An alternative view of Generalized Lagrangian Mean theory

By Rick Salmon

Scripps Institution of Oceanography, La Jolla CA 92093-0213, USA

(Received 14 December 2012)

If the variables describing wave-mean flow interactions are chosen to include a set of fluid-particle labels corresponding to the mean flow, then Generalized Lagrangian Mean (GLM) theory takes the form of an ordinary, classical field theory. Its only truly distinctive features then arise from the distinctive feature of fluid dynamics *as a field theory*, namely, the particle-relabeling symmetry property, which corresponds by Noether's theorem to the many vorticity conservation laws of fluid mechanics. The key feature of the formulation is that *all* the dependent variables depend on a common set of space-time coordinates. This feature permits an easy and transparent derivation of the GLM equations by use of the energy-momentum tensor formalism. The particle-relabeling symmetry property leads to the GLM potential vorticity law in which pseudomomentum is the only wave activity term present. Thus the particle-relabeling symmetry explains the prominent importance of pseudomomentum in GLM theory.

1. Introduction

Since its introduction by Andrews & McIntyre (1978a,b)—see also Bühler (2009)— the theory of the Generalized Lagrangian Mean (hereafter GLM) has enjoyed a special status among wave/mean-flow interaction theories. Although most applications use one of its many abridgments, no theory more general than GLM has been found. This fact, and the evident beauty and obvious physical importance of GLM theory—including, especially, the important role played therein by pseudomomentum—virtually guarantee that the theory will continue to occupy its central position. However, the original derivation of GLM, which seems hardly to have changed, is somewhat lacking in transparency and motivation. It is the purpose of this paper to offer an alternative derivation that is, in many respects, much simpler. We shall show that if the variables describing the fluid motion are chosen to include a set of fluid-particle labels corresponding to the mean flow, then GLM theory takes the form of an ordinary, classical field theory. Its only truly distinctive features then arise from the distinctive feature of fluid dynamics *as a field theory*, namely, the particle-relabeling symmetry property, which corresponds by Noether's theorem to the many vorticity conservation laws of fluid mechanics (e.g. Salmon, 1988, sec. 4).

In classical field theory, the governing equations result from the requirement that

$$\delta \iiiint dt\, d\mathbf{x}\, L(\phi^i, \phi^i_j, x^j) = 0 \tag{1.1}$$

for arbitrary variations $\delta\phi^i$ in the field variables $\{\phi^i(x^j), i = 1, \ldots, n\}$, where n is the number of fields, $(t, \mathbf{x}) = (t, x, y, z) = (x^0, x^1, x^2, x^3)$, and $\phi^i_j = \partial\phi^i/\partial x^j$. The Lagrangian density L is a prescribed function of its arguments. We assume that L does not depend on derivatives higher than the first, but the theory is easily generalized. The variational principle (1.1) yields the Euler-Lagrange equations

$$\delta\phi^i : \qquad \frac{\partial}{\partial x^j}\left(\frac{\partial L}{\partial \phi^i_j}\right) = \frac{\partial L}{\partial \phi^i}, \tag{1.2}$$

where repeated indices are summed. It follows from (1.2) that

$$\frac{\partial}{\partial x^j}\left(T^{ij}\right) = -\left.\frac{\partial L}{\partial x^i}\right|_{\text{explicit}}, \tag{1.3}$$

where

$$T^{ij} = \phi^r_i \frac{\partial L}{\partial \phi^r_j} - L\delta^{ij} \tag{1.4}$$

is the energy-momentum tensor and δ^{ij} is the Kronecker delta. The right-hand side of (1.3) is a derivative with the ϕ^i and ϕ^i_j held constant; it arises from the explicit dependence of L on space or time. If L does not depend explicitly on x^i then (1.3) takes the form of a conservation law. Although (1.3) follows directly from (1.2), it is more revealing to obtain it as the result of a symmetry property: the invariance of L with respect to translation in space or time.

The equations of fluid mechanics fit the form of a field theory in which the field variables are the locations $\mathbf{x}(\mathbf{a},\tau)$ at time τ of the fluid particle labeled by $\mathbf{a} = (a,b,c)$. In its simplest form, the variational principle is

$$\delta \iiiint d\tau \; d\mathbf{a} \; L[\mathbf{x}(\mathbf{a},\tau)] = 0 \tag{1.5}$$

for arbitrary $\delta\mathbf{x}(\mathbf{a},\tau)$. The square bracket denotes functional dependence. It seems best to think of $\delta\mathbf{x}(\mathbf{a},\tau)$ as a variation in the time-dependent mapping $\mathbf{a} \to \mathbf{x}$ from the space of labels to the space of locations. Since this mapping is one-to-one, it is equivalent to require that the integral in (1.5) be stationary with respect to variations in the inverse mapping $\mathbf{a} \leftarrow \mathbf{x}$. Then (1.5) takes the form

$$\delta \iiiint dt \; d\mathbf{x} \; \frac{\partial(\mathbf{a})}{\partial(\mathbf{x})} L[\mathbf{a}(\mathbf{x},t)] = 0 \tag{1.6}$$

for arbitrary $\delta\mathbf{a}(\mathbf{x},t)$. Although the Euler-Lagrange equations corresponding to (1.6) are more complicated than those which result from (1.5), (1.6) fits the form of (1.1) exactly. It is better to apply the energy-momentum formalism to (1.6) than to (1.5), because the physics is not generally translation-invariant in the space of particle labels.

In GLM theory we consider a collection—an ensemble—of flow realizations. To each member of the ensemble we assign a unique value of the *ensemble parameter* μ. We regard the fluid particle locations in ensemble member μ as the sum

$$\mathbf{x}(\mathbf{a},\tau,\mu) = \mathbf{X}(\mathbf{a},\tau) + \boldsymbol{\xi}(\mathbf{X},T,\mu) \tag{1.7}$$

of an arbitrary *reference flow* $\mathbf{X}(\mathbf{a},\tau)$, which is independent of μ and therefore the same for all ensemble members, and a deviation $\boldsymbol{\xi}(\mathbf{X},T,\mu)$ therefrom. Here, $t = \tau = T$ but the use of different symbols for time will allow us to distinguish easily between time derivatives in the various systems. Now suppose that μ takes N discrete values corresponding to an ensemble with N members. Then the left-hand side of (1.7) corresponds to N time-dependent mappings, while the right-hand side of (1.7) corresponds to $N+1$ mappings. Thus the decomposition (1.7) is degenerate until a further condition is imposed. In GLM theory, the condition chosen is that the μ-average of $\boldsymbol{\xi}$ be zero. Then the reference flow

$\mathbf{X}(\mathbf{a},\tau)$ is a mean flow, and the degeneracy is removed. Other choices for removing the degeneracy might conceivably be preferable.

Regardless how $\mathbf{X}(\mathbf{a},\tau)$ is defined, the variational principle corresponding to the entire ensemble—or metasystem—is

$$\delta \int d\mu \iiiint d\tau \; d\mathbf{a} \; L[\mathbf{x}(\mathbf{a},\tau,\mu)] = 0. \tag{1.8}$$

Substituting (1.7) into (1.8), we obtain the mean-flow equations by varying $\mathbf{X}(\mathbf{a},\tau)$ with $\boldsymbol{\xi}(\mathbf{X},T,\mu)$ held fixed, and we obtain the equations for the waves or fluctuations by varying $\boldsymbol{\xi}(\mathbf{X},T,\mu)$ with $\mathbf{X}(\mathbf{a},\tau)$ held fixed. The mean-flow variations are complicated by the fact that $\mathbf{X}$ enters (1.7) as both a dependent and an independent variable. It is our contention that most of the awkwardness and unconventional character of the standard GLM derivation arises from this fact.

To see the better course, we first realize that (1.7) inside (1.8) can best be viewed as variations in the *composite mapping*,

$$\mathbf{a} \rightarrow \mathbf{X} \rightarrow \boldsymbol{\xi}, \tag{1.9}$$

from the space of labels, to the space of mean locations, to the space of deviations. However, independent variations in the mappings represented by the arrows in (1.9) are obviously equivalent to variations in the mappings represented by

$$\mathbf{a} \leftarrow \mathbf{X} \rightarrow \boldsymbol{\xi}, \tag{1.10}$$

in which the first arrow has been reversed. To make use of (1.10), we rewrite (1.8) as

$$\delta \int d\mu \iiiint dT \; d\mathbf{X} \; \frac{\partial(\mathbf{a})}{\partial(\mathbf{X})} L[\mathbf{a}(\mathbf{X},T), \boldsymbol{\xi}(\mathbf{X},T,\mu)] = 0 \tag{1.11}$$

and regard L as a function of $\mathbf{a}(\mathbf{X},T)$, $\boldsymbol{\xi}(\mathbf{X},T,\mu)$, and their derivatives. (For simplicity, in the remainder of this section we ignore the possibility of explicit dependence on $\mathbf{X}$ and T.) The change of viewpoint from (1.9) to (1.10), which imposes a single set of independent variables $(\mathbf{X},T,\mu)$, thus has the effect of converting GLM theory to a standard classical field theory; the general fields $\phi^i(x^j)$ in (1.1) identify with $\mathbf{a}(\mathbf{X},T)$ and $\boldsymbol{\xi}(\mathbf{X},T,\mu)$ in (1.11). The ensemble parameter μ acts merely as an additional independent variable. As we shall see, the GLM equations are simply the energy-momentum-action equations associated with (1.11).

