad \text{and} \quad \operatorname{div} \bar{b} = 0, \tag{3.34c,d}$$

without explicitly appealing to the two-form B.

4. Definitions of the mean flow

4.1. *Preliminaries*

In the previous section, we derived a generic form for the average of the Euler equations and their relatives and showed how key features, central to which is the simplicity and natural averaging of Kelvin's circulation when expressed in terms of $\overline{v}^L$, hold regardless of the specific definition of the mean flow map. In this section we consider possible definitions of the mean flow map $\bar{\phi}$ from which the mean velocity $\bar{u}$ is deduced.

As already discussed, there are two guiding principles in the choice of this definition. First, the mean flow map must be defined, if possible, in a way that ensures that the image of particles $\bar{\phi}a$ remains close to the images $\phi^\alpha a$ in each flow realisation, as one expects from the intuitive notion of mean. We expect this to be possible for wave-like, as distinct from turbulent, flows. 'Remains close' can be made precise in the perturbative context: if the ϕ^α are close together, in the sense that the maps $(\phi^\alpha)^{-1} \circ \phi^\beta$ are close to the identity map for all α and β, then $(\phi^\alpha)^{-1} \circ \bar{\phi}$ should be close to the identity for long times (say times $O(\varepsilon^{-1})$ or longer if the

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

maps are ε-close). Second, the definition of the mean should be geometrically sound in the sense that it does not depend on a particular coordinate choice and relies on the same geometric structure as the Euler equations themselves.

We remark that the first principle rules out simple definitions based on constraining the velocity perturbations w^α in (3.4), $\xi^{\alpha *}w^\alpha$ or $\tilde{w}^\alpha$ in (3.16) to have zero average. The constraints such definitions impose can ensure that the mean velocity is close to the ensemble of velocities but they do not prevent the drift of the mean trajectory away from the ensemble of trajectories over long times.

To illustrate this point, we consider a very simple set-up in which the velocities of the ensemble of flows consist of phase-shifted versions of a single, time-periodic, zero-mean velocity U: $u^\alpha(x,t) = \varepsilon U(x, t-\alpha)$, where $\varepsilon \ll 1$ and $\alpha \in \mathbb{R}$. Then, writing the perturbation map in coordinates as the expansion

$$\xi^{\alpha i}(x) = x^i + \varepsilon \xi_1^{\alpha i}(x) + \varepsilon^2 \xi_2^{\alpha i}(x) + \cdots, \tag{4.1}$$

we examine the implications for $\xi_2^{\alpha i}$ of constraining w^α or $\tilde{w}^\alpha$. The perturbation velocity w^α associated with (4.1) has the form

$$w^{\alpha i} = \varepsilon \dot{\xi}_1^{\alpha i} + \varepsilon^2 (\dot{\xi}_2^{\alpha i} - \xi_1^{\alpha j} \partial_j \dot{\xi}_1^{\alpha i}) + \cdots, \tag{4.2}$$

where $\dot{\xi}_1^{\alpha i} = U^{\alpha i}$. Imposing that $\langle w^\alpha \rangle = 0$ then gives

$$\langle \dot{\xi}_2^{\alpha i} \rangle = \langle \xi_1^{\alpha j} \partial_j U^{\alpha i} \rangle, \tag{4.3}$$

where the right-hand side, which can be recognised as a Stokes drift, does not vanish in general. This indicates that $\langle \xi_2^{\alpha i} \rangle$ grows secularly and hence that the condition $\langle w^\alpha \rangle = 0$ does not lead to a satisfactory definition of the mean. To the order considered, it is equivalent to the use of the Eulerian mean which corresponds to $\langle \tilde{w}^\alpha \rangle = 0$, hence $\bar{u} = \langle U \rangle = 0$ (see (3.16)), and whose deficiencies are well known.

An alternative is to impose that $\langle \xi^{\alpha *} w^\alpha \rangle = 0$, leading, as mentioned in § 3.2, to the attractive feature that $\bar{u} = \overline{u}^L$ in the sense of (3.9). However, this too is unsatisfactory, for the same reason. Indeed, a short calculation shows that

$$(\xi^{\alpha *} w^\alpha)^i = \varepsilon \dot{\xi}_1^{\alpha i} + \varepsilon^2 (\dot{\xi}_2^{\alpha i} - U^{\alpha j} \partial_j \xi_1^{\alpha i}) + \cdots, \tag{4.4}$$

so that

$$\langle \dot{\xi}_2^{\alpha i} \rangle = \langle U^{\alpha j} \partial_j \xi_1^{\alpha i} \rangle, \tag{4.5}$$

which again indicates secular growth.

These failures show that the definition of the mean flow map must be based on constraining particle positions rather than particle velocities. For this purpose we freeze the evolution of the flows and consider the ensemble of fluctuating flow maps ϕ^α at a fixed time t. We want to define a mean flow map $\bar{\phi}$ that in some sense is as close as possible to the maps ϕ^α. We consider several possibilities. Two are based on a distance between diffeomorphisms that is an alternative to the distance (2.25) used to derive the Euler equations, namely

$$\hat{D}^2(\phi, \psi) = \int_M d^2(\phi, \psi)\omega = \inf_{\gamma_s : [0,1] \to \mathrm{Diff}(M)} \int_0^1 \int_M g(q,q) \gamma_{s*} \omega \, \mathrm{d}s, \quad \gamma_0 = \phi, \gamma_1 = \psi, \tag{4.6}$$

where $\gamma_s : [0,1] \to \mathrm{Diff}(M)$ denotes an s-dependent diffeomorphism and $q = \gamma_s' \circ \gamma_s^{-1}$ is the corresponding velocity field. Note that this distance differs from (2.25) in that

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

it relaxes the assumption of volume preservation imposed on the maps γ_s (even when the endpoint maps ϕ and ψ are volume preserving).

We now discuss four possible definitions of the mean flow map $\bar{\phi}$ which we term 'extended GLM', 'optimal transport', 'geodesic' and, following Soward & Roberts (2010), 'glm'. We keep the discussion at a formal level, ignoring for the most part the technical issues associated with the existence, smoothness and invertibility of $\bar{\phi}$ and only point out the most striking difficulties of some definitions in § 4.6.

4.2. *Extended GLM*

We define the mean flow as

$$\bar{\phi} = \underset{\phi \in \mathrm{Diff}(M)}{\arg\min} \langle \hat{D}^2(\phi, \phi^\alpha) \rangle, \tag{4.7}$$

where arg min denotes the minimising map. This definition is coordinate independent and can be recognised as the standard definition of the centre of mass on a Riemannian manifold, also termed the Fréchet mean or Karcher mean (e.g. Berger 2003, Pennec 2006; see appendix D). It will turn out – see below (4.13) – that this choice delivers a mean velocity field $\bar{u}$ that coincides with GLM's $\overline{\boldsymbol{u}}^L$ when M is Euclidean.

The minimiser $\bar{\phi}$ is obtained by expanding the distance in (4.7) using (4.6) and seeking minimising paths γ_s^α joining $\bar{\phi}$ to ϕ^α when s runs from 0 to 1. This is best done using the associated s-dependent velocity fields

$$q^\alpha = (\gamma_s^\alpha)' \circ (\gamma_s^\alpha)^{-1}, \tag{4.8}$$

where we recall that the prime denotes differentiation with respect to s. We stress that s is a fictitious time variable: when measuring the distance between the two t-dependent diffeomorphisms $\bar{\phi}$ and ϕ^α, the real time variable t is frozen and treated as a fixed parameter for the paths γ_s^α, $0 \leqslant s \leqslant 1$. The vector field q^α associated with these paths is similarly unrelated to the velocities $\bar{u}$ and u^α; it depends on both s and t, though we do not make this dependence explicit. The perturbation maps ξ^α are simply given by γ_s^α at the endpoint $s = 1$: since $\phi^\alpha = \gamma_1^\alpha \circ \bar{\phi}$,

$$\xi^\alpha = \phi^\alpha \circ \bar{\phi}^{-1} = \gamma_1^\alpha. \tag{4.9}$$

A computation sketched in appendix E shows that the one-forms $q_\flat^\alpha = g(q^\alpha, \cdot)$ satisfy the equation of a pressureless fluid in fictitious time, that is,

$$\partial_s q_\flat^\alpha + \mathcal{L}_{q^\alpha} q_\flat^\alpha = \tfrac{1}{2}\, \mathrm{d} g(q^\alpha, q^\alpha), \tag{4.10}$$

or equivalently,

$$\partial_s q^\alpha + \nabla_{q^\alpha} q^\alpha = 0. \tag{4.11}$$

At the endpoints $s = 0$ and $s = 1$, the paths γ_s^α are constrained to satisfy $\gamma_0^\alpha = \bar{\phi}$ and $\gamma_1^\alpha = \phi^\alpha$, respectively. Minimisation over the α-independent but otherwise arbitrary $\bar{\phi}$ imposes

$$\langle q_\flat^\alpha \rangle = 0 \quad \text{for } s = 0, \tag{4.12}$$

or, equivalently,

$$\langle q^\alpha \rangle = 0 \quad \text{for } s = 0, \tag{4.13}$$

since the inertia operator mapping $q_\flat^\alpha|_{s=0}$ to $q^\alpha|_{s=0}$ is linear and α-independent.

