Geophysical and Astrophysical Fluid Dynamics
Vol. 00, No. 00, 00 Month 20xx, 1–22

Geodesic motion on groups of diffeomorphisms with H^1 metric as geometric generalised Lagrangian mean theory

M. OLIVER†* and S. VASYLKEVYCH‡

†School of Engineering and Science, Jacobs University, 28759 Bremen, Germany
‡Center for Earth System Research and Sustainability (CEN), Universität Hamburg, 20146 Hamburg, Germany

(*July 1, 2019*)

Generalized Lagrangian mean theories are used to analyze the interactions between mean flows and fluctuations, where the decomposition is based on a Lagrangian description of the flow. A systematic geometric framework was recently developed by Gilbert and Vanneste (J. Fluid Mech., 2018, **839**) who cast the decomposition in terms of intrinsic operations on the group of volume preserving diffeomorphisms or on the full diffeomorphism group. In this setting, the mean of an ensemble of maps can be defined as the Riemannian center of mass on either of these groups. We apply this decomposition in the context of Lagrangian averaging where equations of motion for the mean flow arise via a variational principle from a mean Lagrangian, obtained from the kinetic energy Lagrangian of ideal fluid flow via a small amplitude expansion for the fluctuations.

We show that the Euler-α equations arise as Lagrangian averaged Euler equations when using the L^2-geodesic mean on the volume preserving diffeomorphism group of a manifold without boundaries, imposing a "Taylor hypothesis", which states that first order fluctuations are transported as a vector field by the mean flow, and assuming that fluctuations are statistically nearly isotropic. Similarly, the EPDiff equations arise as the Lagrangian averaged Burgers' equations using the same argument on the full diffeomorphism group. A serious drawback of this construction is that the assumptions of Lie transport of the fluctuation vector field and isotropy of fluctuations cannot persist except for an asymptotically vanishing interval of time. To remedy the problem of persistence of isotropy, we suggest adding strong mean-reverting stochastic term to the Taylor hypothesis and identify a scaling regime in which the inclusion of the stochastic term leads to the same averaged equations up to a constant.

Keywords: Geodesic flow, flow on manifolds, groups of diffeomorphisms, Lagrangian averaging, geodesic mean, Euler-α equations, EPDiff equations

1. Introduction

Averaging, in particular the description of the time evolution of averaged quantities, is a perennial theme in fluid dynamics. The motivation derives from two initially disconnected themes: first, the necessity to model turbulent flows in terms of Reynolds averaging or large-eddy simulation (see, e.g., Alfonsi 2009 and Sagaut 2006 for surveys and detailed references) and second, the study of wave-mean flow interactions (see, e.g., Bühler 2014, and references therein).

While much of the theory and simulation of turbulence uses a decomposition into mean and fluctuations (or coarse scale and fine scale structure) in the Eulerian description of the flow, the wave-mean flow community has looked at the problem from a Lagrangian point of view for a long time. In particular, Andrews and McIntyre (1978a) formulated a framework, the Generalized Lagrangian Mean (GLM), which leads to nonlinear equations of motion for a suitably defined Lagrangian mean of an ensemble of flows. It has since become a central

*Corresponding author. Email: m.oliver@jacobs-university.de

ingredient for the theory of wave-mean flow interactions.

The idea of employing a Reynolds-type decomposition into mean flow and turbulent fluctuations in the Lagrangian description of the fluid was initially developed by Holm (1999) and Marsden and Shkoller (2001, 2003) who, under certain closure assumptions, obtain the Euler-α (also known as the Lagrangian averaged Euler) equations as the resulting mean flow model. Soward and Roberts (2008), also see Roberts and Soward (2009), obtain a similar, but not identical set of equations using a different variational principle.

A recent paper by Gilbert and Vanneste (2018) clarifies two crucial aspects about Lagrangian mean theories. First, such theories can only be fully consistent when they are written in geometrically intrinsic terms: as noted by McIntyre (1988), the Andrews and McIntyre (1978a) generalized Lagrangian mean of a divergence free vector field is generally not divergence free. Thus, GLM theories should be formulated intrinsically. (We note that this has been done in the work of Holm as well as Marsden and Shkoller, without spelling out the general framework explicitly.) Second, and most crucially, Gilbert and Vanneste point out that once the notion of averaged map is specified, for example as the Riemannian center of mass of an ensemble of maps, the fluctuations of an ensemble of maps are fully determined by an ensemble of vector fields; the maps can be reconstructed by integration along geodesics on the group of maps. This observation let Oliver (2017) reconsider the derivation of the Euler-α equations and find that, for flows in Euclidean space, it can be based on the following minimal set of assumptions:

(a) The averaged map is the minimizer of L^2-geodesic distance,
(b) first order fluctuations are statistically nearly isotropic, and
(c) first order fluctuations are transported by the mean flow as a vector field.

Hypothesis (c) was already used by Marsden and Shkoller (2001, 2003) who refer to it as the "generalized Taylor hypothesis". The second order closure stated by Marsden and Shkoller (2003) is not assumed, but arises as a necessary consequence of the geometric notion of averaged map (a) together with (b) and (c). Therefore, only the assumption of isotropy of fluctuations (b) and the first order closure (c) are modeling hypotheses which requires theoretical or empirical verification.

In this paper, we show that these ideas extend to flows on manifolds without boundaries and can be formulated in fully intrinsic terms. We also show that the same concept extends to the derivation of the EPDiff equations as the Lagrangian averaged Burgers' equations. The term "EPDiff equations" was introduced by Holm and Marsden (2005); the system is also known as averaged template matching equations (Hirani *et al.* 2001) and n-dimensional Camassa–Holm equations (Gay-Balmaz 2009). In both cases, the key ingredient leading to a fully intrinsic derivation is the correct interpretation of isotropy in the context of a non-flat manifold. It turns out that setting the fluctuation covariance tensor to be a multiple of the inverse metric tensor results in all curvature-induced terms in the average Lagrangian combine into the Ricci Laplacian.

An important issue for the derivation is the question of mutual consistency of assumptions (a)–(c). It turns out that Lie transport of the fluctuation vector field by the mean flow is compatible with the assumption of isotropy only over short time intervals: initially isotropic fluctuations will develop an anisotropic component at a rate proportional to the deformation tensor of a mean flow velocity. In all of the references mentioned above, restoration of isotropy is assumed to happen by some process outside of the modeling framework, or is simply not discussed. Following a suggestion by J. Vanneste (personal communication), we show that restoration of isotropy can be included into formulation of the "Taylor hypothesis" as a mean-reverting stochastic term akin to an Ornstein–Uhlenbeck process. We identify the correct scaling so that the mean reversion is sufficiently strong, yet introduces only another copy of the deterministic second-order term to the averaged Lagrangian. In other words, the stochastic modification changes only a numerical coefficient in the final result, yet ensures a fully consistent set of assumptions for as long as the mean flow remains uniformly smooth.

The significance of our results is twofold. First, nontrivial manifolds such as the sphere or spherical shells naturally arise in geophysical fluid dynamics. Second, it shows that the result of Oliver (2017) is structurally robust and not tied to special properties of Euclidean geometry. Thus, for the first time, we have achieved a fully intrinsic derivation of the Euler-α equations on non-Euclidean manifolds and have identified a possible self-consistent set of modeling assumptions.

We make no claim about the validity of the deterministic or stochastic Taylor hypothesis and about the usefulness of the Euler-α equations as a momentum closure for turbulence. Our intent here is to clarify the minimal ingredients necessary to achieve a consistent derivation. Within this framework, a computational verification of these ingredients appears feasible since only the dynamics of first order fluctuations would need to be tracked. We note that in the deterministic version of the Taylor hypothesis the ensemble and ensemble mean are not defined *a priori* and still need to be specified before any such verification, whereas in the stochastic version the required specification is already made.

Our stochastic Taylor hypotheses even provides an explicit notion of an ensemble mean, for the deterministic version, the ensemble is not defined *a priori*, so that further specifics would need to be determined.

The remainder of the paper is organized as follows. In section 2, we recall some basic notions from differential geometry and the variational framework leading to the Euler, the Euler-α, Burgers', the EPDiff, and the Camassa–Holm equations. Section 3 defines the geodesic mean of an ensemble of maps. In section 4, we explain the concept of Lagrangian averaging, largely following the setup of Marsden and Shkoller (2001). The main closure assumption, the generalized Taylor hypothesis, is introduced and applied to the variational principle in section 5. The following section 6 shows that this closure, under the assumption of statistical near-isotropy and using the L^2-geodesic mean on the full diffeomorphism group, implies the Euler-α or EPDiff equations when considering, respectively, the group of volume preserving diffeomorphisms (which we will often shorten to “volumorphisms”) or the full diffeomorphism group as underlying configuration manifold. For the Euler-α equations, it is arguably more natural to use the geodesic mean with respect to volume preserving geodesics, consistent with its underlying configuration manifold. This constraint introduces an additional fictitious pressure term. In section 7, we demonstrate that this additional term does not contribute to the final averaged Lagrangian. For the sake of completeness, section 8 recalls the derivation of the Euler-α and the EPDiff equations as the Euler–Poincaré equations of the averaged Lagrangian. In section 9, we discuss mutual consistency of the closure assumptions and show that the deterministic version of the Taylor hypothesis does not preserve isotropy over order-one times. A possible fix via a stochastic modification of the Taylor hypothesis is proposed in section 10. We show that this modification only changes a numerical factor in the resulting equations of motion and discuss the relaxation time scale in comparison with turbulent eddy lifetimes. In section 11, we reformulate the closure in terms of pseudomomentum. Finally, in section 12, we briefly discuss complications arising from boundaries.

2. Notation and preliminaries

Let Ω denote n-dimensional Euclidean space or a compact n-dimensional Riemannian manifold without boundary, let g be a metric tensor on Ω with inverse g^{-1} whose components in a coordinate frame are written g_{ij} and g^{ij}, respectively, and let $\mu = \sqrt{|g|}\,\mathrm{d}x$ be the volume form on Ω induced by the metric.

Let d and δ denote, respectively, the exterior derivative and the co-differential associated with g. We write ∇ for the Levi–Civita connection on (Ω, g) and ∇_v for the covariant derivative in the direction of the vector field v. Our conventions for the Riemannian and Ricci curvature tensors, correspondingly R and Ric, are

$$R(u,v)w \equiv \nabla_u \nabla_v w - \nabla_v \nabla_u w - \nabla_{[u,v]} w\,, \tag{1a}$$

$$\operatorname{Ric}(v,w) = \operatorname{Tr}(u \mapsto R(u,v)w) \tag{1b}$$

for arbitrary vector fields u, v, and w on Ω, or, expressed in charts,

$$[R(u,v)w]^i = R^i_{jkl} u^j v^k w^l\,, \tag{1c}$$

$$\operatorname{Ric}_{kl} = R^i_{ikl}\,, \tag{1d}$$

where summation on repeated indices is implied in accordance with Einstein's convention and Tr denotes the trace operator.