The more *particular* results of GLM theory, especially those involving vorticity, arise from a symmetry property of (1.11) that is not shared by other field theories; it does not correspond to a translation symmetry with respect to $\mathbf{X}$, T, or μ. It is the so-called particle-relabeling symmetry, which in the context of (1.11) corresponds to variations $\delta\mathbf{a}(\mathbf{X},T)$ that leave the Jacobian $\partial(\mathbf{a})/\partial(\mathbf{X})$ unchanged, i.e., to a relabeling of fluid particles that does not affect the distribution of mass. Such variations lead in a motivated fashion to the GLM potential vorticity law, in which pseudomomentum terms are the only 'wave activity' terms present. This explains the prominent importance of pseudomomentum in GLM theory.

The change in viewpoint from (1.9) to (1.10) allows GLM theory to be developed quickly and transparently by standard methods of mathematical physics. As a result, the theory loses much of its special flavor. We obtain the equations for the mean flow by the same general procedure as the equations for the waves. There is no need to distinguish between 'Eulerian variations' and 'Lagrangian variations', or to entertain more than one kind of average. In this paper, there is only one kind of average, and it corresponds to an integration with respect to μ with $\mathbf{X}$ and T held fixed. The potential vorticity law arises from an obvious symmetry property of the Lagrangian and does not require ingenuity to discover. Thus this paper should appeal to readers who understand basic physics and variational methods but who do not easily follow the conventional development of GLM theory. The latter contains many elements—kinematical theorems, Eulerian and Lagrangian averages, Stokes corrections, form stress, radiation stress—that enrich the theory but are not essential to its construction. Because so many of these elements are missing from our derivation, we avoid elaborate notation. In this paper, superscripts always denote component indices, and subscripts always denote derivatives. No averaging symbols besides $\langle\cdot\rangle$ are required.

None of our final results is completely new, and even our methods have strong antecedents in the literature on waves and mean flows. Indeed, the importance of the energy-momentum tensor formalism was already recognized by Andrews & McIntyre (1978b) and strongly emphasized by Grimshaw (1984) in his review of GLM theory. Whitham (1965, 1974) and Bretherton (1971, 1976) showed that many results of the theory could be obtained directly from a variational principle. However, earlier papers used the energy-momentum tensor formalism only to derive the equations governing wave activity; the equations describing the effect of the waves on the mean flow were less gracefully derived by other methods. Gjaja and Holm (1996) considered a Lagrangian of the form (1.11) but they did not employ the energy-momentum tensor formalism as here.

We illustrate our approach by application to two frequently used sets of fluid equations: the rotating shallow-water equations (Sections 2 and 3) and the stratified, incompressible Boussinesq equations (Section 4). The former are a paradigm for rotating compressible flows in 2 or 3 dimensions, and the latter exemplify the special treatment needed to accommodate incompressibility. In both cases, we obtain the fundamental potential vorticity law from the particle-relabeling symmetry of the Lagrangian. The GLM equations are mathematically unclosed in the same sense that the Reynolds-averaged equations of turbulence theory are unclosed. Until closure approximations are invoked, it is impossible to say anything about the mathematical properties of the GLM equations. Section 5 reviews the standard quasilinear closure.

2. Rotating shallow-water dynamics

Let

$$\mathbf{x}(\mathbf{a},\tau,\mu) = (x(\mathbf{a},\tau,\mu), y(\mathbf{a},\tau,\mu)) \tag{2.1}$$

be the location at time τ of the fluid column labeled by $\mathbf{a} = (a, b)$, in the μ-th member of an ensemble of shallow-water flows. The fluid velocity is $(u, v) = (\partial x/\partial\tau, \partial y/\partial\tau)$. It is convenient to assign the labels (a, b) such that areas in label space correspond to fluid volume in physical space. Then

$$h \equiv \frac{\partial(a,b)}{\partial(x,y)} \tag{2.2}$$

is the depth of the fluid. Since the labels are conserved on fluid columns, i.e. $\partial a/\partial\tau = \partial b/\partial\tau = 0$, (2.2) holds at all times, and the τ-derivative of (2.2) yields the continuity equation,

$$\frac{\partial h}{\partial\tau} + h\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) = 0. \tag{2.3}$$

The momentum equations

$$\frac{\partial u}{\partial\tau} - fv = -g\frac{\partial h}{\partial x}, \qquad \frac{\partial v}{\partial\tau} + fu = -g\frac{\partial h}{\partial y} \tag{2.4}$$

with Coriolis parameter $f = \beta y$ and gravity constant g, result from the requirement that

$$\delta \iiiint da\, db\, d\tau\, d\mu \left(\frac{1}{2}\frac{\partial x}{\partial \tau}\left(\frac{\partial x}{\partial \tau} - \beta y^2\right) + \frac{1}{2}\left(\frac{\partial y}{\partial \tau}\right)^2 - \frac{1}{2} g \frac{\partial(a,b)}{\partial(x,y)} \right) = 0 \tag{2.5}$$

for arbitrary independent variations $\delta x(a,b,\tau)$, $\delta y(a,b,\tau)$ in the locations of the fluid columns. The variational principle (2.5) is analogous to (1.5).

Following the method outlined in Section 1, we let

$$\mathbf{x}(\mathbf{a},\tau,\mu) = \mathbf{X}(\mathbf{a},\tau) + \boldsymbol{\xi}(\mathbf{X},T,\mu) \tag{2.6}$$

and require

$$\langle \boldsymbol{\xi} \rangle \equiv \int d\mu\, \boldsymbol{\xi} = 0, \tag{2.7}$$

where $\mathbf{X} = (X,Y)$ and $\boldsymbol{\xi} = (\xi,\eta)$. Then $\mathbf{X}(\mathbf{a},\tau)$ is the mean flow. By the chain rule,

$$\begin{aligned} \frac{\partial x}{\partial \tau} &= U + \frac{D\xi}{DT}, \\ \frac{\partial y}{\partial \tau} &= V + \frac{D\eta}{DT}, \end{aligned} \tag{2.8}$$

where $U \equiv \partial X/\partial \tau$, $V \equiv \partial Y/\partial \tau$, and

$$\frac{D}{DT} \equiv \frac{\partial}{\partial \tau} = \frac{\partial}{\partial T} + U\frac{\partial}{\partial X} + V\frac{\partial}{\partial Y} \tag{2.9}$$

is the time derivative following a fluid column moving with the mean flow. The integral in (2.5) becomes

$$\begin{aligned} \iiiint d\mathbf{a}\, d\tau\, d\mu\ \Big[& \frac{1}{2}\left(U + \frac{D\xi}{DT}\right)\left(U + \frac{D\xi}{DT} - \beta(Y+\eta)^2\right) + \frac{1}{2}\left(V + \frac{D\eta}{DT}\right)^2 \\ & - \frac{1}{2} g \frac{\partial(a,b)}{\partial(X,Y)}\left(1 + \frac{\partial \xi}{\partial X} + \frac{\partial \eta}{\partial Y} + \frac{\partial(\xi,\eta)}{\partial(X,Y)}\right)^{-1} \Big]. \end{aligned} \tag{2.10}$$