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

The constraint (4.13) determines the mean flow map implicitly. It can be made more transparent by noting that, for each $x \in M$ and each ensemble member α, $\gamma_s^\alpha x$ is the geodesic on M starting from $\bar{\phi}x$ at $s = 0$ with speed $q^\alpha|_{s=0}$ to end at $\phi^\alpha x$ at $s = 1$. In the Euclidean case, the geodesics are straight lines, corresponding to s-independent vector fields q^α which can be identified with the displacement vectors $\boldsymbol{\xi}$ of standard GLM. Thus our construction directly generalises GLM's $\boldsymbol{x}^\xi = \boldsymbol{x} + \boldsymbol{\xi}(\boldsymbol{x})$, with the constraint (4.13) generalising $\langle \boldsymbol{\xi} \rangle = 0$. It follows that, in this Euclidean case, the GLM velocity $\overline{\boldsymbol{u}}^L$ corresponds to our mean velocity $\bar{u}$. Note again however that GLM's $\overline{\boldsymbol{u}}^L$, our $\bar{u}$, is not a Lagrangian mean in the sense of (3.9), for the reasons mentioned below (3.10). On curved manifolds, $\bar{u}$ is not obviously related to what may seem like a natural extension of (1.2), namely the velocity obtained by parallel-transporting (using the Levi-Civita connection associated with the metric g, see Frankel (1997, Ch. 9)) the velocities u^α along the geodesics $\gamma_s^\alpha x$ back to the mean position x, then averaging.

Our generalisation of GLM to arbitrary manifolds M respects their geometry in the sense that $\bar{\phi}$ maps M to M and it thus remedies the problem of mean trajectories leaving M encountered in standard GLM. Because there is no constraint that $\bar{\phi}$ be volume preserving, however, the mean velocity $\bar{u}$ is divergent in general (McIntyre 1988), an unavoidable price for the convenience of (4.13). This can be addressed by including a volume-preservation constraint as we now describe.

4.3. *Optimal transport*

An obvious generalisation of the above imposes volume preservation in the definition of $\bar{\phi}$ as a minimiser:

$$\bar{\phi} = \underset{\phi \in \mathrm{SDiff}(M)}{\arg\min} \langle \hat{D}^2(\phi, \phi^\alpha) \rangle. \tag{4.14}$$

This definition relates $\bar{\phi}$ to the problem of optimal transportation (e.g. Brenier 2003; Villani 2003). Indeed, using (4.6), the definition can be rewritten more explicitly as

$$\bar{\phi} = \underset{\phi \in \mathrm{SDiff}(M)}{\arg\min} \left\langle \int_M d^2(\phi, \phi^\alpha)\omega \right\rangle. \tag{4.15}$$

Ignoring smoothness issues (which lead to a distinction between diffeomorphisms and more general rearrangements), we see that $\bar{\phi}$ is chosen among volume-preserving maps to ensure the minimum total cost for the transport from the mean fluid configuration to the perturbed configurations reached by the maps $\xi^\alpha = \phi^\alpha \circ \bar{\phi}^{-1}$.

Carrying out the minimisation in (4.14) leads, as before, to equation (4.10) for a pressureless fluid. The constraint on $\bar{\phi}$ leads to the different end condition

$$\langle q_b^\alpha \rangle = \mathrm{d}\psi \quad \text{for } s = 0, \tag{4.16}$$

where the function $\psi : M \to \mathbb{R}$ is a Lagrange multiplier enforcing the volume preservation of $\bar{\phi}$ (see appendix E). The corresponding condition on the mean of the vector fields q^α now reads

$$\langle q^\alpha \rangle = \nabla\psi \quad \text{for } s = 0. \tag{4.17}$$

Equation (4.17) together with the requirement of geodesic paths is obtained by McCann (2001) in his treatment of optimal transport on a manifold, here with additional averaging and without the technical condition of ψ being $d^2/2$-convex (which is necessary to ensure that $\bar{\phi}$ is a minimiser and not just a critical point).

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

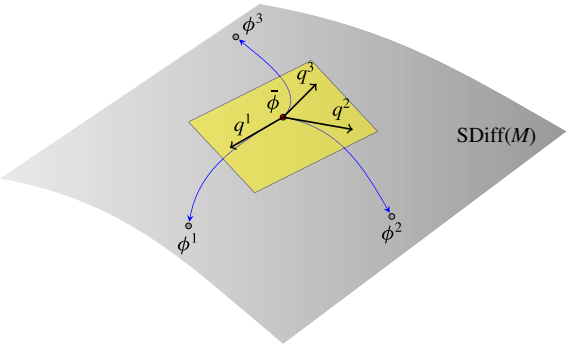


FIGURE 5. (Colour online) Definition of $\bar{\phi}$ as Riemannian centre of mass in the group of volume-preserving diffeomorphisms.

4.4. *Geodesic*

Since the distance (2.25) on SDiff(M) plays a fundamental role in incompressible fluid dynamics through the geodesic interpretation of the Euler equations, the most natural average is arguably that defined as a Riemannian centre of mass using this distance:

$$\bar{\phi} = \underset{\phi \in \mathrm{SDiff}(M)}{\arg\min} \langle D^2(\phi, \phi^\alpha) \rangle. \tag{4.18}$$

The maps ϕ^α are reached from $\bar{\phi}$ (or equivalently the maps ξ^α are reached from id) by integrating the time-dependent divergence-free velocity fields $q^\alpha = q^\alpha_s$ that satisfy the Euler equations in fictitious time s,

$$\partial_s q^\alpha_\flat + \mathcal{L}_{q^\alpha} q^\alpha_\flat = -\mathrm{d}\pi^\alpha, \quad \operatorname{div} q^\alpha = 0, \tag{4.19a,b}$$

with the constraint that

$$\langle q^\alpha_\flat \rangle = 0 \quad \text{for } s = 0, \tag{4.20}$$

hence

$$\langle q^\alpha \rangle = 0 \quad \text{for } s = 0. \tag{4.21}$$

Thus, compared with the extended GLM definition of § 4.2, the only difference is the replacement of paths γ^α_s in Diff(M) satisfying the pressureless fluid equations (and hence unconstrained by volume preservation) by paths in SDiff(M) satisfying the Euler equations for incompressible fluids. Thus the construction of the mean flow map is taking place entirely within SDiff(M), as illustrated in figure 5. Note that the pressure π^α is also fictitious and is simply present to ensure that the flow q^α is divergence free; we spare the reader the introduction of a new symbol here.

While the mean-map definition (4.18) is inspired by the geodesic interpretation of the Euler equations, it can be used for other fluid models, even though the distance on SDiff(M) appearing in (4.18) is then no longer identical to the action functional governing the model's variational structure (for the rotating Boussinesq model, for instance, the action includes rotational and potential-energy terms; see (2.40)). In this case, the auxiliary one-form $q^\alpha_\flat$ continues to satisfy the Euler equations (4.19) in contrast to the dynamical momentum ν^α which satisfies more complicated equations. Alternatively, it may be possible to construct a mean flow map using only the relevant action functional, although sign indefiniteness could make this problematic. We do not explore this possibility in the present paper.

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

4.5. *glm*

The fourth possible definition of $\bar{\phi}$ we discuss is that proposed by Soward & Roberts (2010) in their 'glm' reformulation of GLM. This relies on the group structure of SDiff(M), but not on a metric. In our framework, it simply amounts to choosing q^α to be divergence free and independent of the fictitious time s:

$$\partial_s q^\alpha = 0 \quad \text{with } \operatorname{div} q^\alpha = 0. \tag{4.22}$$

The perturbation maps can then be written as

$$\xi^\alpha = \exp q^\alpha, \tag{4.23}$$

where the exponential is the Lie-group exponential on SDiff(M), defined as the flow map at time $s = 1$ of the s-independent vector field q^α. The glm mean flow map is then defined by imposing

$$\langle q^\alpha \rangle = 0. \tag{4.24}$$

This definition is clearly coordinate independent and guarantees that the mean and perturbation maps are volume preserving. This is achieved by imposing the constraint of vanishing average to linear objects, the vector fields q^α which act as proxies for the perturbation flow maps ξ^α. (Note that we use the term 'glm' for Soward & Roberts' choice (4.23)–(4.24) without implying a restriction to small-amplitude perturbations that accompanied Holm's (2002*b*) original introduction of the term.)

4.6. *Remarks*

We conclude this section with some remarks, first on the limitations of the definitions. It is clear that the four definitions can at best define a mean flow map locally, when the maps ϕ^α are close enough to one another. Any theory can at best work for finite-amplitude displacements from the mean, but within finite limits which we will not attempt to specify here. For the first three, it is enough to note that even in finite dimension, centres of mass often fail to exist globally, as the example of the mean of points uniformly distributed on the sphere demonstrates. Furthermore, it is possible, even in standard GLM in Euclidean geometry, that the mean flow map $\bar{\phi}$ fail to be invertible. Non-smooth solutions are usual in the theory of optimal transport and, while they may provide acceptable solutions in other contexts, it is unclear what their relevance would be for wave–mean-flow interaction. On the other hand, results on the local existence of smooth solutions could be used to check that the definitions of the mean flow map make sense locally.

For the geodesic definition of § 4.4, in particular, the existence of a unique geodesic between sufficiently close volume-preserving diffeomorphisms (Ebin & Marsden 1970) suggests the local existence of a unique Riemannian centre of mass. For the glm definition, given an ensemble of fields q^α satisfying the conditions (4.22), a mean flow map $\bar{\phi}$ and corresponding perturbed maps ϕ^α can be generated, and the governing equations can be averaged as, for example, in dynamo applications (Soward & Roberts 2014). If, however, we regard arbitrary perturbed maps ϕ^α as a starting point, the glm definition is more problematic. This is because the set of flow maps that can be written as the exponential of an s-independent vector field in the manner of (4.23) is limited (see Ebin & Marsden 1970, § 9, for the difference between accessibility by geodesics and accessibility by the Lie-group exponential). To see this, it is enough to notice that the flows generated by time-independent vector fields cannot have isolated

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

periodic points (since the whole orbit of such points consists of periodic points), and so cannot generate general maps.