In the manifold context, it is necessary to distinguish between different Laplace operators. Our reference here is Chow *et al.* (2006). The *rough Laplacian* $\tilde{\Delta} = -\nabla^* \nabla$, where ∇^* is the L^2 adjoint of ∇, takes the form

$$\tilde{\Delta} T = g^{ij} \left(\nabla_{e_i} \nabla_{e_j} - \nabla_{\nabla_{e_i} e_j} \right) T \tag{2}$$

for an arbitrary tensor T. The *Hodge Laplacian* on vector fields is given by

$$\Delta u = [-(\mathrm{d}\delta + \delta \mathrm{d}) u^\flat]^\sharp\,, \tag{3}$$

where $\flat$ is a natural isomorphism between vector fields and 1-forms associate to g and $\sharp = \flat^{-1}$. We recall that, by the Weitzenböck formula (Petersen 2016, Gay-Balmaz and Ratiu 2005),

$$g(\Delta u, v) = g(\tilde{\Delta} u, v) - \operatorname{Ric}(u,v)\,. \tag{4}$$

Finally, we write Δ_R to denote the *Ricci Laplacian*,

$$g(\Delta_R u, v) = g(\tilde{\Delta} u, v) + \operatorname{Ric}(u,v)\,. \tag{5}$$

We remark that in Euclidean space, the differences between $\tilde{\Delta}$, Δ, and Δ_R vanish. The impact of different choices of the Laplacian for modeling viscous fluid flow on manifolds is discussed in Gilbert *et al.* (2014).

We write $\mathcal{D}(\Omega)$ to denote the group of H^s-class diffeomorphisms of Ω and $\mathcal{D}_\mu(\Omega)$ its volume preserving subgroup. We recall that a map is of Sobolev class H^s whenever all of its partial derivatives up to order s are square integrable, so that

$$\mathcal{D}(\Omega) = \{\eta \in H^s(\Omega,\Omega) \colon \eta \text{ is bijective}, \eta^{-1} \in H^s(\Omega,\Omega)\}\,. \tag{6}$$

For $s > n/2 + 1$, these groups are smooth infinite dimensional manifolds in the H^s-topology (Palais 1968, Ebin and Marsden 1970) with tangent spaces at the identity

$$V = \{u \in H^s(\Omega, T\Omega) \colon u(x) \in T_x\Omega \text{ for } x \in \Omega\}\,, \tag{7a}$$

$$V_{\mathrm{div}} = \{u \in V \colon \operatorname{div} u = 0\}\,. \tag{7b}$$

We write $\eta = \eta(x,t)$ to denote the flow of a time-dependent vector field $u(t,\,\cdot\,) \in V$, so that

$$\partial_t \eta(x,t) = u(\eta(x,t),t) \tag{8}$$

or $\dot{\eta} = u \circ \eta$ for short. In this setting, the equations of motions for many continuum theories can be viewed as geodesic motion on one of the diffeomorphism groups with respect to a particular choice of metric. In other words, u is a solution whenever its associated flow η is a stationary point of the action

$$S = \int_{t_1}^{t_2} L(\dot{\eta}, \eta)\,\mathrm{d}t \tag{9}$$

with respect to variations $\delta\eta$ that are fixed at the temporal end points. In the context of this paper, we discuss the following four cases.

As pointed out by Arnold (1966), the *Euler equations* for the motion of an ideal incompressible fluid,

$$\dot{u} + \nabla_u u + \nabla p = 0\,, \tag{10a}$$

$$\operatorname{div} u = 0\,, \tag{10b}$$

where $\nabla p \equiv \mathrm{d}p^\sharp$, are the equations for geodesic flow on $\mathcal{D}_\mu$ with respect to the L^2-metric

$$(u \circ \eta, v \circ \eta)_0 = \int_\Omega g(u,v)\,\mu(x)\,. \tag{11}$$

I.e., u is a solution of (10) whenever η is a stationary point of the action (9) with Lagrangian

$$L(\dot\eta,\eta) = \tfrac{1}{2}\int_\Omega g(u,u)\,\mu(x) \equiv \tfrac{1}{2}\int_\Omega |u|^2\,\mu(x)\,, \tag{12}$$

where $u \subset V_{\text{div}}$ and $\dot\eta \subset T\mathcal{D}_\mu$ are related by (8).

Similarly, *Burgers' equations*

$$\dot{u} + \nabla_u u + (\nabla u)^T \cdot u + u \operatorname{div} u = 0\,, \tag{13}$$

where the $(1,1)$-tensor $(\nabla u)^T$ is defined as the adjoint of ∇u via

$$g((\nabla u)^T \cdot v, w) \equiv g(\nabla_w u, v) \tag{14}$$

for vector fields $u, v, w \in V$, is equivalent to the same variational problem with Lagrangian (12), albeit with configuration space $\mathcal{D}$ rather than $\mathcal{D}_\mu$; see Vizman (2008).

The *Euler-α equations*

$$\dot{m} + \nabla_u m + (\nabla u)^T \cdot m + \nabla p = 0\,, \tag{15a}$$

$$m = u - \varepsilon^2\,\Delta_R u\,, \tag{15b}$$

$$\operatorname{div} u = 0 \tag{15c}$$

are the equations for geodesic flow on the volume-preserving diffeomorphism group $\mathcal{D}_\mu$ with respect to a right-invariant H^1-metric. Their solutions are extremizers of the action S upon replacing the L^2-Lagrangian (12) by

$$L = \tfrac{1}{2}\int_\Omega \left(|u|^2 + 2\,\varepsilon^2\,|\operatorname{Def} u|^2\right)\mu(x)\,, \tag{16}$$

where $\operatorname{Def} u$ is the deformation tensor

$$\operatorname{Def} u = \tfrac{1}{2}(\nabla u + \nabla u^T) \tag{17}$$

and $|\operatorname{Def} u|^2 = g(\operatorname{Def} u, \operatorname{Def} u)$ is defined by extending metric g to arbitrary $(1,1)$-tensors S and R via

$$g(S,R) \equiv \operatorname{Tr}(S^T \cdot R) = g_{ij}\,g^{kl}\,S^i_k\,R^j_l\,, \tag{18}$$

where $[X \cdot Y]^i_j = X^i_k\,Y^k_j$ denotes the contraction of tensors X and Y. The Euler-α equations appeared first in Holm *et al.* (1998) and, in their viscous form, in Chen *et al.* (1999).

Finally, the *EPDiff equations*

$$\dot{m} + \nabla_u m + (\nabla u)^T \cdot m + m \operatorname{div} u = 0\,, \tag{19a}$$

$$m = u - \varepsilon^2\,\Delta_R u\,, \tag{19b}$$

first introduced in Hirani *et al.* (2001), describe geodesic flow on the full diffeomorphism group $\mathcal{D}$ with respect to the right-invariant H^1-metric

$$(u \circ \eta, v \circ \eta)_1 = \int_\Omega \left[|u|^2 + \varepsilon^2 \left(|\nabla u|^2 - \operatorname{Ric}(u,u) \right) \right] \mu(x) \,. \tag{20}$$

Thus, solutions to (19) are extremizers of the action S corresponding to the Lagrangian

$$L(\dot\eta, \eta) = \tfrac{1}{2} \int_\Omega \left[|u|^2 + \varepsilon^2 \left(|\nabla u|^2 - \operatorname{Ric}(u,u) \right) \right] \mu(x) \tag{21}$$

on $\mathcal{D}$. For the sake of completeness, we sketch the derivation of Euler-Poincaré equations (15) and (19) from their respective Lagrangians in section 8. For missing details, we refer the reader to Holm *et al.* (1998), Shkoller (2002), and Gay-Balmaz and Ratiu (2005). A comprehensive overview of variational principles in fluid mechanics is contained in Holm *et al.* (2009) and Badin and Crisciani (2018).

We note that Green's formulae for vector fields $u, v \in V$, in the absence of boundaries, read

$$2 \int_\Omega g(\operatorname{Def} u, \operatorname{Def} v) \, \mu(x) = - \int_\Omega g(\Delta_R u + \nabla \operatorname{div} u, v) \, \mu(x) \tag{22a}$$

and

$$\int_\Omega g(\nabla u, \nabla v) \, \mu(x) = - \int_\Omega g(\tilde\Delta u, v) \, \mu(x) \tag{22b}$$

(see, for example, Gay-Balmaz and Ratiu 2005). Combining these identities with (5), we see that the Euler-α Lagrangian (16) and the EPDiff Lagrangian (21) take the common form

$$L = \tfrac{1}{2} \int_\Omega g(u - \varepsilon^2 \, \Delta_R u, u) \, \mu(x) \,, \tag{23}$$

the difference being that $u \in V_{\mathrm{div}}$ for the Euler-α equations and $u \in V$ for EPDiff. In section 8, we sketch the derivation of the Euler-α and the EPDiff equations from the Lagrangian in the form (23).

On manifolds with boundaries, the two Lagrangians differ and the expressions stated represent their most common form, for the Euler-α equations, e.g., in Marsden and Shkoller (2001), and for the EPDiff equations in Hirani *et al.* (2001) and Gay-Balmaz (2009).

We finally remark that the EPDiff equations on S^1 or $\mathbb{R}$ reduce to the peakon version of the Camassa–Holm equation (see, e.g., Camassa and Holm 1993),

$$u_t - \varepsilon^2 \, u_{xxt} = -3 \, u \, u_x + 2 \, \varepsilon^2 \, u_x \, u_{xx} + \varepsilon^2 \, u \, u_{xxx} \,. \tag{24}$$

3. Geodesic mean

Let $\{\beta\}$ be an abstract set indexing the realizations in an ensemble of fluid flows on Ω, ε be a small parameter, and $u_{\beta,\varepsilon} = u_{\beta,\varepsilon}(x,t)$ denote the velocity field corresponding to a single realization from the ensemble. It generates a flow $\eta_{\beta,\varepsilon} = \eta_{\beta,\varepsilon}(x,t)$ via

$$\dot\eta_{\beta,\varepsilon} = u_{\beta,\varepsilon} \circ \eta_{\beta,\varepsilon} \tag{25}$$

with initial condition $\eta_{\beta,\varepsilon}|_{t=0} = \mathrm{id}$. Now suppose that the realizations can be decomposed into a averaged flow η and a fluctuating part $\xi_{\beta,\varepsilon}$ via

$$\eta_{\beta,\varepsilon} = \xi_{\beta,\varepsilon} \circ \eta \,, \tag{26}$$

Both $\xi_{\beta,\varepsilon} = \xi_{\beta,\varepsilon}(x,t)$ and $\eta = \eta(x,t)$ are again time-dependent maps. Note that all quantities in (26), and any quantities derived from them, also depend on the small parameter ε. In this notation, which follows Marsden and Shkoller (2001), ε is double-used as a formal expansion parameter *only* on the quantities which carry ε as a subscript. Consistency is ensured by requiring that (26) be satisfied, which generally implies an implicit ε-dependence of η. In the following, we suppose that η is generated by a mean velocity field $u = u(x,t)$ via

$$\dot{\eta} = u \circ \eta \tag{27}$$

where $\eta|_{t=0} = \mathrm{id}$. When $u_{\beta,\varepsilon} \in V$, then $\eta_{\beta,\varepsilon} \in \mathcal{D}$ and $\eta \in \mathcal{D}$. When $u_{\beta,\varepsilon} \in V_{\mathrm{div}}$, then $\eta_{\beta,\varepsilon} \in \mathcal{D}_\mu$. In this case, we seek mean flows $\eta \in \mathcal{D}_\mu$ that are also volume preserving.