To express this integral in terms of the functions $a(X,Y,T)$, $b(X,Y,T)$, $\xi(X,Y,T,\mu)$, and $\eta(X,Y,T,\mu)$ we solve the defining equations $Da/DT = Db/DT = 0$ for U and V with the result

$$HU = \frac{\partial(a,b)}{\partial(Y,T)}, \quad HV = \frac{\partial(a,b)}{\partial(T,X)}, \tag{2.11}$$

where

$$H \equiv \frac{\partial(a,b)}{\partial(X,Y)}, \tag{2.12}$$

and we rewrite (2.10) as

$$\iiiint d\mathbf{X}\, dT\, d\mu\ \ L[\mathbf{a}(\mathbf{X},T), \boldsymbol{\xi}(\mathbf{X},T,\mu), Y], \tag{2.13}$$

where, introducing the compact notation $D \equiv D/DT$,

$$\begin{aligned} L =& \frac{H}{2}\left(U + D\xi\right)\left(U + D\xi - \beta(Y+\eta)^2\right) \\ & + \frac{H}{2}\left(V + D\eta\right)^2 - \frac{1}{2} g H^2 \left(1 + \frac{\partial \xi}{\partial X} + \frac{\partial \eta}{\partial Y} + \frac{\partial(\xi,\eta)}{\partial(X,Y)}\right)^{-1}. \end{aligned} \tag{2.14}$$

The symbols H, U, V in (2.14) are to be viewed as abbreviations for the definitions (2.11) and (2.12). Applying the operator $\partial/\partial\tau = D/DT$ to (2.12), we obtain the mean-flow continuity equation in the form

$$\frac{DH}{DT} + H\left(\frac{\partial U}{\partial X} + \frac{\partial V}{\partial Y}\right) = 0. \tag{2.15}$$

Thus (2.15) follows from the definition (2.12) in the same way that (2.3) follows from (2.2). The 'mean depth' H is neither the Lagrangian nor the Eulerian average of h; rather, it is the fluid depth that would occur if the fluid moved at the mean velocity (U, V). Bühler & McIntyre (1998) use the symbol $\tilde{h}$ for H. Many papers on GLM theory use $(\bar{u}^L, \bar{v}^L)$ for (U, V). The remaining dynamical equations may be obtained by requiring that (2.13) be stationary with respect to arbitrary variations $\delta\mathbf{a}(\mathbf{X}, T)$ and $\delta\boldsymbol{\xi}(\mathbf{X}, T, \mu)$.

The physical interpretation of $\mathbf{a}(\mathbf{X}, T)$ and $\boldsymbol{\xi}(\mathbf{X}, T, \mu)$ must be kept constantly in mind: $\mathbf{a}(\mathbf{X}, T)$ is the identity of the fluid particle whose *average* location at time T is $\mathbf{X}$. $\boldsymbol{\xi}(\mathbf{X}, T, \mu)$ is the displacement, from $\mathbf{X}$, of the fluid particle whose average location at time T is $\mathbf{X}$. The upper case independent variables are intended to serve as a powerful reminder of these interpretations.

Arbitrary, independent variations of $\mathbf{a}(\mathbf{X}, T)$ and $\boldsymbol{\xi}(\mathbf{X}, T, \mu)$ yield the Euler-Lagrange equations for the system. However, since we are primarily interested in conservation laws, it is better to use the corresponding energy-momentum equations. Because (2.13)-(2.14) takes the form of a standard field theory with a single set of independent variables $(\mathbf{X}, T, \mu)$, we can apply the formula (1.3)-(1.4) directly, obtaining

$$\begin{aligned}
&\frac{\partial}{\partial T}\left(a_R\frac{\partial L}{\partial a_T} + b_R\frac{\partial L}{\partial b_T} + \xi_R\frac{\partial L}{\partial \xi_T} + \eta_R\frac{\partial L}{\partial \eta_T}\right)\\
&+ \frac{\partial}{\partial X}\left(a_R\frac{\partial L}{\partial a_X} + b_R\frac{\partial L}{\partial b_X} + \xi_R\frac{\partial L}{\partial \xi_X} + \eta_R\frac{\partial L}{\partial \eta_X}\right)\\
&+ \frac{\partial}{\partial Y}\left(a_R\frac{\partial L}{\partial a_Y} + b_R\frac{\partial L}{\partial b_Y} + \xi_R\frac{\partial L}{\partial \xi_Y} + \eta_R\frac{\partial L}{\partial \eta_Y}\right)\\
&+ \frac{\partial}{\partial \mu}\left(a_R\frac{\partial L}{\partial a_\mu} + b_R\frac{\partial L}{\partial b_\mu} + \xi_R\frac{\partial L}{\partial \xi_\mu} + \eta_R\frac{\partial L}{\partial \eta_\mu}\right) - \frac{\partial L}{\partial R} = -\left.\frac{\partial L}{\partial R}\right|_{\text{explicit}}.
\end{aligned} \tag{2.16}$$

Once again, subscripts denote differentiation, and the symbol R, which is analgous to x^i in (1.3), can be any one of the variables (X, Y, T, μ). Thus $a_R \equiv \partial a/\partial R$ stands for $\partial a/\partial X$, $\partial a/\partial Y$, $\partial a/\partial T$, or $\partial a/\partial \mu$. The last term on the left-hand side is a derivative in the (T, X, Y, μ) system. The right-hand side vanishes unless L depends explicitly on R. In the Lagrangian density (2.14), the only explicit dependence arises from the Coriolis term. Hence the right-hand side of (2.16) vanishes unless $R = Y$. Since L does not depend on $a_\mu, b_\mu, \xi_\mu, \eta_\mu$, (2.16) immediately reduces to

$$\begin{aligned}
&\frac{\partial}{\partial T}\left(a_R\frac{\partial L}{\partial a_T} + b_R\frac{\partial L}{\partial b_T} + \xi_R\frac{\partial L}{\partial \xi_T} + \eta_R\frac{\partial L}{\partial \eta_T}\right)\\
&+ \frac{\partial}{\partial X}\left(a_R\frac{\partial L}{\partial a_X} + b_R\frac{\partial L}{\partial b_X} + \xi_R\frac{\partial L}{\partial \xi_X} + \eta_R\frac{\partial L}{\partial \eta_X}\right)\\
&+ \frac{\partial}{\partial Y}\left(a_R\frac{\partial L}{\partial a_Y} + b_R\frac{\partial L}{\partial b_Y} + \xi_R\frac{\partial L}{\partial \xi_Y} + \eta_R\frac{\partial L}{\partial \eta_Y}\right) - \frac{\partial L}{\partial R} = -\left.\frac{\partial L}{\partial R}\right|_{\text{explicit}}.
\end{aligned} \tag{2.17}$$

We are interested in the ensemble average, or μ-integral, of (2.17). Before proceeding, we pause to make four points about ensemble averaging. First, although the introduction of averages introduces considerable subtlety into the usual derivation of GLM, it is unambiguous in our approach: the ensemble-average is simply the μ-integral at *fixed* values of $\mathbf{X}$ and T.

Second, ensemble-averaging introduces the formidable problem of closure. Before ensemble-averaging, the Euler-Lagrange equations, the energy momentum equations, and other equations discussed below are mathematically closed and also highly redundant. After all, only four (scalar) equations are required to predict the evolution of the four fields $\mathbf{a}(\mathbf{X}, T)$ and $\boldsymbol{\xi}(\mathbf{X}, T, \mu)$. However, after ensemble-averaging, every subset of the equations is mathematically unclosed in the same sense that the Reynolds-averaged equations of turbulence theory are unclosed. To make progress, one must introduce some type of closure approximation. We defer further discussion of closure until Section 5. Here we simply note that the vast majority of papers applying wave/mean theory use a *quasi-linear* closure in which $\boldsymbol{\xi}(\mathbf{X}, T, \mu)$ is assumed to be small.

Third, since averaging generally destroys the ability to manipulate equations, the only way to be sure that the conservation laws survive averaging is to average the conservation laws themselves. This is the motivation behind the use of the energy-momentum tensor formalism.

Fourth, closure demands a definition of the ensemble itself. In the vast majority of cases, one either identifies the ensemble parameter μ with one of the variables X, Y, or T so that, for example, the ensemble-average is synonymous with an average in the X-direction; or, one identifies μ with wave phase. In the latter case, one usually also assumes that the mean flow is slowly varying, so that ray-theory methods apply. However, it is not actually necessary to assign a meaning to μ until a relatively late stage in the development of the theory, and we defer further discussion until Section 5.