The constructions in §§ 4.3–4.5 ensure that the mean flow map is volume preserving and are appropriate for the Boussinesq or incompressible MHD equations as well as for the incompressible Euler equations. For a compressible fluid, in contrast, there is no reason to insist that the mean flow map be volume preserving. It may be advantageous to impose other constraints that lead to simplifications of the mean dynamics. We suggest two possibilities. The first is to impose that the Lagrangian-mean density $\overline{\rho}^L = \langle \xi^{\alpha *} \rho^\alpha \rangle$ and the mean mass form in (3.5) satisfy the natural relation $\overline{m}^L = \overline{\rho}^L \omega$. Pulling back by $\bar{\phi}^*$, this constraint can be rewritten as $\langle \phi^{\alpha *} m^\alpha \rangle = \langle \phi^{\alpha *} \rho \rangle \bar{\phi}^* \omega$ and, since (3.5) implies that $\langle \phi^{\alpha *} m^\alpha \rangle \langle \phi^{\alpha *} (\rho^\alpha)^{-1} \rangle = \langle \phi^{\alpha *} \omega \rangle$,

$$\langle \phi^{\alpha *} \omega \rangle = \langle \phi^{\alpha *} \rho^\alpha \rangle \langle \phi^{\alpha *} (\rho^\alpha)^{-1} \rangle \bar{\phi}^* \omega. \tag{4.25}$$

An alternative is to choose $\bar{\phi}$ to simplify the averaging of the equation of state. Writing this equation as $m^\alpha = R(p^\alpha, \theta^\alpha)\omega$, simplifications would ensue if $\bar{\phi}$ could be chosen to ensure that $\overline{m}^L = R(\overline{p}^L, \overline{\theta}^L)\omega$ also; the mean flow would then satisfy the same equation of state as the individual flows.

5. Small-amplitude perturbations

5.1. *Preliminaries*

The definitions of the mean flow map in §§ 4.2–4.5 are somewhat involved. To exploit them in practice requires us to work perturbatively, order-by-order in a small-amplitude parameter $\varepsilon \ll 1$. This is conveniently done in terms of the s-dependent vector fields q^α: the perturbative framework is introduced simply by expanding these in powers of s,

$$q^\alpha = q_1^\alpha + s q_2^\alpha + \cdots, \tag{5.1}$$

where the q_i^α are independent of s, and by taking the maps ξ^α as the flows of q^α at $s = \varepsilon$ instead of $s = 1$. By repeated differentiation with respect to s of the identity

$$\frac{\mathrm{d}}{\mathrm{d}s}(\xi_s^* \tau) = \xi_s^* \mathcal{L}_{q_s} \tau, \tag{5.2}$$

where τ is an arbitrary tensor, we obtain the Taylor expansion

$$\xi^{\alpha *} \tau = \tau + \varepsilon \mathcal{L}_{q_1^\alpha} \tau + \tfrac{1}{2}\varepsilon^2 (\mathcal{L}^2_{q_1^\alpha} + \mathcal{L}_{q_2^\alpha})\tau + \cdots. \tag{5.3}$$

In particular, applying this to the coordinate function x^i gives

$$\xi^{\alpha i}(x) = x^i + \varepsilon \xi_1^{\alpha i} + \varepsilon^2 \xi_2^{\alpha i} + \cdots = x^i + \varepsilon q_1^{\alpha i} + \tfrac{1}{2}\varepsilon^2 (q_2^{\alpha i} + q_1^{\alpha j} \partial_j q_1^{\alpha i}) + \cdots. \tag{5.4}$$

Introducing (5.1) into the equation satisfied by q^α in each mean-flow definition leads to

$$\text{GLM, optimal transport:} \quad q_2^\alpha = -\nabla_{q_1^\alpha} q_1^\alpha, \tag{5.5a}$$

$$\text{geodesic:} \quad q_2^\alpha = -P \nabla_{q_1^\alpha} q_1^\alpha, \tag{5.5b}$$

where P denotes projection onto divergence-free vector fields, and

$$\text{glm:} \quad q_2^\alpha = 0. \tag{5.5c}$$

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

The constraint at the endpoint $s = 0$ simply implies that $\langle q_1^\alpha \rangle = 0$ for the GLM, geodesic and glm definitions. For optimal transport, this only holds to leading order in ε: the $O(\varepsilon)$ function ψ in (4.17) is determined by the volume-preservation constraint $\bar{\phi}^* \omega = \omega$, hence $\langle \xi^{*\alpha} \omega \rangle = \omega$; using (5.3) this imposes that

$$\operatorname{div} \langle q_1^\alpha + \tfrac{1}{2} \varepsilon q_2^\alpha \rangle = O(\varepsilon^2). \tag{5.6}$$

Since $\operatorname{div} q_1 = O(\varepsilon)$ and, from (5.5*a*), q_2 is divergent, the $O(\varepsilon)$ part of $\langle q_1^\alpha \rangle$ introduced by the gradient in (4.17) must be chosen such that

$$\text{optimal transport:} \quad \langle q_1^\alpha + \tfrac{1}{2} \varepsilon q_2^\alpha \rangle = -\tfrac{1}{2} \varepsilon P \langle \nabla_{q_1^\alpha} q_1^\alpha \rangle + O(\varepsilon^2). \tag{5.7}$$

5.2. *Mean flow and Lagrangian-mean momentum*

We can now express the mean velocity $\bar{u}$, Lagrangian-mean momentum $\overline{v}^L$, and hence pseudomomentum p perturbatively in terms of the q_i^α. We assume that the velocity and momentum have the dual expansions

$$u^\alpha = u_0 + \varepsilon u_1^\alpha + \varepsilon^2 u_2^\alpha + \cdots \quad \text{and} \quad v^\alpha = v_0 + \varepsilon v_1^\alpha + \varepsilon^2 v_2^\alpha + \cdots, \tag{5.8a,b}$$

with $\langle u_1^\alpha \rangle = \langle v_1^\alpha \rangle = 0$. The mean velocity is deduced from (3.6), which requires the expansion

$$w^\alpha = \varepsilon \partial_t q_1^\alpha + \tfrac{1}{2} \varepsilon^2 (\partial_t q_2^\alpha - \mathcal{L}_{q_1^\alpha} \partial_t q_1^\alpha) + \cdots. \tag{5.9}$$

This is obtained by considering the paths γ_s^α in (4.8)–(4.9) as functions of both s and t. Defining the corresponding velocities $w_s^\alpha = \dot{\gamma}_s^\alpha \circ (\gamma_s^\alpha)^{-1}$ (such that $w^\alpha = w_s^\alpha|_{s=\varepsilon}$), we can apply (2.18) to find

$$\partial_s w_s^\alpha = \partial_t q_s^\alpha - \mathcal{L}_{q^\alpha} w_s^\alpha, \tag{5.10}$$

from which (5.9) follows by Taylor expansion.

Now, for the GLM, geodesic and glm mean flows, we can treat q_1^α and q_2^α as ε-independent. This is not true for the optimal-transport mean flow since (5.7) implies that q_1^α contains an $O(\varepsilon)$ term. We handle this complication with minimal inconvenience by noting that, to the order $O(\varepsilon^2)$ considered here, (5.3) and (5.9) involve q_2^α only in the combination $q_1^\alpha + (\varepsilon q_2^\alpha)/2$. This makes it possible to redefine q_1^α as the value of q_1^α at $\varepsilon = 0$ and to incorporate its $O(\varepsilon)$-part in a redefined q_2^α. This reduces (5.7) to $\langle q_2^\alpha \rangle = -P \langle \nabla_{q_1^\alpha} q_1^\alpha \rangle$. Because this also holds for the geodesic mean flow, see (5.5*b*), and because $\langle q_2^\alpha \rangle$ is the only part of q_2^α that matters to order $O(\varepsilon^2)$, the form of $\bar{u}$ and $\overline{v}^L$ is the same for the optimal-transport and geodesic mean-flow definitions. We take advantage of this to bypass the complication associated with the optimal-transport definition in what follows.

Introducing (5.9) into (3.5) gives to the first three orders $\bar{u}_0 = u_0$,

$$\partial_t q_1^\alpha + \mathcal{L}_{u_0} q_1^\alpha = u_1^\alpha \tag{5.11}$$

and, upon averaging,

$$\bar{u}_2 = \langle u_2^\alpha \rangle + \tfrac{1}{2} \langle \mathcal{L}_{q_1^\alpha} u_1^\alpha \rangle - \tfrac{1}{2} \langle (\partial_t + \mathcal{L}_{u_0}) q_2^\alpha \rangle, \tag{5.12}$$

where we have used that $\langle q_1^\alpha \rangle = 0$ as follows from (5.11).