Gilbert and Vanneste (2018) point out that flow maps $\eta_{\beta,\varepsilon}$ are points on the infinite dimensional group $\mathcal{D}(\Omega)$ or $\mathcal{D}_\mu(\Omega)$, hence it is possible to define the average map η intrinsically, by utilizing the underlying geometric structure on the group. They discuss several constructions for defining such averages. Referring to them, we will use the umbrella term *geometric GLM*. From among them, we select two that remain fully within the variational framework laid out in section 2: the Riemannian center of mass, also known as the *Fréchet mean*, of $\{\eta_{\beta,\varepsilon}\}$ on $\mathcal{D}(\Omega)$ or $\mathcal{D}_\mu(\Omega)$. We recall the details of the construction below.

Suppose that we have a procedure $\langle\,\cdot\,\rangle$ for averaging scalar quantities over the set β which commutes with spatial integration. The precise definition does not matter so long as the closure assumptions, which we will introduce in the following sections, are satisfied with respect to the induced notion of the mean. Then, the mean map η on $\mathcal{D}(\Omega)$ is defined as the Fréchet mean

$$\eta = \operatorname*{arg\,min}_{\phi \in \mathcal{D}(\Omega)} \langle d_\varepsilon^2(\phi, \eta_{\beta,\varepsilon}) \rangle \,, \tag{28a}$$

where d_ε is a Riemannian distance function. In principle, the choice of metric is not unique. However, we use the L^2-metric for the reason that it corresponds to the setting in which the Euler equations and Burgers' equations, respectively, describe geodesic flow. Thus, the geodesic distance between two maps $\phi, \psi \in \mathcal{D}(\Omega)$ is given by

$$d_\varepsilon^2(\phi,\psi) = \inf_{\substack{\gamma_s \colon [0,\varepsilon] \to \mathcal{D} \\ \gamma_0 = \phi\,, \gamma_1 = \psi}} \int_0^\varepsilon \int_\Omega g(\gamma_s', \gamma_s')\, \mu(x)\, \mathrm{d}s \,. \tag{28b}$$

Here and in the following, the prime symbol denotes a derivative with respect to s, which we think of as an arclength-like parameter. Thus, the scaling introduced into (28b) indicates that we will consider small fluctuations lying on a sphere of Riemannian radius $\mathrm{O}(\varepsilon)$ about the mean. In the terminology introduced by Gilbert and Vanneste (2018), this notion of mean is called *extended GLM*. They show that a single realizations $\eta_{\beta,\varepsilon}$ is reached from η by integrating the transport equation

$$w_{\beta,s}' + \nabla_{w_{\beta,s}} w_{\beta,s} = 0 \,, \tag{29a}$$

in fictitious time s from $s = 0$ to $s = \varepsilon$, together with a constraint on the initial condition,

$$\langle w_{\beta,s} \rangle \big|_{s=0} = 0 \,. \tag{29b}$$

The geodesic $\eta_{\beta,s}$ connecting η and $\eta_{\beta,\varepsilon}$ then is the curve in $\mathcal{D}(\Omega)$ satisfying

$$\eta_{\beta,s}' = w_{\beta,s} \circ \eta_{\beta,s} \tag{30}$$

with the initial condition $\eta_{\beta,s}|_{s=0} = \eta$.

When the configuration space is the volumorphism group $\mathcal{D}_\mu(\Omega) \subset \mathcal{D}(\Omega)$, there are two options to define the mean. We can either use the Fréchet mean with the Riemannian distance inherited from $\mathcal{D}$,

$$\eta = \operatorname*{arg\,min}_{\phi \in \mathcal{D}_\mu(\Omega)} \langle d_\varepsilon^2(\phi, \eta_{\beta,\varepsilon}) \rangle \,, \tag{31}$$

or use Riemannian distance intrinsic to $\mathcal{D}_\mu$, so that

$$\eta = \operatorname*{arg\,min}_{\phi \in \mathcal{D}_\mu(\Omega)} \langle d^2_{\varepsilon,\mu}(\phi, \eta_{\beta,\varepsilon}) \rangle \tag{32a}$$

with

$$d^2_{\varepsilon,\mu}(\phi,\psi) = \inf_{\substack{\gamma_s \colon [0,\varepsilon] \to \mathcal{D}_\mu \\ \gamma_0 = \phi\,, \gamma_1 = \psi}} \int_0^\varepsilon \int_\Omega g(\gamma_s', \gamma_s')\, \mu(x)\, \mathrm{d}t\,. \tag{32b}$$

In the first case, *optimal transport* in the terminology of Gilbert and Vanneste (2018), the fluctuation vector fields $w_{\beta,s}$ satisfy the same transport equation (29a) together with the constraint on the initial condition

$$\langle w_{\beta,s} \rangle|_{s=0} = \nabla \psi \tag{33}$$

for some scalar function ψ. In the second case, termed *geodesic*, the fluctuation vector fields satisfy an incompressible Euler equation in fictitious time s,

$$w_{\beta,s}' + \nabla_{w_{\beta,s}} w_{\beta,s} + \nabla \phi_{\beta,s} = 0\,, \tag{34a}$$

$$\operatorname{div} w_{\beta,s} = 0\,, \tag{34b}$$

with initial conditions constrained by (29b). As we shall demonstrate, both choices lead to the same averaged Lagrangian. This is consistent with the findings of Gilbert and Vanneste (2018) who showed, using a direct asymptotic expansion, that optimal transport and geodesic averaging lead to identical mean velocity fields up to the second order.

4. Lagrangian averaging of geodesic flows on diffeomorphism groups

The advantage of using geometric GLM as the definition of mean flow is that the averaged equations inherit material conservation laws from the underlying system, while retaining geometric constraints such as incompressibility. Gilbert and Vanneste (2018) derive averaged equations of motion by using the map ξ_β to pull back the momentum one-form to the the mean flow, then applying averaging. They also show that the averaged equations arise from an averaged variational principle in the spirit of Salmon (2013b). The resulting equations still need modeling in the form of a relation between the averaged momentum one-form and the mean velocity. In this paper, we first average the underlying system Lagrangian over the set of fluctuations to second order in a small fluctuation expansion and then compute the Euler–Poincaré equations from the resulting averaged Lagrangian. This approach has been pioneered by Holm (1999) and Marsden and Shkoller (2001, 2003) without reference to the concept of geodesic mean. Our approach differs from the earlier works in that we average the Lagrangian over an ensemble of fluctuations around the Riemannian center of mass while Marsden and Shkoller average over a set of initial conditions. It should be noted that both approaches assume an abstract ensemble-dependent averaging procedure, for scalars in the present article and for vector fields in the treatment of Marsden and Shkoller, the details of which need to be specified additionally. The stochastic approach in section 10 provides an example of such a procedure.

We proceed perturbatively, with the amplitude of fluctuations ε as small parameter. It is convenient to work in the Eulerian representation. Let $L_\varepsilon \equiv L(\eta_{\beta,\varepsilon}, \dot\eta_{\beta,\varepsilon})$ denote the L^2-Lagrangian for the Euler equations or Burgers' equations for a *single realization* of the flow, defined, respectively, on $\mathcal{D}_\mu$ or $\mathcal{D}$. We treat both cases in parallel, pointing out important differences along the way. We recall the underlying kinetic energy Lagrangian,

$$L_\varepsilon = \tfrac{1}{2} \int_\Omega g(u_{\beta,\varepsilon}, u_{\beta,\varepsilon})\, \mu(x)\,, \tag{35}$$

and expand u in powers of ε, writing

$$u_{\beta,\varepsilon} = u + \varepsilon\, u'_\beta + \tfrac{1}{2}\,\varepsilon^2\, u''_\beta + \mathrm{O}(\varepsilon^3)\,. \tag{36}$$

Note that, to simplify notation, we read the absence of the index ε as evaluation at $\varepsilon = 0$ so that, in particular, $w_\beta \equiv w_{\beta,s}|_{s=0} = w_{\beta,\varepsilon}|_{\varepsilon=0}$. Then,

$$\begin{aligned} L_\varepsilon &= \tfrac{1}{2} \int_\Omega \bigl[|u|^2 + 2\,\varepsilon\, g(u, u'_\beta) + \varepsilon^2\,(|u'|^2 + g(u, u''_\beta)\bigr]\, \mu(x) + \mathrm{O}(\varepsilon^3) \\ &\equiv L_0 + \varepsilon\, L_1 + \tfrac{1}{2}\,\varepsilon^2\, L_2 + \mathrm{O}(\varepsilon^3)\,. \end{aligned} \tag{37}$$

Truncating terms at $\mathrm{O}(\varepsilon^2)$ and taking the average, we introduce an *averaged Lagrangian* $\bar{L}$,

$$\begin{aligned} \bar{L} &\equiv \tfrac{1}{2} \biggl\langle \int_\Omega \bigl[|u|^2 + 2\,\varepsilon\, u \cdot u'_\beta + \varepsilon^2\,\bigl(|u'_\beta|^2 + g(u, u''_\beta)\bigr)\bigr]\, \mu(x) \biggr\rangle \\ &= \tfrac{1}{2} \int_\Omega \bigl[|u|^2 + 2\,\varepsilon g(\, u, \langle u'_\beta \rangle) + \varepsilon^2\,\bigl(\langle |u'_\beta|^2 \rangle + g(u, \langle u''_\beta \rangle)\bigr)\bigr]\, \mu(x)\,. \end{aligned} \tag{38}$$

This form of the averaged Lagrangian needs closure, i.e., we need to express the averaged quantities in terms of mean quantities. To so so, we first note that s-derivatives of $u_{\beta,s}$ are not independent of the perturbation vector fields $w_{\beta,s}$. Indeed, recall that

$$\dot\eta_{\beta,s} = u_{\beta,s} \circ \eta_{\beta,s} \tag{39}$$

with the initial condition $\eta_{\beta,s}|_{t=0} = \mathrm{id}$. Differentiating (30) with respect to t, (39) with respect to s and equating the resulting mixed partial derivatives, we obtain

$$u'_{\beta,s} = \dot w_{\beta,s} + \nabla_{u_{\beta,s}} w_\beta\, - \nabla_{w_{\beta,s}} u_{\beta,s}\, = \dot w_{\beta,s} + \mathcal{L}_{u_{\beta,s}} w_{\beta,s}\,, \tag{40}$$

where we write $\mathcal{L}_u w$ to denote the Lie derivative of the vector field w in the direction of u. Differentiating (40) and evaluating at $\varepsilon = 0$, we obtain the following expressions for the coefficients of the $u_{\beta,\varepsilon}$-expansion in terms of the fluctuation vector fields w_β:

$$u'_\beta = \dot w_\beta + \mathcal{L}_u w_\beta\,, \tag{41a}$$

$$u''_\beta = \dot w'_\beta + \mathcal{L}_u w'_\beta + \mathcal{L}_{u'} w_\beta\,. \tag{41b}$$

These relations show that once a notion of mean map is imposed, represented by (28), (31), or (32), the problem remains in need of a single closure condition: we are still free to choose an evolution equation for the *first order* fluctuation vector field w_β. This will be discussed in the next section.

