YALE
PEABODY
MUSEUM

JOURNAL OF MARINE RESEARCH

The *Journal of Marine Research,* one of the oldest journals in American marine science, published important peer-reviewed original research on a broad array of topics in physical, biological, and chemical oceanography vital to the academic oceanographic community in the long and rich tradition of the Sears Foundation for Marine Research at Yale University.

An archive of all issues from 1937 to 2021 (Volume 1–79) are available through EliScholar, a digital platform for scholarly publishing provided by Yale University Library at https://elischolar.library.yale.edu/.

Requests for permission to clear rights for use of this content should be directed to the authors, their estates, or other representatives. The *Journal of Marine Research* has no contact information beyond the affiliations listed in the published articles. We ask that you provide attribution to the *Journal of Marine Research.*

Yale University provides access to these materials for educational and research purposes only. Copyright or other proprietary rights to content contained in this document may be held by individuals or entities other than, or in addition to, Yale University. You are solely responsible for determining the ownership of the copyright, and for obtaining permission for your intended use. Yale University makes no warranty that your distribution, reproduction, or other use of these materials will not infringe the rights of third parties.

Journal of
MARINE RESEARCH

Isentropic averaging

by Roland A. de Szoeke[1]

ABSTRACT

The equations of motion, thermodynamics and scalar concentration are averaged separately over infra-grid scales (comprising roughly ocean microstructure scales) and sub-grid scales (encompassing 10–100 km scales), the latter average being carried out on constant potential temperature surfaces. (Since variation of salinity has been neglected, potential temperature is synonymous with entropy.) These methods of averaging are used to lend precision to statements about oceanic turbulent diffusion: that infra-grid scales are primarily responsible for dientropic diffusion; and that sub-grid scales are responsible for along-isentropic diffusion of passive scalars. Equivalent sets of averaged equations expressed either in isentropic coordinates or quasi-Cartesian coordinates can be obtained. Diffusion tensors for potential temperature express only infra-grid effects. For other passive tracers, diffusion caused by sub-grid scales of motion is also felt, whose effects are shown by scale analysis to be oriented principally along infra-grid averaged isentropes.

1. Introduction

The ocean is a turbulent medium, exhibiting random motions over a range of space and time scales spanning many decades. The classical way to treat this randomness is to consider the equations of motion and thermodynamics averaged over an ensemble of realizations (Monin and Yaglom, 1971). This is the Reynolds averaging approach. It generates forces and transports of materials and heat due to covariances among random fluctuations of motion, scalar concentrations and temperature. Usually, to make progress, parameterizations of these forces and transports must be proposed. In the ocean two principal paradigms of diffusive transport processes by random fluctuations are used. The first paradigm concerns near-vertical processes effected by fluctuations on scales of the

---

1. College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis, Oregon, 97330, U.S.A.
email: szoeke@coas.oregonstate.edu

order of meters and smaller. In this paper we refer to these scales as *infra-grid*. The average contributions to eddy scalar variance and eddy kinetic energy from these scales are taken to be nearly stationary in time and homogeneous in space so that balances of production and dissipation pertain (Osborn and Cox, 1972). These hypotheses offer methods of estimating near-vertical eddy exchange coefficients for turbulent diffusion. The other paradigm addresses near-horizontal adiabatic processes by fluctuations on scales of tens to hundreds of kilometers that spread water properties on "density"[2] surfaces or isentropes. These scales are called meso-scale, or *sub-grid*. Combined, these paradigms are considered to furnish an eddy diffusivity tensor that is diagonal when rotated to local isentropic coordinates (Redi, 1982): (D_H, D_H, K_D), where $D_H \sim 10^3 \text{ m}^2\text{s}^{-1}$, $K_D \sim 10^{-5} \text{ m}^2\text{s}^{-1}$ are typical estimates of along-isentropic and dientropic diffusivities. The values for the dientropic parameter K_D are deduced by applying the Osborn-Cox hypothesis to microstructure turbulence measurements, or from purposeful dye-release experiments (Ledwell *et al.,* 1993). The along-isentropic parameter D_H has been calculated from long-term drifting buoy displacement statistics (Freeland *et al.,* 1975; Davis, 1985).

The assumption of isentropic diagonality of the diffusivity tensor reflects the common notion that large-scale ocean property variations should spread along isentropes (or isopycnals [Montgomery, 1938]). Indeed, Veronis (1975) showed how ocean circulation models with diffusivity tensors oriented with respect to purely level surfaces would produce observationally and theoretically unpalatable distributions of active and passive scalars. Still, the isentropic character of diffusivity is far from self-evident. It seems valuable, therefore, to deduce rationally, if possible, the isentropic diffusivity structure from considerations of first principles. That is, can the orientation and parameters of the eddy diffusivity tensor be deduced from consideration of the theory, combined with observations, of the statistical characteristics of the wide range of turbulent eddy motions that constitute it, rather than indirectly from the large-scale output features of a model?

In a comprehensive study of the equations of motion, thermodynamics, and scalar conservation appropriate for seawater, Davis (1994a) reviewed the molecular transport processes for heat, salt and momentum, including the various physical coefficients of molecular viscosity and diffusivity, etc., necessary to describe them, and determined an adequate set of approximate equations for the ocean. Davis (1994b) considered the averaging of these equations and the eddy transport processes that then arise. He questioned several aspects of the theoretical underpinnings of the combined view outlined above. Careful distinctions were made between theoretical ensemble averages and practically realizable averages over time and space intervals. The former have attractive properties, useful in manipulation, that the latter may lack, particularly if the spectra of the variables being averaged have a "red" character (increasing power toward low frequencies), and lack "spectral gaps" at the averaging cut-off scales. Because of this lack, it has

---

2. In this paper we shall neglect influences of varying salinity in the ocean, postponing their consideration to a subsequent paper. "Density" is then synonymous with potential temperature or entropy, and we shall refer to potential temperature surfaces as "Isentropes."

been concluded that certain inferences, known as the Osborn-Cox hypothesis (Osborn and Cox, 1972), drawn from ensemble-averaged temperature variance balances regarding quasi-vertical eddy heat transport cannot be strictly valid for practical averages (Gargett and Holloway, 1984). Davis (1994b) argued that for the same reasons averaging data over several years is a minimum requirement to replicate the statistical-theoretic effect of ensemble averaging.

These questions are re-opened in this paper. Averaging in two stages is proposed of the equations of motion, etc., to separate the effects of the infra-grid from the sub-grid (de Szoeke and Bennett, 1993). The infra-grid average is taken up in Section 2. An essential feature of the proposal is that the second, sub-grid, average (Section 4) be performed on isentropes already averaged on infra-grid scales (Section 3, Appendix B). The Osborn-Cox hypothesis and a similar proposition for infra-grid eddy kinetic energy balance are reviewed in Section 5 for their utility in estimating vertical (dientropic) eddy coefficients. In Appendix A we will consider ways to reconcile the Osborn-Cox hypothesis regarding temperature variance balance with practically realizable averages, even when power spectra of fluctuations are red. In Section 6 and Appendix C parameterizations of eddy diffusion of passive tracers by sub-grid motions constrained to be nearly along isentropes are considered (Lundgren, 1981; Bennett, 1996). In Section 7 and Appendix D we show how to transform the equations of motion again into quasi-Cartesian coordinates. The paper is concluded with a summary and discussion (Section 8). A glossary of symbols and operators is included at the end (Appendix E).

2. Infra-grid- or microstructure-averaged equations

We begin with the equations of motion and scalar concentration averaged over infra-grid scales on level (fixed $z = x_3^*$) surfaces. By *infra-grid* we shall mean scales ranging from those determined by molecular viscosity and diffusion (of order centimeters) up to several or tens of meters, which encompass the microstructure turbulence of the ocean. The averaging should be thought of as a spatio-temporal low-pass filter, replicating at least schematically, first, the practical averaging behavior of standard oceanographic instruments such as CTDs, current meters, etc., and second, the statistical data-processing methods usually applied to reduce high-frequency microstructure data. We will assume in this section that the Reynolds averaging rules that are applicable to statistical ensemble averages apply also to the practical average here employed (Monin and Yaglom, 1971). Thus the averaged equations resemble the Reynolds ensemble-averaged equations. However, as the specific averaging rule, $\overline{\bar{\phi}\psi} = \bar{\phi}\bar{\psi}$ (and even more particularly, $\bar{\bar{\phi}} = \bar{\phi}$) can be questioned (Gargett and Holloway, 1984; Davis, 1994a,b), we address the point in the supplementary Appendix A, where corrected conservation laws not relying on this rule are given.

The conservation equation, so averaged, for an arbitrary scalar concentration C in the ocean, of which potential temperature θ, or salinity S, are examples, is

$$\frac{DC}{Dt} = \frac{1}{\rho} \frac{\partial}{\partial x_j^*} (\rho B_j^{*C}) \equiv q^C. \tag{2.1}$$

The notations used in this paper are listed fully in Appendix E. Un-averaged variables to which the fundamental Navier-Stokes equations apply are indicated where necessary by tildes. Infra-grid mean vectors and tensors relative to the Cartesian coordinates are indicated by asterisks; bars over simple variables, conventionally indicating averages, are omitted. The advective derivative in (2.1) employs the infra-grid averaged motion. Infra-grid averages of mass-specific variables are weighted by total density $\tilde{\rho}$, while for density itself, or pressure, no such weighting is used. (Averaged, weighted, specific volume is $V = \overline{\tilde{\rho}\tilde{V}}/\rho = 1/\rho$, thus preserving the reciprocal relation between averaged specific volume and density.) Density-weighted averages of fluctuations of mass-specific variables vanish, while un-weighted averages of fluctuations of density or pressure vanish. In Eq. (2.1) and similar equations, density-weighted infra-grid averaging of motion and scalar fluctuations produces eddy transports, such as B_j^{*C}.

Infra-grid averages are taken over scales ranging from millimeters to several meters. At the smaller scales it is plausible that turbulent eddying motions are three-dimensionally isotropic, while at the larger scales considerable horizontal-to-vertical anisotropy may be felt owing to the stabilizing effects of buoyancy. An eddy with a vertical scale of one meter might have a horizontal scale of, for argument's sake, 10–100 meters. Fine-scale interleaving in the ocean, as in the laboratory, is observed to have this character. Hence one should expect horizontal elements of the diffusivity tensor to be larger than the vertical element, which is often given as $K_V^* \sim 10^{-5}$ m^2 s^{-1} (see below), even by several orders of magnitude, say $K_H^* \sim 10^{-2}$ m^2 s^{-1}. This disparity is still much narrower than would be brought about were the averaging to encompass in the horizontal dimension mesoscales of tens to hundreds of kilometers (more about that below, in Section 4).