The Lagrangian-mean momentum is computed from its definition: using (5.3) we find that

$$\xi^* v^\alpha = v_0 + \varepsilon (v_1^\alpha + \mathcal{L}_{q_1^\alpha} v_0^\alpha) + \varepsilon^2 (v_2^\alpha + \mathcal{L}_{q_1^\alpha} v_1 + \tfrac{1}{2} \mathcal{L}_{q_1^\alpha}^2 v_0 + \tfrac{1}{2} \mathcal{L}_{q_2^\alpha} v_0) + \cdots. \tag{5.13}$$

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

Averaging and using that $\langle q_1^\alpha \rangle = 0$, we obtain

$$\overline{v}^L = \nu_0 + \varepsilon^2 \overline{v}_2^L + \cdots , \tag{5.14}$$

where

$$\overline{v}_2^L = \langle \nu_2^\alpha \rangle + \langle \mathcal{L}_{q_1^\alpha} \nu_1^\alpha \rangle + \tfrac{1}{2} \langle \mathcal{L}^2_{q_1^\alpha} \nu_0 \rangle + \tfrac{1}{2} \langle \mathcal{L}_{q_2^\alpha} \nu_0 \rangle . \tag{5.15}$$

Equations (5.12) and (5.15) give explicit, coordinate-free expressions for the first non-trivial term in the expansion of the mean velocity and Lagrangian-mean momentum. Using (5.5), we can specialise these expressions for each of the mean-flow definitions to obtain:

$$\text{GLM:} \quad \bar{u}_2 = \langle u_2^\alpha \rangle + \tfrac{1}{2} \langle \mathcal{L}_{q_1^\alpha} u_1^\alpha \rangle + \tfrac{1}{2} (\partial_t + \mathcal{L}_{u_0}) \langle \nabla_{q_1^\alpha} q_1^\alpha \rangle , \tag{5.16a}$$

$$\overline{v}_2^L = \langle \nu_2^\alpha \rangle + \langle \mathcal{L}_{q_1^\alpha} \nu_1^\alpha \rangle + \tfrac{1}{2} \langle \mathcal{L}^2_{q_1^\alpha} \nu_0 \rangle - \tfrac{1}{2} \mathcal{L}_{\langle \nabla_{q_1^\alpha} q_1^\alpha \rangle} \nu_0 , \tag{5.16b}$$

$$\text{geodesic, optimal transport:} \quad \bar{u}_2 = \langle u_2^\alpha \rangle + \tfrac{1}{2} \langle \mathcal{L}_{q_1^\alpha} u_1^\alpha \rangle + \tfrac{1}{2} (\partial_t + \mathcal{L}_{u_0}) P \langle \nabla_{q_1^\alpha} q_1^\alpha \rangle , \tag{5.16c}$$

$$\overline{v}_2^L = \langle \nu_2^\alpha \rangle + \langle \mathcal{L}_{q_1^\alpha} \nu_1^\alpha \rangle + \tfrac{1}{2} \langle \mathcal{L}^2_{q_1^\alpha} \nu_0 \rangle - \tfrac{1}{2} \mathcal{L}_{P \langle \nabla_{q_1^\alpha} q_1^\alpha \rangle} \nu_0 , \tag{5.16d}$$

$$\text{glm:} \quad \bar{u}_2 = \langle u_2^\alpha \rangle + \tfrac{1}{2} \langle \mathcal{L}_{q_1^\alpha} u_1^\alpha \rangle , \tag{5.16e}$$

$$\overline{v}_2^L = \langle \nu_2^\alpha \rangle + \langle \mathcal{L}_{q_1^\alpha} \nu_1^\alpha \rangle + \tfrac{1}{2} \langle \mathcal{L}^2_{q_1^\alpha} \nu_0 \rangle . \tag{5.16f}$$

These are key results of this paper. Together with the Lagrangian-averaged momentum equation (3.11), they govern the mean dynamics, providing an evolution equation for $\langle \nu_2^\alpha \rangle$. The Eulerian-mean momentum $\langle \nu_2^\alpha \rangle$ appears as part of the advected Lagrangian-mean momentum $\overline{v}^L$, while the Eulerian-mean velocity $\langle u_2^\alpha \rangle$ appears as part of the advecting mean velocity $\bar{u}$. They are explicitly related through any of the pairs of equations in (5.16), depending on the choice of mean-flow definition, and by the identity $\langle \nu_2^\alpha \rangle = \langle u_2^\alpha \rangle_\flat$. The equation for $\langle \nu_2^\alpha \rangle$ is not closed: a model for the dynamics of the ensemble of perturbations, linking q_1^α or u_1^α to the mean quantities (and consistent with (5.11)), is necessary to evaluate the quadratic perturbation terms in (5.16). This can be done in a heuristic manner, as in the derivation of α- and related models (e.g. Holm 2002*b*; Marsden & Shkoller 2003), or based on the $O(\varepsilon)$ dynamical equations. We do not consider this closure problem further in this paper.

It is interesting to express the condition satisfied by $\langle q_2^\alpha \rangle$ in terms of the coordinate-dependent $\xi_2^{\alpha i}$ to connect our results with standard GLM. From (5.4), (5.5*a*) and the form of the covariant derivative we find that

$$\text{GLM:} \quad \langle \xi_2^{\alpha i} \rangle = \tfrac{1}{2} \langle q_1^{\alpha j} \partial_j q_1^{\alpha i} \rangle - \tfrac{1}{2} \langle \nabla_{q_1^\alpha} q_1^\alpha \rangle^i = -\tfrac{1}{2} \Gamma^i_{jk} \langle q_1^{\alpha j} q_1^{\alpha k} \rangle . \tag{5.17}$$

In the case where $M = \mathbb{R}^n$ with Cartesian coordinates, $\Gamma^i_{jk} = 0$ and the standard GLM condition $\langle \xi_1^{\alpha i} \rangle = \langle \xi_2^{\alpha i} \rangle = 0$ is recovered. On general manifolds, however, $\langle \xi_2^{\alpha i} \rangle \neq 0$ as is required to ensure that the mean flow is tangent to the manifold. The mean flow is not volume preserving, however, since $\operatorname{div} \langle q^\alpha \rangle \neq 0$. This is remedied for the other mean flows for which we obtain similarly

$$\text{geodesic, optimal transport:} \quad \langle \xi_2^{\alpha i} \rangle = \tfrac{1}{2} \langle q_1^{\alpha j} \partial_j q_1^{\alpha i} \rangle - \tfrac{1}{2} P \langle \nabla_{q_1^\alpha} q_1^\alpha \rangle^i \tag{5.18}$$

and

$$\text{glm:} \quad \langle \xi_2^{\alpha i} \rangle = \tfrac{1}{2} \langle q_1^{\alpha j} \partial_j q_1^{\alpha i} \rangle . \tag{5.19}$$

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

The latter expression recovers a key result in Soward & Roberts (2010). Observe the similarities between expressions (5.17), (5.18) and (5.19) for $\langle \xi_2^{\alpha i} \rangle$ obtained for the various definitions of the mean: the glm term (5.19) is common to all expressions; the other three expressions contain an additional term that depends on the metric and is rendered divergence free by projection in the case of the optimal transport and geodesic definitions. For $M = \mathbb{R}^n$, $\nabla_{\xi_1^\alpha} \xi_1^{\alpha i} = \xi_1^{\alpha j} \partial_j \xi_1^{\alpha i}$, and the difference between glm on the one hand, and optimal transport and geodesic on the other hand, is the projection $(I - P)$ on gradient vector fields (spanning the complement to divergence-free vector fields); thus, in this sense, $\langle \xi_2^{\alpha i} \rangle$ is larger for glm than for the optimal transport and geodesic definitions of the mean.

6. Application: wave–mean-flow interaction in Boussinesq fluid

We illustrate the theory developed in this paper with a perturbative treatment of the Lagrangian-mean dynamics of a Boussinesq fluid derived in § 3.3. We focus on the rapid-rotation, strong-stratification regime, where the mean flow is balanced, hence controlled by potential vorticity, while the perturbations represent fast inertia–gravity waves. Averaging is then naturally over the fast time scale of the waves.

Our aim is to relate the transported Lagrangian potential vorticity $\overline{\mathcal{Q}}^L$ in (3.22) to the advecting velocity $\bar{u}$ under quasi-geostrophic scaling. We make the 'strong-wave' assumption $u_0 = \nu_0 = 0$. This leads to $O(\varepsilon^2)$ mean quantities and reduces (5.11), (5.12) and (5.15) to

$$\partial_t q_1^\alpha = u_1^\alpha, \quad \bar{u}_2^\alpha = \tfrac{1}{2} \langle \mathcal{L}_{q_1^\alpha} u_1^\alpha \rangle \quad \text{and} \quad \overline{v}_2^L = \langle \mathcal{L}_{q_1^\alpha} \nu_1^\alpha \rangle, \tag{6.1a–c}$$

using that $\langle \partial_t \cdot \rangle = 0$. Correspondingly, the leading-order pseudomomentum is given by

$$-p = \langle \mathcal{L}_{q_1^\alpha} \nu_1^\alpha \rangle - \tfrac{1}{2} \langle \mathcal{L}_{q_1^\alpha} u_1^\alpha \rangle_\flat, \tag{6.2}$$

and is identical for all mean-flow definitions. In components and in the Euclidean case on which we focus this is just

$$-p_i = \langle \nu_j^\alpha \partial_i q_1^{\alpha j} \rangle, \tag{6.3}$$

again since $\langle \partial_t \cdot \rangle = 0$, consistent with standard results (Bühler 2014). Note that we retain $p = \bar{u}_\flat - \overline{v}^L$ as the definition of the pseudomomentum, when we could equally well have included background-rotation effects by changing $\overline{v}^L$ to $\overline{v}_a^L$.

