J. Fluid Mech. (2013), *vol.* 729, *pp.* 69–102.
doi:10.1017/jfm.2013.259

The Eliassen–Palm flux tensor

J. R. Maddison† and D. P. Marshall

Atmospheric, Oceanic and Planetary Physics, Department of Physics, University of Oxford, Oxford OX1 3PU, UK

(Received 10 July 2012; revised 16 February 2013; accepted 20 May 2013; first published online 16 July 2013)

The aim of this paper it to derive general coordinate-invariant forms of the Eliassen–Palm flux tensor and thereby characterize the true geometric nature of the eddy–mean-flow interaction in hydrostatic Boussinesq rotating fluids. In the quasi-geostrophic limit previous forms of the Eliassen–Palm flux tensor are shown to be related to each other via a gauge transformation; a general form is stated and its geometric properties are discussed. Similar methodology is applied to the hydrostatic Boussinesq Navier–Stokes equations to re-derive the residual-mean equations in a coordinate-invariant form. Thickness-weighted averaging in buoyancy coordinates is carefully described, via the definition of a volume-form-weighted average, constructed so as to commute with the covariant divergence of a vector. The procedures leading to the thickness-weight averaged equation are discussed, and forms of the Eliassen–Palm flux tensor which arise are identified.

Key words: geostrophic turbulence, ocean processes, quasi-geostrophic flows

1. Introduction

Residual-mean theory enables one to recast the directionally averaged thermodynamic and momentum equations so that only diabatic eddy buoyancy fluxes appear in the thermodynamic equation, and the eddy interaction in the momentum equation appears as the divergence of an Eliassen–Palm flux vector (Eliassen & Palm 1961; Andrews & McIntyre 1976). The approach yields a simple geometric description in terms of a vector eddy flux of horizontal momentum, and has been successful in describing and generalizing earlier theoretical results concerning eddy–mean-flow interaction (Andrews & McIntyre 1978; Andrews, Holton & Conway 1987). In particular, the Eliassen–Palm flux vector sets both interaction properties, appearing as a momentum stress, and eddy propagation properties, appearing as a flux of a wave activity.

A range of approaches have been suggested in order to extend residual-mean theory to more general averages, with three-dimensional averaged fields. In the context of the hydrostatic primitive equations one traditionally introduces two Eliassen–Palm flux vectors, one for each component of the horizontal velocity (see, for example, Gent & McWilliams 1996; Smith 1999; Young 2012). Subject to an appropriate definition of a residual circulation one may, as in the directionally averaged case, remove eddy interaction terms from the thermodynamic equation. This yields a description in which only diabatic eddy buoyancy fluxes appear explicitly in the thermodynamic equation,

† Email address for correspondence: maddison@atm.ox.ac.uk

and in which eddy terms appear in the residual-mean horizontal momentum equations as the divergence of the two Eliassen–Palm flux vectors. In Cronin (1996) a rank-two Eliassen–Palm flux tensor is derived for the statistically steady quasi-geostrophic equations. More generally, in Miyahara (2006) an Eliassen–Palm flux tensor is derived for the Boussinesq momentum equations.

When averaging the buoyancy equation a diabatic eddy diffusivity may emerge, even after defining a residual-mean advection. Higher-order residual-mean velocities can be defined to form a diffusivity that is dependent only upon the true diabatic forcing (Eden, Greatbatch & Olbers 2007), although a dynamical equation for the resulting transport velocity appears obscure. The diabatic fluxes can instead be removed by considering an average based upon flow fields and, in particular, by considering an average at fixed buoyancy (Andrews 1983; de Szoeke & Bennett 1993). Only true diabatic forcing remains in the average buoyancy equation and apparent diabatic forcing, due to averaging, is entirely avoided. Averaging in this manner can preserve sharp, but moving, buoyancy gradients in the averaged system.

A particularly general treatment of this type of averaging is presented in Young (2012), which extends the thickness-weighted-average description of Andrews (1983) and de Szoeke & Bennett (1993) (see also Tung 1986; Lee & Leach 1996; Smith 1999 for similar approaches). It is shown that thickness-weighted averaging the Boussinesq hydrostatic primitive equations in buoyancy coordinates yields a set of equations from which one may, by inspection, identify Eliassen–Palm flux vectors. Initial steps are made towards a more geometric treatment of the problem, with more formal coordinate-invariant operators defined.

The details of the formulation of the residual-mean equations are of particular importance in the context of mesoscale eddy parameterization in the ocean. The widely used Gent and McWilliams parameterization (Gent & McWilliams 1990; Gent *et al.* 1995) can be considered a parameterization constructed in a residual-mean context (Greatbatch & Lamb 1990; Gent & McWilliams 1996). The scheme forms a closure for the adiabatic eddy buoyancy fluxes, which directly yields the residual-mean velocity. This is typically described as a modification of the Eulerian-mean tracer advection terms by an eddy-induced velocity, yielding a residual-mean tracer advection. Alternatively, one may formulate a model entirely in terms of residual-mean quantities, as in Ferreira & Marshall (2006) and Zhao & Vallis (2008).

In Marshall, Maddison & Berloff (2012) it is shown that, by constructing a geometric decomposition of the Plumb (1986) quasi-geostrophic momentum flux matrix, whose divergence is the eddy potential vorticity flux, one can construct a framework for eddy parameterization which conserves momentum and observes energetic constraints. Key to the enforcement of the conservation principles is an identification of the fundamental geometric object describing the nature of the eddy–mean-flow interaction as a rank-two momentum flux tensor. This contrasts with the more conventional picture, such as is assumed by a down-gradient potential vorticity flux closure, classifying the interaction in terms of a rank-one potential vorticity flux vector. The former satisfies momentum conservation while the latter, in general, does not (Welander 1973; Stewart & Thomson 1977; Marshall 1981). This can be viewed as an extension of the Prandtl ‘momentum mixing’ versus Taylor ‘vorticity mixing’ debate (Taylor 1915, 1932; Prandtl 1942). The vorticity approach is naturally free of complications associated with the pressure field, at the cost of the introduction of an implicit dynamical constraint, associated with momentum conservation, on the eddy vorticity fluxes. In the momentum approach the eddy vorticity fluxes are a derived quantity, and momentum conservation can be satisfied by construction, but the situation is complicated by the appearance of a greater number of tensor components, combined with the dynamical influence of the pressure field. A resolution of this lengthy discussion is not provided here; rather, the aim is to yield an appropriate description of the relationship between the two approaches.

The purpose of this paper is to describe the fundamental geometric nature of the eddy–mean-flow interaction. Specifically, residual-mean theory and thickness-weighted averaging in buoyancy coordinates are described using a geometric formalism, and in several cases the approaches are generalized. It follows that the fundamental object describing the eddy–mean-flow interaction is a rank-two momentum flux tensor, termed here the 'Eliassen–Palm' flux tensor. Possible forms for the Eliassen–Palm flux tensor are derived for the quasi-geostrophic equations and for the hydrostatic Boussinesq primitive equations. Combined, these descriptions generalize the geometric picture of the eddy–mean-flow interaction provided by Marshall *et al.* (2012).

The Eliassen–Palm flux tensor is here referred to as the momentum flux tensor appearing in the horizontal momentum equation, consistent with a modification of an equation for a tracer γ (which may correspond to the density, the neutral density, or the buoyancy), via a redefinition of the advection operator or the averaging operator. Furthermore, in all cases, in the absence of mechanical forcing and γ forcing, the Eliassen–Palm flux tensor, combined (where required) with a residual eddy γ flux (which may for example correspond to a diabatic eddy buoyancy flux) completely describes the influence of the eddies on the averaged flow. In the case of directional averaging the Eliassen–Palm flux vector additionally appears in the Eliassen–Palm relation as a flux of a wave activity, and leads to the Eliassen–Palm theorem, stating that the Eliassen–Palm flux vector has zero divergence for small-amplitude conservative waves. A corollary of this theorem is the non-acceleration theorem, which states that small-amplitude conservative waves do not accelerate the directionally averaged flow field (Charney & Drazin 1961). The relationships of specific forms of the Eliassen–Palm flux tensor to generalized forms of the Eliassen–Palm relation are described in Plumb (1986), Cronin (1996) and Miyahara (2006). The Eliassen–Palm theorem and non-acceleration theorem have not been extended for more general averaging operators (McDougall & McIntosh 2001; Young 2012). As a result of this it has been suggested that one should refrain from the use of the term 'Eliassen–Palm flux' when discussing more general averaging operators (McDougall & McIntosh 2001). A complete generalization of these theorems is not shown in this paper although, in a quasi-geostrophic context, there is some discussion of propagation properties. In particular, the Eliassen–Palm flux tensor is associated with eddy propagation properties, and the Eliassen–Palm and non-acceleration theorems appear when taking appropriate limits.

The paper proceeds as follows. In § 2 the Eliassen–Palm flux tensor subject to the quasi-geostrophic approximation is derived. We further discuss the relationship to the Plumb (1986) flux matrix, the Marshall *et al.* (2012) geometric description, and the Eliassen–Palm flux vector that arises subject to a directional average. In § 3 forms of the Eliassen–Palm flux tensor for the Boussinesq hydrostatic primitive equations are derived. These are considered initially via Eulerian averaging and residual-mean theory. A very general averaging operator is then considered, and the thickness-weighted-average treatment is shown to arise when applying this operator to the Boussinesq hydrostatic primitive equations. The paper concludes in § 4, including a discussion of the relevance of the Eliassen–Palm flux tensor when developing a physically consistent eddy closure.

2. Quasi-geostrophic case

In this section the quasi-geostrophic Eliassen–Palm flux tensor is described. The quasi-geostrophic equations are stated in a slightly modified form in § 2.1, by deriving an equation whose divergence is the quasi-geostrophic potential vorticity equation. This is restated in a coordinate-invariant form in § 2.2. An Eulerian averaging operator is defined in § 2.3, and this is used to derive the eddy momentum flux tensor in § 2.4. Gauge freedom is exploited in § 2.5 to derive the quasi-geostrophic residual-mean equations and forms of the quasi-geostrophic Eliassen–Palm flux tensor. In § 2.6 the geometric nature of the flux tensor is exploited to derive a geometric decomposition as per Marshall *et al.* (2012). Finally, in § 2.7, the relationship to the traditional Eliassen–Palm flux vector, arising from a directional average, is discussed.

2.1. *Quasi-geostrophic equations*

The quasi-geostrophic equations arise from an asymptotic expansion of the primitive equations, retaining terms of up to first order in Rossby number (Pedlosky 1987). The quasi-geostrophic horizontal momentum equation is:

$$(\partial_t + u_g\partial_x + v_g\partial_y)u_g - f_0 v_{ag} - \beta y v_g = -\frac{1}{\rho_0}\partial_x p_{ag} + F_x, \qquad (2.1a)$$

$$(\partial_t + u_g\partial_x + v_g\partial_y)v_g + f_0 u_{ag} + \beta y u_g = -\frac{1}{\rho_0}\partial_y p_{ag} + F_y, \qquad (2.1b)$$

where u_g and v_g are the zonal and meridional components of the velocity to leading order in Rossby number, $(u_g + u_{ag})$ and $(v_g + v_{ag})$ are the zonal and meridional components of the velocity retaining contributions up to first order in Rossby number, p_{ag} is the ageostrophic pressure, ρ_0 is a constant reference density, $f = f_0 + \beta y$ is the Coriolis parameter subject to the β-plane approximation, F_x and F_y are additional forcing, x and y are the zonal and meridional coordinates respectively, and t is time. The quasi-geostrophic buoyancy (thermodynamic) equation is:

$$(\partial_t + u_g\partial_x + v_g\partial_y)b + N_0^2 w_{ag} = B, \qquad (2.2)$$

where b is the buoyancy, w_{ag} is the vertical component (z-component) of velocity retaining contributions up to first order in Rossby number, N_0 is the buoyancy frequency, and B represents additional buoyancy forcing. We now note that since the two asymptotic components of the velocity, $\boldsymbol{u_g}$ and $\boldsymbol{u_{ag}}$, are formed from an asymptotically expanded vector, they are each formal three-dimensional vectors which, expressed in the x, y, z coordinate system, have components:

$$\boldsymbol{u_g} = \begin{pmatrix} u_g \\ v_g \\ 0 \end{pmatrix}, \quad \boldsymbol{u_{ag}} = \begin{pmatrix} u_{ag} \\ v_{ag} \\ w_{ag} \end{pmatrix}. \qquad (2.3)$$

These are each non-divergent:

$$\nabla \cdot \boldsymbol{u_g} = 0, \quad \nabla \cdot \boldsymbol{u_{ag}} = 0. \qquad (2.4)$$

By re-arranging (2.1) and (2.2) for $f_0\boldsymbol{u_{ag}}$ one can identify a third vector which, expressed in the x, y, z coordinate system, has components:

$$\boldsymbol{D} = \begin{pmatrix} v_g \\ -u_g \\ \dfrac{f_0}{N_0^2} b \end{pmatrix}, \tag{2.5}$$

with dynamical equation:

$$\partial_t \boldsymbol{D} + (\boldsymbol{u_g} \cdot \nabla)\boldsymbol{D} + f_0 \boldsymbol{u_{ag}} + \beta y \boldsymbol{u_g} = \frac{1}{\rho_0} \hat{\boldsymbol{z}} \times \nabla p_{ag} + \boldsymbol{R}, \tag{2.6}$$

where $\hat{\boldsymbol{z}}$ is the unit vertical vector and, in the x, y, z coordinate system:

$$\boldsymbol{R} = \begin{pmatrix} F_y \\ -F_x \\ \dfrac{f_0}{N_0^2} B \end{pmatrix}. \tag{2.7}$$

The divergence of $\boldsymbol{D}$ is the relative quasi-geostrophic potential vorticity:

$$\nabla \cdot \boldsymbol{D} = q - \beta y, \tag{2.8}$$

where q is the quasi-geostrophic potential vorticity (QGPV). This vector is directly related to the vector $\boldsymbol{D}$ that appears in the field equation description in Schneider, Held & Garner (2003). Utilizing the analogy with Maxwell's equations drawn in Schneider *et al.* (2003) $\boldsymbol{D}$ is here referred to as the 'QGPV induction vector'. A form of this vector also appears in Muraki, Snyder & Rotunno (1999), in which a Helmholtz decomposition of the corresponding vector is used to derive higher-order asymptotic extensions to the quasi-geostrophic equations. The QGPV induction equation (2.6) encapsulates both components of the horizontal momentum equation as well as the buoyancy equation. In addition, taking the divergence of the QGPV induction equation (2.6) yields the QGPV equation:

$$\partial_t q + \nabla \cdot (\boldsymbol{u_g} q) = \nabla \cdot \boldsymbol{R}. \tag{2.9}$$

2.2. *Coordinate-invariant quasi-geostrophic equations*

We now write the QGPV induction equation (2.6) in a coordinate-invariant form using tensor calculus notation. While such a formalism may not strictly be required here, it will be necessary in the more general cases to follow in § 3. In this section only coordinate systems which are fixed in time are considered, and hence time is not treated as a coordinate. Subscripts already have a well-established meaning in geophysical fluid dynamics. In order to avoid confusion non-index subscripts and superscripts will be enclosed within square brackets.