The μ-average of (2.17) is

$$\begin{aligned}
&\frac{\partial}{\partial T}\left(a_R\frac{\partial\langle L\rangle}{\partial a_T}+b_R\frac{\partial\langle L\rangle}{\partial b_T}+\langle\xi_R\frac{\partial L}{\partial\xi_T}+\eta_R\frac{\partial L}{\partial\eta_T}\rangle\right)\\
&+\frac{\partial}{\partial X}\left(a_R\frac{\partial\langle L\rangle}{\partial a_X}+b_R\frac{\partial\langle L\rangle}{\partial b_X}+\langle\xi_R\frac{\partial L}{\partial\xi_X}+\eta_R\frac{\partial L}{\partial\eta_X}\rangle\right)\\
&+\frac{\partial}{\partial Y}\left(a_R\frac{\partial\langle L\rangle}{\partial a_Y}+b_R\frac{\partial\langle L\rangle}{\partial b_Y}+\langle\xi_R\frac{\partial L}{\partial\xi_Y}+\eta_R\frac{\partial L}{\partial\eta_Y}\rangle\right)-\frac{\partial\langle L\rangle}{\partial R}=-\left.\frac{\partial\langle L\rangle}{\partial R}\right|_{\text{explicit}}. \qquad (2.18)
\end{aligned}$$

Note that the averaging commutes with derivatives with respect to a_T, b_T, etc., because $\mathbf{a}(\mathbf{X}, T)$ does not depend on μ. Note also that many of the terms in (2.18) involve the *averaged Lagrangian*,

$$\langle L\rangle=\frac{H}{2}\left(U^2+V^2+\langle(D\xi)^2\rangle+\langle(D\eta)^2\rangle\right)-\frac{\beta H}{2}\langle(U+D\xi)(Y+\eta)^2\rangle-\frac{1}{2}gH\langle h\rangle. \qquad (2.19)$$

In the case $R=\mu$, (2.18) reduces to the *action equation*

$$\frac{\partial}{\partial T}\langle\xi_\mu\frac{\partial L}{\partial\xi_T}+\eta_\mu\frac{\partial L}{\partial\eta_T}\rangle+\frac{\partial}{\partial X}\langle\xi_\mu\frac{\partial L}{\partial\xi_X}+\eta_\mu\frac{\partial L}{\partial\eta_X}\rangle+\frac{\partial}{\partial Y}\langle\xi_\mu\frac{\partial L}{\partial\xi_Y}+\eta_\mu\frac{\partial L}{\partial\eta_Y}\rangle=0. \qquad (2.20)$$

The equations (2.18) apply to the whole system; they involve both $\mathbf{a}(\mathbf{X}, T)$ and $\boldsymbol{\xi}(\mathbf{X}, T, \mu)$. However, according to the variational principle, variations with respect to $\mathbf{a}(\mathbf{X}, T)$ and $\boldsymbol{\xi}(\mathbf{X}, T, \mu)$ are to be taken independently. This means that we can apply the energy-momentum tensor formalism to either $\mathbf{a}(\mathbf{X}, T)$ or $\boldsymbol{\xi}(\mathbf{X}, T, \mu)$ *by itself* provided that we treat the other variable as a part of the *medium*—as a source of explicit dependence on the independent variables. If we regard $\boldsymbol{\xi}(\mathbf{X},T,\mu)$ as the active variable, we obtain

$$\frac{\partial}{\partial T}\langle \xi_R \frac{\partial L}{\partial \xi_T} + \eta_R \frac{\partial L}{\partial \eta_T}\rangle + \frac{\partial}{\partial X}\langle \xi_R \frac{\partial L}{\partial \xi_X} + \eta_R \frac{\partial L}{\partial \eta_X}\rangle + \frac{\partial}{\partial Y}\langle \xi_R \frac{\partial L}{\partial \xi_Y} + \eta_R \frac{\partial L}{\partial \eta_Y}\rangle - \frac{\partial \langle L\rangle}{\partial R} = - \left.\frac{\partial \langle L\rangle}{\partial R}\right|_{\text{explicit}}^{\mathbf{a}(\mathbf{X},T)}. \tag{2.21}$$

The derivative on the right-hand side of (2.21) is a derivative with respect to explicit dependence on R, now *including* the R-dependence that arises from the mean flow $\mathbf{a}(\mathbf{X},T)$. Similarly, if we regard $\mathbf{a}(\mathbf{X},T)$ as the active variable, we obtain

$$\frac{\partial}{\partial T}\left(a_R \frac{\partial \langle L\rangle}{\partial a_T} + b_R \frac{\partial \langle L\rangle}{\partial b_T}\right) + \frac{\partial}{\partial X}\left(a_R \frac{\partial \langle L\rangle}{\partial a_X} + b_R \frac{\partial \langle L\rangle}{\partial b_X}\right) + \frac{\partial}{\partial Y}\left(a_R \frac{\partial \langle L\rangle}{\partial a_Y} + b_R \frac{\partial \langle L\rangle}{\partial b_Y}\right) - \frac{\partial \langle L\rangle}{\partial R} = - \left.\frac{\partial \langle L\rangle}{\partial R}\right|_{\text{explicit}}^{\boldsymbol{\xi}(\mathbf{X},T)}. \tag{2.22}$$

However, (2.21) and (2.22) sum to give (2.18). To see this, suppose that $R = Y$. Then the right-hand sides of (2.21) and (2.22) sum as

$$- \left.\frac{\partial \langle L\rangle}{\partial Y}\right|_{\text{explicit}}^{\mathbf{a}(\mathbf{X},T)} - \left.\frac{\partial \langle L\rangle}{\partial Y}\right|_{\text{explicit}}^{\boldsymbol{\xi}(\mathbf{X},T)} = -\frac{\partial \langle L\rangle}{\partial Y} - \left.\frac{\partial \langle L\rangle}{\partial Y}\right|_{\text{explicit}}, \tag{2.23}$$

where the first term is a derivative with respect to the Y-dependence of the mean flow and the Coriolis term; the second term is a derivative with respect to the Y-dependence of the wave field and the Coriolis term; the third term is an exact derivative with respect to Y (holding X and T constant), and is thus affected by the mean flow, the wave field, *and* the Coriolis term; and the last term is a derivative with respect to the Coriolis term *only*. Thus only two of the three equation sets (2.18), (2.21) and (2.22) can be considered independent. Note that the action equation (2.20) can also be considered a special case of (2.21).

The equations (2.18)-(2.22), with $R = T$, $\mathbf{X}$, and μ, are the basic equations of GLM theory. To complete the theory, we need only take the required derivatives and interpret the results. We find that the equations (2.21) take the form

$$\frac{\partial(-Hp^R)}{\partial T} + \frac{\partial F^R}{\partial X} + \frac{\partial G^R}{\partial Y} - \frac{\partial \langle L\rangle}{\partial R} = - \left.\frac{\partial \langle L\rangle}{\partial R}\right|_{\text{explicit}}^{\mathbf{a}(\mathbf{X},T)}, \tag{2.24}$$

where

$$-Hp^R \equiv \langle \xi_R \frac{\partial L}{\partial \xi_T} + \eta_R \frac{\partial L}{\partial \eta_T}\rangle = H\langle \xi_R\, D\xi\rangle + H\langle \eta_R\, D\eta\rangle - \frac{H}{2}\beta\langle 2Y\eta\xi_R + \eta^2\xi_R\rangle, \tag{2.25}$$

$$F^R \equiv \langle \xi_R \frac{\partial L}{\partial \xi_X} + \eta_R \frac{\partial L}{\partial \eta_X}\rangle = -HUp^R + \frac{1}{2}g\langle h^2\xi_R\rangle + \frac{1}{2}g\langle h^2(\xi_R\eta_Y - \xi_Y\eta_R)\rangle, \tag{2.26}$$

$$G^R \equiv \langle \xi_R \frac{\partial L}{\partial \xi_Y} + \eta_R \frac{\partial L}{\partial \eta_Y}\rangle = -HVp^R + \frac{1}{2}g\langle h^2\eta_R\rangle + \frac{1}{2}g\langle h^2(\xi_X\eta_R - \xi_R\eta_X)\rangle, \tag{2.27}$$

and

$$h \equiv H\left(1 + \frac{\partial \xi}{\partial X} + \frac{\partial \eta}{\partial Y} + \frac{\partial(\xi,\eta)}{\partial(X,Y)}\right)^{-1}. \tag{2.28}$$

In particular, the action equation (2.20) becomes

$$\frac{\partial(-Hp^\mu)}{\partial T} + \frac{\partial F^\mu}{\partial X} + \frac{\partial G^\mu}{\partial Y} = 0. \tag{2.29}$$

We call p^T the pseudo-energy and $\mathbf{p} = (p^X, p^Y)$ the pseudomomentum. The minus sign that appears at the beginning of the definition (2.25) is a well-established convention. Setting $R = X$ in (2.21), we obtain the evolution equation

$$\frac{\partial(-Hp^X)}{\partial T} + \frac{\partial}{\partial X}\left(-HUp^X + \frac{1}{2}g\langle h^2\xi_X\rangle + \frac{1}{2}g\langle h^2(\xi_X\eta_Y - \eta_X\xi_Y)\rangle\right) + \frac{\partial}{\partial Y}\left(-HVp^Y + \frac{1}{2}g\langle h^2\eta_X\rangle\right) - \frac{\partial\langle L\rangle}{\partial X} = -\left.\frac{\partial\langle L\rangle}{\partial X}\right|_{\text{explicit}}^{\mathbf{a}(\mathbf{X},T)} \tag{2.30}$$

for X-direction pseudomomentum. The right-hand side of (2.30) arises from the X-dependence of the mean flow variables U, V, and H. If the mean flow depends on X, then the right-hand side of (2.30) is not an exact derivative, and (2.30) does not have the form of a conservation law. In contrast, unless $R = Y$, (2.18) has the form of a conservation law.