Because of the density-weighted averaging, the continuity equation applies in an identical way to total random density and velocity and to their infra-grid averages:

$$\frac{\partial \tilde{\rho}}{\partial x_0^*} + \frac{\partial}{\partial x_j^*} (\tilde{\rho}\tilde{u}_j) = 0, \quad \frac{\partial \rho}{\partial x_0^*} + \frac{\partial}{\partial x_j^*} (\rho u_j^*) = 0. \tag{2.2}$$

The average horizontal momentum balance is similar in each horizontal component to (2.1), except for the addition of the Coriolis and pressure-gradient forces,

$$\frac{Du_\alpha^*}{Dt} + f\varepsilon_{3j\alpha} u_j^* + \frac{1}{\rho} \frac{\partial p}{\partial x_\alpha^*} = \frac{1}{\rho} \frac{\partial}{\partial x_j^*} (\rho M_{j\alpha}^*) \equiv q_\alpha^M \tag{2.3}$$

---

3. Infra-grid fluctuations around mean density ρ, and indeed spatial and temporal variations of ρ, may usually be safely neglected in (2.2) and elsewhere. (The conventional Boussinesq approximation exploits this property.) However, the inclusion of density variations requires little additional effort here (de Szoeke and Samelson, 2002), and provides an instructive preview of thickness-weighted averaging when isentropic coordinates are used (Section 4, below).

($\alpha = 1, 2$). In the vertical, mean hydrostatic balance is assumed:

$$\frac{\partial p}{\partial x_3^*} = -g\rho. \tag{2.4}$$

The equation of state,

$$\rho^{-1} = V = \check{V}(p, \theta), \tag{2.5}$$

will be taken to link infra-grid-averaged density, pressure and potential temperature. The dependence of seawater density on salinity will be neglected here, to be taken up in a sequel paper. Though there are additional contributions to (2.5) of terms like $\check{V}_{\theta\theta}G_0^\theta$. arising from equation-of-state nonlinearity combined with infra-grid temperature variance (see below), such terms are negligible (Davis, 1994a). Only infra-grid averaged pressure p (though calculated without density weighting) appears in (2.3) and (2.4). The infra-grid pressure fluctuations do not appear in the averaged equations.

The equation for infra-grid scalar concentration variance is

$$\rho \frac{DG_0^C}{Dt} = \frac{\partial}{\partial x_j^*}(\rho G_j^{*C}) + \rho B_j^{*C}\frac{\partial C}{\partial x_j^*} - \rho\chi^C. \tag{2.6}$$

(See Appendix E for definitions.) A similar equation can be obtained for the variance of turbulent infra-grid velocity fluctuations,

$$\rho \frac{DT}{Dt} = \frac{\partial}{\partial x_j^*}(\rho T_j^*) + \rho M_{ji}^*\frac{\partial u_i^*}{\partial x_j^*} - \rho g a B_3^{*\theta} - \rho\Phi. \tag{2.7}$$

3. Transformation to isentropic coordinates

Suppose that $x_3 = \theta$ (infra-grid averaged potential temperature) is used as independent coordinate in place of $x_3^* = z$ (height); otherwise $x_j = x_j^*$, $j = 0, 1, 2$. Since potential temperature is an alias of entropy when variation of salinity is neglected, we call these *isentropic coordinates*. The third coordinate of a "point" indicates the isentrope on which it falls. The mathematical details of the transformation to isentropic coordinates are summarized in appendix B. The average mass conservation equation, second of (2.2), becomes

$$\frac{\partial h}{\partial x_0} + \frac{\partial}{\partial x_j}(hu_j) = 0. \tag{3.1}$$

Here h is the thickness of isentropic layers,

$$h = \rho z_\theta = -g^{-1}p_\theta, \tag{3.2}$$

the second equality following from the hydrostatic relation, Eq. (2.4). Eqs. (2.1), (2.3) and (2.4) become

$$\frac{DC}{Dt} = \frac{1}{h} \frac{\partial}{\partial x_j} (h B_j^C) = q^C(\mathbf{x}, t), \tag{3.3}$$

$$\frac{Du_\alpha}{Dt} + f\varepsilon_{3\beta\alpha} u_\beta + \frac{\partial M}{\partial x_\alpha} = \frac{1}{h} \frac{\partial}{\partial x_j} (h M_{j\alpha}) = q_\alpha^M(\mathbf{x}, t), \quad \alpha = 1, 2, \tag{3.4}$$

$$\frac{\partial M}{\partial x_3} = \Pi. \tag{3.5}$$

Exner function Π is a thermodynamic variable, a function only of p and θ, determined once and for all from the equation of state; M is the Montgomery function. Both variables are described in Appendix B and Appendix E (glossary). The third component of generalized velocity is diagnostically specified with replacement of C by θ in eq. (3.3):

$$u_3 \equiv \dot\theta \equiv \frac{D\theta}{Dt} = \frac{1}{h} \frac{\partial}{\partial x_\alpha} (h B_\alpha^\theta) + \frac{1}{h} \frac{\partial}{\partial \theta} (h B_3^\theta) = q^\theta(\mathbf{x}, t). \tag{3.6}$$

(McDougall (2003) avers that there should be additional source terms in this equation. We shall neglect such effects.)

The equation (2.6) for the infra-grid variance of scalar C becomes

$$\frac{\partial}{\partial x_0} (h G_0^C) + \frac{\partial}{\partial x_j} (h u_j G_0^C - h G_j^C) = h B_j^C \frac{\partial C}{\partial x_j} - h\chi^C. \tag{3.7}$$

In particular, for the temperature variance, where $C = \theta = x_3$, this is

$$\frac{\partial}{\partial x_0} (h G_0^\theta) + \frac{\partial}{\partial x_j} (h u_j G_0^\theta - h G_j^\theta) = h B_3^\theta - h\chi^\theta. \tag{3.8}$$

A similar equation for infra-grid kinetic energy is

$$\frac{\partial}{\partial x_0} (h T) + \frac{\partial}{\partial x_j} (h u_j T - h T_j) = h M_{ji} \frac{\partial u_i}{\partial x_j} - h g a B_3^{*\theta} - h\Phi. \tag{3.9}$$

4. Sub-grid (mesoscale) isentropic averaging

Suppose the equations and variables, (3.1)–(3.9), expressed in isentropic coordinates, are further averaged over larger scales at fixed potential temperature θ,[4] i.e., on isentropes. Such averages will be indicated by angle brackets. This averaging is supplementary to the infra-grid averaging already inherent in the equations of Sections 2 and 3. The precise definition of the term *sub-grid* will depend on the intended application. For example, in a discrete numerical model, the sub-grid scale might be the smallest resolved length of the

---

4. When necessary, averages of variables at fixed $\theta = x_3$ are indicated by $\langle f(\mathbf{x}, t) \rangle$ to distinguish them from averages at fixed $z = x_3^*$, i.e. $\langle f(\mathbf{x}*, t) \rangle$. When the abbreviation $\langle f \rangle$ is written, without elaboration, the former is intended.

model. In displaying synoptic oceanographic data, the sub-grid scale should be determined by the sampling frequency (spacing) and subsequent data processing. Typically, sub-grid averages will be taken over tens to hundreds of kilometers. Their essential feature is that they should be performed on isentropes, not level surfaces, as the infra-grid averaging was done. (Otherwise the two averages might as well be telescoped into a single procedure.)

Averaging on isentropes reflects observational data-reduction practices that are often followed. On the other hand, some compilations of hydrographic data, such as Levitus's (1982), have been horizontally binned and averaged at fixed pressure levels. Yet in the vicinity of a strong current system such as the Gulf Stream where isentropes and other properties undergo large spatial and temporal vertical excursions, level averaging may undesirably confound distinct oceanographic regimes, and averaging within isentropes should be preferred (Lozier *et al.,* 1994). Similarly, in the standard preparation and presentation of property-property diagrams, such as θ-S diagrams, individual "bottle" samples (or CTD observations) are plotted as pinprick points which constitute clouds in aggregate (e.g., Fuglister, 1960; Talley, 2007). The heuristic averaging invited by visual inspection of these diagrams and the resulting mean property relations inferred suggest isentropic averaging, certainly not level averaging.

Averages of variables which are intensive (mass-specific), such as scalar concentration C, velocity u_j (momentum/mass), specific volume V, etc., are thickness weighted and indicated by carets, $\hat{C}$, $\hat{u}_j$, $\hat{V}$ (de Szoeke and Bennett, 1993; Greatbatch and McDougall, 2003). Sub-grid fluctuations from these averages are indicated by single primes. On the other hand, variables which are not mass-specific, such as p, z, M, Π, are averaged without thickness weighting: $\langle p \rangle \equiv P$, $\langle z \rangle \equiv Z$, $\langle M \rangle$, $\langle \Pi \rangle$. The relations among sub-grid-averaged isentrope height Z, pressure P, and thickness $\langle h \rangle$ are obtained as follows. Because $\rho V = 1$,

$$Z_\theta = \langle z_\theta \rho V \rangle = \langle h \rangle \hat{V};$$

and, from the hydrostatic relation (2.4),

$$P_\theta = -\langle g \rho z_\theta \rangle = -g \langle h \rangle.$$

These equations may be recombined as

$$P_\theta = -g\hat{V}^{-1}Z_\theta, \quad \text{and} \quad \langle h \rangle = \langle \rho z_\theta \rangle = \hat{V}^{-1}Z_\theta. \tag{4.1}$$

The first is to be taken as the hydrostatic relation averaged over sub-grid scales on isentropic surfaces. The *exact* establishment of the second equation for average thickness is also notable. The reciprocal of thickness-weighted average specific volume arises spontaneously as the natural form of sub-grid average density.