The QGPV induction equation (2.6) can therefore be written:

$$\partial_t D^a + ([u_g]^b D^a)_{;b} + [f_0][u_{ag}]^a + [f_1][u_g]^a = \frac{1}{[\rho_0]} \varepsilon^{abc} Z_b [p_{ag}]_{,c} + R^a, \tag{2.10}$$

where $[f_1] = \beta y$, Z^a denotes the unit vertical vector, a comma indicates a partial derivative with respect to the indexed coordinate, a semi-colon indicates a covariant derivative with respect to the indexed coordinate, and Einstein summation convention is assumed; ε^{abc} denotes the three-dimensional contravariant Levi–Civita tensor:

$$\varepsilon^{abc} = \frac{1}{\sqrt{G}}\epsilon^{abc}, \tag{2.11}$$

where G is the determinant of the covariant metric tensor:

$$G = |g_{ab}|, \tag{2.12}$$

and where ϵ^{abc} is the three-dimensional Levi–Civita symbol:

$$\epsilon^{abc} = \begin{cases} 0 & \text{if any indices are equal} \\ +1 & \text{if } abc \text{ is an even permutation of } 123 \\ -1 & \text{if } abc \text{ is an odd permutation of } 123. \end{cases} \tag{2.13}$$

$[f_0]$, $[f_1]$, and $[\rho_0]$ are treated as scalars (invariants). The continuity equation (2.4) becomes:

$$[u_g]^a_{;a} = 0, \quad [u_{ag}]^a_{;a} = 0. \tag{2.14}$$

It follows from the definition of the QGPV induction vector that the QGPV is an invariant:

$$q = D^a_{;a} + [f_1]. \tag{2.15}$$

2.3. *Averaging and eddy operators*

We introduce coordinate-invariant averaging and eddy operators. For the discussion of § 3 a particularly careful definition is required, and hence this section provides a precise definition of the averaging and eddy operators and their required properties. On a first reading one may wish to move directly to § 2.4. Note that these properties are typically assumed implicitly when deriving averaged equations, particularly when performing more general averaging such as in the thickness-weighted-average formulation of Andrews (1983), de Szoeke & Bennett (1993) and Young (2012).

The averaging operator $\overline{(\cdots)}$ is understood to be a projection operator for an arbitrary tensor from a higher dimensional space onto a space of equal or lower dimension. The former is termed the 'unaveraged space' and the latter the 'averaged space'. Conversely, the operator $\widehat{(\cdots)}$ is understood to be an extrusion operator for an arbitrary tensor from the averaged space to the unaveraged space. The eddy operator is thus defined as:

$$\phi^{i_1 i_2 \cdots \prime}_{j_1 j_2 \cdots} = \phi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots} - \widehat{\overline{\phi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}}}, \tag{2.16}$$

where $\phi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}$ is an arbitrary tensor in the unaveraged space. Tensors in the averaged system with no overbar are understood to be tensors for which $(\cdots)' = 0$. As a special case, the metric tensor in the averaged system may be denoted $[g_A]^{ab}$ if it is distinct from $\overline{g^{ab}}$.

It is assumed that the averaging operator and the extrusion operator commute with differentiation with respect to time and with the covariant divergence:

$$\overline{\partial_t(\phi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots})} = \partial_t(\overline{\phi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}}), \quad \overline{(\phi^{i_1 i_2 \cdots i_n \cdots}_{j_1 j_2 \cdots})_{;i_n}} = (\overline{\phi^{i_1 i_2 \cdots i_n \cdots}_{j_1 j_2 \cdots}})_{;i_n}, \tag{2.17a}$$

$$\widehat{\partial_t(\phi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots})} = \partial_t(\widehat{\phi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}}), \quad \widehat{(\phi^{i_1 i_2 \cdots i_n \cdots}_{j_1 j_2 \cdots})_{;i_n}} = (\widehat{\phi^{i_1 i_2 \cdots i_n \cdots}_{j_1 j_2 \cdots}})_{;i_n}. \tag{2.17b}$$

The following properties are also assumed, where a is a general constant and $\phi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}$ and $\psi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}$ are general tensors of the same type:

$$\overline{\widehat{\overline{\phi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}}}} = \overline{\phi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}}, \quad \overline{\phi^{i_1 i_2 \cdots \prime}_{j_1 j_2 \cdots}} = 0, \tag{2.18a}$$

$$\overline{a\phi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}} = a\overline{\phi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}}, \quad \overline{\phi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots} + \psi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}} = \overline{\phi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}} + \overline{\psi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}}, \tag{2.18b}$$

$$\widehat{\overline{a\phi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}}} = a\widehat{\overline{\phi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}}}, \quad \widehat{\overline{\phi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots} + \psi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}}} = \widehat{\overline{\phi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}}} + \widehat{\overline{\psi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}}}. \tag{2.18c}$$

We now insist that $g^{ab\prime} = 0$ and $[g_A]^{ab} = \overline{g^{ab}}$. One may, for example, consider $\overline{(\cdots)}$ to be an ensemble average, where the average is taken at a fixed point in space and at a fixed time (an Eulerian average):

$$\overline{\phi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}} = \overline{[\phi_\alpha]^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}} = \frac{1}{N}\sum_{\alpha=1}^{N} [\phi_\alpha]^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}, \tag{2.19a}$$

$$\overset{\alpha}{\widehat{\overline{\phi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}}}} = \overline{\phi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}}. \tag{2.19b}$$

Here $[\phi_\alpha]^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}$ corresponds to a general tensor for ensemble member α of an N-member ensemble, contrasting with $\phi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}$ which represents the tensor for all ensemble members (the full unaveraged space). Similarly $\overset{\alpha}{\widehat{(\cdots)}}$ denotes the extruded tensor for a particular ensemble member α, contrasting with $\widehat{(\cdots)}$ which denotes the extruded tensor over all ensemble members. The extrusion operator appears somewhat redundant here, although this will not be the case in later discussions. The averaging operator (2.19*a*) and the associated eddy operator satisfy the properties (2.18). The averaging operator also commutes with partial derivatives with respect to the coordinates in any coordinate system, and hence the operator commutes with respect to the covariant divergence. The operator further commutes with differentiation with respect to time.

2.4. *Quasi-geostrophic momentum flux tensor*

Applying the operator defined in the previous section to the QGPV induction equation (2.10), and assuming $[f_0]' = 0$, $[f_1]' = 0$ and $Z^{a\prime} = 0$ (as is the case for the ensemble average defined above), yields the average QGPV induction equation:

$$\partial_t \overline{D^a} + (\overline{[u_g]^b}\,\overline{D^a})_{;b} + [f_0]\overline{[u_{ag}]^a} + [f_1]\overline{[u_g]^a} = \frac{1}{[\rho_0]}\varepsilon^{abc} Z_b \overline{[p_{ag}]}_{,c} + \overline{R^a} - T^{ab}_{;b}, \tag{2.20}$$

with continuity equations:

$$\overline{[u_g]^a}_{;a} = 0, \quad \overline{[u_{ag}]^a}_{;a} = 0, \tag{2.21}$$

and where a QGPV induction eddy flux tensor has been introduced:

$$T^{ab} = \overline{[u_g]^{b\prime} D^{a\prime}}. \tag{2.22}$$

It follows that the average QGPV equation is:

$$\partial_t \overline{q} + (\overline{[u_g]^a}\,\overline{q})_{;a} = \overline{R^a}_{;a} - T^{ab}_{;ab}. \tag{2.23}$$

https://doi.org/10.1017/jfm.2013.259 Published online by Cambridge University Press

Since the double divergence of the eddy flux tensor $T^{ab}_{;ab}$ is the eddy QGPV tendency, the divergence of the eddy flux tensor $T^{ab}_{;b}$ is equal to the eddy QGPV flux plus a rotational term.

In the x, y, z coordinate system the eddy flux tensor T^{ab} has components:

$$T_a^b = g_{ac}T^{cb} = \begin{pmatrix} N & M-K & R \\ M+K & -N & S \\ 0 & 0 & 0 \end{pmatrix} \tag{2.24}$$

where the contravariant index indicates the row and the covariant index the column, and where:

$$\left.\begin{aligned} M &= \frac{1}{2}(\overline{[v_g]'^2} - \overline{[u_g]'^2}), \quad N = \overline{[u_g]'[v_g]'}, \\ K &= \frac{1}{2}(\overline{[v_g]'^2} + \overline{[u_g]'^2}), \\ R &= \frac{[f_0]}{[N_0]^2}\overline{[u_g]'b'}, \quad S = \frac{[f_0]}{[N_0]^2}\overline{[v_g]'b'}. \end{aligned}\right\} \tag{2.25}$$

M and N are the eddy Reynolds stresses (the horizontal fluxes of horizontal momentum), R and S are (proportional to) the eddy buoyancy fluxes, and K is the eddy kinetic energy.

2.5. *Eliassen–Palm flux tensor*

Since the eddy tendency in the QGPV equation appears as the double divergence of the eddy flux tensor T^{ab}, two forms of gauge freedom may be exploited. In particular, one may add rotational terms to either the columns or rows of the component expansion (2.24). Hence T^{ab} can be replaced with $[T^*]^{ab}$ where:

$$[T^*]^{ab} = T^{ab} + \varepsilon^{acd}U^b_{d;c} + \varepsilon^{bcd}V^a_{d;c}, \tag{2.26}$$

where U^a_b and V^a_b are arbitrary mixed-type tensors. Gauge freedom can, for example, be exploited to move eddy momentum fluxes between the momentum and buoyancy equations, and thereby replace horizontal momentum fluxes with horizontal buoyancy fluxes. Note that one may not, in general, delete the horizontal eddy momentum fluxes from the momentum equation without permitting an (arguably non-physical) eddy momentum flux through either the upper or lower boundaries.

In particular, the QGPV equation, and the resulting dynamics, are entirely unaffected by the addition of any antisymmetric components to the eddy flux tensor T^{ab}:

$$\partial_t\overline{q} + (\overline{[u_g]^a}\,\overline{q})_{;a} = \overline{R^a}_{;a} - [T^*]^{ab}_{;ab}, \tag{2.27}$$

where:

$$[T^*]^{ab} = T^{ab} + A^{ab}, \tag{2.28}$$

and where A^{ab} is any antisymmetric tensor, $A^{ab} = -A^{ba}$. This introduces a non-divergent (rotational) term into the QGPV induction equation. For any A^{ab} the divergence of the eddy flux tensor $T^{ab}_{;b}$ is the eddy QGPV flux, plus a rotational term. One may define a residual-mean ageostrophic velocity:

$$[u^*_{ag}]^a = \overline{[u_{ag}]^a} - \frac{1}{[f_0]}A^{ab}_{;b}, \tag{2.29}$$

where the second term corresponds to a curl of a vector in Cartesian coordinates. The QGPV induction equation therefore becomes:

$$\partial_t \overline{D^a} + (\overline{[u_g]^b}\,\overline{D^a})_{;b} + [f_0][u^*_{ag}]^a + [f_1]\overline{[u_g]^a} = \frac{1}{[\rho_0]}\varepsilon^{abc} Z_b \overline{[p_{ag}]}_{,c} + \overline{R^a} - [T^*]^{ab}_{;b}. \quad (2.30)$$

In the following subsections a number of natural gauge choices are described.

2.5.1. *Residual-mean momentum flux tensor*

If a tensor A^{ab} is chosen appropriately then $[T^*]^{ab}Z_a = 0$, and the buoyancy equation will contain no eddy contributions. In particular, one may choose:

$$A^{ab} = [A_R]^{ab} = T^{ba} - T^{ab}, \quad (2.31)$$

leading to:

$$\begin{aligned} [T^*]^{ab} = [T^*_R]^{ab} &= T^{ba} \\ &= \overline{[u_g]^{a\prime} D^{b\prime}}. \end{aligned} \quad (2.32)$$

In the x, y, z coordinate system this has components:

$$[T^*_R]^b_a = g_{ac}[T^*_R]^{cb} = \begin{pmatrix} N & M+K & 0 \\ M-K & -N & 0 \\ R & S & 0 \end{pmatrix}. \quad (2.33)$$

Hence one may, in the average QGPV induction equation (2.20), replace the eddy flux tensor T^{ab} with its transpose. This moves the eddy buoyancy fluxes from the buoyancy equation to the horizontal momentum equation, yielding a system of dynamical equations in which no eddy terms appear in the buoyancy equation. With this choice for A^{ab} the averaged horizontal momentum equation becomes (reached by taking the cross-product of the averaged QGPV induction equation with $-Z^a$):

$$\begin{aligned} &\partial_t \overline{[u_g]^a} + (\overline{[u_g]^b}\,\overline{[u_g]^a})_{;b} - [f_0]\varepsilon^{abc}[u^*_{ag}]_b Z_c - [f_1]\varepsilon^{abc}\overline{[u_g]_b} Z_c \\ &= -\frac{1}{[\rho_0]} h^{ab}\overline{[p_{ag}]}_{,b} + \overline{F^a} - [E_R]^{ab}_{;b}, \end{aligned} \quad (2.34)$$

where $h^{ab} = g^{ab} - g^{cb}Z_c Z^a$, while the averaged buoyancy equation, scaled by $[f_0]/[N_0]^2$, can be written (reached by taking the inner product of the averaged QGPV induction equation with Z^a):

$$\partial_t(\overline{D^b}Z_b) + (\overline{[u_g]^a}\,\overline{D^b}Z_b)_{;a} + [f_0][u^*_{ag}]^a Z_a = \overline{R^a}Z_a, \quad (2.35)$$

where $[E_R]^{ab}$ is the Eliassen–Palm flux tensor:

$$[E_R]^{ab} = -\varepsilon^{acd} g_{ce}[T^*_R]^{eb} Z_d. \quad (2.36)$$

In the x, y, z coordinate system the Eliassen–Palm flux tensor has components:

$$[E_R]^b_a = g_{ac}[E_R]^{cb} = \begin{pmatrix} -M-K & N & 0 \\ N & M-K & 0 \\ -S & R & 0 \end{pmatrix}. \quad (2.37)$$

Hence, for the quasi-geostrophic equations, the residual-mean equations are reached via a transpose of the eddy flux tensor T^{ab}.

https://doi.org/10.1017/jfm.2013.259 Published online by Cambridge University Press

2.5.2. *Cronin (1996) momentum flux tensor*

One may also choose A^{ab} such that, in the x, y, z coordinate system, the eddy flux tensor $[T^*]^{ab} = [T^*_C]^{ab}$ and has components:

$$[T^*_C]_a^b = g_{ac}[T^*_C]^{cb} \begin{pmatrix} N & M-K & 0 \\ M+K & -N & 0 \\ R & S & 0 \end{pmatrix}. \tag{2.38}$$

In this case the Eliassen–Palm flux tensor $[E_R]^{ab}$ is replaced by a tensor with components, in the x, y, z coordinate system:

$$[E_C]_a^b = g_{ac}[E_C]^{cb} = \begin{pmatrix} -M+K & N & 0 \\ N & M+K & 0 \\ -S & R & 0 \end{pmatrix}. \tag{2.39}$$

In Cronin (1996) an Eliassen–Palm flux tensor of this form is derived by applying residual-mean theory directly, with the exact form differing only in that in Cronin (1996) rotational buoyancy fluxes are removed (equivalent to a slightly modified gauge choice).

2.5.3. *Plumb (1986) momentum flux tensor*

Alternatively, one may choose A^{ab} such that, in the x, y, z coordinate system, the eddy flux tensor $[T^*]^{ab} = [T^*_P]^{ab}$ and has components:

$$[T^*_P]_a^b = g_{ac}[T^*_P]^{cb} \begin{pmatrix} N & M-P & 0 \\ M+P & -N & 0 \\ R & S & 0 \end{pmatrix} \tag{2.40}$$

where $P = \overline{b'^2}/(2[N_0]^2)$ is the eddy potential energy. This is exactly the Plumb (1986) flux matrix, and the divergence $[T^*_P]^{ab}_{;b} = \overline{[u_g]^{a\prime} q'}$ is exactly the eddy QGPV flux. The Plumb (1986) flux matrix is itself a version of the Taylor–Bretherton identity (Taylor 1915; Bretherton 1966), relating eddy momentum fluxes to eddy vorticity fluxes.

In this case the Eliassen–Palm flux tensor $[E_R]^{ab}$ is replaced by a tensor with components, in the x, y, z coordinate system:

$$E_a^b = g_{ac}E^{cb} = \begin{pmatrix} -M+P & N & 0 \\ N & M+P & 0 \\ -S & R & 0 \end{pmatrix}. \tag{2.41}$$

2.5.4. *Hoskins, James & White (1983) E-vector*

The Hoskins *et al.* (1983) E-vector is arrived at by choosing A^{ab} such that, in the x, y, z coordinate system, the eddy flux tensor $[T^*]^{ab} = [T^*_H]^{ab}$ and has components:

$$[T^*_H]_a^b = g_{ac}[T^*_H]^{cb} \begin{pmatrix} N & 2M & 0 \\ 0 & -N & 0 \\ R & S & 0 \end{pmatrix}. \tag{2.42}$$

The Hoskins E-vector is then the y-component of this tensor:

$$[E_H]^a = \begin{pmatrix} 2M \\ -N \\ S \end{pmatrix}. \tag{2.43}$$

In Hoskins *et al.* (1983) the limiting case $\partial_x N + \partial_z R \approx 0$ is considered, in which case the E-vector captures the dynamically significant components of the eddy flux tensor. Since the Hoskins *et al.* (1983) E-vector is formed from a component of the tensor $[T_H^*]^{ab}$ it is not a formal vector. It is noted in Hoskins *et al.* (1983) that the E-vector fails to transform as a vector, and hence it is termed a 'quasi-vector'. The correct geometric object characterizing the eddy–mean-flow interaction is the full rank-two tensor.