5. Generalized Taylor hypothesis

We choose a closure condition in the form

$$\dot w_\beta + \mathcal{L}_u w_\beta = 0\,. \tag{42}$$

The expressions for the first and second order fluctuations of the velocity field (41) then reduce to

$$u'_\beta = 0\,, \tag{43a}$$

$$u''_\beta = \dot w'_\beta + \mathcal{L}_u w'_\beta\,. \tag{43b}$$

Up until this point the procedure for Euler and Burgers' equation was completely identical and it did not matter whether the map averaging is defined by (28), (31), or (32). In all cases, the average Lagrangian is given by (38) and the expansion vector fields are expressed in terms of fluctuations by (43). In the following, we make a choice that allows for further simultaneous treatment of the Euler equations and Burgers' equations. Below, we only assume that the fluctuation vector fields satisfy the transport equation (29a). This is compatible with both definitions of the mean map, equations (28) and (31). The case when the mean map is defined by (32) is considered in section 7.

To simplify notation, we drop the β indexes from now on, writing e.g. u' for u'_β and w for w_β, as no confusion can result from such simplification. Further, differentiating (29a) in time, setting $\varepsilon = 0$, and substituting for $\dot{w}$ from (42), we can eliminate $\dot{w}'$ from (43b) to obtain

$$u'' = \nabla_w(\mathcal{L}_u w) + \nabla_{\mathcal{L}_u w} w - \mathcal{L}_u(\nabla_w w)\,. \tag{44}$$

Regrouping terms and recalling the standard geometric identity

$$\mathcal{L}_u w \equiv [u, w] = \nabla_u w - \nabla_w u\,, \tag{45}$$

we further simplify (44) as follows:

$$u'' = -R(u, w)w + \nabla_{\nabla_w w} u - \nabla_w \nabla_w u\,. \tag{46}$$

Then, substituting (43a) and (46) into (38), we obtain

$$\begin{aligned} \bar{L} &= \tfrac{1}{2} \int_\Omega \bigl[|u|^2 + \varepsilon^2\, g\bigl(\langle -R(u, w)w + \nabla_{\nabla_w w} u - \nabla_w \nabla_w u\rangle, u\bigr)\bigr]\, \mu(x) \\ &\equiv L_0 + \tfrac{1}{2}\, \varepsilon^2\, L_2\,. \end{aligned} \tag{47}$$

6. Isotropy of fluctuations

The final simplification of the averaged Lagrangian L_2 comes from the near-isotropy assumption. Let $\{e_i = \partial/\partial x_i\}$ be a set of coordinate vector fields and write

$$w = w^i e_i\,. \tag{48}$$

Statistical near-isotropy of fluctuations shall be expressed by the condition

$$\langle w^i w^j \rangle = g^{ij} + \mathrm{O}(\varepsilon^a)\,, \tag{49}$$

where g^{ij} are the components of the inverse metric tensor and $a > 0$. It is important to emphasize that while the inclusion of order ε^a term in (49) does not affect the $\mathrm{O}(\varepsilon^2)$ averaged Lagrangian $\bar{L}$, it is crucial for the consistency of closure assumptions: As we shall show in section 9, strict isotropy where $\langle w^i w^j \rangle = g^{ij}$ coupled with the generalized Taylor hypothesis leads to unphysical restrictions on the mean flow. Therefore, the isotropy condition in previous works utilizing these two assumptions (e.g. Marsden and Shkoller 2003, Oliver 2017) must also be interpreted as near-isotropy in the sense of (49).

Assuming near-isotropy, we simplify the terms in (47) which contribute to the L_2-Lagrangian as follows. First, using the symmetries of the Riemannian curvature tensor and Bianchi's identity, we compute

$$\begin{aligned} g(\langle R(u, w)w\rangle\rangle, u) &= \langle g(R(w, u)u, w)\rangle \\ &= \langle g_{ij} R^i_{klm} w^k u^l u^m w^j \rangle \\ &= R^i_{ilm} u^l u^m + \mathrm{O}(\varepsilon^a) \\ &= \mathrm{Ric}(u, u) + \mathrm{O}(\varepsilon^a)\,. \end{aligned} \tag{50}$$

Second, we find by direct computation that

$$
\begin{aligned}
\langle \nabla_w \nabla_w u - \nabla_{\nabla_w w} u \rangle &= \langle w^i \, \nabla_{e_i} (w_j \nabla_{e_j} u) - w^i \, \nabla_{\nabla_{e_i}(w^j e_j)} u \rangle \\
&= \langle w^i w^j \, \nabla_{e_i} \nabla_{e_j} u + w^i \, (\nabla_{e_i} w_j) \, (\nabla_{e_j} u) - w^i w^j \, \nabla_{\nabla_{e_i} e_j} u - w^i \, \nabla_{(\nabla_{e_i} w^j) e_j} u \rangle \\
&= \langle w^i w^j \rangle \, (\nabla_{e_i} \nabla_{e_j} - \nabla_{\nabla_{e_i} e_j}) u \\
&= \tilde{\Delta} u + \mathrm{O}(\varepsilon^a) \, .
\end{aligned}
\tag{51}
$$

Noting that the right hand sides of (50) and (51) in the metric inner product with u add up to a quadratic form involving the Ricci Laplacian, see (5), we find that the averaged Lagrangian to second order in ε reads

$$
\bar{L} = \tfrac{1}{2} \int_\Omega \big[|u|^2 - \varepsilon^2 \, g(\Delta_R u, u) \big] \, \mu(x) \, . \tag{52}
$$

This is precisely the Lagrangian (23) of the EPDiff and of the Euler-α equations. The Camassa–Holm equations are the EPDiff equations on a one-dimensional manifold. For the latter, it is easier, of course, to verify the passage from (47) to (52) directly in Euclidean coordinates.

7. Intrinsic derivation of the Euler-α equations

The derivation of Euler-α equations in sections 3–6 uses the notion of mean flow arising from connecting elements of $\mathcal{D}_\mu(\Omega)$ by curves lying in $\mathcal{D}(\Omega)$. A more natural definition would use the notion of distance intrinsic to $\mathcal{D}_\mu(\Omega)$. The argument below shows that this intrinsic definition of the geodesic mean also leads to the Euler-α equations.

From now on, we assume that η is the Fréchet mean of $\eta_{\beta,\varepsilon}$ in $\mathcal{D}_\mu(\Omega)$ as specified by (32), so that the fluctuations satisfy the incompressible Euler equation (34) (see Gilbert and Vanneste 2018). The "pressure" field ϕ_ε is recovered by solving the Poisson equation

$$
\Delta \phi_\varepsilon = - \operatorname{div}(\nabla_{w_\varepsilon} w_\varepsilon) \quad \text{ in } \Omega \, , \tag{53a}
$$

which we will write as $\phi_\varepsilon = -\Delta^{-1} \operatorname{div}(\nabla_{w_\varepsilon} w_\varepsilon)$.

Assuming the Taylor hypothesis (42) and the isotropy of fluctuations (49), the calculation from sections 3–6 are modified as follows. Fluctuations now satisfy the Euler equations (34a) rather than the transport equation (29a) so that (44) is replaced by

$$
u'' = \nabla_w (\mathcal{L}_u w) + \nabla_{\mathcal{L}_u w} w - \mathcal{L}_u (\nabla_w w) - \mathcal{L}_u \nabla \phi - \nabla \dot{\phi} \, . \tag{54}
$$

Therefore, the expression of the L_2-Lagrangian derived in section 6 must be augmented with two extra terms, so that

$$
\begin{aligned}
L_2 &= - \Big\langle \int_\Omega g\big(\Delta_R u + \mathcal{L}_u \nabla \phi + \nabla \dot{\phi}, u\big) \, \mu(x) \Big\rangle \\
&= - \int_\Omega g(\Delta_R u, u) \, \mu(x) - \Big\langle \int_\Omega g(\mathcal{L}_u \nabla \phi, u) \, \mu(x) \Big\rangle \, ,
\end{aligned}
\tag{55}
$$

where the last term in the first line vanishes since gradients are L^2-orthogonal to divergence free vector fields.