Thickness-weighted sub-grid averages of products of mass-specific variables are, for example,

$$\langle h u_j C \rangle = \langle h \rangle \hat{u}_j \hat{C} + \langle h u_j' C' \rangle. \tag{4.2}$$

Eqs. (3.1), (3.3)–(3.6), averaged, become

$$\frac{\partial \langle h \rangle}{\partial x_0} + \frac{\partial}{\partial x_\alpha} (\langle h \rangle \hat{u}_\alpha) + \frac{\partial}{\partial x_3} (\langle h \rangle \hat{\theta}) = 0, \tag{4.3}$$

$$\frac{\hat{D}\hat{C}}{Dt} = -\frac{1}{\langle h \rangle} \frac{\partial}{\partial x_j} \langle h u_j' C' \rangle + \hat{q}^C(\mathbf{x}, t), \tag{4.4}$$

$$\hat{u}_3 = \hat{\theta} = \hat{q}^\theta(\mathbf{x}, t), \tag{4.5}$$

$$\frac{\hat{D}\hat{u}_\alpha}{Dt} + f\varepsilon_{3\beta\alpha}(\hat{u}_\beta - u_\beta^+) + \frac{\partial \langle M \rangle}{\partial x_\alpha} = -\frac{1}{\langle h \rangle} \frac{\partial}{\partial x_j} \langle h u_j' u_\alpha' \rangle + \hat{q}^M_\alpha(\mathbf{x}, t), \tag{4.6}$$

$$\frac{\partial \langle M \rangle}{\partial x_3} = \langle \Pi \rangle. \tag{4.7}$$

The total eddy transport of average scalar $\hat{C}$ is composed of the sum of covariance of sub-grid motion and concentration, $-\langle h u_j' C' \rangle$, and infra-grid eddy transport, $\langle h B_j^C \rangle$—the divergence of the latter constituting $\hat{q}^C$. This is presumably dominated by the transport across isentropes, $\langle h B_3^C \rangle$, which has now been further averaged over sub-grid scales. Eq. (4.5), compared with (4.4), contains contributions only from infra-grid fluxes (because $\theta' \equiv 0$, trivially). In isentropic coordinates, this equation serves as a diagnostic specification of the third component of average generalized velocity, $\hat{u}_3 \equiv \hat{\theta}$, thus describing the dientropic mass flux, $\langle h \rangle \hat{\theta}$, with units of kg m^{-2} s^{-1}.

In (4.6), some particular terms have arisen involving thickness and Montgomery gradient covariances, namely,

$$u_\beta^+ = f^{-1} \varepsilon_{3\alpha\beta} \frac{1}{\langle h \rangle} \left\langle h' \frac{\partial M'}{\partial x_\alpha} \right\rangle. \tag{4.8}$$

They are written in this form because they will tend to force additional contributions to average $\hat{u}_\beta$. Certain specific parameterizations have been proposed for these terms (Gent and McWilliams, 1990).

The role of $\langle p'^2 \rangle$ when the equation of state is sub-grid averaged needs to be examined (McDougall and McIntosh, 1996, 2001). The averaged Montgomery and Exner functions appearing in (4.6), (4.7) are

$$\langle M \rangle = \langle H(p, \theta) \rangle + gZ, \qquad \langle \Pi \rangle = \left\langle \frac{\partial H(p, \theta)}{\partial \theta} \right\rangle. \tag{4.9}$$

The latter is

$$\langle \Pi \rangle = \left\langle \frac{\partial H(P + p', \theta)}{\partial \theta} \right\rangle = \left\langle \frac{\partial H(P, \theta)}{\partial \theta} + p' \breve{V}_\theta(P, \theta) + \frac{1}{2} p'^2 \breve{V}_{P\theta}(P, \theta) + \cdots \right\rangle$$
$$= \Pi(P, \theta) + \frac{1}{2} \langle p'^2 \rangle \breve{V}_{P\theta} + \cdots \tag{4.10}$$

Illustrating with a simple equation of state that exhibits nonzero $\check{V}_{p\theta}$, namely, $\check{V}(p, \theta) = V_1(p) + \Delta V(1 + \mu p)a\theta$, one sees that this would be

$$\langle \Pi \rangle = \Delta V a \{ P + \tfrac{1}{2}\mu P^2 + \tfrac{1}{2}\mu \langle p'^2 \rangle \}. \tag{4.11}$$

What is the effect of $\langle p'^2 \rangle$? Suppose that as an archetypal model of the vertical distribution of sub-grid pressure variance associated with meso-scale disturbances one takes the WKB form of the first baroclinic mode function, namely,

$$\langle p'^2 \rangle \sim p_0^2 |\psi_1(P)|^2, \quad \text{where } \psi_1(P) \approx \sin \left[\pi \frac{\int_0^P N(p)dp}{\int_0^{P_B} N(p)dp} \right], \tag{4.12}$$

For the vertical buoyancy frequency dependence, suppose $N(p) \approx N_0 e^{\sigma p}$. Typical isentrope displacements of order $p_0 \sim 100$ dbar might be expected in the main thermocline at $p \sim \sigma^{-1} \sim 1000$ dbar. Then

$$\frac{\langle p'^2 \rangle}{P^2} \rightarrow p_0^2 \sigma^2 \pi^2 \sim 0.1, \quad \text{as } P \rightarrow 0. \tag{4.13}$$

This estimate near $P \sim 0$ is an upper bound of the ratio. Near $P = 1000$ dbar, the ratio reduces to 0.01. Thus the $\tfrac{1}{2}\mu \langle p'^2 \rangle$ term is negligible compared to the $\tfrac{1}{2}\mu P^2$ term (it is even smaller compared to the P term), and we shall neglect it. Thus, simply,

$$\langle \Pi \rangle = \Pi(P, \theta). \tag{4.14}$$

Similarly, the averaged heat function is

$$\langle H(p, \theta) \rangle = \langle H(P + p', \theta) \rangle = \langle H(P, \theta) + p'H_P(P, \theta) + \tfrac{1}{2}p'^2 H_{PP}(P, \theta) + \cdots \rangle$$

$$= H(P, \theta) + \tfrac{1}{2}\langle p'^2 \rangle \check{V}_P(P, \theta) + \cdots \tag{4.15}$$

Hence the leading neglected term in approximating mean isentrope height by

$$Z = g^{-1}\{\langle M \rangle - H(P, \theta)\} \tag{4.16}$$

is

$$\Delta Z = \tfrac{1}{2}g^{-1}\langle p'^2 \rangle \check{V}_P = \tfrac{1}{2}\langle z'^2 \rangle g/c^2. \tag{4.17}$$

This may be estimated as $\Delta Z \sim \tfrac{1}{2}(100 \text{ m})^2 (9.8 \text{ m s}^{-2})/(2 \times 10^6 \text{ m}^2 \text{ s}^{-2}) \sim 2.5$ cm, which is quite negligible.

Eqs. (4.3) and (4.6) are solved prognostically for $\langle h \rangle = -P_\theta/g$ and $\hat{u}_\alpha$. From the former, P may be obtained by a vertical integration, and $\Pi(P, \theta)$ determined by substitution. A

second vertical integration, of (4.7), yields $\langle M \rangle$, which is required by the next prognostic step for $\langle h \rangle$ and $\hat{u}_\alpha$; and so on. Eq. (4.16) is required only for an offline diagnostic calculation of mean isentrope height Z, which does not otherwise enter the prognostic equations. Were the assessment of the importance of the $\langle p'^2 \rangle$ contributions to $\langle \Pi \rangle$ and $\langle H \rangle$ different, a prognostic equation for $\langle p'^2 \rangle$—effectively a meta-state variable created by the sub-grid averaging—would have to be developed.

The role of $\langle p'^2 \rangle$ when specific volume is averaged is also negligible. The sub-grid averaged, thickness-weighted specific volume is calculated as follows:

$$\hat{V} = \frac{\langle p_\theta V \rangle}{P_\theta} = \frac{\langle (P_\theta + p'_\theta) \check{V}(P + p', \theta) \rangle}{P_\theta}$$

$$= \check{V}(P, \theta) + \frac{1}{2P_\theta} \frac{\partial}{\partial \theta} \langle p'^2 \rangle \check{V}_P + \frac{1}{2} \langle p'^2 \rangle \check{V}_{PP} + \cdots .$$

(4.18)

For estimation purposes, suppose that $(1/P_\theta)\partial/\partial\theta \sim 1/(1000 \text{ dbar})$, and $\langle p'^2 \rangle \sim (100 \text{ dbar})^2$; also $\check{V}_P/\check{V} = -\check{V}/c^2 = 4 \times 10^{-6} \text{ dbar}^{-1}$, $\check{V}_{PP}/\check{V} = 1.3 \times 10^{-10} \text{ dbar}^{-2}$. Then the second and third terms as a proportion of $\hat{V}$ are $2 \times 10^{-5}, 7 \times 10^{-7}$ respectively. The second term represents a contribution to *density* of 0.02 kg m^{-3}, which is barely significant. The third term is smaller than 0.001 kg m^{-3}, and is quite negligible. Hence the approximation, $\hat{V}^{-1} \approx \check{V}(P, \theta)^{-1}$, incurs an error no more than 0.02 kg m^{-3}. However that may be, $\hat{V}$, as it occurs in the equations, is not required to nearly this accuracy, so that this error is of no consequence. All pertinent dynamical buoyancy effects are expressed accurately and in detail through the mean Exner function $\langle \Pi \rangle$ in (4.7).