2.5.5. *Symmetric momentum flux tensor*

Finally, one may choose:

$$A^{ab} = [A_S]^{ab} = \tfrac{1}{2}(T^{ba} - T^{ab}), \tag{2.44}$$

leading to:

$$[T^*]^{ab} = [T_S^*]^{ab} = \tfrac{1}{2}(T^{ab} + T^{ba}). \tag{2.45}$$

In the x, y, z coordinate system this has components:

$$[T_S^*]_a^{\ b} = g_{ac}[T_S^*]^{cb} = \begin{pmatrix} N & M & \tfrac{1}{2}R \\ M & -N & \tfrac{1}{2}S \\ \tfrac{1}{2}R & \tfrac{1}{2}S & 0 \end{pmatrix}. \tag{2.46}$$

Hence one may, in the average QGPV induction equation (2.20), replace the eddy flux tensor T^{ab} with its symmetric part. This 'half-residual-mean' formulation, in which half of the buoyancy fluxes are transferred to the momentum equation and half are retained in the buoyancy equation, filters out trivially rotational eddy QGPV fluxes, and yields a QGPV induction eddy flux tensor with (among the possibilities associated with choices of A^{ab}) minimum Frobenius norm.

2.6. *Geometric decomposition*

The QGPV induction eddy flux tensor T^{ab} is an inherently geometric object. Hence this represents an entirely geometric description of the influence of the eddies on the mean flow. This is necessarily equivalent to the geometric description in Marshall *et al.* (2012).

In order to demonstrate this, first identify two invariants, the eddy kinetic energy:

$$K = \tfrac{1}{2}\overline{[u_g]^{a\prime}[u_g]_a^{\prime}}, \tag{2.47}$$

and a weighted sum of the eddy kinetic and the eddy potential energies:

$$\begin{aligned} L &= \tfrac{1}{2}\overline{D^{a\prime}D_a^{\prime}} \\ &= K + \frac{[f_0]^2}{[N_0]^2}P. \end{aligned} \tag{2.48}$$

Proceeding in a similar manner to Marshall *et al.* (2012) one can derive, via the triangle inequality:

$$\overline{[u_g]^{a\prime}[u_g]^{b\prime}}\,\overline{[u_g]'_a[u_g]'_b} \leqslant 4K^2, \tag{2.49a}$$

$$\overline{[u_g]^{a\prime}D^{b\prime}}\,\overline{[u_g]'_aD'_b} - \overline{[u_g]^{a\prime}[u_g]^{b\prime}}\,\overline{[u_g]'_a[u_g]'_b} \leqslant 4\frac{[f_0]^2}{[N_0]^2}KP. \tag{2.49b}$$

Hence one may define two additional invariants via:

$$\overline{[u_g]^{a\prime}[u_g]^{b\prime}}\,\overline{[u_g]'_a[u_g]'_b} = 4[\gamma_m^*]^2K^2, \tag{2.50a}$$

$$\overline{[u_g]^{a\prime}D^{b\prime}}\,\overline{[u_g]'_aD'_b} - \overline{[u_g]^{a\prime}[u_g]^{b\prime}}\,\overline{[u_g]'_a[u_g]'_b} = 4[\gamma_b^*]^2\frac{[f_0]^2}{[N_0]^2}KP, \tag{2.50b}$$

where $[\gamma_m^*]$ and $[\gamma_b^*]$ are non-dimensional and bounded between zero and unity. In the x, y, z coordinate system this leads directly to a general decomposition for the components of the Eliassen–Palm flux tensor:

$$\left.\begin{aligned} &M = -L[\gamma_m^*]\cos^2[\lambda^*]\cos 2[\phi_m], \quad N = L[\gamma_m^*]\cos^2[\lambda^*]\sin 2[\phi_m],\\ &R = L[\gamma_b^*]\sin 2[\lambda^*]\cos[\phi_b], \quad S = L[\gamma_b^*]\sin 2[\lambda^*]\sin[\phi_b],\\ &K = L\cos^2[\lambda^*], \quad \frac{[f_0]^2}{[N_0]^2}P = L\sin^2[\lambda^*]. \end{aligned}\right\} \tag{2.51}$$

$[\lambda^*]$ is an invariant expressing the partitioning between eddy kinetic and eddy potential energies, and $[\phi_m]$ and $[\phi_b]$ are coordinate-dependent angles expressing the orientation of the Eliassen–Palm flux tensor: $[\phi_m]$ is the Reynolds' stress angle and $[\phi_b]$ is the eddy buoyancy flux angle.

In the context of the atmosphere or ocean the vertical coordinate is typically given a special status. The coordinate system is often identified purely by the choice of the vertical coordinate, and some orthonormal coordinate system in the horizontal is assumed. Hence it is meaningful to consider the case in which only horizontal transformations are considered, and in which $x^3 = z$ (a 'z-coordinate' system). Subject to this constraint one can consider the scaled eddy flux tensor (a tensor under horizontal transformations) which, in the x, y, z coordinate system, has components:

$$[T_W]_a^b = g_{ac}[T_W]^{cb} = \begin{pmatrix} N & M-K & \dfrac{[N_0]}{[f_0]}R \\ M+K & -N & \dfrac{[N_0]}{[f_0]}S \\ 0 & 0 & 0 \end{pmatrix}. \tag{2.52}$$

Proceeding as before leads directly to the Marshall *et al.* (2012) geometric decomposition:

$$\left.\begin{aligned} &M = -E[\gamma_m]\cos^2\lambda\cos 2[\phi_m] \quad N = E[\gamma_m]\cos^2\lambda\sin 2[\phi_m],\\ &R = \frac{[f_0]}{[N_0]}E[\gamma_b]\sin 2\lambda\cos[\phi_b], \quad S = \frac{[f_0]}{[N_0]}E[\gamma_b]\sin 2\lambda\sin[\phi_b],\\ &K = E\cos^2\lambda, \quad P = E\sin^2\lambda, \end{aligned}\right\} \tag{2.53}$$

where $E = K + P$ is the total eddy energy and λ expresses the partitioning between eddy kinetic and eddy potential energies.

The geometric decomposition (2.51) can be further related to the Eliassen–Palm flux tensor (2.41) by noting that the singular-value decomposition of this tensor can be represented by an oriented ellipse. The (squared) Frobenius norm $E^{ab}E_{ab}$ is bounded above by $(2K^2 + 2P^2 + 4([f_0]^2/[N_0]^2)KP)$. This yields one dimensional parameter and one non-dimensional parameter $\mathscr{E}$ bounded between zero and unity and defined such that $E^{ab}E_{ab} = \mathscr{E}^2(2K^2 + 2P^2 + 4([f_0]^2/[N_0]^2)KP)$. Given the two largest singular values of E^{ab}, $[\sigma_1]$ and $[\sigma_2]$, with $|[\sigma_1]| \geqslant |[\sigma_2]|$, one can define an eccentricity that is non-dimensional and bounded between zero and unity:

$$e = \frac{|[\sigma_1]| - |[\sigma_2]|}{|[\sigma_1]| + |[\sigma_2]|}. \tag{2.54}$$

Finally the singular vectors of E^{ab} define directions. Since they are orthogonal for distinct singular values the singular vectors are associated with three independent angles. As in the geometric decomposition (2.51) the singular-value decomposition of E^{ab} is associated with six free parameters: one dimensional and five non-dimensional and bounded in magnitude.

A key parameter that arises in the singular-value decomposition of E^{ab} is the determinant of the symmetric eddy flux tensor $[T^*_S]^{ab}$. Expressed in terms of the decomposition (2.51) this is given by:

$$|[T^*_S]^{ab}| = \tfrac{1}{4}L^3[\gamma_m][\gamma_b]^2\cos^2[\lambda^*]\sin^2 2[\lambda^*]\sin 2([\phi_m] - [\phi_b]). \tag{2.55}$$

This signed quantity is an inherently three-dimensional property of the eddy interaction. A positive value indicates that the eddy Reynolds stress angle $[\phi_m]$ is oriented to the left of the eddy buoyancy flux angle $[\phi_b]$.

2.7. *Eliassen–Palm flux vector*

Let $\overline{(\cdots)}$ correspond to a directional average and, without loss of generality, let the average be taken along the x-direction. Then averaged quantities have no x-derivative and, subject to zero-flux boundary conditions, the y-component of the velocity vanishes. Hence the residual-mean eddy flux tensor $[T^*_R]^{ab}$ can be replaced with a tensor which, in the x, y, z coordinate system, has components:

$$\widehat{[T^*_R]}_a{}^b = g_{ac}\widehat{[T^*_R]}^{cb} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -N & 0 \\ 0 & S & 0 \end{pmatrix}. \tag{2.56}$$

The second column of this tensor is the Eliassen–Palm flux 'vector'. Since this appears as a component of a rank-two tensor, the Eliassen–Palm flux vector is not, here, a formal vector. A formal vector emerges if considering only transformations of the y-coordinate, or if one considers the inner product of $[T^*_R]^{ab}$ with a unit vector in the y-direction.

A key feature of the Eliassen–Palm flux vector that arises when considering directional averaging is that it describes both the eddy interaction and the eddy propagation properties. The Eliassen–Palm flux vector acts both as a momentum flux and, via the Eliassen–Palm relation, as a flux of wave activity $\overline{\varLambda}/\partial_y\overline{q}$ with $2\varLambda = q'q'$. A consequence of the Eliassen–Palm relation is the Eliassen–Palm theorem: for a steady eddy field of small amplitude and neglecting mechanical or diabatic forcing the Eliassen–Palm flux vector has zero divergence. A direct extension of the Eliassen–Palm relation to more general averages appears elusive and, indeed, appears as something of a contradiction: the eddy momentum stress appearing in the momentum equation must necessary be a rank-two tensor, while the flux of any scalar wave activity must necessarily be rank one.

Nevertheless, some general eddy propagation properties can be identified. Consider, for example, the eddy enstrophy equation:

$$\partial_t \overline{\Lambda} + (\overline{[u_g]^a}\,\overline{\Lambda})_{;a} + (\overline{[u_g]^{a\prime}\Lambda})_{;a} = -\overline{[u_g]^{a\prime}q'}\,\overline{q}_{,a} + \overline{R^{a\prime}_{;a}q'}. \tag{2.57}$$

Utilizing the Plumb (1986) gauge choice (2.40) this can be written:

$$\partial_t \overline{\Lambda} + (\overline{[u_g]^a}\,\overline{\Lambda})_{;a} + (\overline{[u_g]^{a\prime}\Lambda})_{;a} = -([T^*_P]^{ac}\overline{q}_{,a})_{;c} + [T^*_P]^{ac}\overline{q}_{;ac} + \overline{R^{a\prime}_{;a}q'}. \tag{2.58}$$

Hence the flux tensor $[T^*_P]^{ab}$ (related to the Eliassen–Palm flux tensor via $E = -\varepsilon^{acd}g_{ce}[T^*_P]^{eb}Z_d$) leads to an eddy enstrophy flux term (the first right-hand-side term in (2.58)) and a term corresponding to conversion between mean and eddy enstrophy (the second right-hand-side term in (2.58)). Note that, due to the symmetry of the Hessian tensor $\overline{q}_{;ab}$, the conversion term is invariant under the addition of any antisymmetric components to the stress tensor. Hence, more generally:

$$\partial_t \overline{\Lambda} + (\overline{[u_g]^a}\,\overline{\Lambda})_{;a} + (\overline{[u_g]^{a\prime}\Lambda})_{;a} = -([T^*_P]^{ac}\overline{q}_{,a})_{;c} + [T^*]^{ac}\overline{q}_{;ac} + \overline{R^{a\prime}_{;a}q'}, \tag{2.59}$$

where, as before, $[T^*]^{ab} = T^{ab} + A^{ab}$, $T^{ab} = \overline{[u_g]^{b\prime}D^{a\prime}}$, and $A^{ab} = -A^{ba}$ is an antisymmetric tensor. The eddy enstrophy flux term is not invariant under such a transformation. In the linear limit and with a constant stratification it follows that:

$$[T^*_P]^{ba}\overline{q}_{,b} = (c^a - \overline{[u_g]^a})\overline{\Lambda}, \tag{2.60}$$

where c^a is the group velocity (see also Hoskins *et al.* 1983; Plumb 1985, 1986 for related discussions).

Hence while no generalization to the Eliassen–Palm theorem is derived here, the Eliassen–Palm flux tensor does describe both eddy interaction and eddy propagation properties. The Eliassen–Palm flux vector appears in an appropriate limit, and hence the Eliassen–Palm relation, Eliassen–Palm theorem, and non-acceleration theorem also emerge in an appropriate limit.

3. Primitive equation case

In this section the coordinate-invariant residual-mean equations are derived. The general approach encapsulates a number of previous residual-mean formulations, and allows the associated Eliassen–Palm flux tensors to be identified and formally defined. This provides a precise definition of the fundamental geometric nature of the eddy–mean-flow interaction. Section 3.1 describes the coordinate-invariant form of the Boussinesq hydrostatic primitive equations. Section 3.2 details the application of an Eulerian average. A very general averaging operator is then considered in § 3.3, and the thickness-weighted average approach is shown to arise when applying this operator to the Boussinesq hydrostatic primitive equations in §§ 3.4 and 3.5.

3.1. *Coordinate-invariant primitive equations*

In this and the following subsections time-dependent coordinate transformations are considered, and hence time is treated as an independent coordinate. The coordinate-invariant Boussinesq hydrostatic momentum equation is:

$$(u^b v^a)_{;b} + \varepsilon^{abcd} f_b v_c T_d = -\frac{1}{[\rho_0]}h^{ab}p_{,b} + \frac{\rho}{[\rho_0]}h^{ab}\Phi_{,b} + F^a, \tag{3.1}$$

where the velocity is assumed to be incompressible:

$$u^a_{;a}. \tag{3.2}$$

In the x, y, z, t coordinate system, where x is the zonal coordinate, y is the meridional coordinate, z is the vertical coordinate, and t is the time coordinate, the velocity vector has components:

$$u^a = \begin{pmatrix} u \\ v \\ w \\ 1 \end{pmatrix}. \tag{3.3}$$

For full dimensional consistency one may scale the time component of u^a by an arbitrary constant velocity scale. The factors that emerge from such a scaling are neglected here. Also, v^a is the horizontal velocity, $v^a = u^a - u^b Z_b Z^a$, where Z^a is the unit vertical vector; g^{ab} is the metric tensor. For simplicity $g^{ab} = \delta^{ab}$ in the x, y, z, t coordinate system is assumed. The four-dimensional contravariant Levi–Civita tensor is ε^{abcd} :

$$\varepsilon^{abcd} = \frac{1}{\sqrt{G}} \epsilon^{abcd}, \tag{3.4}$$

where G is the determinant of the covariant metric tensor:

$$G = |g_{ab}|, \tag{3.5}$$

and where ϵ^{abcd} is the four-dimensional Levi–Civita symbol:

$$\epsilon^{abcd} = \begin{cases} 0 & \text{if any indices are equal} \\ +1 & \text{if } abcd \text{ is an even permutation of } 1234 \\ -1 & \text{if } abcd \text{ is an odd permutation of } 1234. \end{cases} \tag{3.6}$$

Also, in (3.1) f^b is the Coriolis vector, ρ is the density, $[\rho_0]$ is a constant reference density, p is the pressure, Φ is the gravitational potential, T^a is the unit time vector, and F^a includes additional forcing. F^a has no vertical or time components, with $F^a Z_a = 0$ and $F^a T_a = 0$. In addition, h^{ab} is defined to be $h^{ab} = g^{ab} - g^{cb} T_c T^a$, so that the time component of (3.1) is the continuity equation (3.2) (i.e. (3.2) is arrived at by taking the inner product of (3.1) with T_a). Note that here the $(u^b v^a)_{;b}$ term represents the full material derivative, including the time derivative. For example, in the x, y, z, t coordinate system:

$$(u^b v^a)_{;b} = \partial_t v^a + u\partial_x v^a + v\partial_y v^a + w\partial_z v^a. \tag{3.7}$$

A further equation for some scalar tracer γ is assumed:

$$(u^a \gamma)_{;a} = \Theta, \tag{3.8}$$

where Θ includes additional forcing. This equation could correspond to an evolution equation for the buoyancy b, or more rigorously for the neutral density (Jackett & McDougall 1997). For generality the scalar γ is, at this stage, left unspecified.