We compute the last term in (55) by noting that due to the Hodge decomposition, the operator $\nabla \Delta^{-1} \operatorname{div}$ is L^2 symmetric, i.e., for arbitrary sufficiently smooth vector fields v and w,

$$
\int_\Omega g(\nabla \Delta^{-1} \operatorname{div}(v), w) \, \mu(x) = \int_\Omega g(v, \nabla \Delta^{-1} \operatorname{div}(w)) \, \mu(x) \, . \tag{56}
$$

Since u is necessarily divergence free as a vector field generating $\eta \in \mathcal{D}_\mu(\Omega)$, integrating by parts, we have

$$\begin{aligned}\int_\Omega g(\mathcal{L}_u \nabla\phi, u)\, \mu(x) &= \int_\Omega \big[g(\nabla_u \nabla\phi, u) - g(\nabla_{\nabla\phi} u, u) \big]\, \mu(x) \\ &= -\int_\Omega g\big(\nabla\phi\,, \nabla_u u + \tfrac{1}{2}\nabla|u|^2\big)\, \mu(x) \\ &= \int_\Omega g\big(\nabla_w w, \nabla\Delta^{-1} \operatorname{div}(\nabla_u u + \tfrac{1}{2}\nabla|u|^2)\big)\, \mu(x) \\ &= -\int_\Omega g\big(w, \nabla_w \nabla\Delta^{-1} \operatorname{div}(\nabla_u u + \tfrac{1}{2}\nabla|u|^2)\big)\, \mu(x)\,. \end{aligned} \tag{57}$$

For an arbitrary vector field v, discarding $\mathrm{O}(\varepsilon^\alpha)$ terms,

$$\begin{aligned}\left\langle \int_\Omega g(w, \nabla_w v)\, \mu(x) \right\rangle &= \int_\Omega g_{ij}\, \langle w^i w^k \rangle \left(\frac{\partial v^j}{\partial x_k} + \Gamma^j_{ks} v^s \right) \mu(x) \\ &= \int_\Omega \left(\frac{\partial v^j}{\partial x_j} + \Gamma^j_{js} v^s \right) \mu(x) \\ &= \int_\Omega \operatorname{div} v\, \mu(x)\,, \end{aligned} \tag{58}$$

where the last equality follows from the standard expression for the divergence of a vector field,

$$\operatorname{div} v = \frac{1}{\sqrt{g}} \frac{\partial}{\partial x_i} \big(\sqrt{g}\, v^i \big)\,. \tag{59}$$

Now, combining (57) and (58), we obtain

$$\left\langle \int_\Omega g(\mathcal{L}_u \nabla\phi, u)\, \mu(x) \right\rangle = -\int_\Omega \operatorname{div}(\nabla_u u + \tfrac{1}{2}\nabla|u|^2)\, \mu(x) = 0\,. \tag{60}$$

Substituting (60) into (55), we obtain

$$L_2 = -\int_\Omega g(\Delta_R u, u)\, \mu(x)\,, \tag{61}$$

so that the full averaged Lagrangian $\bar{L}$ coincides with the Euler-α Lagrangian (23).

8. Averaged equations of motion

In this section, we derive Euler-α equations (15) and the EPDiff equations (19) as the Euler–Poincaré equations for the averaged Lagrangian $\bar{L}$ on $\mathcal{D}_\mu(\Omega)$ and $\mathcal{D}(\Omega)$, respectively. To do so, we must compute the stationary points of the averaged action

$$\bar{S} = \int_{t_1}^{t_2} \bar{L}(\dot\eta, \eta)\, \mathrm{d}t \tag{62}$$

with respect to variations of the flow map $\delta\eta$ in the respective configuration spaces which vanish at the temporal endpoints.

First, we note that variations in the flow map $\delta\eta = w \circ \eta$ and the fluid velocity $u = \dot\eta \circ \eta^{-1}$ are related by the Lin constraint (Bretherton 1970)

$$\delta u = \dot w + \mathcal{L}_u w\,, \tag{63}$$

which is proved analogously to (41a). Next, due to the symmetry of the Ricci tensor, the averaged Lagrangian $\bar{L}$ is of the form

$$\bar{L} = \tfrac{1}{2} \int_{\Omega} g(Au, u)\, \mu(x)\,, \tag{64}$$

where

$$A = \mathrm{id} - \varepsilon^2 \Delta_R \tag{65}$$

is a linear $L^2(\Omega, g)$-self-adjoint operator on vector fields. Therefore,

$$\delta L = \int_{\Omega} g(Au, \delta u)\, \mu(x) = \int_{\Omega} g(m, \delta u)\, \mu(x)\,, \tag{66}$$

where the circulation velocity m is given by

$$m = Au = u - \varepsilon^2 \Delta_R u\,. \tag{67}$$

From this point on, while the overall strategy remains similar, the details of computation depend on the configuration space. Therefore, we will treat both cases separately.

On $\mathcal{D}(\Omega)$, w is an arbitrary vector field in V. Using (63), (66), and integration by parts, we compute

$$\begin{aligned}
\delta \bar{S} &= \int_{t_1}^{t_2} \delta \bar{L}\, \mathrm{d}t = \int_{t_1}^{t_2} \int_{\Omega} g(m, \delta u)\, \mu(x)\, \mathrm{d}t \\
&= \int_{t_1}^{t_2} \int_{\Omega} g(m, \dot{w} + \mathcal{L}_u w)\, \mu(x)\, \mathrm{d}t \\
&= \int_{t_1}^{t_2} \int_{\Omega} \bigl[-g(\dot{m}, w) + g(m, \nabla_u w) - g(m, \nabla_w u) \bigr]\, \mu(x)\, \mathrm{d}t \\
&= -\int_{t_1}^{t_2} \int_{\Omega} \bigl[g(\dot{m}, w) + g(\nabla_u m, w) - \nabla_u g(m, w) + g((\nabla u)^T m, w) \bigr]\, \mu(x)\, \mathrm{d}t \\
&= -\int_{t_1}^{t_2} \int_{\Omega} g(\dot{m} + \nabla_u m + m \operatorname{div} u + (\nabla u)^T m, w)\, \mu(x)\, \mathrm{d}t \equiv 0\,.
\end{aligned} \tag{68}$$

Since w is arbitrary, m must satisfy the EPDiff momentum equation (19a).

On $\mathcal{D}_\mu(\Omega)$, w is an arbitrary divergence-free vector field in V_{div}. Moreover, the velocity u is a curve in V_{div}. Therefore, the computation in (68) implies that

$$\int_{t_1}^{t_2} \int_{\Omega} g(\dot{m} + \nabla_u m + (\nabla u)^T m, w)\, \mu(x)\, \mathrm{d}t = 0\,. \tag{69}$$

Since, by the Hodge decomposition, the space of vector fields orthogonal to V_{div} in $L^2(\Omega, g)$ consists of gradients, the circulation velocity m satisfies the Euler-α momentum equation (15a).

9. Mutual consistency of closure assumptions

Let us recall that we derived the EPDiff and the Euler-α equations under two closure conditions for first-order fluctuations: the generalized Taylor hypothesis (42) and the near-isotropy of fluctuation condition (49). A third requirement, which also imposes a restriction on first order fluctuation, is

$$\langle w \rangle = 0 \qquad \text{or} \qquad \langle w \rangle = \nabla \psi\,. \tag{70a,b}$$

We recall that conditions (70) arise as a natural geometric constraint from the GLM construction itself, where the first alternative pertains to the Euler-α derived from the metric (32b) or to the EPDiff equations, while the second alternative pertains to the Euler-α derived from the metric (28b).

In the following, we show that our Taylor hypothesis is easily consistent with (70), but not necessarily with persistence of isotropy. Indeed, let $\mathrm{D}_t \equiv \partial_t + \nabla_u$ denote the material derivative associated with the mean flow. Assuming the Taylor hypothesis and supposing that averaging commutes with D_t, the evolution of $\langle w \rangle$ is governed by

$$\mathrm{D}_t \langle w \rangle = \langle \mathrm{D}_t w \rangle = \nabla_{\langle w \rangle} u \,, \tag{71}$$

Hence, condition (70a) persists for all times provided it holds initially. To study consistency of the alternative condition (70b) with the Taylor hypothesis, note that

$$\nabla_u \nabla \psi = \nabla(\nabla_u \psi) - (\nabla u^T) \cdot \nabla \psi + (\nabla_u g^{-1})^\flat \cdot \nabla \psi \,, \tag{72}$$

which is straightforward to check in charts. Therefore, plugging $\langle w \rangle = \nabla \psi$ into (71) yields

$$\nabla \left(\dot{\psi} + \nabla_u \psi \right) = \left[2 \operatorname{Def} u - (\nabla_u g^{-1})^\flat \right] \cdot \nabla \psi \,, \tag{73}$$

so that ψ must satisfy the consistency condition

$$\mathrm{D}_t \psi = \Delta^{-1} \operatorname{div} \left([2 \operatorname{Def} u - (\nabla_u g^{-1})^\flat] \cdot \nabla \psi \right) \,. \tag{74}$$

Since (74) has many solutions, it is possible that (70b) is consistent with the Taylor hypothesis. However, it appears that this is not guaranteed: the Taylor hypothesis alone does not imply that condition (70b) persists at all times when it holds initially. Instead, the restrictions on fluctuations imposed by (74) are in addition to those implied by the definition of the mean flow via (31).

In contrast, the notion of the mean (32), intrinsic to $\mathcal{D}_\mu$, leads to no such restrictions. In this regard, the intrinsic derivation of Euler-α provides a clear conceptual advantage. We note, however, that the set of fluctuation vector fields satisfying the non-intrinsic geometric mean conditions (29a), (33), and (74) is not a subset of vector fields satisfying the intrinsic version (34) and (29b), and vice versa. Thus, none of the two derivations is a special case of the other.

We now turn to discussing consistency of the Taylor hypothesis with the isotropy condition. Traditionally, the isotropy condition was formulated as an exact equality (see Marsden and Shkoller 2001 and Oliver 2017), which on a Riemannian manifold reads

$$\langle w^i w^j \rangle = g^{ij} \,. \tag{75}$$

However, it turns out that a literal interpretation of (75) has undesirable consequences: Tensoring the Taylor hypothesis with w on the right and on the left, adding the results and taking the average, we infer that

$$\langle (\partial_t + \mathcal{L}_u)(w \otimes w) \rangle = 0 \,. \tag{76}$$

Now, assuming that $\partial_t + \mathcal{L}_u$ commutes with averaging, the exact isotropy (75) and the requirement (76) imply that

$$\mathcal{L}_u g^{-1} = 0 \,, \tag{77}$$

i.e., u is a Killing field. This condition is clearly too restrictive for any useful model of a mean flow; for instance, it would preclude any shear flow as a mean flow on a flat manifold.

We remark that in our derivation, we do not assume that averaging and differentiation commute. In the general case, the Taylor hypothesis and exact isotropy impose the restriction

$$\langle \mathrm{D}_t (w \otimes w) \rangle = 2 \, (\operatorname{Def} u)^\sharp \,. \tag{78}$$

Indeed,

$$\langle(\partial_t + \mathcal{L}_u)(w \otimes w)\rangle = \langle \mathrm{D}_t(w \otimes w)\rangle - \langle \nabla_w u \otimes w + w \otimes \nabla_w u\rangle \,. \tag{79a}$$

Computing in charts, noting that $(\nabla u^T)^i_k = g_{km}\, g^{il}\, (\nabla u)^m_l$, and applying the isotropy condition (75), we obtain

$$\begin{aligned}\big\langle \nabla_w u \otimes w + w \otimes \nabla_w u\big\rangle^{ij} &= \big\langle w^k\, (\nabla u)^i_k\, w^j + w^i\, (\nabla u)^j_k\, w^k\big\rangle \\ &= g^{kj}\, (\nabla u)^i_k + g^{ik}\, (\nabla u)^j_k \\ &= g^{kj}\, (\nabla u)^i_k + g^{jk}\, g_{km}\, g^{il}\, (\nabla u)^m_l \\ &= (\nabla u)^\sharp + (\nabla u^T)^\sharp \\ &= 2\, (\mathrm{Def}\, u)^\sharp \,, \end{aligned} \tag{79b}$$

hence, assuming isotropy, (78) is equivalent to (76).