The equations (2.18) demand derivatives of the averaged Lagrangian (2.19) with respect to the derivatives of $\mathbf{a}(\mathbf{X}, T)$. For these it seems best to regard (2.19) as a function of U, V, H, Y, and then to use (2.11)-(2.12) and the chain rule. We find that

$$a_R\frac{\partial\langle L\rangle}{\partial a_T} + b_R\frac{\partial\langle L\rangle}{\partial b_T} = A\frac{\partial(a,b)}{\partial(Y,R)} + B\frac{\partial(a,b)}{\partial(R,X)}, \tag{2.31}$$

$$a_R\frac{\partial\langle L\rangle}{\partial a_X} + b_R\frac{\partial\langle L\rangle}{\partial b_X} = B\frac{\partial(a,b)}{\partial(T,R)} + (C - UA - VB)\frac{\partial(a,b)}{\partial(R,Y)}, \tag{2.32}$$

$$a_R\frac{\partial\langle L\rangle}{\partial a_Y} + b_R\frac{\partial\langle L\rangle}{\partial b_Y} = A\frac{\partial(a,b)}{\partial(R,T)} + (C - UA - VB)\frac{\partial(a,b)}{\partial(X,R)}, \tag{2.33}$$

where

$$A \equiv \frac{1}{H}\frac{\partial\langle L\rangle}{\partial U} = U - p^X - \frac{\beta}{2}(Y^2 + \langle\eta^2\rangle), \tag{2.34}$$

$$B \equiv \frac{1}{H}\frac{\partial\langle L\rangle}{\partial V} = V - p^Y, \tag{2.35}$$

$$C \equiv \frac{\partial\langle L\rangle}{\partial H} = \frac{1}{2}\left(U^2 + V^2 + \langle(D\xi)^2\rangle + \langle(D\eta)^2\rangle\right) - \frac{\beta}{2}\langle(U + D\xi)(Y+\eta)^2\rangle - g\langle h\rangle. \tag{2.36}$$

Setting $R = X$, in (2.31)-(2.33) we obtain

$$a_X\frac{\partial\langle L\rangle}{\partial a_T} + b_X\frac{\partial\langle L\rangle}{\partial b_T} = -H\left(U - p^X - \frac{\beta}{2}(Y^2 + \langle\eta^2\rangle)\right), \tag{2.37}$$

$$a_X\frac{\partial\langle L\rangle}{\partial a_X} + b_X\frac{\partial\langle L\rangle}{\partial b_X} = H\frac{\partial\langle L\rangle}{\partial H} - HU\left(U - p^X - \frac{\beta}{2}(Y^2 + \langle\eta^2\rangle)\right), \tag{2.38}$$

$$a_X\frac{\partial\langle L\rangle}{\partial a_Y} + b_X\frac{\partial\langle L\rangle}{\partial b_Y} = -HV\left(U - p^X - \frac{\beta}{2}(Y^2 + \langle\eta^2\rangle)\right). \tag{2.39}$$

Then substituting (2.37)-(2.39) and (2.25)-(2.27) into (2.18) and cancelling terms, we obtain the X-direction momentum equation

$$\frac{\partial(HM)}{\partial T} + \frac{\partial}{\partial X}\left(HUM + \frac{1}{2}gH\langle h\rangle - \frac{1}{2}g\langle h^2\xi_X\rangle - \frac{1}{2}g\langle h^2(\xi_X\eta_Y - \eta_X\xi_Y)\rangle\right) + \frac{\partial}{\partial Y}\left(HVM - \frac{1}{2}g\langle h^2\eta_X\rangle\right) = 0, \tag{2.40}$$

where

$$M \equiv U - \frac{\beta}{2}(Y^2 + \langle \eta^2 \rangle). \tag{2.41}$$

Again, because the right-hand side of (2.18) vanishes when $R = X$, (2.40) has the form of a conservation law. In deriving (2.40) it is handy to realize that

$$H\frac{\partial \langle L \rangle}{\partial H} - \langle L \rangle = -\frac{1}{2} gH\langle h \rangle. \tag{2.42}$$

In a typical application one identifies the ensemble average with an average over one of the variables X, Y or T. Suppose, for example, that we identify ensemble averages with averages over X. Then all averaged quantities, including the mean flow itself, are independent of X, and the GLM equations corresponding to $R = X$ simplify dramatically. In particular, the last two terms vanish in each of the three equations (2.18), (2.21) and (2.22). The X-direction pseudomomentum equation (2.30) reduces to

$$\frac{\partial}{\partial T}(-Hp^X) + \frac{\partial}{\partial Y}(-HVp^X + \frac{1}{2} g\langle h^2 \eta_X \rangle) = 0, \tag{2.43}$$

and the mean momentum equation (2.40) reduces to

$$\frac{\partial (HM)}{\partial T} + \frac{\partial}{\partial Y}\left(HVM - \frac{1}{2} g\langle h^2 \eta_X \rangle \right) = 0. \tag{2.44}$$

We note that the last term in (2.44)—the divergence of the *Eliassen-Palm* (hereafter EP) flux—occurs with opposite sign in (2.43). Hence, summing (2.43) and (2.44), we obtain the simpler result

$$\frac{\partial (HA)}{\partial T} + \frac{\partial}{\partial Y}(HVA) = 0, \tag{2.45}$$

where A is defined by (2.34).

In GLM theory, the EP flux is always the pressure-related flux of momentum. The cancellation of EP fluxes between (2.43) and (2.44) corresponds to the cancellation of the terms

$$\langle \xi_X \frac{\partial L}{\partial \xi_Y} + \eta_X \frac{\partial L}{\partial \eta_Y} \rangle \tag{2.46}$$

when, setting $R = X$, the X-average of (2.21) is subtracted from the X-average of (2.18) to give the X-average of (2.22). The latter is equivalent to (2.45). In the following section we obtain an equation free of EP fluxes without the need to average in a particular direction.

3. The particle-relabeling symmetry

Now we consider variations of the particle labels that leave the mean depth (2.12) unchanged in the Lagrangian density (2.14). These correspond to a relabeling of fluid columns that does not affect the distribution of mass. Since $\delta H = 0$ implies

$$\frac{\partial \delta a}{\partial a} + \frac{\partial \delta b}{\partial b} = 0, \tag{3.1}$$

these variations must take the form

$$\delta a = -\frac{\partial \delta\psi}{\partial b}, \quad \delta b = +\frac{\partial \delta\psi}{\partial a}, \tag{3.2}$$

where $\delta\psi$ is an arbitrary infinitesimal function. Then the variational principle implies