3.2. *Eulerian averaging and the Eliassen–Palm flux tensor*

In this section the Eliassen–Palm flux tensor arising from an Eulerian average is derived. The derivation follows Gent & McWilliams (1996) and Nurser & Lee (2004), although here emphasis is placed on a general coordinate-invariant geometric formalism. A similar development is also presented in Plumb & Ferrari (2005).

We again make use of the coordinate-invariant averaging operator $\overline{(\cdots)}$ and the coordinate-invariant eddy operator $(\cdots)'$ defined as per (2.16). It is assumed that the averaging operator and the extrusion operator commute with respect to the covariant divergence:

$$\overline{(\phi^{i_1 i_2 \ldots i_n \ldots}_{j_1 j_2 \ldots})_{;i_n}} = (\overline{\phi^{i_1 i_2 \ldots i_n \ldots}_{j_1 j_2 \ldots}})_{;i_n}, \quad \overline{\widehat{(\phi^{i_1 i_2 \ldots i_n \ldots}_{j_1 j_2 \ldots})_{;i_n}}} = (\overline{\widehat{\phi^{i_1 i_2 \ldots i_n \ldots}_{j_1 j_2 \ldots}}})_{;i_n}, \tag{3.9}$$

where $\phi^{i_1 i_2 \ldots}_{j_1 j_2 \ldots}$ is any tensor. The properties (2.18) are also assumed and, as before, it is assumed that $g^{ab'} = 0$ and $[g_A]^{ab} = \overline{g^{ab}}$. One may, for example, consider $\overline{(\cdots)}$ to be an ensemble average taken at fixed point (an Eulerian average), as per (2.19*a*). Assuming $f^{a'} = 0$ and $\Phi' = 0$, the average momentum equation is:

$$(\overline{u^b}\,\overline{v^a})_{;b} + \varepsilon^{abcd} f_b \overline{v_c} T_d = -\frac{1}{[\rho_0]} h^{ab} \overline{p}_{,b} + \frac{\overline{\rho}}{[\rho_0]} h^{ab} \Phi_{,b} + \overline{F^a} - (\overline{u^{b\prime} v^{a\prime}})_{;b}, \tag{3.10}$$

and the average tracer equation is:

$$(\overline{u^a}\,\overline{\gamma})_{;a} = \overline{\Theta} - (\overline{u^{a\prime}\gamma'})_{;a}. \tag{3.11}$$

Introduce the residual-mean velocity:

$$[u^*]^a = \overline{u^a} - [u_R]^a, \tag{3.12}$$

where:

$$[u_R]^a = [\Psi_R]^{ab}_{;b}, \tag{3.13a}$$

$$[\Psi_R]^{ab} = \overline{u^{a\prime}\gamma'} \frac{H^b}{H^c \overline{\gamma}_{,c}} - \overline{u^{b\prime}\gamma'} \frac{H^a}{H^c \overline{\gamma}_{,c}}. \tag{3.13b}$$

It is assumed that the $[\Psi_R]^{ab}$ thus defined is finite. H^a is an as-yet-unspecified vector field which may vary in space and time. Note that in Cartesian coordinates, and if $H^a T_a = 0$, (3.13) states that:

$$\boldsymbol{u}_R = \nabla \times \left(\overline{\boldsymbol{u}'\gamma'} \times \frac{\boldsymbol{H}}{\boldsymbol{H} \cdot \nabla \overline{\gamma}} \right). \tag{3.14}$$

By definition $[u^*]^a$ and $[u_R]^a$ are non-divergent. Via the identity (A 4) derived in appendix A it follows that, for any vector field H^a:

$$\overline{u^{a\prime}\gamma'} = ([\Psi_R]^{ab}\overline{\gamma})_{;b} - [u_R]^a \overline{\gamma} + \overline{u^{b\prime}\gamma'}\, \overline{\gamma}_{,b} \frac{H^a}{H^c \overline{\gamma}_{,c}}. \tag{3.15}$$

Hence the eddy flux can be decomposed into a non-divergent component, an advective component, and a residual component in the direction of H^a (see Nurser & Lee 2004, particularly figure 2 and § 2). Taking the divergence leads to:

$$(\overline{u^{a\prime}\gamma'})_{;a} = \left(\overline{u^{b\prime}\gamma'}\, \overline{\gamma}_{,b} \frac{H^a}{H^c \overline{\gamma}_{,c}} \right)_{;a} - ([u_R]^a \overline{\gamma})_{;a}. \tag{3.16}$$

The average tracer equation (3.11) can therefore be written:

$$([u^*]^a \overline{\gamma})_{;a} = \overline{\Theta} - \left(\overline{u^{b\prime}\gamma'}\, \overline{\gamma}_{,b} \frac{H^a}{H^c \overline{\gamma}_{,c}} \right)_{;a}. \tag{3.17}$$

Note that the above analysis makes no assumptions regarding the time component of H^a. Hence one can consider 'space–time residual-mean' formulations in which $[u^*]^a T_a \neq 1$. In this case the full material derivative operator is modified, including the time derivative. Note that this terminology should not be confused with the 'temporal-residual mean' formulations of McDougall & McIntosh (1996) or McDougall & McIntosh (2001), which specifically refer to the use of a time average.

The residual-mean momentum equation is therefore:

$$([u^*]^b[v^*]^a)_{;b} + \varepsilon^{abcd} f_b [v^*]_c T_d = -\frac{1}{[\rho_0]} h^{ab} \overline{p}_{,b} + \frac{\overline{\rho}}{[\rho_0]} h^{ab} \Phi_{,b} + \overline{F^a} - E^{ab}_{;b}, \tag{3.18}$$

where the Eliassen–Palm flux tensor is given by:

$$E^{ab} = \overline{u^{b\prime} v^{a\prime}} + \overline{u^b}[v_R]^a + [u_R]^b [v^*]^a + \varepsilon^{acde} f_c g_{di} ([\Psi_R]^{ib} - [\Psi_R]^{jb} Z_j Z^i) T_e, \tag{3.19}$$

and where:

$$\begin{aligned} [v^*]^a &= \overline{v^a} - [v_R]^a \\ &= [u^*]^a - [u^*]^b Z_b Z^a \end{aligned} \tag{3.20}$$

is the horizontal component of the residual-mean velocity. Equation (3.18) follows directly from the average momentum equation (3.10) and the definition of the residual-mean velocity (3.12), together with $T_{a;b} = 0$, $g_{ab:c} = 0$, $\Phi' = 0$, and the assumption that:

$$\varepsilon^{acde} f_{c;b} g_{di} ([\Psi_R]^{ib} - [\Psi_R]^{jb} Z_j Z^i) T_e = 0. \tag{3.21}$$

This residual-mean formulation yields a residual-mean tracer equation (3.17) in which only eddy fluxes in the H^a direction appear explicitly, and a residual-mean momentum equation (3.18) in which the eddy forcing appears as the divergence of a rank-two Eliassen–Palm flux tensor.

The usual form of the residual-mean equations, for example as discussed in Gent & McWilliams (1996) and McDougall & McIntosh (1996), is reached if one chooses H^a to be equal to the unit vertical vector $H^a = Z^a$. This yields a residual-mean velocity with components in the x, y, z, t coordinate system:

$$[u^*]^a = \begin{pmatrix} \overline{u} - \partial_z \left(\dfrac{\overline{u'\gamma'}}{\partial_z \overline{\gamma}} \right) \\ \overline{v} - \partial_z \left(\dfrac{\overline{v'\gamma'}}{\partial_z \overline{\gamma}} \right) \\ \overline{w} + \partial_x \left(\dfrac{\overline{u'\gamma'}}{\partial_z \overline{\gamma}} \right) + \partial_y \left(\dfrac{\overline{v'\gamma'}}{\partial_z \overline{\gamma}} \right) \\ 1 \end{pmatrix}. \tag{3.22}$$

With $H^a = Z^a$, $\gamma = b$, and $\partial_z \overline{b} = [N_0]^2$, the Eliassen–Palm flux tensor (3.19) reduces to (2.39) in the quasi-geostrophic limit (Gent & McWilliams 1996).

The choice $H^a = Z^a$ can lead to an ill-defined residual-mean velocity where the vertical gradient of $\overline{\gamma}$ is small. One may reduce such issues by, as in Andrews & McIntyre (1978) and Holton (1981), choosing H^a to be equal to the average (spatial) gradient of $\overline{\gamma}$, $H^a = h^{ab} \overline{\gamma}_{,b}$. With this choice the average tracer equation (3.17) takes the form:

$$([u^*]^a \overline{\gamma})_{;a} = \overline{\Theta} + (\kappa h^{ab} \overline{\gamma}_{,b})_{;a}, \tag{3.23}$$

https://doi.org/10.1017/jfm.2013.259 Published online by Cambridge University Press

where the eddy diffusivity is given by:

$$\kappa = -\frac{\overline{u^{a\prime}\gamma^{\prime}}\,\overline{\gamma}_{,a}}{h^{bc}\overline{\gamma}_{,b}\overline{\gamma}_{,c}}. \tag{3.24}$$

While such a diffusivity may be defined in principle (and, in the higher-order formulation of Eden *et al.* 2007, higher-order diffusivities may be defined similarly), this does not imply that the diffusivity is well behaved or that this is a useful interpretation of the residual flux.

In general one can encounter problems if, at domain boundaries, H^a is not perpendicular to the boundary normal n^a. At the boundary the normal component of the eddy flux vanishes. Hence, on domain boundaries, one has from (3.15):

$$0 = \left[([\Psi_R]^{ab}\overline{\gamma})_{;b} - [u_R]^a\overline{\gamma}\right] n_a + \overline{u^{b\prime}\gamma^{\prime}}\,\overline{\gamma}_{,b}\frac{H^a n_a}{H^c\alpha_{,c}}. \tag{3.25}$$

As discussed in detail in Nurser & Lee (2004) and Plumb & Ferrari (2005), if $H^a n_a \neq 0$ then this constraint can lead to non-trivial boundary conditions for the residual-mean velocity. For example, in Held & Schneider (1999) H^a is chosen to be equal to the unit meridional vector precisely in order to align with the lower boundary of the atmosphere. However, as noted by Nurser & Lee (2004) and Plumb & Ferrari (2005), one has complete freedom in the choice of H^a and hence one can, locally at domain boundaries, enforce $H^a n_a = 0$.

In the case of a directional average, for linearized perturbations and in steady state, the remaining right-hand-side eddy term in (3.17) can be shown to vanish when $\Theta = 0$ (Holton 1981). For more general averaging operators this result need not hold (McDougall & McIntosh 1996). This issue can be addressed by defining a higher-order residual-mean velocity as per the temporal-residual-mean I formulation of McDougall & McIntosh (1996) or the arbitrary-order residual-mean formulation of Eden *et al.* (2007), and such formulations are described in appendix B. However, while one can construct a residual-mean tracer equation to higher-order in perturbation amplitude, attempting to derive an associated residual-mean momentum equation is troublesome. Additional terms involving time derivatives arise, meaning that it is unclear how one derives a useful dynamical equation for the higher-order residual-mean velocity. This motivates the development of the thickness-weighted-average formulation of Andrews (1983), de Szoeke & Bennett (1993) and Young (2012), utilized in the temporal-residual-mean II formulation of McDougall & McIntosh (2001) and described in the following sections.

3.3. *Volume-form-weighted averaging*

In previous sections only averaging operators with $g^{ab\prime} = 0$ and $[g_A]^{ab} = \overline{g^{ab}}$ have been considered, such as arise from an Eulerian average. These assumptions are now relaxed, and somewhat more general averaging operators are considered. The motivation for this treatment is to formalize and generalize averaging in coordinates based upon non-averaged fields, and in particular to generalize the thickness-weighted average of Andrews (1983), de Szoeke & Bennett (1993) and Young (2012). Key to the discussions of the previous sections have been the properties of the averaging operator (2.18), and commutation of the averaging operator with the covariant divergence (3.9). We therefore seek an averaging operator which preserves these properties. In particular, as in the discussion of Young (2012), we seek to define a single averaging operator which can be applied consistently to both the momentum and thermodynamic equations. This is a departure, for example, from the approach of Lee & Leach (1996) and Smith (1999).

Consider an N-member ensemble, and for each ensemble member α define an independent coordinate system $[\breve{x}_\alpha]^a$. From a coordinate system $\tilde{x}^a$ let the contravariant transformation matrices to the $[\breve{x}_\alpha]^a$ coordinate systems be given by $[\Lambda_\alpha]^a_b$, which are assumed non-singular, and let the corresponding covariant transformation matrices be given by $[M_\alpha]^a_b$:

$$[\Lambda_\alpha]^a_b = \frac{\partial [\breve{x}_\alpha]^a}{\partial \tilde{x}^b}, \quad [M_\alpha]^a_b = \frac{\partial \tilde{x}^a}{\partial [\breve{x}_\alpha]^b}. \tag{3.26}$$

Define an ensemble average for an arbitrary quantity $\mathscr{F}$ such that, in the $\tilde{x}^a$ coordinate system:

$$\langle \mathscr{F} \rangle|_{\tilde{x}^1,\tilde{x}^2,\ldots} = \langle [\mathscr{F}_\alpha] \rangle|_{\tilde{x}^1,\tilde{x}^2,\ldots} = \frac{1}{N}\sum_{\alpha=1}^{N} [\mathscr{F}_\alpha]|_{[\breve{x}_\alpha]^1=\tilde{x}^1,[\breve{x}_\alpha]^2=\tilde{x}^2\ldots}, \tag{3.27}$$

where $[\mathscr{F}_\alpha]$ indicate the values of $\mathscr{F}$ for each ensemble member α and where the summation argument is evaluated at the indicated locations. It follows by definition that in the $\tilde{x}^a$ coordinate system and for an arbitrary contravariant tensor $\phi^{i_1 i_2\cdots}$:

$$\frac{\partial}{\partial \tilde{x}^a}\left\langle [\Lambda_\alpha]^{i_1}_{j_1}[\Lambda_\alpha]^{i_2}_{j_2}\ldots[\phi_\alpha]^{j_1 j_2\cdots}\right\rangle = \left\langle \frac{\partial}{\partial [\breve{x}_\alpha]^a}([\Lambda_\alpha]^{i_1}_{j_1}[\Lambda_\alpha]^{i_2}_{j_2}\ldots[\phi_\alpha]^{j_1 j_2\cdots})\right\rangle. \tag{3.28}$$