In principle, it could be possible to satisfy the restriction (78) by choosing a special measure for the averaging operation. However, it is easy to see that averaging commutes with both space and time differentiation in the most elementary case when the measure is independent of space and time. Thus, imposing a non-commuting averaging process is not a reasonable way out of the inconsistency between the Taylor hypothesis and the isotropy condition.

Using the relaxed version of isotropy, the near-isotropy condition (49), instead of (75), equation (78) reads

$$\langle \mathrm{D}_t(w \otimes w)\rangle = 2\, (\mathrm{Def}\, u)^\sharp + \mathrm{O}(\varepsilon^a) \,. \tag{80}$$

In this case, near-isotropy is guaranteed to persist at least on the time scale $\mathrm{O}(\varepsilon^a)$. Marsden and Shkoller (2001, 2003) combine an $\mathrm{O}(\varepsilon)$ approximate Taylor hypothesis with exact isotropy, which leads to the same result, which is not completely satisfactory, since the time of consistency reduces to zero when $\varepsilon \to 0$. We conclude that interpreting modeling assumptions approximately without further modifications of the argument cannot resolve the question of long-time consistency.

In the next section, we propose a stochastic modification of the Taylor hypothesis which maintains persistence of near-isotropy over arbitrary intervals of time. At the same time, it imposes minimal changes—no more than a factor 2 in front of the second order averaged Lagrangian—on the derivation of the Euler-α equation.

10. Stochastic Taylor hypothesis

In this section, we replace the Taylor hypothesis with a stochastic modification to obtain an alternative derivation of the Euler-α model. The advantage of this approach is that the stochastic Taylor hypothesis ensures that near isotropy persists for long time intervals without imposing additional conditions on fluctuations. Our stochastic Taylor hypothesis assumes that fluctuations w_β are independently generated realizations of the process w satisfying the stochastic differential equation

$$\mathrm{d}w + \mathcal{L}_u w\, \mathrm{d}t = -\varepsilon^{-a}\, w\, \mathrm{d}t + \sqrt{2}\, \varepsilon^{-a/2}\, \mathrm{d}\mathcal{W} \,, \tag{81}$$

where $0 < a < 1/2$ and $\mathcal{W}(t)$ is a time-dependent Wiener process with

$$\mathbb{E}[\mathrm{d}\mathcal{W} \otimes \mathrm{d}\mathcal{W}] = g^{-1}\, \mathrm{d}t \,. \tag{82}$$

Since divergence free vector fields form a Lie algebra with respect to bracket (45), (81) requires, moreover, that $\operatorname{div} \mathrm{d}\mathcal{W} = 0$.

Furthermore, we define $\langle \cdot \rangle$ to be the statistical average, so that for arbitrary $q = q(w, \mathcal{W})$,

$$\langle q_\beta \rangle \equiv \mathbb{E}[q] \,. \tag{83}$$

Note that this notion of averaging commutes both with space and time-differentiation. Hence, (81) implies persistence of condition (29) under the time evolution.

We now compute the evolution equation governing the covariance tensor

$$\kappa = \langle w_\beta \otimes w_\beta \rangle = \mathbb{E}[w \otimes w] \,. \tag{84}$$

The Itô formula (see, e.g., Jacobs 2010, section 3.8.2) implies

$$\mathrm{d}(w \otimes w) = \mathrm{d}w \otimes w + w \otimes \mathrm{d}w + \mathrm{d}w \otimes \mathrm{d}w \,, \tag{85a}$$

where, by (81),

$$\mathrm{d}w \otimes \mathrm{d}w = 2\varepsilon^{-a} \, \mathrm{d}\mathcal{W} \otimes \mathrm{d}\mathcal{W} \tag{85b}$$

and all other product of infinitesimals are $o(\mathrm{d}t)$, as is standard in Itô calculus. Further substituting $\mathrm{d}w$ from (81) into (85), taking the expectation, and abbreviating $\gamma = 2\,\varepsilon^{-a}$, we obtain altogether

$$\dot\kappa + \mathcal{L}_u \kappa = \gamma \, (g^{-1} - \kappa) \,. \tag{86}$$

Hence,

$$\kappa(t) - g^{-1} = \mathrm{e}^{-\gamma t} \, (\kappa(0) - g^{-1}) - \int_0^t \mathrm{e}^{\gamma(s-t)} \, \mathcal{L}_{u(s)} \kappa(s) \, \mathrm{d}s \,. \tag{87}$$

Thus, the near-isotropy condition (49) persists for arbitrary large times so long as the mean flow u remains uniformly smooth, which we suppose throughout this paper.

To proceed, we abbreviate the right hand side of the stochastic Taylor hypothesis (81) by $\mathrm{d}\tau$, so that

$$\mathrm{d}w = -\mathcal{L}_u w \, \mathrm{d}t + \mathrm{d}\tau \,, \tag{88}$$

and insert it into (41) to obtain

$$u' \, \mathrm{d}t = \mathrm{d}w + \mathcal{L}_u w \, \mathrm{d}t = \mathrm{d}\tau \,, \tag{89a}$$

$$u'' \, \mathrm{d}t = \mathrm{d}w' + \mathcal{L}_u w' \, \mathrm{d}t + \mathcal{L}_{\mathrm{d}\tau} w \,. \tag{89b}$$

Thus, the Lin contraint implies that $u' \, \mathrm{d}t$ and $u'' \, \mathrm{d}t$ are rough in time, with the smoothness of a Wiener increment. However, the fluctuation vector field still evolves smoothly in fictitious time s. Following the intrinsic definition of the geodesic distance, we express w' via (34a), differentiate using the Itô formula,

$$\begin{aligned} \mathrm{d}w' &= -\nabla_{\mathrm{d}w} w - \nabla_w \mathrm{d}w - \nabla_{\mathrm{d}w} \mathrm{d}w - \nabla \mathrm{d}\phi \\ &= -\nabla_{\mathrm{d}w} w - \nabla_w \mathrm{d}w - \gamma \, \nabla_{\mathrm{d}\mathcal{W}} \mathrm{d}\mathcal{W} - \nabla \mathrm{d}\phi \,, \end{aligned} \tag{90}$$

and substitute into (89b) to obtain

$$\begin{aligned} u'' \, \mathrm{d}t = {} & \big(\nabla_w \mathcal{L}_u w + \nabla_{\mathcal{L}_u w} w - \mathcal{L}_u \nabla_w w - \mathcal{L}_u \nabla \phi \big) \, \mathrm{d}t - \nabla \mathrm{d}\phi \\ & - \nabla_{\mathrm{d}\tau} w - \nabla_w \mathrm{d}\tau + \mathcal{L}_{\mathrm{d}\tau} w - \gamma \, \nabla_{\mathrm{d}\mathcal{W}} \mathrm{d}\mathcal{W} \,. \end{aligned} \tag{91}$$

We now consider the averaged Lagrangian, truncated to $\mathrm{O}(\varepsilon^2)$,

$$\mathrm{d}\bar{L}_{\text{stoch}} = \tfrac{1}{2} \left\langle \int_\Omega \big[|u|^2 \, \mathrm{d}t + 2\varepsilon \, g(u, u' \, \mathrm{d}t) + \varepsilon^2 \, \big(|u'|^2 \, \mathrm{d}t + g(u, u'' \, \mathrm{d}t) \big) \big] \, \mu(x) \right\rangle , \tag{92}$$

insert the expressions for u' and u'' from (89a) and (91), respectively, and observe that the terms from the first line of (91) combine with the O(1) term to yield the deterministic Lagrangian L_{det} by the argument from section 7. We obtain

$$\mathrm{d}\bar{L}_{\text{stoch}} = L_{\text{det}}\,\mathrm{d}t + \varepsilon\,\mathrm{d}I_1 + \tfrac{1}{2}\,\varepsilon^2\,\mathrm{d}I_2 + \tfrac{1}{2}\,\varepsilon^2\,\mathrm{d}I_3 \tag{93a}$$

where the contribution from the deterministic terms reads

$$L_{\text{det}} = -\int_\Omega g(\Delta_R, u)\,\mu(x) + \mathrm{O}(\varepsilon^{2+a}) \tag{93b}$$

as before, and the stochastic contributions read

$$\mathrm{d}I_1 = \int_\Omega g(u, \langle \mathrm{d}\tau\rangle)\,\mu(x) = 0 \tag{93c}$$

since, by (29), $\langle w\rangle = 0$ so that $\langle \mathrm{d}\tau\rangle = 0$,

$$\begin{aligned} \mathrm{d}I_2 &= \biggl\langle \int_\Omega g(\mathrm{d}\tau, \mathrm{d}\tau)\,(\mathrm{d}t)^{-1}\,\mu(x) \biggr\rangle \\ &= \biggl\langle \int_\Omega \bigl[\varepsilon^{-2a}\,|w|^2\,\mathrm{d}t + 2\,\varepsilon^{-a}\,g(\mathrm{d}\mathcal{W}, \mathrm{d}\mathcal{W})(\mathrm{d}t)^{-1}\bigr]\,\mu(x) \biggr\rangle \end{aligned} \tag{93d}$$

due to the statistical independence of w and $\mathrm{d}\mathcal{W}$, and

$$\begin{aligned} \mathrm{d}I_3 &= \biggl\langle \int_\Omega g(u, \mathcal{L}_{\mathrm{d}\tau} w - \nabla_{\mathrm{d}\tau} w - \nabla_w \mathrm{d}\tau - \gamma \nabla_{\mathrm{d}\mathcal{W}} \mathrm{d}\mathcal{W})\,\mu(x) \biggr\rangle \\ &= \biggl\langle \int_\Omega \bigl[-2\,g(u, \nabla_w \mathrm{d}\tau) + \gamma\,g(u, \mathrm{d}\mathcal{W})\,\operatorname{div}\mathrm{d}\mathcal{W} + \gamma\,g(\nabla_{\mathrm{d}\mathcal{W}} u, \mathrm{d}\mathcal{W})\bigr]\,\mu(x) \biggr\rangle \\ &= \biggl\langle \int_\Omega \bigl[\gamma\,g(\nabla_w u, w\,\mathrm{d}t) - 2\sqrt{\gamma}\,g(\nabla_w u, \mathrm{d}\mathcal{W})\bigr]\,\mu(x) \biggr\rangle = 0\,, \end{aligned} \tag{93e}$$

where we have used that $\operatorname{div}\mathrm{d}\mathcal{W} = 0$, the statistical independence of w and $\mathrm{d}\mathcal{W}$, and that $\langle g(\nabla_{\mathrm{d}\mathcal{W}} u, \mathrm{d}\mathcal{W})\rangle$ and $\langle g(\nabla_w u, w)\rangle$ vanish due to (58).