$$
\begin{aligned}
0 &= \iiint d\mathbf{X}\, dT \left(\frac{\partial \langle L \rangle}{\partial U} \delta U + \frac{\partial \langle L \rangle}{\partial V} \delta V \right) \\
&= \iiint d\mathbf{X}\, dT \left(A\, \delta \frac{\partial(a,b)}{\partial(Y,T)} + B\, \delta \frac{\partial(a,b)}{\partial(T,X)} \right) \\
&= \iiint d\mathbf{X}\, dT \left(A\, \frac{\partial(\delta a,b)}{\partial(Y,T)} + A\, \frac{\partial(a,\delta b)}{\partial(Y,T)} + B\, \frac{\partial(\delta a,b)}{\partial(T,X)} + B\, \frac{\partial(a,\delta b)}{\partial(T,X)} \right) \\
&= \iiint d\mathbf{X}\, dT \left(A\, \frac{\partial(X,\delta a,b)}{\partial(X,Y,T)} + A\, \frac{\partial(X,a,\delta b)}{\partial(X,Y,T)} - B\, \frac{\partial(\delta a,Y,b)}{\partial(X,Y,T)} - B\, \frac{\partial(a,Y,\delta b)}{\partial(X,Y,T)} \right) \\
&= \iiint d\mathbf{a}\, d\tau \left(A\, \frac{\partial(X,\delta a,b)}{\partial(a,b,\tau)} + A\, \frac{\partial(X,a,\delta b)}{\partial(a,b,\tau)} - B\, \frac{\partial(\delta a,Y,b)}{\partial(a,b,\tau)} - B\, \frac{\partial(a,Y,\delta b)}{\partial(a,b,\tau)} \right) \\
&= \iiint d\mathbf{a}\, d\tau \left(-\delta a\, \frac{\partial(X,A,b)}{\partial(a,b,\tau)} - \delta b\, \frac{\partial(X,a,A)}{\partial(a,b,\tau)} + \delta a\, \frac{\partial(B,Y,b)}{\partial(a,b,\tau)} + \delta b\, \frac{\partial(a,Y,B)}{\partial(a,b,\tau)} \right) \\
&= \iiint d\mathbf{a}\, d\tau \left(-\frac{\partial \delta\psi}{\partial b}\, \frac{\partial(X,A)}{\partial(a,\tau)} + \frac{\partial \delta\psi}{\partial a}\, \frac{\partial(X,A)}{\partial(b,\tau)} + \frac{\partial \delta\psi}{\partial b}\, \frac{\partial(B,Y)}{\partial(a,\tau)} - \frac{\partial \delta\psi}{\partial a}\, \frac{\partial(B,Y)}{\partial(b,\tau)} \right) \\
&= \iiint d\mathbf{a}\, d\tau\, \delta\psi \left(\frac{\partial}{\partial b}\, \frac{\partial(X,A)}{\partial(a,\tau)} - \frac{\partial}{\partial a}\, \frac{\partial(X,A)}{\partial(b,\tau)} - \frac{\partial}{\partial b}\, \frac{\partial(B,Y)}{\partial(a,\tau)} + \frac{\partial}{\partial a}\, \frac{\partial(B,Y)}{\partial(b,\tau)} \right) \\
&= \iiint d\mathbf{a}\, d\tau\, \delta\psi\, \frac{\partial}{\partial \tau} \left(\frac{\partial(X,A)}{\partial(a,b)} - \frac{\partial(B,Y)}{\partial(a,b)} \right),
\end{aligned}
\tag{3.3}
$$

where A and B are defined by (2.34) and (2.35). Since $\delta\psi$ is arbitrary, we must have

$$\frac{\partial}{\partial \tau} \left(\frac{\partial(X,A)}{\partial(a,b)} - \frac{\partial(B,Y)}{\partial(a,b)} \right) = 0, \tag{3.4}$$

which is equivalent to

$$\frac{D}{DT} \left(\frac{1}{H} \left(\frac{\partial B}{\partial X} - \frac{\partial A}{\partial Y} \right) \right) = 0. \tag{3.5}$$

Thus

$$\frac{DQ}{DT} = 0, \tag{3.6}$$

where

$$Q = H^{-1} \left(\frac{\partial V}{\partial X} - \frac{\partial U}{\partial Y} - \frac{\partial p^Y}{\partial X} + \frac{\partial p^X}{\partial Y} + \beta Y + \frac{\beta}{2} \frac{\partial \langle \eta^2 \rangle}{\partial Y} \right) \tag{3.7}$$

is the potential vorticity.

Suppose that the pseudomomenta (p^X, p^Y) and mean square displacement $\langle \eta^2 \rangle$ are known, either by solution of the ray equations for the wave statistics, or from direct measurements of the wave field. Then (2.15) and (3.6) represent two of the three equations needed to predict the evolution of the mean flow. A handy third equation would be the divergence equation, which could be derived from (2.40) and its counterpart in the Y-direction. However, Bühler & McIntyre (1998) argue that mean flows are typically 'balanced flows' in which the potential vorticity field determines the whole flow. For such flows, (3.6) is *by itself* a sufficient description of the mean flow dynamics. Even if the waves are small, they argue, continual dissipation or irreversible mixing of pseudomomentum can lead to a secular growth in the mean velocity by continually 'trapping' potential vorticity in the mean vorticity $V_X - U_Y$. In contrast, the portion of the mean-flow response determined by the mean divergence equation is typically small and likely to disappear when the waves cease to exist.

In any case, as descriptors of mean flow evolution, (2.15) and (3.6) have a striking simplicity. No wave effects appear in (2.15) at all, and in (3.6) they are confined to the numerator in (3.7). The operator D/DT refers to advection by the mean flow alone. This confinement of wave effects to the numerator of the potential vorticity is in sharp contrast to what would result if ensemble averaging were applied directly to the original, primitive equations for mass and potential vorticity conservation, and it is a primary advantage of GLM theory.

The potential vorticity equation (3.6)-(3.7) provides a physical explanation for the absence of EP fluxes in (2.45). Consider the region north of some fixed northern-hemisphere latitude line. Let X be eastward distance measured along this line. Then, integrating (3.6) over this 'polar cap', applying the Stokes and divergence theorems, and using (2.15), we obtain the 'zonal' momentum equation (2.45). In short, no EP fluxes appear in (2.45) because none are present in (3.6)-(3.7).

Our derivation of the GLM potential vorticity equation was motivated by the rather obvious symmetry property that most of (2.14) is unaffected by particle-label variations that do not affect H. The *unaffected* terms are the very terms that differ the most from one fluid model to another; the *affected* terms—the kinetic energy and Coriolis terms—always take the same form. This explains the generality of results like (3.6)-(3.7). The following section offers another example.

4. Boussinesq dynamics

Let $\mathbf{x}(a,b,\theta,\tau,\mu)$ be the three-dimensional Cartesian location of the fluid particle labeled by (a,b,θ) at time τ in ensemble member μ. We choose the third label θ to be the buoyancy. Then buoyancy is conserved,

$$\frac{\partial \theta}{\partial \tau} = 0, \tag{4.1}$$

as a matter of definition. The two remaining labels a and b are assigned such that

$$\frac{\partial(x,y,z)}{\partial(a,b,\theta)} = 1. \tag{4.2}$$

The τ-derivative of (4.2) yields

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0, \tag{4.3}$$

where $\mathbf{v} = (u,v,w) = (\partial x/\partial\tau, \partial y/\partial\tau, \partial z/\partial\tau)$ is the fluid velocity. The requirement that

$$\int d\mu \iiiint d\tau\; da\; db\; d\theta \left(\frac{1}{2}\frac{\partial \mathbf{x}}{\partial \tau}\cdot\frac{\partial \mathbf{x}}{\partial \tau} + z\theta + p(\frac{\partial(x,y,z)}{\partial(a,b,\theta)} - 1) \right) \tag{4.4}$$

be stationary for arbitrary variations $\delta\mathbf{x}(a,b,\theta,\tau,\mu)$, $\delta p(a,b,\theta,\tau,\mu)$ yields the incompressibility condition (4.2) and the Boussinesq momentum equations

$$\frac{\partial \mathbf{v}}{\partial \tau} = -\nabla p + \theta\mathbf{k}, \tag{4.5}$$

where $\mathbf{k}$ is the vertical unit vector. The Lagrange multiplier p turns out to be the pressure (divided by the constant mass density). To obtain the hydrostatic form of the Boussinesq equations, one simply omits the $\partial z/\partial\tau$-term from (4.4).

As in the shallow-water case, we let

$$\mathbf{x}(a,b,\theta,\tau,\mu)=\mathbf{X}(a,b,\theta,\tau)+\boldsymbol{\xi}(\mathbf{X},T,\mu), \tag{4.6}$$

where now $\mathbf{X}=(X,Y,Z)$ and $\boldsymbol{\xi}=(\xi,\eta,\zeta)$. Again we assume that $\langle\boldsymbol{\xi}\rangle=0$ so that $\mathbf{X}(a,b,\theta,\tau)$ is the mean flow. The incompressibility constraint (4.2) may be rewritten as

$$\frac{\partial(x,y,z)}{\partial(a,b,\theta)}=\frac{\partial(X,Y,Z)}{\partial(a,b,\theta)}\frac{\partial(X+\xi,Y+\eta,Z+\zeta)}{\partial(X,Y,Z)}=\frac{1}{J}\left(1+\Lambda[\xi,\eta,\zeta]\right)=1, \tag{4.7}$$

where

$$J[a,b,\theta]\equiv\frac{\partial(a,b,\theta)}{\partial(X,Y,Z)} \tag{4.8}$$

and

$$\Lambda[\xi,\eta,\zeta]\equiv\xi_X+\eta_Y+\zeta_Z+\frac{\partial(\xi,\eta)}{\partial(X,Y)}+\frac{\partial(\xi,\zeta)}{\partial(X,Z)}+\frac{\partial(\eta,\zeta)}{\partial(Y,Z)}+\frac{\partial(\xi,\eta,\zeta)}{\partial(X,Y,Z)}. \tag{4.9}$$