Hence:

$$\frac{1}{\sqrt{\tilde{G}}}\frac{\partial}{\partial \tilde{x}^a}\left(\sqrt{\tilde{G}}\,\overline{\phi^{i_1 i_2\cdots}}\right) = \overline{[M_\alpha]^{i_1}_{j_1}[M_\alpha]^{i_2}_{j_2}\cdots\frac{1}{\sqrt{[\breve{G}_\alpha]}}\frac{\partial}{\partial [\breve{x}_\alpha]^a}\left(\sqrt{[\breve{G}_\alpha]}[\Lambda_\alpha]^{j_1}_{k_1}[\Lambda_\alpha]^{j_2}_{k_2}\cdots[\phi_\alpha]^{k_1 k_2\cdots}\right)}, \tag{3.29}$$

where $\overline{(\cdots)}$ is a volume-form-weighted average, defined for an arbitrary mixed-type tensor $\phi^{i_1 i_2\ldots}_{j_1 j_2\ldots}$ such that, in the $\tilde{x}^a$ coordinate system:

$$\overline{\phi^{i_1 i_2\ldots}_{j_1 j_2\ldots}} = \overline{[\phi_\alpha]^{i_1 i_2\ldots}_{j_1 j_2\ldots}} = \frac{1}{\sqrt{\tilde{G}}}\left\langle [\Lambda_\alpha]^{i_1}_{k_1}[\Lambda_\alpha]^{i_2}_{k_2}\ldots[M_\alpha]^{l_1}_{j_1}[M_\alpha]^{l_2}_{j_2}\ldots\sqrt{[\breve{G}_\alpha]}[\phi_\alpha]^{k_1 k_2\ldots}_{l_1 l_2\ldots}\right\rangle. \tag{3.30}$$

$\sqrt{\tilde{G}}$ is the square root of the determinant of the covariant metric tensor in the $\tilde{x}^a$ coordinate system and $\sqrt{[\breve{G}_\alpha]}$ is the square root of the determinant of the covariant metric tensor in each coordinate system $[\breve{x}_\alpha]^a$ (the volume forms for the coordinate systems $\tilde{x}^a$ and $[\breve{x}_\alpha]^a$). It follows from (3.29) that for a contravariant vector:

$$\frac{1}{\sqrt{\tilde{G}}}\frac{\partial}{\partial \tilde{x}^i}\left(\sqrt{\tilde{G}}\,\overline{\phi^i}\right) = \overline{\frac{1}{\sqrt{[\breve{G}_\alpha]}}\frac{\partial}{\partial [\breve{x}_\alpha]^i}\left(\sqrt{[\breve{G}_\alpha]}[\Lambda_\alpha]^i_j[\phi_\alpha]^j\right)}, \tag{3.31}$$

and hence that any volume-form-weighted average operator commutes with the covariant divergence of a vector:

$$(\overline{\phi^i})_{;i} = \overline{(\phi^i)_{;i}}. \tag{3.32}$$

Finally, define an extrusion operator $\widehat{(\cdots)}$ such that, for each ensemble member α, and in each $[\breve{x}_\alpha]^a$ coordinate system:

$$\left.\overset{\alpha}{\widehat{\overline{\phi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}}}}\right|_{[\breve{x}_\alpha]^1, [\breve{x}_\alpha]^2, \ldots} = \left\{ [M_\alpha]^{i_1}_{k_1} [M_\alpha]^{i_2}_{k_2} \ldots [\Lambda_\alpha]^{l_1}_{j_1} [\Lambda_\alpha]^{l_2}_{j_2} \ldots \overline{\phi^{k_1 k_2 \cdots}_{l_1 l_2 \cdots}} \right\}\Big|_{\tilde{x}^1 = [\breve{x}_\alpha]^1, \tilde{x}^2 = [\breve{x}_\alpha]^2 \ldots}, \tag{3.33}$$

and thus define an eddy operator:

$$\phi^{i_1 i_2 \ldots \prime}_{j_1 j_2 \ldots} = \phi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots} - \widehat{\overline{\phi^{i_1 i_2 \cdots}_{j_1 j_2 \cdots}}}. \tag{3.34}$$

Any volume-form-weighted average operator (3.30) and the associated eddy operator also satisfy the properties (2.18). For following discussions it is useful to also define an operator $(\cdots)^\#$ such that, for an arbitrary quantity $\mathscr{F}$ and in each $[\breve{x}_\alpha]^a$ coordinate system:

$$[\mathscr{F}^\#_\alpha]|_{[\breve{x}_\alpha]^1, [\breve{x}_\alpha]^2, \ldots} = [\mathscr{F}_\alpha]|_{[\breve{x}_\alpha]^1, [\breve{x}_\alpha]^2, \ldots} - \langle \mathscr{F}_\alpha \rangle|_{\tilde{x}^1 = [\breve{x}_\alpha]^1, \tilde{x}^2 = [\breve{x}_\alpha]^2 \ldots}, \tag{3.35}$$

where $[\mathscr{F}^\#_\alpha]$ indicates the values of $\mathscr{F}^\#$ for each ensemble member α.

Hence a volume-form-weighted average is a linear projection operator that, for a contravariant vector, commutes with the covariant divergence. It is important to note that these properties are preserved for any choice of coordinate systems $[\breve{x}_\alpha]^a$ and $\tilde{x}^a$ (for which the $[\Lambda_\alpha]^a_b$ are non-singular). Eulerian averaging appears as a special case with $[\breve{x}_{\alpha_1}]^a = [\breve{x}_{\alpha_2}]^a = \tilde{x}^a \, \forall \alpha_1, \alpha_2 \in \{1 \ldots N\}$. One may restrict the choice of coordinate system by insisting that the volume-form-weighted average of a constant scalar yields the same constant scalar. By considering the definition (3.30) for a unity-valued scalar this leads to the requirement that:

$$\sqrt{\tilde{G}} = \left\langle \sqrt{[\breve{G}_\alpha]} \right\rangle. \tag{3.36}$$

Since the properties (2.18) and the commutation relation (3.32) are satisfied by a volume-form-weighted average, one may generalize the residual-mean tracer equations defined in the previous section to the application of a volume-form-weighted average. However, while a volume-form-weighted average commutes with the divergence of a contravariant vector, the average will in general not commute with the divergence of higher-order tensors (the relevant Christoffel symbol terms will not in general vanish: see example 23 of De, Shaikh & Sengupta 2005). Hence the residual-mean momentum equations derived in previous sections, containing the divergence of rank-two tensors with symmetric components, do not generalize directly to an arbitrary volume-form-weighted average. Furthermore, the geometrical contortions involved in defining a volume-form-weighted averaging operator mean that the metric tensor can have a non-zero eddy component, $g^{ab\prime} \neq 0$, and that the average of the metric tensor need not be the metric tensor in the averaged system $[g_A]^{ab} \neq \overline{g^{ab}}$.

3.4. *Thickness-weighted averaging in buoyancy coordinates*

When specializing the volume-form-weighted average operator to thickness-weighted averaging in buoyancy coordinates then one arrives at the thickness-weighted-average formulation of de Szoeke & Bennett (1993) and Young (2012). The remaining steps required to derive the resulting Eliassen–Palm tensor, which we sketch here, are described in detail in Young (2012).

Start with the Boussinesq hydrostatic primitive momentum equation (3.1). Choose the $[\breve{x}_\alpha]^a$ so that they form a set of buoyancy coordinates in each ensemble member, with $[\breve{x}_\alpha]^1 = [\breve{x}_\alpha] = x$, $[\breve{x}_\alpha]^2 = [\breve{y}_\alpha] = y$, $[\breve{x}_\alpha]^4 = [\breve{t}_\alpha] = t$, and $[\breve{x}_\alpha]^3 = [\breve{z}_\alpha] = [b_\alpha]$. This choice requires that the buoyancy varies monotonically with height in all ensemble members. Also choose $\tilde{x}^1 = \tilde{x} = x$, $\tilde{x}^2 = \tilde{y} = y$, $\tilde{x}^4 = \tilde{t} = t$, and $\tilde{x}^3 = \tilde{z} = \overline{b} = b$. By the condition (3.36) the vertical coordinate in the averaged system is defined such that $z = \langle \zeta \rangle$, where ζ is a scalar defined such that $[\zeta_\alpha] = [z_\alpha]$. The volume forms become:

$$\sqrt{[\breve{G}_\alpha]} = \frac{1}{\partial_z [b_\alpha]} = \partial_{[\breve{z}_\alpha]} [\zeta_\alpha], \qquad \sqrt{\tilde{G}} = \frac{1}{\partial_z b} = \partial_{\tilde{z}} \langle \zeta \rangle, \tag{3.37}$$

and the volume-form-weighted average now takes the form of a thickness-weighted average.

In de Szoeke & Bennett (1993) and Young (2012) physical quantities are thickness-weight averaged in density or buoyancy coordinates, while the vertical coordinate is defined via an unweighted average. In McDougall & McIntosh (2001) it is shown that such a definition of the averaged vertical coordinate allows one to relate the thickness-weight-averaged velocity to the Eulerian-averaged velocity, with the difference defining a 'quasi-Stokes stream function'. Here this definition for the vertical coordinate arises directly from the definition of the volume-form-weighted average, combined with the constraint (3.36).

Let $[\varepsilon_A]^{abcd}$ be the Levi–Civita tensor in the averaged system. Then although $\overline{\varepsilon^{abcd}} = [\varepsilon_A]^{abcd}$, nevertheless $\varepsilon^{abcd\prime} \neq 0$. Hence the thickness-weighted average in buoyancy coordinates does not commute with the cross-product or curl operators. Similarly, if $[Z_A]^a$ and $[T_A]^a$ are the unit vertical and time vectors in the averaged system, then $\overline{Z^a} = [Z_A]^a$ and $\overline{T_a} = [T_A]_a$, but $Z^{a\prime} \neq 0$ and $T^{a\prime} \neq 0$.

Choose $\gamma = b$ with $(\rho/[\rho_0]) h^{ac} \Phi_{,c} = b Z^a$, where it is assumed that a dynamical equation for the buoyancy exists:

$$(u^a b)_{;a} = \Theta. \tag{3.38}$$

Then applying the thickness-weighted average with the coordinate systems as defined above, $b' = 0$ by definition, and the average buoyancy equation becomes:

$$(\overline{u^a} b)_{;a} = \overline{\Theta}. \tag{3.39}$$

No eddy terms appear in the average buoyancy equation, and hence no additional residual-mean velocity need be defined. The average continuity equation is:

$$\overline{u^a}_{;a} = 0. \tag{3.40}$$

To derive the thickness-weight-averaged momentum equation, consider the *covariant* momentum equation:

$$(u^c v_a)_{;c} + g_{ac} \varepsilon^{cdei} f_d v_e T_i - F_i = -\frac{1}{[\rho_0]} g_{ac} h^{cd} p_{,d} + b Z_a. \tag{3.41}$$

While the thickness-weighted average operator does not commute with the divergence of a general rank-two tensor, the average operator does commute with the divergence of a mixed-type tensor ϕ^a_b provided $\phi^a_b Z^b = 0$:

$$\phi^a_b Z^b = 0 \quad \Rightarrow \quad \overline{(\phi^a_{b;a})} = (\overline{\phi^a_b})_{;a}. \tag{3.42}$$

In particular $u^a v_b Z^b = 0$. Hence the thickness-weight-averaged covariant momentum equation becomes:

$$(\overline{u^c}\,\overline{v_a})_{;c} + [g_A]_{ac} [\varepsilon_A]^{cdei} f_d \overline{v_e} [T_A]_i - \overline{F_i} = -\frac{1}{[\rho_0]} [g_A]_{ac} [h_A]^{cd} \overline{p_{,d}} + b \overline{Z_a} - (\overline{u^{c\prime} v_a{}'})_{;c}. \tag{3.43}$$

Taking the inner product with $[Z_A]^a$ yields the hydrostatic balance relation:

$$0 = -\frac{1}{[\rho_0]}\langle p\rangle_{,a}[Z_A]^a + b. \tag{3.44}$$

Evaluating the $\tilde{x}$ and $\tilde{y}$ components and applying the hydrostatic balance relation then leads to the result:

$$\begin{aligned}(\overline{u^c}\,\overline{v_a})_{;c} + [g_A]_{ac}[\varepsilon_A]^{cdei}f_d\overline{v_e}[T_A]_i - \overline{F_i} \\ = -\frac{1}{[\rho_0]}[g_A]_{ac}[h_A]^{cd}\langle p\rangle_{,d} + b[Z_A]_a - [g_A]_{ac}E^{cd}_{;d},\end{aligned} \tag{3.45}$$

where the Eliassen–Palm flux tensor is given by:

$$E^{ab} = \overline{u^{b\prime}v^{a\prime}} + P^{ab}, \tag{3.46}$$

and where P^{ab} has components, in the $\tilde{x}$, $\tilde{y}$, $\tilde{z}$, $\tilde{t}$ coordinate system:

$$P^b_a = [g_A]_{ac}P^{cb} = \frac{1}{[\rho_0]\sqrt{G}}\begin{pmatrix} \frac{1}{2}[\rho_0]\langle\zeta^{\#}\zeta^{\#}\rangle & 0 & 0 & 0 \\ 0 & \frac{1}{2}[\rho_0]\langle\zeta^{\#}\zeta^{\#}\rangle & 0 & 0 \\ \langle\zeta^{\#}\partial_{[\check{x}_\alpha]}m^{\#}\rangle & \langle\zeta^{\#}\partial_{[\check{y}_\alpha]}m^{\#}\rangle & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{3.47}$$

where m is the Montgomery potential $m = p - \rho_0 b\zeta$. The columns of this tensor are exactly the Eliassen–Palm flux 'vectors' derived in Young (2012); these vectors form the components of a mixed-type rank-two tensor.

The Eliassen–Palm flux tensor thus derived can be directly related to the flux of potential vorticity. Taking the cross-product of the average momentum equation with the average buoyancy gradient (applying the operator $[\varepsilon_A]^{acde}\ldots b_{,d}[T_A]_e$), assuming f^a is parallel to Z^a, and then applying standard vector calculus identities, yields a form of the potential vorticity induction equation (the Boussinesq primitive equation generalization of (2.10)). Taking the divergence then yields an equation for the potential vorticity substance (Haynes & McIntyre 1990) q^* associated with the average dynamics:

$$\begin{aligned}(\overline{u^a}q^*)_{;a} &= Q - ([\varepsilon_A]^{acde}[g_A]_{ci}E^{ij}_{;j}b_{,d}[T_A]_e)_{;a} \\ &= Q - [\varepsilon_A]^{acde}[g_A]_{ci}E^{ij}_{;ja}b_{,d}[T_A]_e,\end{aligned} \tag{3.48}$$

where q^* is given by

$$\begin{aligned}q^* &= ([\varepsilon_A]^{acde}\overline{v}_c b_{,d}[T_A]_e + f^a b)_{;a} \\ &= [\varepsilon_A]^{acde}\overline{v}_{c;a}b_{,d}[T_A]_e + f^a b_{,a},\end{aligned} \tag{3.49}$$

and the mechanical and diabatic forcing Q is given by:

$$Q = ([\varepsilon_A]^{acde}\overline{F}_c b_{,d}[T_A]_e + [\varepsilon_A]^{acde}\overline{v}_c\overline{\Theta}_{,d}[T_A]_e + f^a\overline{\Theta})_{;a}. \tag{3.50}$$

Hence $[\varepsilon_A]^{acde}[g_A]_{ci}E^{ij}_{;j}b_{,d}[T_A]_e$ is equal to the eddy flux of potential vorticity substance q^*, plus a rotational term. Since the thickness-weighted average does not commute with the cross-product operator, the potential vorticity associated with the average system is not equal to the average of the potential vorticity associated with the unaveraged system.

Finally, the Eliassen–Palm flux tensor (3.46) is related to the quasi-geostrophic flux tensors described in § 2.5. Consider the geostrophic limit where:

$$\partial_x p = \partial_{\tilde{x}} m = [\rho_0][f_0]v, \quad \partial_y p = \partial_{\tilde{y}} m = -[\rho_0][f_0]u, \quad w = 0. \tag{3.51}$$

Assume that the perturbations $\zeta^{\#}$ are small so that the Eulerian buoyancy anomaly at the height $\overline{\zeta}$ can be approximated by:

$$b^{\flat} = -\partial_z \tilde{z} \zeta^{\#}. \tag{3.52}$$

Also assume that $\partial_{\tilde{x}} z = \partial_{\tilde{y}} z = 0$, corresponding to an assumption of small dynamical aspect ratio. This latter approximation certainly holds in the quasi-geostrophic approximation. Then, to leading order in Rossby number and aspect ratio, the Eliassen–Palm flux tensor (3.46) has components, in the x, y, z, t coordinate system:

$$E_a^b = g_{ac}E^{cb} = \begin{pmatrix} \overline{u'u'} + \dfrac{1}{2[N_0]^2}\langle b^{\flat}b^{\flat}\rangle & \overline{u'v'} & 0 & 0 \\ \overline{u'v'} & \overline{v'v'} + \dfrac{1}{2[N_0]^2}\langle b^{\flat}b^{\flat}\rangle & 0 & 0 \\ -\dfrac{[f_0]}{[N_0]^2}\langle v^{\#}b^{\flat}\rangle & \dfrac{[f_0]}{[N_0]^2}\langle u^{\#}b^{\flat}\rangle & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{3.53}$$

where $[N_0]^2 = \partial_z \tilde{z} = \partial_z b$. This corresponds to a combination of the gauge choices leading to the quasi-geostrophic Eliassen–Palm flux tensors (2.39) and (2.41). Hence, to leading order in Rossby number and aspect ratio, one can derive an energy which bounds the Frobenius norm of the Eliassen–Palm flux tensor (3.46).