5. Parameterizing infra-grid eddy transports

For eddy transport of a scalar by infra-grid scales of motion, one may adopt a parameterization similar to the common one given in Appendix B, Eqs. (B13)–(B18). After further sub-grid averaging and thickness-weighting, one may take this to be:

$$\langle hB_i^C \rangle = \langle h \rangle K_{ij} \frac{\partial \hat{C}}{\partial x_j} = \begin{cases} \hat{V}^{-1} K_H^* \left(Z_\theta \frac{\partial \hat{C}}{\partial x_i} - Z_{,i} \frac{\partial \hat{C}}{\partial \theta} \right) & (i = 1, 2) \\ \hat{V}^{-1} \frac{K_D}{Z_\theta} \frac{\partial \hat{C}}{\partial \theta} - \hat{V}^{-1} K_H^* Z_{,\alpha} \frac{\partial \hat{C}}{\partial x_\alpha} & (i = 3) \end{cases}$$

(5.1)

(Upon sub-grid averaging, z has been replaced by Z, and ρ by $\hat{V}^{-1}$ in formula (B16), appendix B). Horizontal and vertical eddy diffusivities due to motions of the infra-grid scales are K_H^*, K_V^*. The coefficient appearing in the vertical component of (5.1) is the dientropic diffusivity,

$$K_D = K_V^* + K_H^* \{(Z_{,1})^2 + (Z_{,2})^2\}$$

(5.2)

[cf. (B17)]. (Note the distinction between K_D and K_V^*.) In particular, the average infra-grid eddy heat transport is taken to be

$$\langle hB_i^\theta \rangle = \begin{cases} -\hat{V}^{-1}K_H^*Z_{,i} & (i = 1, 2) \\ \hat{V}^{-1}\dfrac{K_D}{Z_\theta} & (i = 3) \end{cases} \tag{5.3}$$

[cf. (B19)]. The divergences of these transports give the scalar and temperature infra-grid source functions, $\hat{q}^C$, $\hat{q}^\theta$. In particular the temperature source gives the average dientropic mass transport,

$$\langle h \rangle \hat{\dot{\theta}} = \langle h \rangle \hat{q}^\theta = \frac{\partial}{\partial x_\alpha}(-\hat{V}^{-1}K_H^*Z_{,\alpha}) + \frac{\partial}{\partial \theta}\left(\hat{V}^{-1}\frac{K_D}{Z_\theta}\right). \tag{5.4}$$

When inserted into the prognostic equation for average thickness, (4.3), this gives

$$\frac{\hat{D}\langle h \rangle}{Dt} + \langle h \rangle \frac{\partial \hat{u}_\alpha}{\partial x_\alpha} = \frac{\partial}{\partial x_\alpha}\left(K_H^*\frac{\partial \langle h \rangle}{\partial x_\alpha}\right) - \frac{\partial^2}{\partial \theta^2}\left(\hat{V}^{-2}\frac{K_D}{\langle h \rangle}\right). \tag{5.5}$$

This is a modified nonlinear diffusion equation for average thickness $\langle h \rangle$. There is along-isentropic linear diffusion with coefficient K_H^*, and nonlinear cross-isentropic diffusion with coefficient K_D. For the latter one would take a value suggested by microstructure turbulence observations in the ocean; say $K_D \sim 10^{-5}$ m^{2}s^{-1} (see below). Even on infra-grid scales, O(1–10 m), one expects motions to be significantly anisotropic so that K_H^* may well be several orders of magnitude larger than K_D. Not, however, 10^8 times larger, as is suggested from observations of the effects of meso-scale stirring of large-scale ocean properties. This is simply because the infra-grid eddying motions that comprise the heat transport $\langle hB_j^\theta \rangle$ do not include the effects of these larger scales.

a. Osborn-Cox hypothesis

In equations like (3.7) for infra-grid scalar variance the time-evolution and transport divergence terms on the left side are often neglected. One is then left with the balance of production and dissipation, in particular for temperature variance, Eq. (3.8), $B_3^\theta = \chi^\theta$. This proposition is known as the Osborn-Cox hypothesis (Osborn and Cox, 1972, Davis, 1994b). The temperature variance dissipation thus furnishes the *dientropic* component of infra-grid eddy heat transport. In this paper this hypothesis is invoked only after further thickness-weighted sub-grid averaging:

$$\langle hB_3^\theta \rangle \approx \langle h\chi^\theta \rangle. \tag{5.6}$$

The neglect of the time-evolution and transport divergence terms in (5.6) seems far better justified after such additional averaging. Indeed, in practice the hypothesis is never applied to a single small segment of data from which χ^θ might have been estimated, but to averages of ensembles of numerous such segments (Davis, 1996). The innovation in (5.6) is to weight the averages with thickness. Substituting the third of (5.3), one obtains an estimate of the dientropic diffusivity:

$$K_D \approx \hat{V}Z_\theta \langle \rho z_\theta \chi^\theta \rangle \approx \hat{V}Z_\theta \langle z_\theta \overline{\tilde{\rho}\kappa^\theta(\partial\theta/\partial x_j^*)^2} \rangle \equiv \kappa^\theta \text{Cx}. \tag{5.7}$$

The last equality defines the Cox number,

$$\text{Cx} = (Z_\theta)^2 \frac{\langle h\chi^\theta \rangle}{\kappa^\theta \langle h \rangle} \approx \frac{K_D}{\kappa^\theta}. \tag{5.8}$$

This definition differs from Osborn and Cox's (1972) in its employment of thickness weighting in the average.

Estimates such as (5.7), if considered typical of the turbulent regimes in which they were made, such as the near-surface ocean, the main thermocline, or the abyssal ocean, are often taken to be canonical for their respective regimes, and applicable not only to temperature diffusion but to the diffusion of any passive scalar.

The neglect of cross-scale contributions to the eddy covariances, B_j^θ, etc., has been scutinized and criticized (Gargett and Holloway, 1984; Davis, 1994b). (These extra contributions have been explicitly written out in Appendix A.) This neglect is dubious because of the absence of gaps in the turbulent spectra (usually "red," i.e., rising power towards low wavenumbers and frequencies). On the other hand, because the spectrum of temperature *gradient* rises toward high wavenumber, where it is eventually cut off by molecular diffusion, this objection does not apply to estimates of temperature variance dissipation. Estimators of dissipation, even if noisy, are fair and unbiased (Davis, 1996). The Osborn-Cox hypothesis (5.6) furnishes the dientropic component of the eddy heat transport enhanced by cross-scale contributions [Appendix A, Eq. (A4)]. It is precisely this enhanced transport that appears when averaged eddy transport effects are considered. (Appendix A may be consulted for more details.)

b. Turbulent kinetic energy balance

As an alternative to the Osborn-Cox hypothesis of temperature-variance production vs. dissipation, though similar to it, one may reason from the microstructure turbulent kinetic energy production-destruction balance, modified by buoyancy consumption of energy, Eq. (3.9), thickness-weighted and averaged:

$$P - ag\langle hB_3^{*\theta} \rangle \approx \langle h\Phi \rangle, \tag{5.9}$$

where P and $\langle h\Phi \rangle$ are sub-grid averaged, thickness weighted, kinetic energy production and dissipation; $agB_3^{*\theta}$ is the vertical transport of buoyancy. The average *vertical* heat flux appearing in (5.9) is

$$\frac{\langle hB_3^{*\theta} \rangle}{\langle h \rangle} \approx \frac{K_V^*}{Z_\theta} \tag{5.10}$$

[cf. (B14), Appendix B]. Employing the flux Richardson number, defined by

$$\text{Rf} = \frac{ag\langle hB_3^{*\theta} \rangle}{P}, \tag{5.11}$$

one may express (5.9) as

$$K_V^* \approx \frac{\text{Rf}}{1 - \text{Rf}} \frac{\hat{V}\langle h\Phi\rangle}{ag}. \tag{5.12}$$

Thus, measurements of kinetic energy dissipation Φ have been employed to give estimates of vertical diffusivity K_V^*, analogous to the Osborn-Cox estimate of dientropic diffusivity K_D. The distinction between the two parameters, noted in Eq. (5.2), is customarily overlooked. The two estimation methods, when compared in the field, have given similar values for the respective parameters (Gregg, 1987, 1989, 1999; Klymak and Nash, 2009). Because the turbulent shear spectra that constitute $\hat{\Phi}$ peak at longer scales than the temperature gradient spectra, their routine estimation is more practicable in the field.

6. Sub-grid diffusion

With the empirical specification of the dientropic eddy coefficient K_D, determined from microstructure turbulence, the specification of the mean dientropic flow $\langle h\rangle\hat{\theta}$ [Eq. (5.4)] is closed. Any other mean scalar $\hat{C}$, or mean velocity $\hat{u}_\alpha$, is affected also by transports due to sub-grid eddy turbulence, $-\langle hu_j'C'\rangle$. For a strictly passive scalar, a plausible theory can be developed as shown in Appendix C for the sub-grid eddy covariance in (4.4) that gives the scalar conservation equation,

$$\frac{\hat{D}\hat{C}}{Dt} \approx \text{div}\,\{\mathbf{D} : \text{grad}\,\hat{C}\}, \tag{6.1}$$

where the "div" and "grad" operators and eddy diffusion tensor $\mathbf{D}$ are given by (C14) and (C15), with (C12) and (B16).

a. Scaling analysis of the sub-grid diffusivity tensor

Scaling the sub-grid covariances that occur in (C12) suggests the following form for the total diffusivity tensor for average passive scalar $\hat{C}$:

$$
\begin{aligned}
\mathbf{D} \equiv (D_{ik}) &\sim
\begin{pmatrix}
\tau_H\sigma_1^2 & a_{12}'\tau_H\sigma_1\sigma_2 & a_{13}'\sqrt{\tau_H\tau_V}\,\sigma_1\sigma_3 \\
a_{21}'\tau_H\sigma_1\sigma_2 & \tau_H\sigma_2^2 & a_{23}'\sqrt{\tau_H\tau_V}\,\sigma_2\sigma_3 \\
a_{31}'\sqrt{\tau_H\tau_V}\,\sigma_1\sigma_3 & a_{32}'\sqrt{\tau_H\tau_V}\,\sigma_2\sigma_3 & \tau_V\sigma_3^2 + K_D
\end{pmatrix} \\
&\sim
\begin{pmatrix}
L_H\sigma_1 & a_{12}L_H\sigma_1 & a_{13}\sqrt{L_HL_V\sigma_1\sigma_3} \\
a_{21}L_H\sigma_1 & a_{22}L_H\sigma_1 & a_{23}\sqrt{L_HL_V\sigma_1\sigma_3} \\
a_{31}\sqrt{L_HL_V\sigma_1\sigma_3} & a_{32}\sqrt{L_HL_V\sigma_1\sigma_3} & L_V\sigma_3 + K_D
\end{pmatrix} \\
&\sim D_H
\begin{pmatrix}
1 & a_{12} & a_{13}\delta \\
a_{21} & a_{22} & a_{23}\delta \\
a_{31}\delta & a_{32}\delta & \delta^2 + \varepsilon
\end{pmatrix},
\end{aligned}
\tag{6.2}
$$

with $a_{12} = a_{12}'\sigma_2/\sigma_1$, $a_{21} = a_{21}'\sigma_2/\sigma_1$, $a_{32} = a_{32}'\sigma_2/\sigma_1$, $a_{23} = a_{23}'\sigma_2/\sigma_1$, $a_{22} = (\sigma_2/\sigma_1)^2$. In these formulas, $\sigma_i = \sqrt{\langle hv_i'^2\rangle}/\langle h\rangle$ (no summation implied), where $v_i' =$

$\Delta_{il}u'_l$. Thus the third component, $v'_3 = Z_\theta\dot\theta'$, has units of velocity, like the other components. Also τ_H, τ_V are decorrelation times for along- and cross-isopycnal fluctuating motions—the latter much smaller than the former (note the assumed intermediate decorrelation time $\sqrt{\tau_H\tau_V}$ for horizontal vs. cross-isentropic covariances); $L_H = \tau_H\sigma_1$, $L_V = \tau_V\sigma_3$ are corresponding mixing lengths; $D_H = \tau_H\sigma_1^2 = L_H\sigma_1$, $D_V = \tau_V\sigma_3^2 = L_V\sigma_3$, $\delta = \sqrt{D_V/D_H}$, $\varepsilon = K_D/D_H$. It is assumed that $a_{ij} \gtrsim O(1)$.