Then, regarding $\mathbf{a}=(a,b,\theta)=(a^1,a^2,a^3)$ and $\boldsymbol{\xi}=(\xi,\eta,\zeta)=(\xi^1,\xi^2,\xi^3)$ as functions of the common set of independent variables $(\mathbf{X},T,\mu)$, we rewrite (4.4) as

$$\int d\mu\ \iiiint dT\ d\mathbf{X}\ \ L[\mathbf{a}(\mathbf{X},T),\boldsymbol{\xi}(\mathbf{X},T,\mu),p(\mathbf{X},T,\mu),Z], \tag{4.10}$$

where

$$L=\frac{J}{2}(\mathbf{U}+D\boldsymbol{\xi})\cdot(\mathbf{U}+D\boldsymbol{\xi})+J(Z+\zeta)\theta+p(1+\Lambda-J) \tag{4.11}$$

and

$$D\equiv\frac{D}{DT}\equiv\frac{\partial}{\partial T}+U\frac{\partial}{\partial X}+V\frac{\partial}{\partial Y}+W\frac{\partial}{\partial Z}. \tag{4.12}$$

The average Lagrangian

$$\langle L\rangle=\frac{J}{2}\mathbf{U}\cdot\mathbf{U}+\frac{J}{2}\langle D\boldsymbol{\xi}\cdot D\boldsymbol{\xi}\rangle+JZ\theta+\langle p(1+\Lambda-J)\rangle. \tag{4.13}$$

Here, just as in Section 2, the symbol J is to be viewed as an abbreviation for (4.8), and $(U,V,W)\equiv(U^1,U^2,U^3)$ are to be viewed as abbreviations for

$$U^i=\frac{1}{J}\frac{\partial(X^i,a,b,\theta)}{\partial(T,X,Y,Z)}, \tag{4.14}$$

where now $(X^0,X^1,X^2,X^3)=(T,X,Y,Z)$. Equation (4.14) is analogous to (2.11), and (4.8) is analogous to (2.12). The ensemble-averaged energy-momentum equation for the system as a whole is

$$\frac{\partial}{\partial X^j}\left(a_i^r\frac{\partial\langle L\rangle}{\partial a_j^r}+\langle\xi_i^r\frac{\partial L}{\partial\xi_j^r}\rangle\right)-\frac{\partial\langle L\rangle}{\partial X^i}=-\left.\frac{\partial\langle L\rangle}{\partial X^i}\right|_{\text{explicit}}, \tag{4.15}$$

and the average energy-momentum equation for the fluctuations by themselves is

$$\frac{\partial}{\partial X^j}\langle\xi_i^r\frac{\partial L}{\partial\xi_j^r}\rangle-\frac{\partial\langle L\rangle}{\partial X^i}=-\left.\frac{\partial\langle L\rangle}{\partial X^i}\right|_{\text{explicit}}^{\mathbf{a}(\mathbf{X},T)}. \tag{4.16}$$

The action equation is

$$\frac{\partial}{\partial X^j}\langle\xi_\mu^r\frac{\partial L}{\partial\xi_j^r}\rangle=0. \tag{4.17}$$

Our notation is $a_i^r=\partial a^r/\partial X^i$ and $\xi_i^r=\partial\xi^r/\partial X^i$. Once again, the summation convention is in effect. The implied summations in j are summations over the 4 spacetime coordinates (T, X, Y, Z). The implied summations in r are summations over the 3 dependent variables, (a, b, θ) or (ξ, η, ζ), in each group.

As in Section 2, we obtain the GLM equations by using (4.11) and (4.13) in (4.15)-(4.17). We are primarily interested in the mean momentum equations, corresponding to $i = 1, 2, 3$ in (4.15). By steps similar to those in Section 2, we find that these take the form

$$J\frac{DU^i}{DT} + \frac{\partial P}{\partial X^i} - J\theta\delta^{i3} = \frac{\partial E^{ij}}{\partial X^j}, \tag{4.18}$$

where $P \equiv J\langle p \rangle$,

$$E^{ij} \equiv \langle p\xi_i^r \frac{\partial \Lambda}{\partial \xi_j^r}\rangle = \langle p\xi_i^r K^{rj}\rangle, \tag{4.19}$$

is the EP flux, and K^{rj} is the cofactor of the matrix $\delta^{rj} + \xi_j^r$. In typical applications, μ is identified with a particular coordinate, say X, and interest attaches to the X-component of (4.18). The terminology 'EP flux' is then applied to the off-diagonal components E^{XY} and E^{XZ}. In this section, by 'EP flux' we mean the complete tensor E^{ij}.

The τ-derivative of (4.8) implies

$$\frac{\partial J}{\partial T} + \frac{\partial (JU)}{\partial X} + \frac{\partial (JV)}{\partial Y} + \frac{\partial (JW)}{\partial Z} = 0, \tag{4.20}$$

and we may rewrite (4.1) as

$$\frac{D\theta}{DT} = 0. \tag{4.21}$$

Except for the presence of the density-like factor J, the left-hand sides of (4.18), (4.20), and (4.21) are the same as in the primitive, unaveraged form of the Boussinesq equations. The right-hand side of (4.18) contains the divergence of the EP flux. To complete the mean-flow equations, we take the ensemble average of the constraint equation (4.7), obtaining

$$J = 1 + \langle \Lambda \rangle. \tag{4.22}$$

Bühler & McIntyre (1998) use the symbol $\tilde{\rho}$ for J.

Suppose that the wave statistics are somehow known in advance. That is, suppose that E^{ij} and $\langle \Lambda \rangle$ are prescribed. Then (4.18) and (4.20)-(4.22) represent 6 equations in the six mean-flow variables U, V, W, J, θ, and P. The 'mean flow pressure' P is determined by the requirement that (4.20) be consistent with (4.22) and the prescribed evolution of $\langle \Lambda \rangle$. In reality, of course, the wave statistics are *not* prescribed; instead the mean-flow equations (4.18)-(4.22) and the wave equations (4.16)-(4.17) comprise a set of equations that are exact but mathematically unclosed in the same sense that the Reynolds-averaged equations of turbulence theory are exact but unclosed. Until we introduce approximations that *close* the GLM equations, it is meaningless to pose questions about their mathematical properties. For example, one would certainly like to know the circumstances under which J remains positive: a negative J would imply that the mapping $\mathbf{a} \to \mathbf{X}$ has become multi-valued, signifying a complete breakdown of the theory. However, it is meaningless to inquire about the sign of a solution to a set of unclosed equations. We take up the issue of closure in Section 5.

As in the shallow-water case, the EP flux divergence that appears in (4.18) appears with opposite sign in the pseudomomentum equations (4.16). Once again, the best way to eliminate the EP flux divergence is *not* to take the difference between (4.15) and (4.16); rather, it is to exploit the particle-relabeling symmetry of (4.13). Consider particle-label variations that do not affect J or θ. These variations must satisfy $\delta\theta = 0$ and (3.2), where $\delta\psi$ is an arbitrary function and derivatives are taken with θ held constant. For such variations,

$$\begin{aligned} 0 &= \iiiint dT\, d\mathbf{X}\, \delta\langle L\rangle = \iiiint dT\, d\mathbf{X}\, J\,(\mathbf{U}\cdot\delta\mathbf{U} + \langle D\boldsymbol{\xi}\cdot\delta D\boldsymbol{\xi}\rangle) \\ &= \iiiint dT\, d\mathbf{X}\, JA^i\delta U^i = \iiiint dT\, d\mathbf{X}\, A^i\, \delta\frac{\partial(X^i,a,b,\theta)}{\partial(T,X,Y,Z)}, \end{aligned} \tag{4.23}$$

where $A^i = U^i - p^i$, $p^i = -\langle \xi^j_i D\xi^j\rangle$ is the pseudomomentum, and the summation is over $i = 1,2,3$. The rest of the calculation proceeds in a manner similar to (3.3). We find, by the arbitrariness of $\delta\psi$, that

$$\frac{\partial}{\partial\tau}\frac{\partial(A^i,X^i)}{\partial(a,b)} = 0, \tag{4.24}$$

which is equivalent to

$$\frac{D}{DT}\left(\frac{\nabla\theta\cdot(\nabla\times(\mathbf{U}-\mathbf{p}))}{J}\right) = 0. \tag{4.25}$$

Bühler & McIntyre (1998) argue that the potential vorticity equation (4.25) plays a similar role with respect to Boussinesq dynamics as does (3.6) for shallow water dynamics: For sufficiently balanced mean flows, one may compute the effects of waves on the mean flow by use of (4.20), (4.21), and (4.25), thus avoiding (4.18) with its EP fluxes. Salmon (1998, pp. 309-313) derived a result similar to (4.25) by methods similar to those of the present paper.