3.5. *Generalized thickness-weighted averaging*

The thickness-weighted-average formulation can be generalized via a redefinition of $[\check{z}_\alpha]$ and $\tilde{z}$, with an associated redefinition of the average and eddy operators. In the general case diabatic eddy fluxes once again appear in the buoyancy equation:

$$(\overline{u^a}\,\overline{b})_{;a} = \overline{\Theta} - (\overline{u^{a\prime}b'})_{;a}, \tag{3.54}$$

while the average momentum equation becomes:

$$\begin{aligned}(\overline{u^c}\,\overline{v_a})_{;c} &+ [g_A]_{ac}[\varepsilon_A]^{cdei} f_d \overline{v_e}[T_A]_i - \overline{F_i} \\ &= -\frac{1}{[\rho_0]}[g_A]_{ac}[h_A]^{cd}\langle p\rangle_{,d} + \overline{b}[Z_A]_a - [g_A]_{ac}E^{cd}_{;d},\end{aligned} \tag{3.55}$$

with Eliassen–Palm flux tensor:

$$E^{ab} = \overline{u^{b\prime}v^{a\prime}} + P^{ab}, \tag{3.56}$$

and where P^{ab} now has components, in the $\tilde{x}$, $\tilde{y}$, $\tilde{z}$, $\tilde{t}$ coordinate system:

$$\begin{aligned} P_a^b &= [g_A]_{ac}P^{cb} \\ &= \frac{1}{[\rho_0]\sqrt{\tilde{G}}} \begin{pmatrix} \frac{1}{2}\langle p^{\#}\partial_{[\check{z}_\alpha]}\zeta^{\#} - \zeta^{\#}\partial_{[\check{z}_\alpha]}p^{\#}\rangle & 0 & 0 & 0 \\ 0 & \frac{1}{2}\langle p^{\#}\partial_{[\check{z}_\alpha]}\zeta^{\#} - \zeta^{\#}\partial_{[\check{z}_\alpha]}p^{\#}\rangle & 0 & 0 \\ \frac{1}{2}\langle \zeta^{\#}\partial_{[\check{x}_\alpha]}p^{\#} - p^{\#}\partial_{[\check{x}_\alpha]}\zeta^{\#}\rangle & \frac{1}{2}\left\langle \zeta^{\#}\partial_{[\check{y}_\alpha]}p^{\#} - p^{\#}\partial_{[\check{y}_\alpha]}\zeta^{\#}\right\rangle & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned} \tag{3.57}$$

This form for the Eliassen–Palm flux tensor is derived in appendix C. Diabatic eddy fluxes of buoyancy are (in general) introduced into regions wherever the average deviates from a thickness-weighted average in buoyancy coordinates. Where the buoyancy varies non-monotonically with height one typically considers a two-step averaging approach as described by de Szoeke & Bennett (1993), where monotonicity is enforced via a preliminary averaging step. However, in principle at least, one may instead redefine $[\breve{z}_\alpha]$ and $\tilde{z}$ in regions where buoyancy varies non-monotonically with height, thereby (in general) introducing diabatic eddy fluxes where this redefinition differs from thickness-weighted averaging in buoyancy coordinates. One may further transition from an average in buoyancy coordinates in the interior to an average in Eulerian or surface-following coordinates at upper and lower boundaries. These conceptual steps enable one, for example, to interpret an ocean model containing adiabatic eddy processes in the interior as a representation of the dynamics subject to thickness-weighted averaging in buoyancy coordinates (as described by McDougall & McIntosh 2001), except where fundamentally diabatic processes are active.

4. Conclusions

The purpose of residual-mean theory is to reduce the magnitude of the explicit eddy fluxes in a prognostic tracer equation (for example the eddy buoyancy flux in the thermodynamic equation) via a modified advection. Physically this is achieved by recognizing that the eddy fluxes are both advective and diffusive in nature, and by decomposing the eddy flux accordingly. Further approaches seek to reduce the eddy buoyancy fluxes to zero by an appropriately defined averaging operator. By either redefining the advection operator, as in the former case, or redefining the averaging operator, as in the latter, one must also modify the momentum equation, and thereby produce modified eddy momentum flux tensors. This has previously been described in terms of varying forms of Eliassen–Palm flux vectors, or occasionally in terms of rank-two Eliassen–Palm flux tensors. In this paper the eddy–mean-flow interaction problem has been treated using a geometric formalism. This has enabled the explicit identification of forms of the rank-two Eliassen–Palm flux tensor.

In the quasi-geostrophic limit the eddy–mean-flow interaction is conveniently explored by deriving a new equation, whose divergence is the quasi-geostrophic potential vorticity equation. This enables one to characterize the eddy–mean-flow interaction via a momentum flux tensor with two forms of gauge freedom and whose divergence is the eddy quasi-geostrophic potential vorticity flux plus a rotational flux. Particular choices for the momentum flux tensor yield the residual-mean quasi-geostrophic equations, from which one may identify corresponding Eliassen–Palm flux tensors.

Forms of the Eliassen–Palm flux tensor were derived for the more general Boussinesq hydrostatic primitive equations. The residual-mean velocity was left in a general form, allowing one to consider alternative decompositions of the eddy fluxes as described by Nurser & Lee (2004) and Plumb & Ferrari (2005). The precise form of the resulting Eliassen–Palm flux tensor was explicitly identified. The general approach permitted 'space–time residual-mean' formulations, modifying the full material derivative (including the time derivative). This was utilized (in appendix B) to derive the higher-order residual-mean formulation of Eden *et al.* (2007).

A volume-form-weighted average was introduced, constructed such that the averaging operator commutes with the covariant divergence of a vector. The thickness-weighted-average approach of de Szoeke & Bennett (1993) and Young (2012) arises from a particular choice of volume-form-weighted average when applied to the Boussinesq hydrostatic primitive equations. The procedure of 'averaging in buoyancy coordinates' implicitly involves the definition of a complicated averaging operator, involving transformation of tensors into coordinates in which to perform averaging, and the implicit definition of an averaged coordinate system. This paper describes the detailed steps which such a procedure involves, and identifies the geometric objects which result when applying this procedure to the hydrostatic Boussinesq primitive equations. While the averaging operator commutes with the covariant divergence operator, it does not commute with the cross-product or curl operators. As a result the average of the potential vorticity is not equivalent to the potential vorticity associated with the averaged system.

A correct physical interpretation of the eddy–mean-flow interaction problem is crucial if one wishes to construct an eddy parameterization. A vector flux of a scalar is physically a very different quantity from a tensor flux of a vector. Difficulties arising from an incorrect characterization of the eddy–mean-flow interaction problem are acute in the context of the quasi-geostrophic equations. Attempting to parameterize the eddy quasi-geostrophic potential vorticity fluxes (a vector quantity), as is implied by assuming a down-gradient eddy flux of potential vorticity (Green 1970; Rhines & Young 1982), violates momentum conservation in general (Marshall 1981). Indeed, attempting to parameterize the eddy quasi-geostrophic potential vorticity tendency (a scalar quantity) violates both potential vorticity conservation and momentum conservation in general. Implicit physical constraints appear in a down-gradient potential vorticity closure, and hence one is unlikely to preserve the true geometric nature of the eddy–mean-flow interaction with such an approach. Preservation of the geometric structure of the eddy interaction, as well as conservation of both momentum and potential vorticity, can be achieved by instead forming a parameterization for the rank-two Eliassen–Palm flux tensor itself (Marshall *et al.* 2012), and not its divergence.

Since the Eliassen–Palm flux tensors thus derived are inherently geometric objects, one may decompose the tensors and construct geometric frameworks characterizing the eddy–mean-flow interaction. In particular, in a quasi-geostrophic context, geometric decompositions of the flux tensors are equivalent to the geometric parameterization framework of Marshall *et al.* (2012). For the more general Boussinesq momentum and hydrostatic primitive equations the Frobenius norm of the Eliassen–Palm flux tensor always has the dimensions of energy per mass, and one may always consider a singular-value decomposition of the tensor. Such a decomposition could be useful in interpretation of model and observational data, or in the construction of more general geometric parameterization frameworks. In particular, a general and fundamental picture of the true nature of the eddy–mean-flow interaction problem is essential if one wishes to construct a physically consistent eddy closure.

Acknowledgements

The authors are grateful to W. R. Young for insightful discussions, particularly regarding the invariance properties of the potential vorticity, to P. Mana for invaluable assistance with mathematical details, and to D. Ferreira, T. J. McDougall, J. C. McWilliams, D. R. Munday, G. Shutts and R. G. Williams for useful comments and references. Financial support was provided by the UK Natural Environment Research Council (NE/H020454/1).

Appendix A. Tensor calculus identities

Given two vectors A^a and B^b and a scalar α, define a rank-two antisymmetric tensor:

$$\Psi^{ab} = A^a \frac{B^b}{B^c \alpha_{,c}} - A^b \frac{B^a}{B^c \alpha_{,c}}, \tag{A 1}$$

where it is assumed that this tensor is finite. Thus define a non-divergent vector:

$$u^a = \Psi^{ab}_{;b}. \tag{A 2}$$

Then, by definition:

$$\begin{aligned} u^a \alpha &= \Psi^{ab}_{;b} \alpha \\ &= (\Psi^{ab}\alpha)_{;b} - A^a + A^b \alpha_{,b} \frac{B^a}{B^c \alpha_{,c}}. \end{aligned} \tag{A 3}$$

Hence, via re-arrangement, one arrives at the identity:

$$A^a = (\Psi^{ab}\alpha)_{;b} - u^a \alpha + A^b \alpha_{,b} \frac{B^a}{B^c \alpha_{,c}}. \tag{A 4}$$

This can be expressed:

$$A^a = S^a + \frac{A^b \alpha_{,b}}{B^c \alpha_{,c}} B^a, \tag{A 5}$$

where S^a is defined:

$$\begin{aligned} S^a &= (\Psi^{ab}\alpha)_{;b} - u^a \alpha \\ &= \Psi^{ab} \alpha_{,b}. \end{aligned} \tag{A 6}$$

Note that, by the antisymmetry of Ψ^{ab}:

$$\begin{aligned} S^a \alpha_{,a} &= \Psi^{ab} \alpha_{,a} \alpha_{,b} \\ &= 0, \end{aligned} \tag{A 7}$$

and hence S^a is perpendicular to the gradient of α. Hence a vector A^a can be decomposed into two components: a component S^a which is perpendicular to the gradient of α, and a component in the direction of B^a. The former further decomposes into a non-divergent vector field and the product of α with a non-divergent vector field. This is a tensor calculus generalization of the decomposition of § 2 of Nurser & Lee (2004).

Taking the divergence of (A 4) leads to the identity:

$$A^a_{;a} = \left(A^b \alpha_{,b} \frac{B^a}{B^c \alpha_{,c}} \right)_{;a} - (u^a \alpha)_{;a}. \tag{A 8}$$

Hence the divergence of a vector A^a can be decomposed into a term which takes the form of a non-divergent advection of a scalar α, combined with a term that is the divergence of a field oriented in the direction of B^a.

A corollary to the identity (A 8) is that, if a vector field R^a is non-divergent, $R^a_{;a} = 0$, then:

$$\left((R^b \alpha)_{;b} \frac{C^a}{C^c \alpha_{,c}} \right)_{;a} = (v^a \alpha)_{;a}, \tag{A 9}$$

where v^a is a non-divergent vector field defined via:

$$v^a = \left(R^a \frac{C^b}{C^c \alpha_{,c}} - R^b \frac{C^a}{C^c \alpha_{,c}} \right)_{;b}, \tag{A 10}$$

where C^a is some given vector field and where it is assumed that v^a is finite.

Appendix B. Residual-mean tracer equation to arbitrary order

In this appendix the residual-mean formulation of Eden *et al.* (2007) is derived, yielding a residual-mean tracer equation to arbitrary order in perturbation amplitude. Start with the dynamical equation for a scalar tracer γ of the form:

$$(u^a \gamma)_{;a} = \Theta, \tag{B 1}$$

where Θ is a given scalar and u^a is a non-divergent velocity with $u^a_{;a} = 0$. The velocity u^a may be considered to be the four-dimensional velocity with $u^a T_a = 1$, where T^a is equal to the unit time vector, although one may consider more general non-divergent velocity vectors. Introduce the averaging operator $\overline{(\cdots)}$, and the associated projection operator $\widehat{(\cdots)}$ and eddy operator $(\cdots)'$, which satisfy the properties (2.18), and the commutation relation (3.9) (see § 2.3). Then (B 1) can be written:

$$u^a \widehat{\overline{\gamma}}_{,a} + (u^a \gamma')_{;a} = \Theta. \tag{B 2}$$

Let non-index subscripts and superscripts be contained within square brackets. One can define a series of scalar quantities $[\Pi^n]$ via:

$$n[\Pi^n] = (\gamma')^n, \tag{B 3}$$

where, on the right-hand side of (B 3), n is used as an exponent, and where n is a positive integer. Then multiplying (B 2) by $n[\Pi^n]$ and applying the averaging operator leads to:

$$n\overline{u^a[\Pi^n]}\,\overline{\gamma}_{,a} + (\overline{u^a[\Pi^{n+1}]})_{;a} = n\overline{\Theta[\Pi^n]}. \tag{B 4}$$

Hence:

$$\overline{u^a[\Pi^n]}\,\overline{\gamma}_{,a} = \overline{\Theta[\Pi^n]} - \left(\frac{1}{n} \overline{u^a[\Pi^{n+1}]} \right)_{;a}. \tag{B 5}$$

Introduce a series of vectors $[H^n]^a$ associated with each scalar $[\Pi^n]$ for each n. Thus define a series of rank-two antisymmetric tensors $[X^n]^{ab}$ for each $n > 1$:

$$[X^n]^{ab} = \frac{1}{n-1} \overline{u^a[\Pi^n]} \frac{[H^n]^b}{[H^n]^c \overline{\gamma}_{,c}} - \frac{1}{n-1} \overline{u^b[\Pi^n]} \frac{[H^n]^a}{[H^n]^c \overline{\gamma}_{,c}}, \tag{B 6}$$

and non-divergent vectors $[v^n]^a$:

$$[v^n]^a = [X^n]^{ab}_{;b}, \tag{B 7}$$

where it is assumed that all $[X^n]^{ab}$ thus defined are finite. Then, via the identity (A 8), (B 5) becomes:

$$\overline{u^a[\Pi^n]}\,\overline{\gamma}_{,a} = \overline{\Theta[\Pi^n]} - \left(\frac{1}{n} \overline{u^a[\Pi^{n+1}]}\,\overline{\gamma}_{,a} \frac{[H^{n+1}]^b}{[H^{n+1}]^c \overline{\gamma}_{,c}} \right)_{;b} + ([v^{n+1}]^a \overline{\gamma})_{;a}. \tag{B 8}$$

Also, applying the averaging operator to (B 1) yields:

$$(\overline{u^a}\,\overline{\gamma})_{;a} = \overline{\Theta} - (\overline{u^{a\prime}\gamma^\prime})_{;a}. \tag{B 9}$$

Via the identity (A 8), this leads to the residual-mean formulation (Andrews & McIntyre 1976, 1978; Holton 1981; Nurser & Lee 2004):

$$\left((\overline{u^a} - [u_R^1]^a)\overline{\gamma}\right)_{;a} = \overline{\Theta} - \left(\overline{u^{a\prime}\gamma^\prime}\,\overline{\gamma}_{,a}\frac{[H^1]^b}{[H^1]^c\overline{\gamma}_{,c}}\right)_{;b}, \tag{B 10}$$

where:

$$[u_R^1]^a = [\Psi_R^1]^{ab}_{;b}, \tag{B 11a}$$

$$[\Psi_R^1]^{ab} = \overline{u^{a\prime}\gamma^\prime}\frac{[H^1]^b}{[H^1]^c\overline{\gamma}_{,c}} - \overline{u^{b\prime}\gamma^\prime}\frac{[H^1]^a}{[H^1]^c\overline{\gamma}_{,c}}. \tag{B 11b}$$

Hence:

$$\left((\overline{u^a} - [u_R^1]^a)\overline{\gamma}\right)_{;a} = \overline{\Theta} + O(e^2), \tag{B 12}$$

where $O(e^n)$ denotes a quantity of order n or greater in perturbed quantities.