We now consider the Lagrangian-mean momentum equations (3.21) in the quasi-geostrophic regime. We use the familiar notation: (x, y, z) for Cartesian coordinates, $(u, v, w) = (u^{\alpha 1}, u^{\alpha 2}, u^{\alpha 3})$ for the velocity components, and $(\xi, \eta, \zeta) = (\xi^{\alpha 1}, \xi^{\alpha 2}, \xi^{\alpha 3})$ for the displacement components, dropping the superscript α in the process. The dual use of u and ξ for a vector and its first component should not lead to confusion. Taking $\Phi = z$ and $\theta_0(z) = N^2 z$, with constant N for simplicity, (3.21) reduces to the geostrophic and hydrostatic balances

$$-f\bar{v} = -\overline{\pi}_x^L + \overline{\theta}^L \langle \zeta_2 \rangle_x, \tag{6.4a}$$

$$f\bar{u} = -\overline{\pi}_y^L + \overline{\theta}^L \langle \zeta_2 \rangle_y, \tag{6.4b}$$

$$0 = -\overline{\pi}_z^L + \overline{\theta}^L (1 + \langle \zeta_2 \rangle_z), \tag{6.4c}$$

where the average of the second-order vertical displacement ζ_2 arises as a result of the pull-back of $\mathrm{d}\Phi = \mathrm{d}z$ (cf. Xie & Vanneste 2015). Note how the mean

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

velocity components $\bar{u}$ and $\bar{v}$ appear on the left-hand side, and not the corresponding components of the momentum $\overline{v}^L$. The Lagrangian-mean buoyancy can be written as

$$\overline{\theta}^L = N^2 z + N^2 \langle \zeta_2 \rangle + \bar{\psi}_z / f, \tag{6.5}$$

where $\bar{\psi}$ is an $O(\varepsilon^2)$ streamfunction. Substituting into (6.4*c*) and retaining only the $O(1)$ and $O(\varepsilon^2)$ terms leads to

$$\overline{\pi}^L = \tfrac{1}{2} N^2 z^2 + N^2 z \langle \zeta_2 \rangle + \bar{\psi} / f. \tag{6.6}$$

Introducing this into (6.4*a*)–(6.4*b*) gives the geostrophic balance

$$\bar{u} = -\bar{\psi}_y \quad \text{and} \quad \bar{v} = \bar{\psi}_x. \tag{6.7}$$

We compute $\overline{\mathcal{Q}}^L$ from its definition (3.22), using that

$$\overline{v}^L_a = \langle \xi^{\alpha *} A \rangle + \bar{u}\, \mathrm{d}x + \bar{v}\, \mathrm{d}y - p, \tag{6.8}$$

where only the first term is $O(1)$. Using (6.3), we obtain that

$$-\mathrm{d}p \wedge \mathrm{d}z = \langle \partial(u_1, \xi_1) + \partial(v_1, \eta_1) \rangle \, \mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z \tag{6.9}$$

where $\partial(\cdot, \cdot)$ is the Jacobian in (x, y). Also,

$$\mathrm{d}\langle \xi^{\alpha *} A \rangle = f \langle \xi^{\alpha *} (\mathrm{d}x \wedge \mathrm{d}y) \rangle = f \left(1 + \langle \partial(\xi_1, \eta_1) \rangle + \langle \xi_2 \rangle_x + \langle \eta_2 \rangle_y \right) \mathrm{d}x \wedge \mathrm{d}y. \tag{6.10}$$

Introducing into (3.22) and using the quasi-geostrophic scaling to retain only the leading-order terms, we obtain

$$\begin{aligned} \overline{\mathcal{Q}}^L \;=\; & N^2 \left(f + (\partial_{xx} + \partial_{yy} + f^2 N^{-2} \partial_{zz}) \bar{\psi} \right. \\ & \left. + \langle \partial(u_1, \xi_1) + \partial(v_1, \eta_1) \rangle + f \langle \partial(\xi_1, \eta_1) \rangle + f \langle \partial_i \xi_2^{\alpha i} \rangle \right) \mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z. \end{aligned} \tag{6.11}$$

For the optimal-transport, geodesic and glm mean-flow definitions, (5.4) and the property $\partial_i q_2^i = 0$ resulting from the volume preservation of ξ^α imply that

$$\langle \partial_i \xi_2^{\alpha i} \rangle = \tfrac{1}{2} \partial_i \langle \xi_1^{\alpha j} \partial_j \xi_1^{\alpha i} \rangle = \tfrac{1}{2} \partial_{ij} \langle \xi_1^{\alpha i} \xi_1^{\alpha j} \rangle. \tag{6.12}$$

The material conservation of $\overline{\mathcal{Q}}^L = \overline{Q}^L \omega$ can finally be written as

$$\partial_t \overline{Q}^L + \partial(\bar{\psi}, \overline{Q}^L) = 0, \tag{6.13}$$

where

$$\overline{Q}^L = \left(\partial_{xx} + \partial_{yy} + \frac{f^2}{N^2} \partial_{zz} \right) \bar{\psi} + \langle \partial(u_1, \xi_1) + \partial(v_1, \eta_1) \rangle + f \langle \partial(\xi_1, \eta_1) \rangle + \tfrac{1}{2} f \partial_{ij} \langle \xi_1^{\alpha i} \xi_1^{\alpha j} \rangle. \tag{6.14}$$

Equations (6.13)–(6.14) have been obtained by Holmes–Cerfon *et al.* (2011), Wagner & Young (2015) and Salmon (2016) using, respectively, standard GLM, multi-scale asymptotics and a variational approach. They describe the advection of the Lagrangian-mean quasi-geostrophic potential vorticity which includes, in addition to the familiar (scaled) Laplacian of the streamfunction $\bar{\psi}$, a wave-induced contribution that is quadratic in the wave fields.

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

The derivation of (6.13)–(6.14) in the GLM framework is somewhat different. In this case, the terms $\langle \xi_2^{\alpha i} \rangle$ disappear from $\overline{\mathcal{Q}}^L$ in (6.11), but there is a compensating effect stemming from the divergence of $\bar{u}$. As noted in § 3.3, $\overline{Q}^L$ is obtained from $\overline{\mathcal{Q}}^L$ by dividing by the density $\bar{\rho}$ such that $\bar{\phi}_* \omega = \bar{\rho}\omega$, and given by $\bar{\rho} = 1 - \partial_{ij}\langle \xi_1^{\alpha i} \xi_1^{\alpha j} \rangle/2 + \cdots$. A Taylor expansion then leads to the form (6.14) of $\overline{Q}^L$.

It should not come as a surprise that the four mean-flow definitions lead to the same result: they all describe the same dynamics, specifically the same cluster of trajectories moving by $O(1)$ over a long time scale. The four Lagrangian-mean theories predict the position of this cluster using notions of the centre of the cluster that differ by $O(\varepsilon^2)$ only and indistinguishable to the order considered.

7. Wave action

In the case where α is a phase-like variable ($\alpha \in \mathbb{S}^1$, where $\mathbb{S}^1$ is the circle) the flows ϕ^α describe a closed loop in SDiff(M). A conservation law is then associated with the invariance of the loop with respect to shifts in α; this is the wave-action conservation law generalising the action conservation of classical mechanics (e.g. Arnold, Kozlov & Neishdadt 1988). We derive this conservation from the equations of motion, thus extending the derivation of Andrews & McIntyre (1978*b*) (see also Grimshaw (1984) and § 10.3 of Bühler (2014)).

Let z^α denote the divergence-free vector field generating shifts in α of the perturbation maps ξ^α, that is,

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\xi^\alpha x = z^\alpha(\xi^\alpha x). \tag{7.1}$$

An application of (2.18) shows that

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}u^\alpha = \partial_t z^\alpha + \mathcal{L}_{u^\alpha} z^\alpha. \tag{7.2}$$

Now, we pair the Euler equation (2.28) for the flow $v = v^\alpha$ with z^α to obtain

$$\partial_t v^\alpha(z^\alpha) - v^\alpha(\partial_t z^\alpha) + \mathcal{L}_{u^\alpha} v^\alpha(z^\alpha) - v^\alpha(\mathcal{L}_{u^\alpha} z^\alpha) = -\mathcal{L}_{z^\alpha}\pi^\alpha. \tag{7.3}$$

Using (7.2), we compute

$$v^\alpha(\partial_t z^\alpha) = v^\alpha\left(\frac{\mathrm{d}}{\mathrm{d}\alpha}u^\alpha - \mathcal{L}_{u^\alpha} z^\alpha\right) = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}\alpha}v^\alpha(u^\alpha) - v^\alpha(\mathcal{L}_{u^\alpha} z^\alpha). \tag{7.4}$$

Introducing this into (7.3) leads to

$$(\partial_t + \mathcal{L}_{u^\alpha})v^\alpha(z^\alpha) - \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}\alpha}v^\alpha(u^\alpha) = -\mathcal{L}_{z^\alpha}\pi^\alpha. \tag{7.5}$$

Taking the average eliminates the derivative with respect to α and multiplying by the volume form ω gives

$$\partial_t\langle v^\alpha(z^\alpha)\rangle\omega + \langle \mathcal{L}_{u^\alpha} v^\alpha(z^\alpha)\omega\rangle + \langle \mathcal{L}_{z^\alpha}\pi^\alpha\omega\rangle = 0, \tag{7.6}$$

using also that z^α and u^α are divergence free. Applying Cartan's formula (2.20) to write for example $\mathcal{L}_{z^\alpha}\pi^\alpha\omega = \mathrm{d}(\pi^\alpha \iota_{z^\alpha}\omega)$, we finally obtain

$$\partial_t\langle v^\alpha(z^\alpha)\rangle\omega + \mathrm{d}\langle v^\alpha(z^\alpha)\iota_{u^\alpha}\omega + \pi^\alpha\iota_{z^\alpha}\omega\rangle = 0. \tag{7.7}$$

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

This is the local form of a conservation law for the action

$$A = \int_M \langle \nu^\alpha(z^\alpha) \rangle \, \omega, \tag{7.8}$$

since $u^\alpha \parallel \partial M$ and $z^\alpha \parallel \partial M$. Note that for small-amplitude perturbations $\xi^\alpha = \mathrm{id} + O(\varepsilon)$, this action is $O(\varepsilon^2)$ (since $\langle z^\alpha \rangle = O(\varepsilon^2)$) and is therefore a 'wave activity' in the sense of Andrews & McIntyre (1978*b*).