Estimates of the horizontal-diagonal elements of **D** have been made from neutrally buoyant drifters (Davis, 1985; Freeland *et al.,* 1975). These support values of $D_H \sim 10^3$ m^{2}s^{-1}. While it is difficult to envisage data sets from which direct estimates might be made of the cross-isentropic to horizontal covariances that constitute the off-diagonal elements in the third row and column of **D**, a scaling estimate of σ_3 can be made as follows. From (3.6),

$$\dot\theta \approx \frac{1}{\rho z_\theta} \frac{\partial}{\partial\theta} (\rho z_\theta B_3^\theta) \sim \frac{1}{\rho z_\theta} \frac{\partial}{\partial\theta} \left(\rho \frac{K_D}{z_\theta}\right), \tag{6.3}$$

having made, in the last replacement, the same estimate of the dientropic heat flux, B_3^θ, before meso-scale averaging, as was employed in (5.3) after averaging. The fluctuating dientropic velocity we therefore take as

$$v'_3 = Z_\theta\dot\theta' \sim \frac{\partial}{\partial\theta}\left(\frac{K_D}{(Z_\theta)^2}(-z'_\theta)\right); \tag{6.4}$$

and the variance of this as

$$\sigma_3^2 = \text{var}(v'_3) \sim \left(\frac{K_D}{L_V}\right)^2 \frac{\langle z'^2\rangle}{L_V^2}. \tag{6.5}$$

Here we estimated $\frac{\partial}{\partial\theta}\left(\frac{1}{Z_\theta}\circ\right) \sim \frac{1}{L_V}\circ \sim \frac{1}{1000\ \text{m}}\circ$, using the scale of variation of the thermocline for the vertical mixing length. Thus

$$D_V = L_V\sigma_3 \sim K_D \frac{\sqrt{\langle z'^2\rangle}}{L_V} \sim K_D \frac{100\ \text{m}}{1000\ \text{m}} \sim 0.1K_D, \tag{6.6}$$

having taken 100 m for the r.m.s. isentrope displacement. Hence

$$\delta^2 = \frac{D_V}{D_H} \sim 0.1\frac{K_D}{D_H} \sim 0.1\varepsilon. \tag{6.7}$$

Accepting this scaling argument, one concludes that the δ-terms in (6.2) are negligible. The eigenvalues, or principal diffusivities, of (6.2) will be $D_H\{O(1) + O(\delta^2)\}$ (two values), and $K_D + D_H O(\delta^2) \sim K_D + O(D_V)$, with respect to principal directions (eigenvectors) differing from along-isentrope directions and the cross-isentrope direction, respectively, by $O(\delta)$. As noted, it is doubtful that these differences are significantly nonzero. The axes of diffusive spreading of passive tracers differ insignificantly from isentropes.

Thus the third off-diagonal row and column of **D** are negligible. With the further simplifying assumption that the along-isentrope sub-grid diffusion is independent of orientation, one obtains the Redi (1982) tensor for scalar diffusion,

$$\mathbf{D} \approx \text{diag}\,(D_H, D_H, K_D). \tag{6.8}$$

The coefficient D_H governs the spreading of passive scalars along isentropes.

b. An alternative: Averaging in level coordinates

A quite particular approach has been taken in this paper by sub-grid averaging on isentropes (de Szoeke and Bennett, 1993). The conventional approach is to consider variables and equations averaged at fixed $z = x_3^*$ levels (Davis, 1994a,b; Bennett, 1996). When the derivations of the diffusion equation (6.1) [Appendix C] for a passive scalar averaged on level surfaces are recapitulated, one obtains

$$\frac{\partial}{\partial x_0^*} \langle C \rangle + \langle u_i^* \rangle \frac{\partial}{\partial x_i^*} \langle C \rangle = \frac{1}{\langle \rho \rangle} \frac{\partial}{\partial x_i^*} \left(\langle \rho \rangle K_{ik}^g \frac{\partial}{\partial x_k^*} \langle C \rangle \right).^5 \tag{6.9}$$

Here $\langle C \rangle = \langle C(\mathbf{x}^*, t) \rangle$, $\langle u_i^* \rangle = \langle u_i^*(\mathbf{x}^*, t) \rangle$, $\langle \rho \rangle = \langle \rho(\mathbf{x}^*, t) \rangle$ (see footnote 4). The sub-grid diffusivity tensor with regard to level coordinates is

$$K_{ik}^g(\mathbf{x}^*, t|r) = \frac{1}{\langle \rho \rangle} \int_r^t \langle \rho(\mathbf{x}^*, t) u_i'^*(\mathbf{x}^*, t) u_k'^*(\mathbf{A}^*(\mathbf{x}^*, t|w), w) \rangle dw \tag{6.10}$$

[or, in matrix notation,

$$\mathbf{K}^g = \frac{1}{\langle \rho \rangle} \int_r^t \langle \rho \mathbf{u}'^*(\mathbf{x}^*, t) \mathbf{u}'^*(\mathbf{A}^*, w)^T \rangle dw], \tag{6.10}'$$

where $\mathbf{A}^*(\mathbf{x}^*, t|w)$ is a Lagrangian trajectory in conventional Cartesian coordinates. The fluctuations here, $u_i'^*(\mathbf{x}^*, t) = u_i^*(\mathbf{x}^*, t) - \langle u_i^*(\mathbf{x}^*, t) \rangle$, viewed from fixed levels, differ from the fluctuations appearing in the sub-grid eddy diffusivity tensor (C12), $u_i'(\mathbf{x}, t) = u_i(\mathbf{x}, t) - \langle h u_i(\mathbf{x}, t) \rangle / \langle h \rangle$ as viewed from isentropes. The third components especially, even when expressed in comparable units, $v_3' = Z_\theta \theta'$ vs. $u_3'^* = \dot{z}'$, which are respectively the fluctuating dientropic flow across an isentrope, and the fluctuating vertical velocity across a level surface, are radically different quantities. The former differs from zero solely because of infra-grid turbulence. Formally, in the limit $K_D \rightarrow 0$, both the mean and fluctuation of dientropic flow would vanish. On the other hand, the vertical component of motion, $u_3'^*$, need not vanish even in the limit of adiabatic flow.

---

5. The effects of infra-grid and sub-grid averaging may be lumped into a single operation, so that the infra-grid q^C term need not be separately distinguished.

It is often asserted that the diffusivity tensor $\mathbf{K}^g$ should be diagonal when rotated[6] into local mean potential temperature surfaces (isentropes),[7] where potential temperature sub-grid averaged at fixed levels—a different field than infra-grid averaged potential temperature—is meant. This assertion is by no means self-evident. If the necessary rotation at a point $\mathbf{x}^*$ and time t is denoted by $\mathbf{R}$, the rotated tensor is

$$\mathbf{K}^I = \mathbf{R}^T \mathbf{K}^g \mathbf{R} = \frac{1}{\langle \rho \rangle} \int_r^t \langle \rho \mathbf{R}^T \mathbf{u}'^*(\mathbf{x}^*) \mathbf{u}'^*(\mathbf{A}^*)^T \mathbf{R} \rangle dw. \tag{6.11}$$

Redi's (1982) assertion is that $\mathbf{K}^I$ is diagonal, say $\mathbf{K}^I = \mathrm{diag}\,(K_I, K_I, K_D)$ (where plausibly $K_I \sim D_H$). This would appear to require arguing that $(\mathbf{R}^T \mathbf{u}'^*)_3$ is small in some sense—at least far smaller than $u_3'^*$—though the justification for this is far from compelling. In contrast, the case we have made above for the near-diagonality of (6.2), based on order-of-magnitude arguments for instantaneous dientropic flow, seems far more persuasive.

7. Sub-grid-averaged pressure coordinates

The equations of mass, scalar concentration, temperature, and momentum, sub-grid-averaged on isentropes, (4.3)–(4.7), may be transformed "back" to quasi-Cartesian pressure coordinates, as described in appendix D. This gives

$$\frac{\partial U_i}{\partial X_i} = 0, \tag{7.1}$$

$$\frac{\hat{D}\hat{C}}{Dt} = \frac{\partial}{\partial X_i}\left(\bar{D}_{ij}\frac{\partial \hat{C}}{\partial X_j}\right), \tag{7.2}$$

where $\bar{D}_{ij} = \bar{E}_{ij} + \mathrm{diag}\,(K_H^*, K_H^*, \hat{V}^{-2}g^2 K_V^*)$,

$$\hat{\hat{\theta}} \equiv \frac{\hat{D}\theta}{Dt} = \frac{\partial}{\partial X_\alpha}\left(K_H^* \frac{\partial \theta}{\partial X_\alpha}\right) + \frac{\partial}{\partial X_3}\left(\hat{V}^{-2}g^2 K_V^* \frac{\partial \theta}{\partial X_3}\right), \tag{7.3}$$

$$\frac{\hat{D}U_\alpha}{Dt} + f\varepsilon_{3\beta\alpha}(U_\beta - u_\beta^+) + g\frac{\partial Z}{\partial X_\alpha} = \frac{\partial}{\partial X_i}\left(-\frac{\langle h U_i' U_\alpha' \rangle}{\langle h \rangle}\right) + \hat{q}_\alpha^M, \tag{7.4}$$

$$g\frac{\partial Z}{\partial X_3} = -\check{V}(P, \theta). \tag{7.5}$$

These have obvious formal similarities to the infra-grid equations (2.2), (2.1), (2.3), (2.4) in naïve Cartesian coordinates. (The similarity is even closer if the latter set is expressed with

---

6. The transformation to isentropic coordinates given in appendix A is not strictly a rotation, such as Redi (1982) gives, though the distinction is slight and immaterial.