Substituting (B 8) with $n = 1$ into (B 10) gives:

$$\begin{aligned}\left((\overline{u^a} - [u_R^1]^a)\overline{\gamma}\right)_{;a} = \overline{\Theta} - \left(\overline{\Theta\gamma^\prime}\frac{[H^1]^a}{[H^1]^b\overline{\gamma}_{,b}}\right)_{;a} \\ + \left(\left(\overline{u^a[\Pi^2]}\,\overline{\gamma}_{,a}\frac{[H^2]^b}{[H^2]^c\overline{\gamma}_{,c}}\right)_{;b}\frac{[H^1]^d}{[H^1]^e\overline{\gamma}_{,e}}\right)_{;d} - \left(([v^2]^a\overline{\gamma})_{;a}\frac{[H^1]^b}{[H^1]^c\overline{\gamma}_{,c}}\right)_{;b}.\end{aligned} \tag{B 13}$$

Since $[v^2]^a$ is non-divergent one can apply the identity (A 9) to the final term in this equation, leading to:

$$\begin{aligned}\left((\overline{u^a} - [u_R^1]^a - [u_R^2]^a)\overline{\gamma}\right)_{;a} = \overline{\Theta} - \left(\overline{\Theta\gamma^\prime}\frac{[H^1]^a}{[H^1]^b\overline{\gamma}_{,b}}\right)_{;a} \\ + \left(\left(\overline{u^a[\Pi^2]}\,\overline{\gamma}_{,a}\frac{[H^2]^b}{[H^2]^c\overline{\gamma}_{,c}}\right)_{;b}\frac{[H^1]^d}{[H^1]^e\overline{\gamma}_{,e}}\right)_{;d},\end{aligned} \tag{B 14}$$

where:

$$[u_R^2]^a = [\Psi_R^2]^{ab}_{;b}, \tag{B 15a}$$

$$[\Psi_R^2]^{ab} = -[v^2]^a\frac{[H^1]^b}{[H^1]^c\overline{\gamma}_{,c}} + [v^2]^b\frac{[H^1]^a}{[H^1]^c\overline{\gamma}_{,c}}. \tag{B 15b}$$

Note that, as defined, if u^a is considered to be the four-dimensional velocity with $u^aT_a = 1$, then $[u_R^2]^aT_a \neq 0$. Hence this forms a space–time residual-mean formulation. Equation (B 15) is a space–time residual-mean version of the formulation of Medvedev & Greatbatch (2004). Medvedev & Greatbatch (2004) choose $[H^1]^a$ and $[H^2]^a$ to be equal to the (spatial) gradient of $\overline{\gamma}$. A space–time residual-mean version of the temporal-residual-mean I formulation of McDougall & McIntosh (1996) is arrived at by replacing $\overline{u^a[\Pi^2]}$ with $\overline{u^a}\,\overline{\Pi}$ in both (B 14) and the definition of $[X^2]^{ab}$ (and hence the definition of $[v^2]^a$), combined with a substitution using (B 5) for $n = 2$.

Via repeated substitution of (B 8) with increasing values of n, and with repeated application of the identity (A 9), one arrives at:

$$\left((\overline{u^a} - [u_R^1]^a - [u_R^2]^a - \cdots - [u_R^N]^a)\overline{\gamma}\right)_{;a} = [\varXi^1] + [\varXi^2] + \cdots + [\varXi^N] + O(e^N). \quad \text{(B 16)}$$

The $[\varXi^n]$ are defined via:

$$[\varXi^1] = \overline{\varTheta}, \quad \text{(B 17a)}$$

$$[\varXi^n] = [\xi^{n-1,n-1}] \quad \text{for } n \geqslant 2, \quad \text{(B 17b)}$$

with:

$$[\xi^{p,1}] = -\left(\overline{\varTheta[\varPi^p]}\frac{[H^p]^a}{[H^p]^b\overline{\gamma}_{,b}}\right)_{;a}, \quad \text{(B 18a)}$$

$$[\xi^{p,q+1}] = -\frac{1}{p-q}\left([\xi^{p,q}]\frac{[H^{p-q}]^a}{[H^{p-q}]^b\overline{\gamma}_{,b}}\right)_{;a} \quad \text{for } p \geqslant 2,\ 1 \leqslant q < p. \quad \text{(B 18b)}$$

The $[u_R^n]^a$ are defined via:

$$[u_R^n]^a = [\varPsi_R^n]^{ab}_{;b}, \quad \text{(B 19)}$$

where:

$$[\varPsi_R^1]^{ab} = \overline{u^{a\prime}\gamma^\prime}\frac{[H^1]^b}{[H^1]^c\overline{\gamma}_{,c}} - \overline{u^{b\prime}\gamma^\prime}\frac{[H^1]^a}{[H^1]^c\overline{\gamma}_{,c}}, \quad \text{(B 20a)}$$

$$[\varPsi_R^n]^{ab} = [s^{n-1,n-1}]^{ab} \quad \text{for } n \geqslant 2, \quad \text{(B 20b)}$$

and where:

$$[s^{p,1}]^{ab} = -[v^{p+1}]^a\frac{[H^p]^b}{[H^p]^d\overline{\gamma}_{,d}} + [v^{p+1}]^b\frac{[H^p]^a}{[H^p]^d\overline{\gamma}_{,d}}, \quad \text{(B 21a)}$$

$$[s^{p,q+1}]^{ab} = -\frac{1}{p-q}[s^{p,q}]^{ac}_{;c}\frac{[H^{p-q}]^b}{[H^{p-q}]^d\overline{\gamma}_{,d}} + \frac{1}{p-q}[s^{p,q}]^{bc}_{;c}\frac{[H^{p-q}]^a}{[H^{p-q}]^d\overline{\gamma}_{,d}} \quad \text{for } p \geqslant 2,\ 1 \leqslant q < p. \quad \text{(B 21b)}$$

Hence one may formulate a residual-mean tracer equation to arbitrary order in perturbation amplitude. Note that all $[\varXi^n]$ depend upon $\varTheta$, and in particular vanish when $\varTheta = 0$. This is the space–time residual-mean formulation of the expansion of Eden *et al.* (2007). Eden *et al.* (2007) choose all $[H^n]^a$ to be equal to the (spatial) gradient of $\overline{\gamma}$. The convergence properties of this expansion are not explored here, although it is noted that, should the expansion fail to converge, one can locally terminate the expansion at arbitrary order by locally choosing $[H^n] = \overline{u^a[\varPi^n]}$.

Appendix C. Thickness-weighted averaging with a general vertical coordinate

In this appendix the thickness-weight averaged primitive equations are derived. This appendix makes no use of tensor calculus notation, and hence square brackets around non-index subscripts and superscripts are removed.

First, express the hydrostatic Boussinesq primitive equations in a coordinate system $\breve{x} = x$, $\breve{y} = y$, $\breve{t} = t$, and with $\breve{z}$ left unspecified, where it is assumed that the transformation from the $\breve{x}$, $\breve{y}$, $\breve{z}$, $\breve{t}$ coordinate system to the Cartesian x, y, z, t coordinate system is non-singular. The equations (which arise directly from the covariant momentum equation (3.41)) are:

$$\partial_z\breve{z}(\partial_{\breve{t}}(\partial_{\breve{z}}zu) + \partial_{\breve{x}}(\partial_{\breve{z}}zu_{\breve{x}}u) + \partial_{\breve{y}}(\partial_{\breve{z}}zu_{\breve{y}}u) + \partial_{\breve{z}}(\partial_{\breve{z}}zu_{\breve{z}}u)) - fv = -\frac{1}{\rho_0}\partial_{\breve{x}}p + b\partial_{\breve{x}}\zeta + F_{\breve{x}}, \tag{C 1a}$$

$$\partial_z\breve{z}(\partial_{\breve{t}}(\partial_{\breve{z}}zv) + \partial_{\breve{x}}(\partial_{\breve{z}}zu_{\breve{x}}v) + \partial_{\breve{y}}(\partial_{\breve{z}}zu_{\breve{y}}v) + \partial_{\breve{z}}(\partial_{\breve{z}}zu_{\breve{z}}v)) + fu = -\frac{1}{\rho_0}\partial_{\breve{y}}p + b\partial_{\breve{y}}\zeta + F_{\breve{y}}, \tag{C 1b}$$

$$0 = -\frac{1}{\rho_0}\partial_{\breve{z}}p + b\partial_{\breve{z}}\zeta, \tag{C 1c}$$

$$\partial_z\breve{z}(\partial_{\breve{t}}(\partial_{\breve{z}}z) + \partial_{\breve{x}}(\partial_{\breve{z}}zu_{\breve{x}}) + \partial_{\breve{y}}(\partial_{\breve{z}}zu_{\breve{y}}) + \partial_{\breve{z}}(\partial_{\breve{z}}zu_{\breve{z}})) = 0. \tag{C 1d}$$

Here, u, v and w are the components of an incompressible velocity in the x, y, z, t, coordinate system, and $u_{\breve{x}}$, $u_{\breve{y}}$ and $u_{\breve{z}}$ are the corresponding (contravariant) components in the $\breve{x}$, $\breve{y}$, $\breve{z}$, $\breve{t}$ coordinate system. It follows by direct evaluation that $u = u_{\breve{x}}$ and $v = u_{\breve{y}}$. Also, f is the Coriolis parameter, ρ_0 is a reference density, p is the pressure, ζ is a scalar field equal to the value of the z coordinate, and $F_{\breve{x}}$ and $F_{\breve{y}}$ are $\breve{x}$ and $\breve{y}$ components of additional terms. It again follows by direct evaluation that $F_{\breve{x}} = F_x$ and $F_{\breve{y}} = F_y$, where F_x and F_y are the x and y components of additional terms.

Now consider a volume-form-weighted average of these equations. Specifically, an ensemble of systems is defined, where each ensemble member α defines its own coordinate system $\breve{x}_\alpha$, $\breve{y}_\alpha$, $\breve{z}_\alpha$, $\breve{t}_\alpha$. Introduce an ensemble-average operator $\langle\cdots\rangle$ defined as per (3.27). Define an averaged coordinate system $\tilde{x} = x$, $\tilde{y} = y$, $\tilde{t} = t$, and with $\tilde{z}$ defined such that $\langle\zeta\rangle = z$. The ensemble-average operator $\langle\cdots\rangle$ commutes with respect to derivatives in the following sense:

$$\langle\partial_{\breve{x}_\alpha}\mathscr{F}_\alpha\rangle = \partial_{\tilde{x}}\langle\mathscr{F}_\alpha\rangle, \quad \langle\partial_{\breve{y}_\alpha}\mathscr{F}_\alpha\rangle = \partial_{\tilde{y}}\langle\mathscr{F}_\alpha\rangle, \tag{C 2a}$$

$$\langle\partial_{\breve{z}_\alpha}\mathscr{F}_\alpha\rangle = \partial_{\tilde{z}}\langle\mathscr{F}_\alpha\rangle, \quad \langle\partial_{\breve{t}_\alpha}\mathscr{F}_\alpha\rangle = \partial_{\tilde{t}}\langle\mathscr{F}_\alpha\rangle, \tag{C 2b}$$

where $\mathscr{F}_\alpha$ is an arbitrary quantity for each ensemble member α. Thus introduce a volume-form-weighted average operator $\overline{(\cdots)}$ defined as per (3.30), where here $\sqrt{\breve{G}_\alpha} = \partial_{\breve{z}_\alpha}\zeta_\alpha$ and $\sqrt{\tilde{G}} = \partial_{\tilde{z}}z$. The volume-form-weighted average operator is therefore a thickness-weighted average operator. Define also an eddy operator $(\cdots)'$ associated with $\overline{(\cdots)}$ as per (3.34), and an eddy operator $(\cdots)^\#$ associated with $\langle\cdots\rangle$ as per (3.35). For notational convenience, let $\langle\mathscr{F}\rangle = \langle\mathscr{F}_\alpha\rangle$ and $\overline{\mathscr{F}} = \overline{\mathscr{F}_\alpha}$.

In this context the thickness-weight-average operator can be considered an operator which, given a series of equations written, for each ensemble member α, in the $\breve{x}_\alpha$, $\breve{y}_\alpha$, $\breve{z}_\alpha$, $\breve{t}_\alpha$ coordinate system, yields a single averaged equation written in the $\tilde{x}$, $\tilde{y}$, $\tilde{z}$, $\tilde{t}$ coordinate system. If it understood that a quantity $\mathscr{F}$ is expressed in the $\breve{x}_\alpha$, $\breve{y}_\alpha$, $\breve{z}_\alpha$, $\breve{t}_\alpha$ coordinate system, one can write:

$$\overline{\mathscr{F}} = \partial_z\tilde{z}\langle\partial_{\breve{z}_\alpha}\zeta_\alpha\mathscr{F}_\alpha\rangle, \tag{C 3}$$

where $\langle\cdots\rangle$ is an ensemble average of an arbitrary quantity. Particular caution is required when defining the thickness-weighted average of tensor components: in this case $\mathscr{F}_\alpha$ is understood to correspond to the tensor component in the $\breve{x}_\alpha$, $\breve{y}_\alpha$, $\breve{z}_\alpha$, $\breve{t}_\alpha$ coordinate system, and $\overline{\mathscr{F}}$ is then a corresponding tensor component in the $\tilde{x}$, $\tilde{y}$, $\tilde{z}$, $\tilde{t}$ coordinate system.

The thickness-weighted average of (C 1*d*) yields simply:

$$\partial_z\tilde{z}(\partial_{\tilde{t}}(\partial_{\tilde{z}}z) + \partial_{\tilde{x}}(\partial_{\tilde{z}}z\overline{u}) + \partial_{\tilde{y}}(\partial_{\tilde{z}}z\overline{v}) + \partial_{\tilde{z}}(\partial_{\tilde{z}}z\overline{u_{\alpha,\check{z}_\alpha}})) = 0, \tag{C 4}$$

which is a consequence of the fact that the thickness-weighted average commutes with the covariant divergence.