8. Discussion

In this paper, we revisit Andrews & McIntyre's GLM theory to remedy the deficiencies caused by its reliance on Euclidean parallel transport and consequent lack of intrinsic geometric meaning. We achieve this by formulating the problem on a general Riemannian manifold and by using the coordinate-free notation of modern differential geometry. Several lessons are learnt in the process. (i) The structure of the averaged equations that are obtained is independent of the exact definition of the mean flow. (ii) These averaged equations are naturally expressed in terms of a Lagrangian-mean momentum $\overline{\nu}^L$, defined as the average of the pull-back of the momentum one-form (dual of the velocity field) from the perturbed to the mean configuration. It corresponds to the quantity $\overline{\boldsymbol{u}}^L - \boldsymbol{p}$ of standard GLM, and is the integrand of the mean Kelvin circulation theorem. The central part played by the momentum one-form reflects its role as the natural dynamical variable of fluid dynamics. (iii) The mean flow is in fact defined as a mean flow map, from which the mean velocity is deduced by time differentiation. The definition of the mean flow map involves a degree of arbitrariness, as is inevitable when attempting to define the 'mean' of an ensemble of points on a manifold (in this instance an ensemble of diffeomorphisms) that does not have a natural linear structure with natural averaging operators. Defining the mean in terms of the flow map ensures that the mean flow retains its interpretation as a representative of an ensemble of nearby flows over long times. (iv) It is clear that there are advantages, for incompressible fluids, in definitions that respect the volume-preservation constraint. For example in the case of MHD, the averaged equations (3.34) retain the structure of the original MHD system. (v) The pseudomomentum has a natural interpretation as a measure of the non-commutativity between two operations applied to the velocity field: Lagrangian averaging, and the transformation of vectors into one-forms by means of the metric. This measure encapsulates the impact of the perturbation on the mean flow. (vi) Our computations of the mean flow and pseudomomentum show that it is possible to obtain explicit expressions perturbatively, order-by-order in the small-amplitude parameter ε; the complexity of the resulting expressions suggests that in general they are probably best dealt with using a systematic method, for example a Lie-series expansion (e.g. Nayfeh 1973). (vii) Among the intrinsic definitions of the mean that we consider, the glm theory of Soward & Roberts appears to the most attractive for its relative ease of implementation. A potential weakness is however the fact that the vector field on which it relies is not always defined, even for near-identity flow, though an approximate solution q^α of (4.23) can computed order-by-order in ε. The geodesic definition is likely more robust but leads to more complicated expressions. It is certainly the most natural of the other three definitions. (viii) In a perturbative context, the various definitions of the mean flow differ by bounded $O(\varepsilon^2)$ terms, as expected since they all track (in somewhat different senses) the same ensemble of nearby flows.

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

The aim of this paper is to introduce geometric generalised mean theories in as simple a set-up as possible. We envision extensions in several directions. A first direction concerns generalisations to fluid models more complex than the incompressible Euler equations on which the paper focuses. Our brief treatment of stratified–rotating Boussinesq fluids, compressible fluids and MHD makes it plain that such generalisations are possible and likely advantageous. It would also be of interest to apply a geometric perspective to commonly used averaging procedures other than GLM such as isentropic (or isopycnal) averaging. Another direction concerns manifolds with symmetries: key results of GLM and related theories apply to the special case where the ensemble average can be identified with an average along one coordinate (e.g. a zonal average), leading to the celebrated non-acceleration theorems, stemming from the mean Kelvin circulation theorem. The geometric understanding of these and their generalisation to a larger class of manifolds than $\mathbb{R}^3$ and $\mathbb{S}^2$ will require us to study how continuous symmetries of manifolds can be exploited for averaging. A third direction concerns the use of geometric methods in the parameterisation of unresolved perturbations in numerical models. The geometric interpretation of the pseudomomentum, which captures this wave impact, can guide its parameterisation in terms of mean quantities as is required for the closure of the mean equations. In particular, it would be of interest to adapt the geometric set-up of the present paper to the short-wave, WKB treatment of the perturbations (e.g. Gjaja & Holm 1996; Bühler & McIntyre 1998, 2005).

Acknowledgements

The authors are grateful to A. M. Soward for many useful discussions and for sending them papers in advance of formal publication, to M. E. McIntyre for extensive comments and for encouraging the development of the Boussinesq example in § 6, to D. D. Holm for helpful comments and references, and to the other (anonymous) referee. They thank M. A. Singer for early conversations on differential geometry, E. Ghys and O. Bühler for helpful remarks, M. Oliver for comments that led to the restructuring of § 5, and S. M. Tobias for pointing out that they were working, at the time independently, along parallel tracks.

Appendix A. Derivation details

We set out geometric proofs of key identities involving Lie derivatives and time-dependent pull-backs; coordinate versions may also be found in Soward & Roberts (2010, 2014).

A.1. *Derivation of (2.14) and (2.15)*

The first equality can be derived by computing

$$\partial_t(\phi_{0,t}^*\tau) = \phi_{0,t}^*(\partial_t\tau + \partial_s|_{s=0}(\phi_{t,s}^*\tau)) = \phi_{0,t}^*(\partial_t\tau + \mathcal{L}_u\tau), \tag{A 1}$$

where, for clarity, we denote by $\phi_{t,s} = \phi_s \circ \phi_t^{-1}$ the map sending positions at time t to positions at time s.

The second equality, which expresses that the pull-back acts on Lie derivatives in a natural way can be derived as follows: denoting by ψ_s the flow map associated with the velocity field v, we compute

$$\phi^*\mathcal{L}_v\tau = \phi^*\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0}(\psi_s^*\tau) = \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0}((\phi^{-1}\circ\psi_s\circ\phi)^*\phi^*\tau) = \mathcal{L}_{\phi^*v}(\phi^*\tau), \tag{A 2}$$

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

where the last line relies on the definition of the Lie derivative and the observation that

$$\frac{\mathrm{d}}{\mathrm{d}s}((\phi^{-1}\circ\psi_s\circ\phi)x) = (\phi^{-1})_*(v(\psi_s\phi x)) = (\phi^* v)(\psi_s x). \tag{A 3}$$

A.2. *Derivation of (3.7)*

This property is readily established using (2.14): applied to the flow u^α and the map ϕ^α this equation implies that

$$(\partial_t + \mathcal{L}_{u^\alpha})\tau^\alpha = \phi^\alpha_*\partial_t(\phi^{\alpha*}\tau^\alpha) = \xi^\alpha_*\bar{\phi}_*\partial_t(\bar{\phi}^*\xi^{\alpha*}\tau^\alpha) \tag{A 4}$$

on using (3.1) and that $(\xi^\alpha\circ\bar{\phi})^* = \bar{\phi}^*\xi^{\alpha*}$. Applying $\xi^{\alpha*}$ yields

$$\xi^{\alpha*}(\partial_t + \mathcal{L}_{u^\alpha})\tau^\alpha = \bar{\phi}_*\partial_t(\bar{\phi}^*\xi^{\alpha*}\tau^\alpha) = (\partial_t + \mathcal{L}_{\bar{u}})(\xi^{\alpha*}\tau^\alpha), \tag{A 5}$$

where the last equality follows again from (2.14), this time applied to the tensor $\xi^{\alpha*}\tau^\alpha$ with the flow $\bar{u}$ and map $\bar{\phi}$. An alternative derivation relying on (2.15) is instructive. Using this equation, the left-hand side of (3.7) can be written as

$$\begin{aligned}
\xi^{\alpha*}(\partial_t + \mathcal{L}_{u^\alpha})\tau^\alpha &= \xi^{\alpha*}\partial_t\tau^\alpha + \mathcal{L}_{\xi^{\alpha*}u^\alpha}(\xi^{\alpha*}\tau^\alpha)\\
&= \partial_t(\xi^{\alpha*}\tau^\alpha) - \xi^{\alpha*}\mathcal{L}_w\tau^\alpha + \mathcal{L}_{\xi^{\alpha*}u^\alpha}(\xi^{\alpha*}\tau^\alpha)\\
&= \partial_t(\xi^{\alpha*}\tau^\alpha) - \mathcal{L}_{\xi^{\alpha*}w}(\xi^{\alpha*}\tau^\alpha) + \mathcal{L}_{\xi^{\alpha*}u^\alpha}(\xi^{\alpha*}\tau^\alpha)\\
&= \partial_t(\xi^{\alpha*}\tau^\alpha) + \mathcal{L}_{\xi^{\alpha*}(u^\alpha - w)}(\xi^{\alpha*}\tau^\alpha) = (\partial_t + \mathcal{L}_{\bar{u}})(\xi^{\alpha*}\tau^\alpha).
\end{aligned} \tag{A 6}$$

Here we have used (2.14), (2.15) and (3.5) to write the second line, third and fourth line, respectively.

Appendix B. Euler equations as geodesics

The Euler equations in the form (2.28) are obtained as geodesic equations on SDiff(M) considered as an (infinite-dimensional) Riemannian manifold when equipped with the metric associated with the distance (2.25). To show this, we consider the augmented action

$$\mathscr{A}[\phi,\pi] = \int_0^T\int_M \left(\frac{1}{2}g(u,u)\omega + \pi(\phi^*\omega - \omega)\right)\mathrm{d}t, \tag{B 1}$$

where π is a scalar field with the role of a Lagrange multiplier enforcing volume preservation. This action is required to be stationary for perturbations $\phi\mapsto\psi_s\circ\phi$ with arbitrary, t-dependent ψ_s such that $\psi_s|_{t=0} = \psi_s|_{t=T} = \mathrm{id}$. This perturbation changes the velocity field u into

$$u_s(x) = \frac{\mathrm{d}}{\mathrm{d}t}(\psi_s\circ\phi)\circ\phi^{-1}\circ\psi_s^{-1}. \tag{B 2}$$

Denoting by v the vector field generating ψ_s as a function of s, in the sense that

$$\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0}\psi_s x = v(x), \tag{B 3}$$

we can apply (2.18) to the two-parameter map $\psi_s\circ\phi$ and obtain the first-order variation of u_s as

$$\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} u_s = \partial_t v - \mathcal{L}_v u = \partial_t v + \mathcal{L}_u v. \tag{B 4}$$

Note that we require $v\parallel\partial M$, as for u, but v need not be divergence free.