5. Closure

All the preceding results are *exact*; they involve no approximations besides those implicit in shallow-water and Boussinesq dynamics. Moreover, all the preceding results have been obtained without defining the ensemble under consideration, and without addressing the problem of closure. However, until we invoke approximations that close the GLM equations, it is meaningless to talk about their solutions.

Most applications of GLM theory employ *quasilinear closure*, in which the wave variables $\boldsymbol{\xi}(\mathbf{X},T,\mu)$ are assumed infinitesimal. If $\boldsymbol{\xi}(\mathbf{X},T,\mu)$ are infinitesimal then they obey linear equations, which are in principle always solvable. Once the $\boldsymbol{\xi}(\mathbf{X},T,\mu)$ are known—and this requires specification of the ensemble—we can compute the statistical quantities E^{ij} and $\langle\Lambda\rangle$ that act as forcing terms in the mean-flow equations (4.18)-(4.22). Alternatively, we may adopt the better approach of Bühler & McIntyre (1998); this requires the computation of J and the pseudomomentum $\mathbf{p}$.

Whatever the method of closure, a logical beginning point is the Euler-Lagrange equations for the wave variables. For Boussinesq dynamics with Lagrangian (4.11) these are

$$\delta\xi_i: \quad JD(U^i + D\xi^i) = -\frac{\partial}{\partial X^j}\left(p\frac{\partial\Lambda}{\partial\xi^i_j}\right) + J\theta\delta^{i3}, \tag{5.1}$$

$$\delta p: \quad J = 1 + \Lambda. \tag{5.2}$$

Subtracting the mean momentum equation (4.18) from (5.1) we obtain

$$J\frac{D^2\xi^i}{DT^2} = -\frac{\partial}{\partial X^j}\left(p\frac{\partial\Lambda}{\partial\xi^i_j}\right) - \frac{\partial E^{ij}}{\partial X^j} + \frac{\partial(J\langle p\rangle)}{\partial X^i}, \tag{5.3}$$

which may be used as an alternative to (5.1). If $\boldsymbol{\xi} = O(\epsilon)$ with $\epsilon \to 0$, then $J = 1 + O(\epsilon^2)$, and (5.3) become

$$\frac{D^2\xi}{DT^2} = -\frac{\partial\pi}{\partial X} - \frac{\partial(\langle p\rangle,\eta)}{\partial(X,Y)} - \frac{\partial(\langle p\rangle,\zeta)}{\partial(X,Z)} + O(\epsilon^2), \tag{5.4}$$

$$\frac{D^2\eta}{DT^2} = -\frac{\partial\pi}{\partial Y} - \frac{\partial(\langle p\rangle,\xi)}{\partial(Y,X)} - \frac{\partial(\langle p\rangle,\zeta)}{\partial(Y,Z)} + O(\epsilon^2), \tag{5.5}$$

$$\frac{D^2\zeta}{DT^2} = -\frac{\partial\pi}{\partial Z} - \frac{\partial(\langle p\rangle,\xi)}{\partial(Z,X)} - \frac{\partial(\langle p\rangle,\eta)}{\partial(Z,Y)} + O(\epsilon^2), \tag{5.6}$$

where $\pi = p - \langle p\rangle = O(\epsilon)$. Similarly (5.2) becomes

$$\xi_X + \eta_Y + \zeta_Z = O(\epsilon^2). \tag{5.7}$$

In this limit the first term on the right-hand side of (5.3) partly cancels the last term, and the E^{ij}-term makes no contribution. If we neglect the $O(\epsilon^2)$ terms, then (5.4)-(5.7) comprise 4 linear equations for the 4 variables ξ, η, ζ, and π. Although these equations are linear, they have nonconstant coefficients involving the mean flow variables U, V, W, and $\langle p\rangle$. Thus (4.18)-(4.22) and (5.4)-(5.7) form a coupled set. We note that the 'wave forcing terms' E^{ij} and $\langle\Lambda\rangle$ that appear in the mean-flow equations are both $O(\epsilon^2)$ terms and hence can be computed consistently from the solutions of (5.4)-(5.7). If the mean pressure $\langle p\rangle$ depends only on Z then (5.4)-(5.6) take the simpler form

$$\frac{D^2\xi^i}{DT^2} = -\frac{\partial\phi}{\partial X^i} - \delta^{i3}N^2\zeta, \tag{5.8}$$

where $\phi \equiv \pi - \theta\zeta$ and $N^2 = d\theta/dZ$, but in general we must face (5.4)-(5.6).

It is simpler to derive the complete set of coupled quasilinear equations by expanding (4.11) in ϵ and keeping only terms of $O(\epsilon^2)$ and larger. The requirement that

$$\begin{aligned}\int d\mu \iiiint dT\, d\mathbf{X}\ \Big[\ &\frac{J}{2}(\mathbf{U} + D\boldsymbol{\xi})\cdot(\mathbf{U} + D\boldsymbol{\xi}) + J(Z+\zeta)\theta + \langle p\rangle(1-J)\\ &+ \langle p\rangle\left(\xi_X + \eta_Y + \zeta_Z + \frac{\partial(\xi,\eta)}{\partial(X,Y)} + \frac{\partial(\xi,\zeta)}{\partial(X,Z)} + \frac{\partial(\eta,\zeta)}{\partial(Y,Z)}\right) + \pi(\xi_X + \eta_Y + \zeta_Z)\ \Big]\end{aligned} \tag{5.9}$$

be stationary with respect to variations of $\mathbf{a}, \langle p\rangle, \boldsymbol{\xi}$ and π yields equations equivalent to (4.18)-(4.22) and (5.4)-(5.7) but with (4.19) and (4.22) replaced by the consistent-order approximations. Thus (5.9) is the logical beginning point for quasilinear theory.

In all of this, the ensemble parameter μ plays an essentially passive role; the definition of the ensemble is needed only to compute the 'wave-forcing terms' E^{ij} and J that appear in the mean flow equations (4.18)-(4.22). However, quasilinear theory is often combined with the assumption that the waves be *slowly varying*. Then the method of choice is the elegant method invented by Whitham (1965, 1974), and in this μ plays a somewhat more active role. One substitutes

$$\boldsymbol{\xi} = \mathbf{A}(\mathbf{X},T)\cos(\Theta(\mathbf{X},T) + \mu) \tag{5.10}$$

into (5.9), and one *performs* the integration with respect to μ. Then (5.9) is required to be stationary with respect to variations $\delta\mathbf{A}(\mathbf{X},T)$ and $\delta\Theta(\mathbf{X},T)$.

This paper has been primarily concerned with the exact GLM equations. Our sketch of slowly-varying, quasi-linear theory is intended only to furnish a single specific example of closure. However, several of the general results given above have been checked by showing that they reduce to more familiar expressions for EP flux and pseudomomentum in the quasilinear limit. The manipulations required are straightforward but surprisingly tedious. This suggests that the general GLM equations are not the best place to start if one intends to make the quasilinear approximation; it seems more efficient to introduce the small-amplitude approximation at a much earlier stage in the development of the theory. The GLM equations are best reserved for closures that do *not* assume that the waves are weakly nonlinear.

I thank three anonymous referees for important suggestions.

REFERENCES

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

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

Bretherton, F. P., 1971 The generalised theory of wave propagation. In Mathematical Problems in the geophysical Sciences, Lectures in Applied Mathematics **13**(1), 61-102. Am. Math. Soc. Providence, RI.

Bretherton, F. P., 1976 Conservation of wave action and angular momentum in a spherical atmosphere. Unpublished manuscript.

Bühler, O., 2009 Waves and Mean Flows. Cambridge, 334 pp.

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

Gjaja, I. & Holm, D. D., 1996 Self-consistent Hamiltonian dynamics of wave mean-flow interaction for a rotating stratified incompressible fluid. Physica D **98**, 343-378.

Grimshaw, R., 1984 Wave action and wave-mean flow interaction, with applications to stratified shear flows. Ann. Rev. Fluid Mech. **16**, 11-44.

Salmon, R., 1988 Hamiltonian fluid mechanics. Ann. Rev. Fluid Mech. **20**, 225-256.

Salmon, R., 1998 Lectures on Geophysical Fluid Dynamics. Oxford, 372 pp.

Whitham, G. B., 1965 A general approach to linear and nonlinear waves using a Lagrangian. J. Fluid Mech. **22**, 273-283.

Whitham, G. B., 1974 Linear and nonlinear waves. John Wiley, 628 pp.