The thickness-weighted average of (C 1*c*) yields:

$$0 = -\frac{1}{\rho_0}\partial_z\tilde{z}\langle\partial_{\check{z}_\alpha}p_\alpha\partial_{\check{z}_\alpha}\zeta_\alpha\rangle + \partial_z\tilde{z}\langle b_\alpha\partial_{\check{z}_\alpha}\zeta_\alpha\partial_{\check{z}_\alpha}\zeta_\alpha\rangle, \tag{C 5}$$

which becomes:

$$0 = -\frac{1}{\rho_0}\partial_{\tilde{z}}\langle p_\alpha\rangle + \langle b_\alpha\partial_{\check{z}_\alpha}\zeta_\alpha\rangle + \partial_z\tilde{z}\left\langle\left(-\frac{1}{\rho_0}\partial_{\check{z}_\alpha}p_\alpha^{\#} + (b_\alpha\partial_{\check{z}_\alpha}\zeta_\alpha)^{\#}\right)\partial_{\check{z}_\alpha}\zeta_\alpha^{\#}\right\rangle, \tag{C 6}$$

where $\rho_0^{\#} = 0$ has been assumed. Considering the application of the $(\cdots)^{\#}$ operator to (C 1*c*):

$$0 = -\frac{1}{\rho_0}\partial_{\check{z}_\alpha}p_\alpha^{\#} + (b_\alpha\partial_{\check{z}_\alpha}\zeta_\alpha)^{\#}, \tag{C 7}$$

the final term in (C 6) vanishes. This leads to the thickness-weight averaged hydrostatic balance relation.

$$0 = -\frac{1}{\rho_0}\partial_{\tilde{z}}\langle p\rangle + \overline{b}\partial_{\tilde{z}}z. \tag{C 8}$$

Now consider the thickness-weighted average of (C 1*a*). First, it follows immediately that the material derivative term becomes:

$$\begin{aligned}
&\overline{\frac{1}{\partial_{\check{z}_\alpha}\zeta_\alpha}(\partial_{\check{t}_\alpha}(\partial_{\check{z}_\alpha}\zeta_\alpha u_\alpha) + \partial_{\check{x}_\alpha}(\partial_{\check{z}_\alpha}\zeta_\alpha u_{\alpha,\check{x}_\alpha}u_\alpha) + \partial_{\check{y}_\alpha}(\partial_{\check{z}_\alpha}\zeta_\alpha u_{\alpha,\check{y}_\alpha}u_\alpha) + \partial_{\check{z}_\alpha}(\partial_{\check{z}_\alpha}\zeta_\alpha u_{\alpha,\check{z}_\alpha}u_\alpha))} \\
&\quad = \partial_z\tilde{z}(\partial_{\tilde{t}}(\partial_{\tilde{z}}z\overline{u}) + \partial_{\tilde{x}}(\partial_{\tilde{z}}z\overline{u}\,\overline{u}) + \partial_{\tilde{y}}(\partial_{\tilde{z}}z\overline{v}\,\overline{u}) + \partial_{\tilde{z}}(\partial_{\tilde{z}}z\overline{u_{\alpha,\check{z}_\alpha}}\,\overline{u})) \\
&\qquad + \partial_z\tilde{z}(\partial_{\tilde{x}}(\partial_{\tilde{z}}z\overline{u'u'}) + \partial_{\tilde{y}}(\partial_{\tilde{z}}z\overline{v'u'}) + \partial_{\tilde{z}}(\partial_{\tilde{z}}z\overline{u'_{\alpha,\check{z}_\alpha}u'})).
\end{aligned} \tag{C 9}$$

Assuming $f' = 0$, the Coriolis term becomes simply:

$$-\overline{f_\alpha v_\alpha} = -f\overline{v}. \tag{C 10}$$

The pressure gradient and buoyancy terms require more careful treatment. Assuming $\rho_0^{\#} = 0$, these can be expressed:

$$\begin{aligned}
-\frac{1}{\rho_0}\overline{\partial_{\check{x}_\alpha}p_\alpha} + \overline{b_\alpha\partial_{\check{x}_\alpha}\zeta_\alpha} &= -\frac{1}{\rho_0}\partial_{\tilde{x}}\langle p_\alpha\rangle + \partial_z\tilde{z}\langle b_\alpha\partial_{\check{z}_\alpha}\zeta_\alpha\rangle\partial_{\tilde{x}}z \\
&\quad -\frac{1}{\rho_0}\partial_z\tilde{z}\langle\partial_{\check{x}_\alpha}p_\alpha^{\#}\partial_{\check{z}_\alpha}\zeta_\alpha^{\#}\rangle + \partial_z\tilde{z}\langle(b_\alpha\partial_{\check{z}_\alpha}\zeta_\alpha)^{\#}\partial_{\check{x}_\alpha}\zeta_\alpha^{\#}\rangle.
\end{aligned} \tag{C 11}$$

Via the definition of the thickness-weighted average of b_α, and using (C 7), this leads to:

$$\begin{aligned}
-\frac{1}{\rho_0}\overline{\partial_{\check{x}_\alpha}p_\alpha} + \overline{b_\alpha\partial_{\check{x}_\alpha}\zeta_\alpha} &= -\frac{1}{\rho_0}\partial_{\tilde{x}}\langle p_\alpha\rangle + \overline{b_\alpha}\partial_{\tilde{x}}z \\
&\quad -\frac{1}{\rho_0}\partial_z\tilde{z}\langle\partial_{\check{x}_\alpha}p_\alpha^{\#}\partial_{\check{z}_\alpha}\zeta_\alpha^{\#} - \partial_{\check{z}_\alpha}p_\alpha^{\#}\partial_{\check{x}_\alpha}\zeta_\alpha^{\#}\rangle,
\end{aligned} \tag{C 12}$$

which can be expressed as:

$$-\frac{1}{\rho_0}\overline{\partial_{\check{x}_\alpha} p_\alpha} + \overline{b_\alpha \partial_{\check{x}_\alpha} \zeta_\alpha} = -\frac{1}{\rho_0}\partial_{\tilde{x}}\langle p\rangle + \overline{b}\partial_{\tilde{x}} z$$
$$-\frac{1}{\rho_0}\partial_z \tilde{z}\left\langle \partial_{\check{x}_\alpha}\left(\frac{1}{2}p^{\#}\partial_{\check{z}_\alpha}\zeta^{\#} - \frac{1}{2}\zeta^{\#}\partial_{\check{z}_\alpha}p^{\#}\right) + \partial_{\check{z}_\alpha}\left(\frac{1}{2}\zeta^{\#}\partial_{\check{x}_\alpha}p^{\#} - \frac{1}{2}p^{\#}\partial_{\check{x}_\alpha}\zeta^{\#}\right)\right\rangle. \quad \text{(C 13)}$$

This leads to the thickness-weight averaged equation:

$$\partial_z\tilde{z}(\partial_{\tilde{t}}(\partial_{\tilde{z}}z\overline{u}) + \partial_{\tilde{x}}(\partial_{\tilde{z}}z\overline{u}\,\overline{u}) + \partial_{\tilde{y}}(\partial_{\tilde{z}}z\overline{v}\,\overline{u}) + \partial_{\tilde{z}}(\partial_{\tilde{z}}z\overline{u_{\alpha,\check{z}_\alpha}}\,\overline{u})) - f\overline{v}$$
$$= -\frac{1}{\rho_0}\partial_{\tilde{x}}\langle p\rangle + \overline{b}\partial_{\tilde{x}}z + \overline{F_x}$$
$$-\partial_z\tilde{z}\left(\partial_{\tilde{x}}\left(\partial_{\tilde{z}}z\left(\overline{u'u'} + \frac{1}{\rho_0\partial_{\tilde{z}}z}\left\langle\frac{1}{2}p^{\#}\partial_{\check{z}_\alpha}\zeta^{\#} - \frac{1}{2}\zeta^{\#}\partial_{\check{z}_\alpha}p^{\#}\right\rangle\right)\right) + \partial_{\tilde{y}}(\partial_{\tilde{z}}z\overline{v'u'})\right.$$
$$\left.+\partial_{\tilde{z}}\left(\partial_{\tilde{z}}z\left(\overline{u'_{\alpha,\check{z}_\alpha}u'} + \frac{1}{\rho_0\partial_{\tilde{z}}z}\left\langle\frac{1}{2}\zeta^{\#}\partial_{\check{x}_\alpha}p^{\#} - \frac{1}{2}p^{\#}\partial_{\check{x}_\alpha}\zeta^{\#}\right\rangle\right)\right)\right). \quad \text{(C 14)}$$

The thickness-weighted average of (C 1*b*) follows similarly. Upon identifying the divergence of a mixed-type rank-two tensor, one arrives at the thickness-weight-averaged momentum equation (3.55) with Eliassen–Palm flux tensor (3.56).

REFERENCES

ANDREWS, D. G. 1983 A finite-amplitude Eliassen–Palm theorem in isentropic coordinates. *J. Atmos. Sci.* **40**, 1877–1883.

ANDREWS, D. G., HOLTON, J. R. & CONWAY, C. B. 1987 *Middle Atmosphere Dynamics*, International Geophysics Series, vol. 40. Academic.

ANDREWS, D. G. & MCINTYRE, M. E. 1976 Planetary waves in a horizontal and vertical shear: the generalized Eliassen–Palm relation and the mean zonal acceleration. *J. Atmos. Sci.* **33**, 2031–2048.

ANDREWS, D. G. & MCINTYRE, M. E. 1978 Generalized Eliassen–Palm and Charney–Drazin theorems for waves on axisymmetric mean flows in compressible atmospheres. *J. Atmos. Sci.* **35** (2), 175–185.

BRETHERTON, F. P. 1966 Critical layer instability in baroclinic flows. *Q. J. R. Meteorol. Soc.* **92** (393), 325–334.

CHARNEY, J. G. & DRAZIN, P. G. 1961 Propagation of planetary-scale disturbances from the lower into the upper troposphere. *J. Geophys. Res.* **66** (1), 83–109.

CRONIN, M. 1996 Eddy-mean flow interaction in the Gulf Stream at 68°W. Part II: eddy forcing on the time-mean flow. *J. Phys. Oceanogr.* **26** (10), 2132–2151.

DE, U. C., SHAIKH, A. A. & SENGUPTA, J. 2005 *Tensor Calculus*. Alpha Science.

EDEN, C., GREATBATCH, R. J. & OLBERS, D. 2007 Interpreting eddy fluxes. *J. Phys. Oceanogr.* **37** (5), 1282–1296.

ELIASSEN, A. & PALM, E. 1961 On the transfer of energy in stationary mountain waves. *Geofysicke Publikasjoner* **22** (3), 1–23.

FERREIRA, D. & MARSHALL, J. 2006 Formulation and implementation of a 'residual-mean' ocean circulation model. *Ocean Model.* **13** (1), 86–107.

GENT, P. R. & MCWILLIAMS, J. C. 1990 Isopycnal mixing in ocean circulation models. *J. Phys. Oceanogr.* **20** (1), 150–155.

GENT, P. R. & MCWILLIAMS, J. C. 1996 Eliassen–Palm fluxes and the momentum equation in non-eddy-resolving ocean circulation models. *J. Phys. Oceanogr.* **26** (11), 2539–2546.

GENT, P. R., WILLEBRAND, J., MCDOUGALL, T. J. & MCWILLIAMS, J. C. 1995 Parameterizing eddy-induced tracer transports in ocean circulation models. *J. Phys. Oceanogr.* **25** (4), 463–474.

GREATBATCH, R. J. & LAMB, K. G. 1990 On parameterizing vertical mixing of momentum in non-eddy resolving ocean models. *J. Phys. Oceanogr.* **20** (10), 1634–1637.

GREEN, J. S. A. 1970 Transfer properties of the large-scale eddies and the general circulation of the atmosphere. *Q. J. R. Meteorol. Soc.* **96** (408), 157–185.

HAYNES, P. H. & MCINTYRE, M. E. 1990 On the conservation and impermeability theorems for potential vorticity. *J. Atmos. Sci.* **47** (16), 2021–2031.

HELD, I. M. & SCHNEIDER, T. 1999 The surface branch of the zonally averaged mass transport circulation in the troposphere. *J. Atmos. Sci.* **56** (11), 1688–1697.

HOLTON, J. R. 1981 An advective model for two-dimensional transport of stratospheric trace species. *J. Geophys. Res.* **86** (C12), 11989–11994.

HOSKINS, B. J., JAMES, I. N. & WHITE, G. H. 1983 The shape, propagation and mean-flow interaction of large-scale weather systems. *J. Atmos. Sci.* **40** (7), 1595–1612.

JACKETT, D. R. & MCDOUGALL, T. J. 1997 A neutral density variable for the world's oceans. *J. Phys. Oceanogr.* **27** (2), 237–263.

LEE, M.-M. & LEACH, H. 1996 Eliassen–Palm flux and eddy potential vorticity flux for a nonquasigeostrophic time-mean flow. *J. Phys. Oceanogr.* **26** (7), 1304–1319.

MARSHALL, D. P., MADDISON, J. R. & BERLOFF, P. S. 2012 A framework for parameterizing eddy potential vorticity fluxes. *J. Phys. Oceanogr.* **42** (4), 539–557.

MARSHALL, J. C. 1981 On the parameterization of geostrophic eddies in the ocean. *J. Phys. Oceanogr.* **11** (2), 257–271.

MCDOUGALL, T. J. & MCINTOSH, P. C. 1996 The temporal-residual-mean velocity. Part I: derivation and the scalar conservation equations. *J. Phys. Oceanogr.* **26** (12), 2653–2665.

MCDOUGALL, T. J. & MCINTOSH, P. C. 2001 The temporal-residual-mean velocity. Part II: isopycnal interpretation and the tracer and momentum equations. *J. Phys. Oceanogr.* **31** (5), 1222–1246.

MEDVEDEV, A. S. & GREATBATCH, R. J. 2004 On advection and diffusion in the mesosphere and lower thermosphere: the role of rotational fluxes. *J. Geophys. Res.* **109**, D07104.

MIYAHARA, S. 2006 A three-dimensional wave activity flux applicable to inertio-gravity waves. *Sci. Online Lett. Atmos.* **2**, 108–111.

MURAKI, D. J., SNYDER, C. & ROTUNNO, R. 1999 The next-order corrections to quasigeostrophic theory. *J. Atmos. Sci.* **56** (11), 1547–1560.

NURSER, A. J. G. & LEE, M.-M. 2004 Isopycnal averaging at constant height. Part II: relating to the residual streamfunction in Eulerian space. *J. Phys. Oceanogr.* **34** (12), 2740–2755.

PEDLOSKY, J. 1987 *Geophysical Fluid Dynamics*, 2nd edn. Springer.

PLUMB, R. A. 1985 On the three-dimensional propagation of stationary waves. *J. Atmos. Sci.* **42** (3), 217–229.

PLUMB, R. A. 1986 Three-dimensional propagation of transient quasi-geostrophic eddies and its relationship with the eddy forcing of the time-mean flow. *J. Atmos. Sci.* **43** (16), 1657–1678.

PLUMB, R. A. & FERRARI, R. 2005 Transformed Eulerian-mean theory. Part I: nonquasigeostrophic theory for eddies on a zonal-mean flow. *J. Phys. Oceanogr.* **35** (2), 165–174.

PRANDTL, L. 1942 Bemerkungen zur Theorie der freien Turbulenz. *Z. Angew. Math. Mech.* **22** (5), 241–243.

RHINES, P. B. & YOUNG, W. R. 1982 Homogenization of potential vorticity in planetary gyres. *J. Fluid Mech.* **122**, 347–367.

SCHNEIDER, T., HELD, I. M. & GARNER, S. T. 2003 Boundary effects in potential vorticity dynamics. *J. Atmos. Sci.* **60** (8), 1024–1040.

SMITH, R. D. 1999 The primitive equations in the stochastic theory of adiabatic stratified turbulence. *J. Phys. Oceanogr.* **29** (8), 1865–1880.

STEWART, R. W. & THOMSON, R. E. 1977 Re-examination of vorticity transfer theory. *Proc. R. Soc. Lond.* A **354** (1676), 1–8.

DE SZOEKE, R. A. & BENNETT, A. F. 1993 Microstructure fluxes across density surfaces. *J. Phys. Oceanogr.* **23** (10), 2254–2264.

TAYLOR, G. I. 1915 Eddy motion in the atmosphere. *Phil. Trans. R. Soc. Lond.* A **215**, 1–26.

TAYLOR, G. I. 1932 The transport of vorticity and heat through fluids in turbulent motion. *Phil. Trans. R. Soc. Lond.* A **135** (828), 685–702.

TUNG, K. T. 1986 Nongeostrophic theory of zonally averaged circulation. Part I: formulation. *J. Atmos. Sci.* **43** (22), 2600–2617.

WELANDER, P. 1973 Lateral friction in the oceans as an effect of potential vorticity mixing. *Geophys. Fluid Dyn.* **5** (1), 173–189.

YOUNG, W. R. 2012 An exact thickness-weighted average formulation of the Boussinesq equations. *J. Phys. Oceanogr.* **42** (5), 692–707.

ZHAO, R. & VALLIS, G. 2008 Parameterizing mesoscale eddies with residual and Eulerian schemes, and a comparison with eddy-permitting models. *Ocean Model.* **23** (1–2), 1–12.