Thus, the only nontrivial stochastic contribution comes from $\mathrm{d}I_2$. The second term in (93d) is infinite, but does not depend on u, so we remove it by renormalizing the Lagrangian. The first term in (93d) can be rewritten as follows. Integrating (87) by parts, we find that

$$\begin{aligned} \kappa(t) - g^{-1} = \mathrm{e}^{-\gamma t}\,(\kappa(0) - g^{-1}) &- \frac{1}{\gamma}\,\bigl(\mathcal{L}_{u(t)}\kappa(t) - \mathrm{e}^{-\gamma t}\,\mathcal{L}_{u(0)}\kappa(0)\bigr) \\ &+ \int_0^t \frac{\mathrm{e}^{\gamma(s-t)}}{\gamma}\,\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{L}_{u(s)}\kappa(s)\,\mathrm{d}s\,. \end{aligned} \tag{94}$$

Further integration by parts shows that the integral remainder is $\mathrm{O}(\gamma^{-2})$ so long as u and κ remain smooth, which we assume throughout. Hence, taking the Lie derivative, we find that

$$\begin{aligned} \mathcal{L}_{u(t)}(\kappa(t) - g^{-1}) &= \mathrm{e}^{-\gamma t}\,\mathcal{L}_{u(t)}(\kappa(0) - g^{-1}) \\ &\quad - \frac{1}{\gamma}\,\bigl(\mathcal{L}_{u(t)}\mathcal{L}_{u(t)}\kappa(t) - \mathrm{e}^{-\gamma t}\,\mathcal{L}_{u(t)}\mathcal{L}_{u(0)}\kappa(0)\bigr) + \mathrm{O}(\gamma^{-2})\,. \end{aligned} \tag{95}$$

To simplify this expression, note that for any $f \in C^1([0,T])$ with $f(0) = 0$,

$$\mathrm{e}^{-\gamma t}\,f(t) = \mathrm{O}(\gamma^{-1})\,. \tag{96}$$

Without loss of generality, we may assume that $f > 0$. Then, for sufficiently large $\gamma > 0$, the bound

$$\sup_{t\in[0,T]} \mathrm{e}^{-\gamma t}\,f(t) \le \gamma^{-1} \sup_{t\in[0,T]} |\dot{f}(t)| \tag{97}$$

follows by maximizing $h(t) = \mathrm{e}^{-\gamma t} f(t)$. Then, using (97) and near-isotropy,

$$\mathrm{e}^{-\gamma t} \left(\mathcal{L}_{u(t)} - \mathcal{L}_{u(0)}\right) \left(\kappa(0) - g^{-1}\right) = \mathrm{O}(\gamma^{-2}) \,. \tag{98}$$

Further applying (96) to (95) with $f(t) = (\mathcal{L}_{u(t)} - \mathcal{L}_{u(0)}) \, \mathcal{L}_{u(0)}\kappa(0)$, we obtain

$$\begin{aligned} \mathcal{L}_{u(t)}(\kappa(t) - g^{-1}) &= \mathrm{e}^{-\gamma t} \, \mathcal{L}_{u(0)}(\kappa(0) - g^{-1}) \\ &\quad - \frac{1}{\gamma} \left(\mathcal{L}_{u(t)}\mathcal{L}_{u(t)}\kappa(t) - \mathrm{e}^{-\gamma t} \, \mathcal{L}_{u(0)}\mathcal{L}_{u(0)}\kappa(0)\right) + \mathrm{O}(\gamma^{-2}) \\ &= \mathrm{e}^{-\gamma t} \, \mathcal{L}_{u(0)}(\kappa(0) - g^{-1}) \\ &\quad - \frac{1}{\gamma} \left(\mathcal{L}_{u(t)}\mathcal{L}_{u(t)}g^{-1} - \mathrm{e}^{-\gamma t} \, \mathcal{L}_{u(0)}\mathcal{L}_{u(0)}g^{-1}\right) + \mathrm{O}(\gamma^{-2}) \end{aligned} \tag{99}$$

where, in the second step, we have used near-isotropy in two places. Now, taking the trace of (87), substituting $\mathcal{L}_{u(t)}\kappa(t)$ from (99), noting the standard Riemannian geometry identity

$$\operatorname{Tr} \mathcal{L}_u g^{-1} = -2 \operatorname{div} u = 0 \,, \tag{100}$$

and integrating by parts, we compute

$$\begin{aligned} \operatorname{Tr} \kappa(t) &= \operatorname{Tr} g^{-1} + \mathrm{e}^{-\gamma t} \, \operatorname{Tr}(\kappa(0) - g^{-1}) \\ &\quad - \int_0^t \mathrm{e}^{\gamma(s-t)} \left(\operatorname{Tr} \mathcal{L}_{u(s)}g^{-1} + \mathrm{e}^{-\gamma s} \, \operatorname{Tr} \mathcal{L}_{u(0)}(\kappa(0) - g^{-1}) + \mathrm{O}(\gamma^{-2})\right) \mathrm{d}s \\ &\quad + \frac{1}{\gamma} \int_0^t \mathrm{e}^{\gamma(s-t)} \, \operatorname{Tr}\left(\mathcal{L}_{u(s)}\mathcal{L}_{u(s)}g^{-1} - \mathrm{e}^{-\gamma s} \, \mathcal{L}_{u(0)}\mathcal{L}_{u(0)}g^{-1}\right) \mathrm{d}s \\ &= \frac{1}{\gamma^2} \, \operatorname{Tr} \mathcal{L}_{u(t)}\mathcal{L}_{u(t)}g^{-1} + G(t) + \mathrm{O}(\gamma^{-3}) \end{aligned} \tag{101}$$

where $G(t)$ summarizes all terms which only depend on t and on the initial conditions, hence do not contribute to the variational principle and will be discarded henceforth. Thus, inserting (101) into the expression (93d) for $\mathrm{d}I_2$ and renormalizing, we obtain

$$\begin{aligned} \mathrm{d}I_2 &= \left\langle \int_\Omega \varepsilon^{-2a} \, |w|^2 \, \mathrm{d}t \, \mu(x) \right\rangle = \varepsilon^{-2a} \int_\Omega \operatorname{Tr} \kappa(t) \, \mu(x) \, \mathrm{d}t \\ &= \int_\Omega \operatorname{Tr} \mathcal{L}_u\mathcal{L}_u g^{-1} \, \mu(x) \, \mathrm{d}t + \mathrm{O}(\varepsilon^a) \\ &= 4 \int_\Omega g(\operatorname{Def} u, \operatorname{Def} u) \, \mu(x) \, \mathrm{d}t + \mathrm{O}(\varepsilon^a) \,. \end{aligned} \tag{102}$$

The last equality is, once again, a Riemannian geometry identity which is lengthy but straightforward to verify in charts; since all the quantities involved are intrinsic, it suffices to verify this identity in an orthonormal frame.

Altogether, collecting all contributions to the stochastic variational principle and recalling the equality between (16) and (23), we obtain

$$\mathrm{d}\bar{L}_{\text{stoch}} = \tfrac{1}{2} \int_\Omega g(u - 2\,\varepsilon^2 \, \Delta_R u, u) \, \mu(x) \, \mathrm{d}t \,. \tag{103}$$

This expression is identical to the Euler-α Lagrangian (23) up to a scaling factor 2 in front of the Laplacian. The derivation above, as in the preceding sections of the paper, is formal and raises questions regarding a rigorous formulation of the Lin constraints (89) under lack of smoothness in time and the necessary renormalization of the resulting Lagrangian.

Subject to such precautions, we find that the inclusion of a strong mean-reverting term into the Taylor hypothesis does not change the equations of motion except for modifying the constant in front of the second-order α-term. The time scale of decay of perturbations toward isotropy, $T_{\rm relax}$, in the stochastic Taylor hypothesis (81) is solely determined by the coefficient of the linear deterministic dissipative term, so that

$$T_{\rm relax} \approx \varepsilon^a \tag{104}$$

where, in order to ensure the assumed ordering of terms in the expansion, $a \in (0, \frac{1}{2})$.

On the other hand, in standard scaling theories of turbulence, the eddy life time is estimated as follows (e.g., Foias *et al.* 2001). The energy of eddies between wave number k and $2k$ is given by

$$\int_k^{2k} E(k)\, \mathrm{d}k \approx k\, E(k)\,, \tag{105}$$

where $E(k)$ is the energy spectrum of the flow, so that the average velocity of those eddies is $(k\, E(k))^{1/2}$. Assuming that eddies break up in the time it takes to travel the distance of their linear size, we find that the eddy life time is given by

$$T_{\rm eddy} \approx \kappa^{-3/2}\, E(\kappa)^{-1/2}\,. \tag{106}$$

In the inertial range of three-dimensional isotropic turbulence, $E(k) \approx k^{-5/3}$, in the enstrophy range of two-dimensional turbulence, $E(k) = k^{-3}$. Thus,

$$T_{\rm eddy} \approx \begin{cases} 1 & \text{for } n = 2\,, \\ k^{-2/3} & \text{for } n = 3\,. \end{cases} \tag{107}$$

Thus, in three-dimensional turbulence, $T_{\rm relax} \gg T_{\rm eddy}$, which is consistent with physical intuition: fluctuations cannot isotropize more rapidly than eddies decay, but they may do so more slowly. In two-dimensional turbulence, $T_{\rm relax} \ll T_{\rm eddy}$. This appears to be contradicting the assumptions of our derivation, as fluctuations at scale ε cannot isotropize within the lifetime for their coherent evolution.

This conclusion is at least consistent with scattered pieces of evidence, e.g. Mohseni *et al.* (2003) who report positive results regarding the tracking of the correct energy spectrum by a particular Navier–Stokes-α model, and Graham and Ringler (2013) who report unphysical enstrophy pileup at small scales in an analogous model for rotating two-dimensional turbulence. It is possible that Euler-α-like models may be re-interpreted for two-dimensional turbulence in a completely different way, namely via consistent tracking of the associated modification of the pattern of triad-interactions in spectral space (see section 5.2 in Danilov *et al.* 2019).

We emphasize that the discussion presented here is largely speculative and in need of more solid physical arguments and careful numerical evidence beyond the scope of this paper.