7. Redi (1982), allowing for salinity, framed the assertion in terms of *density,* meaning, presumably, *potential density,* and *isopycnals,* rather than *potential temperature* and *isentropes.*

pressure p as independent vertical coordinate [de Szoeke and Samelson, 2002].) Where they differ is in the appearance of eddy transport processes, principally along-isentrope from sub-grid motions, and across-isentrope from infra-grid motions, for which the parameterizations shown in (7.2) have been obtained in Appendix D. These parameterizations are strictly justified only for passive scalars so that the eddy transport terms for momentum in (7.4) have not been analogously re-written.

The equation for dientropic flow, (4.5), has transformed into Eq. (7.3). Comparing (7.3) to (7.2), one notices the absence of sub-grid diffusion terms, characterized by the tensor $\bar{E}_{ij}$ for mean scalar $\hat{C}$. Though we have argued for a form of this tensor oriented to directions along local isentropes, this form is quite immaterial to the diffusion terms in (7.3).

The solution of (7.3), along with the other equations, furnishes $\theta(P)$. (The dependence on the other independent variables has been suppressed.) What is its relation to $\theta(p)$? Expanding in a Taylor series,

$$\theta(p) = \theta(P) + \theta_P p' + \tfrac{1}{2} \theta_{PP} p'^2 + \cdots, \tag{7.6}$$

its mean value may be calculated as

$$\langle \theta(p) \rangle = \theta(P) + \tfrac{1}{2} \theta_{PP} \langle p'^2 \rangle + \cdots, \tag{7.7}$$

So $\theta(P)$, when enhanced by the bias $\tfrac{1}{2} \theta_{PP} \langle p'^2 \rangle$, gives an estimate of $\theta = \langle \theta(p) \rangle$ at the mean pressure height $\langle p \rangle = P$. This requires a specification of $\langle p'^2 \rangle$: ideally, a prognostic equation would be formulated and solved for this variance; or, a global atlas of isentrope height variance having been constructed empirically, it could be used to adjust the bias in (7.7). The variance of the estimator (7.7) is

$$\langle \{\theta(p) - \langle \theta(p) \rangle\}^2 \rangle \approx (\theta_P)^2 \langle p'^2 \rangle. \tag{7.8}$$

This is a measure of the error in (7.7) that comes about from the mean-square sub-grid isentrope height variation. For an estimate of scale, suppose $\langle p'^2 \rangle \sim (100 \text{ dbar})^2$. Then (7.7) and (7.8) may give quite significant estimates for bias and error. Despite the obvious similarity of (7.1)–(7.5) to the equations of motion relative to naïve coordinates the distinctions that these factors introduce are substantial.

8. Discussion and summary

There are two principal, complementary paradigms of mixing and diffusion in the ocean. There is the small-scale paradigm in which the ocean is regarded as nearly homogeneous in the horizontal, and where the mixing of passive and active scalars in the vertical is accomplished by microstructure (infra-grid) scales of motion ranging from sub-centimeter to several or even tens of meters. At the other end of the spectrum sub-grid meso-scale motions (tens to hundreds of kilometers) are held to diffuse properties along density surfaces. (In this paper, where, salinity having been neglected, potential density is determined solely by potential temperature, we have taken isopycnal surfaces to be synonymous with isentropic surfaces.) To unify these paradigms, and to appropriate the

distinctive turbulence parameterizations commonly employed in each, we proposed averaging variables and equations in two stages, first over the infra-grid scales, then over the sub-grid scales. Davis (1994b) examined such two-stage averaging, though both averages were performed centered at fixed Cartesian space-time points. A crucial innovation is to perform the second, sub-grid, average centered on isentropes, i.e., at fixed horizontal position, time, and potential temperature (de Szoeke and Bennett, 1993).

In passing, we resolved a difficulty noted by Gargett and Holloway (1984) and Davis (1994b). Because of the redness of the spectra of many geophysical random variables, that is, power increasing towards low wave-numbers and frequencies, statistical estimators of moments of random fields (such as means, standard deviations, etc.) are notoriously unreliable. The contributions of cross-scale terms at the margin of the filter resolution cannot be neglected. Our solution, simply, is not to neglect them. It turns out that, in the turbulent balances of infra-grid temperature variance or kinetic energy, wherever a second-order moment occurs—such as eddy Reynolds stresses or heat transports in the temperature variance or kinetic energy production terms—the usual forms are enhanced with cross-scale contributions. In the first-order average of the temperature or momentum balance, too, precisely the divergences of the enhanced eddy transport or stress occur.

The Gargett-Holloway stricture does not apply to estimators of temperature gradient or velocity shear variances (in effect, the temperature variance or kinetic energy dissipations) because the wavenumber spectra of these quantities, peaked at very high wavenumbers, have a spectral gap at lower wavenumbers within the inertial subrange. If the infra-grid filtering scale is chosen in this subrange, reliable estimates of temperature, scalar, or kinetic energy dissipations can be made. The theoretical links supplied by the Osborn-Cox hypothesis, leading from variance dissipations to infra-grid eddy transports of temperature, the latter enhanced as necessary by cross-scale contributions, are then restored. The validity and applicability of microstructure estimates of quasi-vertical eddy diffusivities to infra-grid averaged first-order fields—temperature, scalar, or velocity—is thereby established.

We recommended that subsequent sub-grid averages should be performed on isentropes. Thus, further averages of potential temperature produce, trivially, no change. The sub-grid averaging of other scalars and vectors produces eddy transports from sub-grid fluctuations. No such sub-grid eddy transport of heat occurs. Only infra-grid eddy transports of heat, further averaged over ensembles of estimates, collected on sub-grid scales, contribute to average heat balance and to cross-isentropic mass exchange. (By contrast, if temperature were sub-grid averaged on *level* surfaces, there would be additional sub-grid eddy heat transports from sub-grid motion and temperature covariances.) The supplementary eddy transports, averaged on isentropes, of any other passive scalar can be related to average gradients which are mediated by an eddy diffusivity tensor with the appearance of a lagged sub-grid covariance tensor of motion component fluctuations relative to isentropes (Lundgren, 1981; Bennett, 1996). This is the classic Fickian diffusion law. Its derivation entails many questionable hypotheses or neglects (Appendix C). The class of flows called sub-mesoscale, such as surface fronts or topographic flow separations (on scales of

hundreds of meters or fewer) may make important contributions to eddy transports—contributions that would be ill-represented in the present more-or-less homogeneous Fickian view. Persevering nonetheless, scale analysis of this diffusivity tensor suggested that the elements of its third row and column, involving the dientropic motion, were negligible. Assuming further that sub-grid eddy transport within isentropes is isotropic, one might arrive at something similar to Redi's (1982) hypothesized eddy diffusivity tensor, diagonal when referred to isentropes. By this reading, the along-isentrope diffusivities would be associated with the sub-grid averaging, and the dientropic diffusivity with infra-grid averaging.

This is the closest we were able to come to the ideal of a rigorous justification for the hypothesis of the eddy diffusivity tensor's being diagonal with respect to isentropes. It would be well to reflect that the case is not completely watertight. The fluctuating dientropic motions might conceivably be energetic enough under some circumstances to engender significant levels in the third, dientropic, row and column of the diffusivity tensor. The principal directions (eigenvectors) of that tensor might not align with the isentropic coordinate system. This means that systematic lagged correlations between dientropic and horizontal motion could lead to diffusive spreading along non-isentropic axes. However that may be, the diffusion of potential temperature itself is never affected by sub-grid diffusivity. (That this is so is a mere tautological consequence of the isentropic averaging approach that has been employed. In Redi's (1982) formulation also, the along-isentrope diffusivity has no effect when applied to temperature.)

What is interesting and notable about the arguments here employed is that they nowhere relied on the notion of the buoyant restoration of water parcels displaced from their initial isentrope. Rather, it is the mere kinematic fact of the near-conservative property of potential temperature that is key. It follows that a different diffusion tensor may be obtained by averaging fluctuations of motions and concentrations referred to isopleth surfaces of any other near-conservative variable (for example, tritium corrected for radio-decay, or one of the chlorofluorocarbons). Just as the surfaces of these variables differ from one another, so the averaging with respect to each differs, and the respective diffusivity tensors and their principal values and axes may differ. Potential temperature is distinguished in this group by its sole possession of an acceleration potential (the Montgomery function), which links it strongly to momentum dynamics.

Standard physical oceanographic measurements reflect the ideal of what we have called infra-grid averaging. For example, traditional hydrographic bottle and thermometer casts furnish discrete measurements of temperature, salinity, etc., which should be thought of as averages in the vertical over at least the size of the bottle (∼1 m)—more likely far larger—and horizontally over the tens to hundreds of meters of the ship motion during the cast. Similar remarks could be made about the edited output of standard CTD lowerings. Moored current meters are usually set to average measurements over several minutes or longer. The different averaging characteristics of such instruments, even if imperfectly understood, have never deterred comparison among the variables that they produce. In the

same spirit, the infra-grid average that we have proposed may stand for the averaging inherent in these instruments.

Difficult microstructure measurements of temperature and velocity are nowadays routinely being made to resolve the smallest, infra-grid scales—limited only by molecular diffusion and viscosity—in the ocean (Klymak and Nash, 2009). These then furnish (infra-grid averaged) estimates of temperature variance dissipation and kinetic energy dissipation. As such dissipation estimates are notoriously intermittent (Davis, 1996), ensembles of numerous realizations of them are assembled and further averaged to produce stable estimates. To these, the Osborn-Cox hypothesis of quasi-steady temperature variance balance, or the hypothesis of eddy kinetic energy balance, may be applied to furnish (infra-grid) eddy heat and momentum transports. The second average, now typifying a larger horizontal domain, represents what we have called the sub-grid, meso-scale average. (One of our innovations is to suggest the use of weighting by infra-grid averaged thickness when sub-grid averaging the dissipation estimates, just as other sub-grid averaged variables are weighted.)

We stressed the crucial consequences that the sub-grid averaging being performed on isentropes has to the theoretical structure of the diffusivity tensor. That averaging done, there seems to be no bar to transforming the equations back to Cartesian coordinates. The natural vertical coordinate to use is not naïve height, but the sub-grid averaged height of fixed isentropes.[8] One then recovers a set of equations whose appearance is exactly like the conventional hydrostatic equations of motion. Notwithstanding the peculiar vertical coordinate, one might simply interpret a solution of the equations as mapped onto a conventional Cartesian grid. We showed that such re-mapped solution fields were biased from the true fields by amounts which could nevertheless be estimated. The bias estimate is proportional to mean-square isentrope height variance.