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

With this result, we compute

$$\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \mathscr{A}[\psi_s \circ \phi, \pi_s] = \int_0^T \int_M (\nu(\partial_t v + \mathcal{L}_u v)\omega + \pi \mathcal{L}_v \omega)\, \mathrm{d}t. \tag{B 5}$$

This is rewritten using integration by parts, and the volume preservation $\phi^*\omega = \omega$ enforced by the Lagrange multiplier π and its consequence $\mathcal{L}_u\omega = 0$. First, we have that

$$\int_0^T \int_M \nu(\partial_t v)\omega\, \mathrm{d}t = -\int_0^T \int_M \partial_t \nu(v)\omega\, \mathrm{d}t \tag{B 6}$$

since $v|_{t=0} = v|_{t=T} = 0$. Second,

$$\int_M \nu(\mathcal{L}_u v)\omega = \int_M \mathcal{L}_u(\nu(v)\omega) - \int_M (\mathcal{L}_u \nu)(v)\omega = \int_{\partial M} \nu(v)\iota_u\omega - \int_M (\mathcal{L}_u\nu)(v)\omega, \tag{B 7}$$

where we have used Cartan's formula (2.20) to write $\mathcal{L}_u(\nu(v)\omega) = \mathrm{d}(\nu(v)\iota_u\omega)$ and Stokes' theorem (e.g. Schutz 1980; Frankel 1997). The boundary integral vanishes since $u \parallel \partial M$. Third, we have similarly

$$\int_M \pi \mathcal{L}_v \omega = \int_M \mathcal{L}_v(\pi\omega) - \int_M (\mathcal{L}_v \pi)\omega = \int_{\partial M} \pi \iota_v \omega - \int_M \mathrm{d}\pi(v)\omega, \tag{B 8}$$

with again a vanishing boundary integral, this time because $v \parallel \partial M$. Introducing (B 6)–(B 8) into (B 5) gives

$$\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{\epsilon=0} \mathscr{A}[\psi_s \circ \phi, \pi_s] = -\int_0^T \int_M (\partial_t \nu + \mathcal{L}_u \nu + \mathrm{d}\pi)(v)\omega\, \mathrm{d}t. \tag{B 9}$$

This variation vanishes for arbitrary v leading to the Euler equations (2.28), with the condition $\mathcal{L}_u\omega = 0$ written as $\operatorname{div} u = 0$ by definition of the divergence.

Note that the derivation in this appendix, in particular the constrained variations (B 4), can be generalised to cover the class of so-called Euler–Poincaré equations that includes, among others, all the fluid models treated in this paper; see Holm, Marsden & Ratiu (1998) for details.

Appendix C. Variational derivation of (3.11)

Substituting the decomposition (3.5) into the action (2.26), enforcing the volume-preservation constraints for ξ^α and $\bar{\phi}$ by means of Lagrange multipliers π^α and $\bar{\pi}$ and averaging leads to

$$\begin{aligned}&\langle \mathscr{A} \rangle[\xi^\alpha, \bar{\phi}, \pi^\alpha, \bar{\pi}] \\ &\quad = \left\langle \int_0^T \int_M \left(\frac{1}{2} g(w^\alpha + \xi_*^\alpha \bar{u}, w^\alpha + \xi_*^\alpha \bar{u})\omega + \pi^\alpha(\xi^{\alpha *}\omega - \omega) + \bar{\pi}(\bar{\phi}\omega - \omega) \right) \mathrm{d}t \right\rangle.\end{aligned} \tag{C 1}$$

Requiring this to vanish for arbitrary variations of ξ^α, $\bar{\phi}$, π^α and $\bar{\pi}$ yields dynamical equations for both the mean and perturbations. First, variations with respect to $\bar{\phi}$ generated by a vector field v as in (B 3) lead to

$$\left\langle \int_0^T \int_M \nu^\alpha(\xi_*^\alpha(\partial_t v + \mathcal{L}_{\bar{u}} v))\omega + \bar{\pi}\mathcal{L}_v\omega \right\rangle = 0. \tag{C 2}$$

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

Now, we can pull-back the first term to write

$$\left\langle \int_M \nu^\alpha(\xi^\alpha_*(\partial_t v + \mathcal{L}_{\bar{u}} v))\omega \right\rangle = \left\langle \int_M (\xi^{\alpha *}\nu^\alpha)(\partial_t v + \mathcal{L}_{\bar{u}} v)\omega \right\rangle = \int_M \overline{v}^L(\partial_t v + \mathcal{L}_{\bar{u}} v)\omega. \quad \text{(C 3)}$$

Introducing this into (C 2) and integrating by parts in the manner of appendix B gives the averaged equation (3.11), with $\bar{\pi}$ replacing $\overline{\pi}^L$. Variations of (C 1) can be shown to lead to the Euler equations (2.28) for each of the momenta ν^α.

Appendix D. Riemannian centre of mass

Consider a manifold M with a distance between points x and y given by $d(x, y)$. Suppose points x^α are distributed according to a measure $\mu(x)\,\mathrm{d}x$. Then the Riemannian centre of mass (Fréchet or Karcher mean) of their position is defined as

$$\bar{x} = \arg\min_{x\in M} \langle d^2(x, x^\alpha)\rangle = \arg\min_{x\in M} \int_M d^2(x, y)\mu(y)\,\mathrm{d}y. \quad \text{(D 1)}$$

The existence and uniqueness of such a minimiser is of course not trivial and depends on both the manifold M and measure $\mu(x)\,\mathrm{d}x$.

One interpretation of the centre of mass is that it provides the best approximation of the measure $\mu(x)\,\mathrm{d}x$ by an atomic measure $\bar{\mu}(x)\,\mathrm{d}x = \delta(x-\bar{x})\,\mathrm{d}x$ when the Wasserstein distance $W^2(\mu, \nu)$ between the measures is used to define the approximation. This distance is defined as

$$W^2(\bar{\mu}, \mu) = \inf_{k(x,y)} \int d^2(x, y)k(x, y)\,\mathrm{d}x\,\mathrm{d}y, \quad \text{(D 2)}$$

with $\int k(x, y)\,\mathrm{d}y = \bar{\mu}(x) = \delta(x-\bar{x})$ and $\int k(x, y)\,\mathrm{d}x = \mu(y)$. It is clear that $k(x, y) = \delta(\bar{x}-x)\mu(y)$, so that the minimisation reduces to

$$W^2(\bar{\mu}, \mu) = \inf_{x\in M} \int_M d^2(x, y)\,\mu(y)\,\mathrm{d}y, \quad \text{(D 3)}$$

in agreement with the definition (D 1) of the centre of mass.

The connection between the optimal transport and geodesic definitions of the mean flow map of §§ 4.3–4.4 is made clear by thinking of the ensemble of flow maps ϕ^α as defining a measure on the space of diffeomorphisms SDiff(M); for example, for a discrete family labelled by $\alpha = 1, 2, \ldots$, this is the atomic measure concentrated on the maps $\phi^\alpha \in \mathrm{SDiff}(M)$. Such a measure on SDiff(M) is in fact a more general concept since it allows for continuous, multidimensional α, for instance, and it arises in the theory of generalised flows (Brenier 1989; Shnirelman 1994). From this point of view the problem of defining the mean flow map $\bar{\phi}$ corresponds to specifying a classical flow that is 'closest' to a given generalised flow.

Appendix E. Derivation of (4.10), (4.12) and (4.16)

The extended GLM formulation (4.7) requires that the action

$$\mathscr{A}[\gamma_s] = \left\langle \int_0^1 \int_M (g(q^\alpha, q^\alpha)\gamma^\alpha_{s*}\omega)\,\mathrm{d}s \right\rangle, \quad \text{(E 1)}$$

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

be stationary for perturbations $\gamma_s^\alpha \mapsto \psi_\varepsilon^\alpha \circ \gamma_s^\alpha$ of the paths joining $\bar\phi$ to ϕ^α. Here the perturbation paths ψ_ε^α are parameterised (implicitly) by s and constrained by the condition $\psi_\varepsilon^\alpha|_{s=1} = \mathrm{id}$ at the endpoint $s=1$. At the other endpoint $s=0$, the constraint is weaker since that starting map is not fixed but an arbitrary perturbation of $\bar\phi$; thus $\psi_\varepsilon^\alpha|_{s=0}$ is only required to be independent of α. Denoting by v^α the vector field generating ψ_ε^α in a manner analogous to (B 3), these constraints become $v^\alpha|_{s=1} = 0$ and $v^\alpha|_{s=0} = v$, where v is an arbitrary vector field independent of α. The velocity fields q^α appearing in (E 1) are perturbed to q_ε^α; using (2.18), we have then that

$$\left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} q_\varepsilon^\alpha = \partial_s v^\alpha + \mathcal{L}_{q^\alpha} v^\alpha. \tag{E 2}$$

Using this, we can compute the first-order variation $\mathrm{d}\mathscr{A}/\mathrm{d}\varepsilon|_{\varepsilon=0}$ and require that it vanishes. The computation is very similar to that in appendix B and we do not detail it here. It leads to the pressureless equation in the form (4.19), and to the boundary term

$$\langle q_\flat^\alpha(v^\alpha)\rangle = \langle q_\flat^\alpha\rangle(v) \quad \text{at } s=0. \tag{E 3}$$

Since v is arbitrary, this gives constraint (4.12) on the initial value of $q_\flat^\alpha$.