11. Pseudomomentum

It is customary to formulate GLM closures in terms of pseudomomentum, which measures the effect of fluctuations on the vortical dynamics (see, e.g., Andrews and McIntyre 1978a,b, Salmon 2013a, and Bühler 2014). Following Gilbert and Vanneste (2018), pseudomomentum for geometric GLM theories is given by

$$-p = \bar{\nu}^L - u^\flat\,, \tag{108}$$

where $\bar{\nu}^L$ is the Lagrangian-mean one-form entering the Kelvin circulation theorem.

Since the Kelvin circulation theorem for a Lagrangian $\bar{L}(u)$ on a diffeomorphism group has a form

$$\frac{\mathrm{d}}{\mathrm{d}t} \oint_{\gamma_t} \frac{\delta \bar{L}}{\delta u} = 0\,, \tag{109}$$

where γ_t is a material loop, see, e.g., the Euler–Poincaré theorem for continua in Holm *et al.* (1998), it follows that

$$\bar{\nu}^L = \frac{\delta \bar{L}}{\delta u} = m^\flat \,, \tag{110}$$

where m is the circulation velocity defined in section 8. Thus,

$$p = \varepsilon^2 \, (\Delta_R u)^\flat \tag{111}$$

both for the Euler-α and for the EPDiff equations. Remarkably, the idea that the Laplacian term in the Euler-α equations represents the Lagrangian-mean closure for the pseudomomentum of the turbulent components of the flow was expressed already by Holm *et al.* (1998), who derived the Euler-α in flat space as abstract Euler–Poincaré equations on the volumorphism group.

We finally remark that the expression for the Lagrangian-mean one-form given by Gilbert and Vanneste (2018) reads $\bar{\nu}^L = \langle \xi^*_{\beta,\varepsilon} \, \nu_{\beta,\varepsilon} \rangle$ with $\nu_{\beta,\varepsilon} = u^\flat_{\beta,\varepsilon}$. In principle, it is feasible to expand $\xi^*_{\beta,\varepsilon} \, \nu_{\beta,\varepsilon}$ and use that

$$\frac{\mathrm{d}}{\mathrm{d}s} (\xi^*_{\beta,s} \, \nu_{\beta,s}) = \xi^*_{\beta,s} \, (\partial_s \nu_{\beta,s} + \mathcal{L}_u \nu_{\beta,s}) \tag{112}$$

but this direct computation is substantially more tedious than the indirect approach used above.

12. Manifolds with boundaries

Our methods are flexible enough to treat manifolds with a boundary. However, it has already been noted in Marsden and Shkoller (2003) that the no-flux boundary conditions $u \cdot n = 0$, where n denotes the outward normal to the boundary $\partial\Omega$, which are the natural boundary conditions for the Euler equations, are incompatible with isotropy of fluctuations.

Indeed, suppose that the normal $n(x) = e_3$ at a point $x \in \partial\Omega$. Then, the no-flux boundary conditions would imply $w^3(x) = 0$ for an arbitrary fluctuation vector field w, so that $\langle w^3(x) w^3(x) \rangle = 0$, whereas isotropy requires $\langle w^3(x) w^3(x) \rangle = 1$. A similar problem emerges for Burgers' equations with the natural no-slip boundary condition $u = 0$ on $\partial\Omega$.

Thus, on manifolds with boundary, one must generally consider anisotropic equations, which are a coupled system of evolution equations for the mean velocity and Taylor diffusivity tensor

$$\kappa = \langle w \otimes w \rangle \tag{113}$$

(see, e.g., Marsden and Shkoller 2003 or Holm 1999).

However, for certain simplified geometries, for instance for a horizontal strip $\Omega = \mathbb{R}^2 \times [0, H]$, the rigid lid boundary conditions are compatible with spatial uniformity of the Taylor diffusivity tensor κ. In such cases, one could still derive analogues of isotropic Euler-α equations on manifolds with boundary by replacing the isotropy with an appropriate form of spatial uniformity in the closure hypothesis. We refer the reader to Badin *et al.* (2018) for the examples of such a construction.

Acknowledgments

We thank Gualtiero Badin and Jacques Vanneste for many interesting discussions, Jacques Vanneste, in particular, for suggesting that we try the stochastic Taylor hypothesis (81). We also thank the two anonymous referees for insightful remarks that led to substantial improvement of the paper. This work contributes to the project "The Interior Energy Pathway" of the Collaborative Research Center TRR 181 "Energy Transfers in Atmosphere and Ocean" funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under project number 274762653.

References

Alfonsi, G., Reynolds-averaged Navier–Stokes equations for turbulence modeling. *Appl. Mech. Rev.* 2009, **62**, 040802.

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

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

Arnold, V., Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. *Ann. Inst. Fourier* 1966, **16**, 319–361.

Badin, G. and Crisciani, F., *Variational Formulation of Fluid and Geophysical Fluid Dynamics – Mechanics, Symmetries and Conservation Laws*, 2018 (Springer-Verlag).

Badin, G., Oliver, M. and Vasylkevych, S., Geometric Lagrangian averaged Euler–Boussinesq and primitive equations. *J. Phys. A: Math. Theor.* 2018, **51**, 455501.

Bretherton, F.P., A note on Hamilton's principle for perfect fluids. *J. Fluid Mech.* 1970, **44**, 19–31.

Bühler, O., *Waves and Mean Flows*, Cambridge Monographs on MechanicsSecond edn, 2014 (Cambridge University Press, Cambridge).

Camassa, R. and Holm, D.D., An integrable shallow water equation with peaked solitons. *Phys. Rev. Lett.* 1993, **71**, 1661–1664.

Chen, S., Foias, C., Holm, D.D., Olson, E., Titi, E.S. and Wynne, S., A connection between the Camassa–Holm equations and turbulent flows in channels and pipes. *Phys. Fluids* 1999, **11**, 2343–2353.

Chow, B., Lu, P. and Ni, L., *Hamilton's Ricci flow*, Vol. 77, Graduate Studies in Mathematics 2006 (American Mathematical Society, Providence, RI; Science Press Beijing, New York).

Danilov, S., Juricke, S., Kutsenko, A. and Oliver, M., Toward consistent subgrid momentum closures in ocean models. In *Energy Transfers in Atmosphere and Ocean*, edited by C. Eden and A. Iske, pp. 145–192, 2019 (Springer Nature: Cham, Switzerland).

Ebin, D.G. and Marsden, J., Groups of diffeomorphisms and the motion of an incompressible fluid. *Ann. of Math. (2)* 1970, **92**, 102–163.

Foias, C., Manley, O., Rosa, R. and Temam, R., *Navier-Stokes equations and turbulence*, Vol. 83, Encyclopedia of Mathematics and its Applications 2001 (Cambridge University Press, Cambridge).

Gay-Balmaz, F., Well-posedness of higher dimensional Camassa–Holm equations. *Bull. Transilv. Univ. Braşov Ser. III* 2009, **2(51)**, 55–58.

Gay-Balmaz, F. and Ratiu, T.S., The Lie–Poisson structure of the LAE-α equation. *Dyn. Partial Differ. Equ.* 2005, **2**, 25–57.

Gilbert, A.D., Riedinger, X. and Thuburn, J., On the form of the viscous term for two dimensional Navier-Stokes flows. *Quart. J. Mech. Appl. Math.* 2014, **67**, 205–228.

Gilbert, A.D. and Vanneste, J., Geometric generalised Lagrangian-mean theories. *J. Fluid Mech.* 2018, **839**, 95–134.

Graham, J.P. and Ringler, T., A framework for the evaluation of turbulence closures used in mesoscale ocean large-eddy simulations. *Ocean Model.* 2013, **65**, 25–39.

Hirani, A.N., Marsden, J.E. and Arvo, J., Averaged Template Matching Equations; in *Energy Minimization Methods in Computer Vision and Pattern Recognition*, edited by M. Figueiredo, J. Zerubia and A.K. Jain, 2001, pp. 528–543.

Holm, D.D., Fluctuation effects on 3D Lagrangian mean and Eulerian mean fluid motion. *Phys. D* 1999, **133**, 215–269.

Holm, D.D. and Marsden, J.E., Momentum maps and measure-valued solutions (peakons, filaments, and sheets) for the EPDiff equation. In *The breadth of symplectic and Poisson geometry*, Progr. Math., Vol. 232 of *Progr. Math.*, pp. 203–235, 2005 (Birkhäuser: Boston, MA).

Holm, D.D., Marsden, J.E. and Ratiu, T.S., The Euler–Poincaré equations and semidirect products with applications to continuum theories. *Adv. Math.* 1998, **137**, 1–81.

Holm, D.D., Schmah, T. and Stoica, C., *Geometric Mechanics and Symmetry: From Finite to Infinite Dimensions*, 2009 (Oxford University Press).

Jacobs, K., *Stochastic processes for physicists*, 2010 (Cambridge University Press, Cambridge).

Marsden, J.E. and Shkoller, S., Global well-posedness for the Lagrangian averaged Navier–Stokes (LANS-α) equations on bounded domains. *R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci.* 2001, **359**, 1449–1468.

Marsden, J.E. and Shkoller, S., The anisotropic Lagrangian averaged Euler and Navier–Stokes equations. *Arch. Ration. Mech. Anal.* 2003, **166**, 27–46.

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

Mohseni, K., Kosović, B., Shkoller, S. and Marsden, J.E., Numerical simulations of the Lagrangian averaged Navier–Stokes equations for homogeneous isotropic turbulence. *Phys. Fluids* 2003, **15**, 524–544.

Oliver, M., Lagrangian averaging with geodesic mean. *Proc. R. Soc. A* 2017, **473**, 20170558, 9.

Palais, R.S., *Foundations of Global Non-Linear Analysis*, 1968 (W. A. Benjamin, Inc., New York-Amsterdam).

Petersen, P., *Riemannian Geometry*, third edn, 2016 (Springer).

Roberts, P.H. and Soward, A.M., The Navier–Stokes-α equations revisited. *Geophys. Astrophys. Fluid Dyn.* 2009, **103**, 303–316.

Sagaut, P., *Large Eddy Simulation for Incompressible Flows: An Introduction*, 2006 (Springer).

Salmon, R., On wave-action and its relatives. *J. Fluid Mech.* 2013a, **719**, 165–182.

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

Shkoller, S., The Lagrangian averaged Euler (LAE-α) equations with free-slip or mixed boundary conditions. In *Geometry, Mechanics, and Dynamics*, edited by P. Newton, P. Holmes and A. Weinstein, pp. 169–180, 2002 (Springer: New York).

Soward, A.M. and Roberts, P.H., On the derivation of the Navier–Stokes-alpha equations from Hamilton's principle. *J. Fluid Mech.* 2008, **604**, 297–323.

Vizman, C., Geodesic equations on diffeomorphism groups. *SIGMA Symmetry Integrability Geom. Methods Appl.* 2008, **4**, Paper 030, 22.