APPENDIX A

A practical infra-grid average

Suppose the infra-grid average is defined, as a practical procedure, by

$$\overline{\phi(\mathbf{x}^*, t)} = \int_{\Omega} \phi(\mathbf{x}^* - \mathbf{x}_1^*, t - t_1) W(\mathbf{x}_1^*, t_1) d\mathbf{x}_1^* dt_1 \qquad (A1)$$

where $W(\mathbf{x}^*, t)$ is a filter kernel centered at $\mathbf{x}^* = 0$, $t = 0$, compactly supported in four-dimensional Cartesian space-time Ω, and normalized so that

---

8. Actually, the averaged *pressure* level of isentropes was used as vertical coordinate. This has the consequence that the mass conservation equation is simplified to a solenoidal equation.

$$\int_\Omega W(\mathbf{x}^*, t)d\mathbf{x}^*dt = 1. \tag{A2}$$

This definition satisfies many of the Reynolds averaging conditions, although generally *not*

$$\overline{\overline{\phi(\mathbf{x}^*, t)}\psi(\mathbf{x}^*, t)} = \overline{\phi(\mathbf{x}^*, t)}\overline{\psi(\mathbf{x}^*, t)}, \ \overline{\overline{\phi(\mathbf{x}^*, t)}} = \overline{\phi(\mathbf{x}^*, t)}, \quad \text{or} \ \overline{(\phi(\mathbf{x}^*, t) - \overline{\phi(\mathbf{x}^*, t)})} = 0. \tag{A3}$$

For such relations to hold accurately, it has to be assumed that in the spectra of the variables ϕ and ψ there is a considerable gap between high-frequency or high-wavenumber contributions and low-frequency or low-wavenumber contributions, and that the cut-off averaging scales inherent in the filter kernel W in (A1) are chosen in this gap (Monin and Yaglom, 1971). However, Gargett and Holloway (1984) argue forcefully that spectra of velocity (and presumably temperature and other scalars), encompassing at the high wavenumber and frequency end three-dimensional turbulence and internal waves, exhibit no such gap. As a consequence, second moments, like the covariance constituting the infra-grid eddy transport appearing in the infra-grid averaged balance (2.1), are more accurately written

$$B_j^{*C} = -\overline{\tilde{\rho}\tilde{u}_j\tilde{C}}/\rho + u_j^*C = -\overline{\tilde{\rho}u_j''C''}/\rho - R_j^{uC} \tag{A4}$$

(Davis, 1994b), where

$$R_j^{uC} = \{\overline{\tilde{\rho}C''u_j^*} + \overline{\tilde{\rho}u_j''C} + \overline{\tilde{\rho}u_j^*C} - \rho u_j^*C\}/\rho. \tag{A5}$$

(Recall that $\rho = \bar{\rho} = \overline{(\rho + \rho'')}$, $u_j^* = \overline{\tilde{\rho}(u_j^* + u_j'')}/\rho$, $C = \overline{\tilde{\rho}(C + C'')}/\rho$, etc., and note that none of $\overline{\rho''}$, $\overline{\tilde{\rho}u_j''}$, $\overline{\tilde{\rho}C''}$ is necessarily zero.) The infra-grid momentum transport, $M_{j\alpha}^*$, figuring in the average momentum balance (2.3), has a similar elaboration. Should the conditions (A3) be true or approximately so, the residual R-terms, (A5), called cross-scale interaction terms, would be zero or negligible (Davis, 1994b). Otherwise, they represent covariances among variables with length and time scales bridging the averaging scale of the filter, and may not be negligible. However, with the modifications of (A4), (A5), and the like, the forms of the infra-grid averaged scalar concentration equation, (2.1), and the momentum equations, (2.3), are otherwise unchanged.

The infra-grid scalar concentration variance equation (2.6) is also unchanged, if the following generalizations, similar to (A4), are adopted:

$$G_0^C = \tfrac{1}{2}\overline{\tilde{\rho}\tilde{C}^2}/\rho - \tfrac{1}{2}C^2 = \tfrac{1}{2}\overline{\tilde{\rho}C''^2}/\rho + \tfrac{1}{2}R^{CC}, \tag{A6a}$$

$$G_j^{*C} = -\tfrac{1}{2}\overline{\tilde{\rho}\tilde{u}_j\tilde{C}^2}/\rho + \tfrac{1}{2}u_j^*C^2 - B_j^{*C}C + u_j^*G_0^C, \tag{A6b}$$

$$\chi^C = \rho^{-1}\overline{\tilde{\rho}\kappa^C\left(\frac{\partial\tilde{C}}{\partial x_j^*}\right)^2} - \kappa^C\left(\frac{\partial C}{\partial x_j^*}\right)^2 = \rho^{-1}\overline{\tilde{\rho}\kappa^C\left(\frac{\partial C''}{\partial x_j^*}\right)^2} + \kappa^C R^{C_x C_x} \tag{A6c}$$

The cross-scale interaction terms, R^{CC}, $R^{C_x C_x}$, are defined in a manner analogous to (A5). The infra-grid kinetic energy equation (2.7) holds if similar modifications are made to T, T_j^*, Φ.

Though the cross-scale interaction terms in (A4), (A6a,b), may not be negligible, the assessment of the cross-scale contribution to scalar dissipation (A6c) is different. The power spectrum of scalar gradient (in particular, temperature gradient), of which the scalar variance dissipation χ^C (or χ^θ) is constituted, exhibits, when turbulence in the ocean is well enough developed, a universal high-wavenumber form peaked at a short wavelength, $\sim$1–10 mm, determined in part by molecular diffusion coefficients, and joined at longer wavelengths, > 1 m, to an inertial subrange, varying like $\sim$(wavenumber)$^{+1/3}$. This decline with decreasing wavenumber is sufficient to ensure that the cross-scale contributions to (A6c) are negligible if the averaging scale in W is chosen to be, say, $\gtrsim O(1 \text{ m})$. That is, the spectral gap condition holds for the evaluation of dissipation. Similar statements could be made about the power spectrum of velocity shear that constitutes infra-grid kinetic energy dissipation Φ. These remarks establish the dominance of the first terms in the second equality of (A6c) for χ^θ, and of similar contributions to Φ, which are precisely what are given in the main text.

APPENDIX B

Transformation to isentropic coordinates

Consider transformation from the Cartesian coordinate set x_j^*, $j = 0, 1, 2, 3$, to the isentropic coordinate system x_j (see appendix E). Partial derivatives transform as

$$\frac{\partial}{\partial x_i^*} = Q_{ik} \frac{\partial}{\partial x_k}, \quad i = 0, 1, 2, 3 \tag{B1}$$

(summation over repeated indices, $k = 0, 1, 2, 3$, in a monomial term is implied), where

$$(Q_{ik}) = \left(\frac{\partial x_k}{\partial x_i^*} \right) = \begin{pmatrix} 1 & 0 & 0 & -z_{,0}/z_\theta \\ 0 & 1 & 0 & -z_{,1}/z_\theta \\ 0 & 0 & 1 & -z_{,2}/z_\theta \\ 0 & 0 & 0 & 1/z_\theta \end{pmatrix} \equiv \mathbf{Q}. \tag{B2}$$

(The third coordinate $z = x_3^*$ is regarded here as the height of an isentrope θ.) Observe that $\det \mathbf{Q} = 1/z_\theta$, and that the inverse of $\mathbf{Q}$ is

$$\mathbf{Q}^{-1} \equiv \tilde{\mathbf{Q}} \equiv (\tilde{Q}_{ki}) = \left(\frac{\partial x_i^*}{\partial x_k} \right) = \begin{pmatrix} 1 & 0 & 0 & z_{,0} \\ 0 & 1 & 0 & z_{,1} \\ 0 & 0 & 1 & z_{,2} \\ 0 & 0 & 0 & z_{,3} \end{pmatrix}. \tag{B3}$$

Thus

$$\frac{\partial}{\partial x_k} = \tilde{Q}_{ki} \frac{\partial}{\partial x_i^*}, \quad k = 0, 1, 2, 3. \tag{B4}$$

An important application of this transformation establishes the identities

$$V \frac{\partial p}{\partial x_3^*} + g = \frac{\partial H}{\partial x_3^*} - \Pi \frac{\partial \theta}{\partial x_3^*} + g = \frac{1}{z_\theta} \left(\frac{\partial M}{\partial x_3} - \Pi \right), \tag{B5}$$

$$V \frac{\partial p}{\partial x_\alpha^*} = V \frac{\partial p}{\partial x_\alpha} - V \frac{z_{,\alpha}}{z_\theta} \frac{\partial p}{\partial x_3} = \frac{\partial M}{\partial x_\alpha} - z_{,\alpha} \left(g + V \frac{\partial p}{\partial x_3^*} \right)$$
$$= \frac{\partial M}{\partial x_\alpha} - \frac{z_{,\alpha}}{z_\theta} \left(\frac{\partial M}{\partial x_3} - \Pi \right), \tag{B6}$$

where the heat function H, Exner function Π, and Montgomery function M are defined in appendix E.

The material derivative of a variable ϕ transforms invariantly as follows:

$$\frac{D\phi}{Dt} \equiv \frac{\partial \phi}{\partial x_0^*} + u_j^* \frac{\partial \phi}{\partial x_j^*} = \frac{\partial \phi}{\partial x_0} + u_j \frac{\partial \phi}{\partial x_j} \tag{B7}$$

(summation is implied over $j = 1, 2, 3$ only), where

$$u_i = Q_{ki} u_k^* - \delta_{i3} \frac{z_{,0}}{z_\theta}, \quad \text{or } u_1 = u_1^*, u_2 = u_2^*, u_3 = \frac{-z_{,0} - u_1^* z_{,1} - u_2^* z_{,2} + u_3^*}{z_\theta} = \frac{D\theta}{Dt} \equiv \dot{\theta}. \tag{B8}$$

(Observe that in (B8) the transpose of $\mathbf{Q}$ is used in the transformation.) The divergence of velocity transforms according to

$$\frac{\partial u_i^*}{\partial x_i^*} = \frac{\partial u_i}{\partial x_i} + \frac{1}{z_\theta} \frac{Dz_\theta}{Dt} \tag{B9}$$

Thus the infra-grid averaged mass conservation equation [second of (2.2)] becomes

$$\frac{Dh}{Dt} + h \frac{\partial u_i}{\partial x_i} = \frac{\partial h}{\partial x_0} + \frac{\partial}{\partial x_i} (hu_i) = 0, \tag{B10}$$

where $h = \rho z_\theta$. A vector $\mathbf{F}$ other than velocity transforms according to

$$F_i = Q_{ki} F_k^*, \quad \text{or } F_1 = F_1^*, F_2 = F_2^*, F_3 = (-F_1^* z_{,1} - F_2^* z_{,2} + F_3^*)/z_\theta, \tag{B11}$$

so that its divergence transforms as

$$\frac{\partial F_i^*}{\partial x_i^*} = \frac{1}{z_\theta} \frac{\partial (z_\theta F_i)}{\partial x_i} \tag{B12}$$

(cf. (B9)).