The optimal transportation formulation differs from the extended GLM one only by the addition of the constraint that $\bar\phi$ be volume preserving. To enforce this, we augment the action (E 1) by the term

$$-\left\langle \int_M \psi(\bar\phi^*\omega - \omega)\right\rangle, \tag{E 4}$$

where ψ is a Lagrange multiplier. Using that $\bar\phi \mapsto \psi^\alpha|_{s=0} \circ \bar\phi$, we find that this extra term combines with the boundary term (E 3) to give

$$\langle q_\flat^\alpha(v^\alpha) - \mathrm{d}\psi(v^\alpha)\rangle = \langle q_\flat^\alpha - \mathrm{d}\psi\rangle(v) \quad \text{at } s=0, \tag{E 5}$$

leading to condition (4.16).

REFERENCES

Andrews, D. G. & McIntyre, M. E. 1978*a* An exact theory of nonlinear waves on a Lagrangian-mean flow. *J. Fluid Mech.* **89**, 609–646.

Andrews, D. G. & McIntyre, M. E. 1978*b* On wave-action and its relatives. *J. Fluid Mech.* **89**, 647–664.

Arnold, V. I. 1966 Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. *Ann. Inst. Fourier* **16**, 316–361.

Arnold, V. I. & Khesin, B. A. 1998 Topological methods in hydrodynamics. In *Applied Mathematical Sciences*, vol. 125. Springer.

Arnold, V. I., Kozlov, V. V. & Neishtadt, A. I. 1988 Mathematical aspects of classical and celestial mechanics. In *Encyclopedia of Mathematical Sciences: Dynamical Systems III*. Springer.

Bhat, H. S., Fetecau, R. C., Marsden, J. E., Mohseni, K. & West, M. 2005 Lagrangian averaging for compressible fluids. *Multiscale Model. Simul.* **3**, 818–837.

Berger, M. 2003 *A Panoramic View of Riemannian Geometry*. Springer.

Besse, N. & Frisch, U. 2017 Geometric formulation of the Cauchy invariants for incompressible Euler flow in flat and curved spaces. *J. Fluid Mech.* **825**, 412–478.

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

BOYLAND, P. 2001 Fluid mechanics and mathematical structures. In *An Introduction to the Geometry and Topology of Fluid Flows* (ed. R. L. Ricca), NATO Science Series, pp. 105–134. Kluwer Academic Publishers.

BRENIER, Y. 1989 The least action principle and the related concept of generalized flows for incompressible perfect fluids. *J. Amer. Math. Soc.* **2**, 225–255.

BRENIER, Y. 2003 Topics in hydrodynamics and volume-preserving maps. In *Handbook of Mathematical Fluid Dynamics* (ed. S. J. Friedlander & D. Serre), vol. 2, pp. 55–86. Elsevier Sience BV.

BRETHERTON, F. P. 1971 The general linearized theory of wave propagation. In *Mathematical Problems in the Geophysical Sciences*, Lect. Appl. Math., vol. 13, pp. 61–102. Pub. Am. Math. Soc.

BRETHERTON, C. S. & SCHÄR, C. 1993 Flux of potential vorticity substance: a simple derivation and uniqueness property. *J. Atmos. Sci.* **50**, 1834–1836.

BÜHLER, O. 2014 *Waves and Mean Flows*, 2nd edn. Cambridge University Press.

BÜHLER, O. & MCINTYRE, M. E. 1998 On non-dissipative wave–mean interactions in the atmosphere or oceans. *J. Fluid Mech.* **354**, 301–343.

BÜHLER, O. & MCINTYRE, M. E. 2005 Wave capture and wave-vortex duality. *J. Fluid Mech.* **354**, 67–95.

DEWAR, R. L. 1970 Interaction between hydromagnetic waves and a time-dependent inhomogeneous medium. *Phys. Fluids* **13**, 2710–2720.

EBIN, D. G. & MARSDEN, J. 1970 Groups of diffeomorphisms and the motion of an incompressible fluid. *Ann. Maths* **92**, 102–163.

FRANKEL, T. 1997 *The Geometry of Physics*. Cambridge University Press.

GJAJA, I. & HOLM, D. D. 1996 Self-consistent wave-mean flow interaction dynamics and its Hamiltonian formulation for rotating stratified incompressible fluids. *Physica* D **98**, 343–378.

GRIMSHAW, R. 1985 Wave action and wave–mean flow interaction, with application to stratified shear flows. *Annu. Rev. Fluid Mech.* **16**, 11–44.

GRIMSHAW, R. H. J. 1975 Nonlinear internal gravity waves in a rotating fluid. *J. Fluid Mech.* **71**, 497–512.

HOLM, D. D. 2002*a* Variational principles for Lagrangian-averaged fluid dynamics. *J. Phys.* A*: Math. Gen.* **35**, 679–688.

HOLM, D. D. 2002*b* Lagrangian averages, averaged Lagrangians, and the mean effects of fluctuations in fluid dynamics. *Chaos* **12**, 518–530.

HOLM, D. D., MARSDEN, J. E. & RATIU, T. 1998 The Euler–Poincaré equations and semi-direct products with applications to continuum theories. *Adv. Math.* **137**, 1–81.

HOLM, D. D., SCHMAH, T. & STOICA, C. 2009 *Geometric Mechanics and Symmetry: From Finite to Infinite Dimensions*. Oxford University Press.

HOLMES–CERFON, M., BÜHLER, O. & FERRARI, R. 2011 Particle dispersion by random waves in the rotating Boussinesq system. *J. Fluid Mech.* **670**, 150–175.

HAWKING, S. W. & ELLIS, G. F. R. 1973 *The Large Scale Structure of Space-time*. Cambridge University Press.

HAYNES, P. H. & MCINTYRE, M. E. 1987 On the evolution of vorticity and potential vorticity in the presence of diabatic heating and frictional or other forces. *J. Atmos. Sci.* **44**, 828–841.

HAYNES, P. H. & MCINTYRE, M. E. 1990 On the conservation and impermeability theorems for potential vorticity. *J. Atmos. Sci.* **47**, 2021–2031.

MARSDEN, J. E. & HUGHES, T. J. R. 1983 *Mathematical Foundations of Elasticity*. Dover.

MARSDEN, J. E. & SHKOLLER, S. 2003 The anisotropic Lagrangian averaged Euler and Navier–Stokes equations. *Arch. Rat. Mech. Anal.* **166**, 27–46.

MCCANN, R. J. 2001 Polar factorization of maps on Riemannian manifolds. *Geom. Funct. Anal.* **11**, 589–608.

MCINTYRE, M. E. 1980 Towards a Lagrangian-mean description of stratospheric circulations and chemical transport. *Phil. Trans. R. Soc. Lond.* A **296**, 129–148.

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press

McIntyre, M. E. 1988 A note on the divergence effect and the Lagrangian-mean surface elevation in periodic water waves. *J. Fluid Mech.* **189**, 235–242.

Nayfeh, A. 1963 *Perturbation Methods*. Wiley.

Pennec, A. 2006 Intrinsic statistics on Riemannian manifolds: basic tools for geometric measurements. *J. Math. Imag. Vis.* **25**, 127–154.

Roberts, P. H. & Soward, A. M. 2006*a* Eulerian-Lagrangian means in rotating magnetohydrodynamic flows I. General results. *Geophys. Astrophys. Fluid Dyn.* **100**, 457–483.

Roberts, P. H. & Soward, A. M. 2006*b* Covariant description of non-relativistic magnetohydrodynamics. *Geophys. Astrophys. Fluid Dyn.* **100**, 485–502.

Salmon, R. 2013 An alternative view of generalized Lagrangian mean theory. *J. Fluid Mech.* **719**, 165–182.

Salmon, R. 2016 Variational treatment of inertia-gravity waves interacting with a quasi-geostrophic mean flow. *J. Fluid Mech.* **809**, 502–529.

Sanders, J. A. & Verhulst, F. 1985 *Averaging Methods in Nonlinear Dynamical Systems*. Springer.

Shnirelman, A. 1994 Generalized fluid flows, their approximation and applications. *Geom. Funct. Anal.* **4**, 586–620.

Schutz, B. 1980 *Geometrical Methods of Mathematical Physics*. Cambridge University Press.

Soward, A. M. 1972 A kinematic theory of large magnetic Reynolds number dynamos. *Phil. Trans. R. Soc. Lond.* A **272**, 431–462.

Soward, A. M. & Roberts, P. H. 2010 The hybrid Euler–Lagrange procedure using an extension of Moffatt's method. *J. Fluid Mech.* **661**, 45–72.

Soward, A. M. & Roberts, P. H. 2014 Eulerian and Lagrangian means in rotating, magnetohydrodynamic flows II. Braginsky's nearly axisymmetric dynamo. *Geophys. Astrophys. Fluid Dyn.* **108**, 269–322.

Villani, C. 2003 *Topics in Optimal Transportation*, Graduate Studies in Mathematics, vol. 58. American Mathematical Society.

Wagner, G. L. & Young, W. R. 2015 Available potential vorticity and wave-averaged quasi-geostrophic flow. *J. Fluid Mech.* **785**, 401–424.

Xie, J.-H. & Vanneste, J. 2015 A generalised-Lagrangian-mean model of the interactions between near-inertial waves and mean flow. *J. Fluid Mech.* **774**, 143–169.

https://doi.org/10.1017/jfm.2017.913 Published online by Cambridge University Press