Illustrative example. Suppose that

$$B_k^{*C} = K_{kl}^* \frac{\partial C}{\partial x_l^*}. \tag{B13}$$

Then

$$B_i^C = Q_{ki}B_k^{*C} = K_{ij}\frac{\partial C}{\partial x_j},$$ (B14)

where

$$K_{ij} = Q_{ki}K_{kl}^*Q_{lj}; \quad \text{with inverse:} \quad K_{nm}^* = \tilde{Q}_{in}K_{ij}\tilde{Q}_{jm}.$$ (B15)

For example, if $(K_{ij}^*) = \text{diag}(K_H^*, K_H^*, K_V^*)$, then

$$(K_{ij}) = \begin{pmatrix} K_H^* & 0 & -K_H^* z_{,1}/z_\theta \\ 0 & K_H^* & -K_H^* z_{,2}/z_\theta \\ -K_H^* z_{,1}/z_\theta & -K_H^* z_{,2}/z_\theta & K_D/(z_\theta)^2 \end{pmatrix}$$ (B16)

where

$$K_D = K_V^* + K_H^*\{(z_{,1})^2 + (z_{,2})^2\}.$$ (B17)

From (B12),

$$q^C = \frac{1}{\rho}\frac{\partial}{\partial x_i^*}(\rho B_i^{*C}) = \frac{1}{\rho}\frac{\partial}{\partial x_i^*}\left(\rho K_{ij}^*\frac{\partial C}{\partial x_j^*}\right) = \frac{1}{\rho z_\theta}\frac{\partial}{\partial x_i}\left(\rho z_\theta K_{ij}\frac{\partial C}{\partial x_j}\right) = \frac{1}{h}\frac{\partial}{\partial x_i}(hB_i^C).$$ (B18)

In particular, for potential temperature, $C = \theta = x_3$, because $\partial\theta/\partial x_j = \delta_{j3}$,

$$hB_i^\theta = -\rho K_H^* z_{,i} \ (i = 1, 2), \ = \frac{\rho K_D}{z_\theta} \ (i = 3).$$ (B19)

APPENDIX C

Sub-grid diffusion closure for passive scalars

The following account of sub-grid diffusion relative to isentropic coordinates parallels the treatments of Lundgren (1981) and Bennett (1996) relative to fixed coordinates.

Suppose that a Lagrangian trajectory is given in isentropic coordinates by $\mathbf{x} = \mathbf{A}(\mathbf{a}, s|t)$ (see appendix E). (Note that $\mathbf{A}(\mathbf{a}, s|s) = \mathbf{a}$.) A closed form for the integration of the scalar concentration equation (3.3) may then be written

$$C(\mathbf{x}, t) = \int_s^t q^C(\mathbf{A}(\mathbf{x}, t|r), r)dr + C(\mathbf{A}(\mathbf{x}, t|s), s) = \int_{s-0}^t dr \int d\mathbf{y}\Psi(\mathbf{y}, r|\mathbf{x}, t)h(\mathbf{y}, r)q^*(\mathbf{y}, r),$$

(C1)

where

$$\Psi(\mathbf{y}, r|\mathbf{x}, t) = \delta(\mathbf{y} - \mathbf{A}(\mathbf{x}, t|r))/h(\mathbf{y}, r), \quad q^*(\mathbf{y}, r) \equiv q^C(\mathbf{y}, r) + \delta(r - s)C(\mathbf{y}, r).$$ (C2)

(The second of these equations incorporates the initial condition at $r = s$ into the source function q^*.) One may verify that the microcanonical pdf Ψ given by (C2) satisfies

$$\frac{\partial}{\partial t} \Psi(\mathbf{y}, r|\mathbf{x}, t) + u_i(\mathbf{x}, t) \frac{\partial}{\partial x_i} \Psi(\mathbf{y}, r|\mathbf{x}, t) = 0, \tag{C3}$$

with initial condition $\Psi(\mathbf{y}, r|\mathbf{x}, r) = \delta(\mathbf{y} - \mathbf{x})/h(\mathbf{y}, r)$.

The thickness-weighted average of the mass-specific scalar C is indicated with a caret, $\hat{C}$, and is given by

$$\hat{C}(\mathbf{x}, t) = \langle h(\mathbf{x}, t)C(\mathbf{x}, t)\rangle/\langle h(\mathbf{x}, t)\rangle$$

$$= \frac{1}{\langle h(\mathbf{x}, t)\rangle} \int_{s-0}^{t} dr \int d\mathbf{y} \langle h(\mathbf{x}, t)\Psi(\mathbf{y}, r|\mathbf{x}, t)h(\mathbf{y}, r)q^*(\mathbf{y}, r)\rangle$$

$$= \frac{1}{\langle h(\mathbf{x}, t)\rangle} \int_{s-0}^{t} dr \int d\mathbf{y} \langle h(\mathbf{x}, t)\Psi(\mathbf{y}, r|\mathbf{x}, t)\rangle\langle h(\mathbf{y}, r)q^*(\mathbf{y}, r)\rangle \tag{C4}$$

$$= \int_{s-0}^{t} dr \int d\mathbf{y} \hat{\Psi}(\mathbf{y}, r|\mathbf{x}, t)\langle h(\mathbf{y}, r)\rangle\hat{q}^*(\mathbf{y}, r)$$

The third equality in this chain, making use of the general proposition $\langle XY \rangle = \langle X \rangle \langle Y \rangle$, depends on the assumption of statistical independence of the sources q^* (including the initial distribution at $t = s$) of C from the motion trajectories $\mathbf{A}$, in other words, that the scalar C is passive (Davis, 1987).

The thickness-weighted average of eq. (C3) (furnishing the macrocanonical pdf $\hat{\Psi}$) is

$$\frac{\partial}{\partial t} \hat{\Psi}(\mathbf{y}, r|\mathbf{x}, t) + \hat{u}_i(\mathbf{x}, t) \frac{\partial}{\partial x_i} \hat{\Psi}(\mathbf{y}, r|\mathbf{x}, t) = -S_1(\mathbf{y}, r|\mathbf{x}, t), \tag{C5}$$

with initial condition $\hat{\Psi}(\mathbf{y}, r|\mathbf{x}, r) = \delta(\mathbf{y} - \mathbf{x})/\langle h(\mathbf{y}, r)\rangle$, where

$$S_1(\mathbf{y}, r|\mathbf{x}, t) \equiv \frac{1}{\langle h(\mathbf{x}, t)\rangle} \frac{\partial}{\partial x_i} \langle h(\mathbf{x}, t)u_i'(\mathbf{x}, t)\Psi'(\mathbf{y}, r|\mathbf{x}, t)\rangle. \tag{C6}$$

Applying the material derivative operator to (C4), and using (C5) and its initial condition, one obtains

$$\frac{\partial}{\partial t} \hat{C}(\mathbf{x}, t) + \hat{u}_i(\mathbf{x}, t) \frac{\partial}{\partial x_i} \hat{C}(\mathbf{x}, t) = -\int_{s-0}^{t} dr \int d\mathbf{y} S_1(\mathbf{y}, r|\mathbf{x}, t)\langle h(\mathbf{y}, r)\rangle\hat{q}^*(\mathbf{y}, r) + \hat{q}^C(\mathbf{x}, t).$$

$$\tag{C7}$$

In the second term, since one is interested in times $t > s$, $\hat{q}^*$ has been replaced by $\hat{q}^C$. Subtracting (C5) from (C3),

$$\frac{\partial}{\partial t} \Psi'(\mathbf{y}, r|\mathbf{x}, t) + u_i(\mathbf{x}, t) \frac{\partial}{\partial x_i} \Psi'(\mathbf{y}, r|\mathbf{x}, t) = S_1 + S_2, \tag{C8}$$

where

$$S_2(\mathbf{y}, r|\mathbf{x}, t) \equiv -u_i'(\mathbf{x}, t) \frac{\partial}{\partial x_i} \hat{\Psi}(\mathbf{y}, r|\mathbf{x}, t). \tag{C9}$$

A closed form for the integral of (C8) [cf. (C1)] is

$$\Psi'(\mathbf{y}, r|\mathbf{x}, t) \approx \sum_{l=1,2} \int_r^t dw S_l(\mathbf{y}, r|\mathbf{A}(\mathbf{x}, t|w), w), \tag{C10}$$

disregarding an initial condition at $t = r$. Substituting (C10) into the right side of (C6), one obtains after a series of approximations (among other things, only the effect of S_2 in (C10) figures principally),

$$S_1 \approx \frac{-1}{\langle h(\mathbf{x}, t) \rangle} \frac{\partial}{\partial x_i} \left(\langle h(\mathbf{x}, t) \rangle E_{ik}(\mathbf{x}, t|s) \frac{\partial}{\partial x_k} \hat{\Psi}(\mathbf{y}, r|\mathbf{x}, t) \right), \tag{C11}$$

where

$$E_{ik}(\mathbf{x}, t|s) = \frac{1}{\langle h(\mathbf{x}, t) \rangle} \int_s^t \langle h(\mathbf{x}, t) u_i'(\mathbf{x}, t) u_k'(\mathbf{A}(\mathbf{x}, t|w), w) \rangle dw. \tag{C12}$$

This is the diffusivity tensor generated by sub-grid eddy motion covariances (Bennett, 1996). The derivation of (C11), (C12) depends on assuming the quasi-Markovian statistical independence of velocity fluctuations from time-lagged concentration fluctuations (Lundgren, 1981). This assumption, and improvements and modifications thereof, have been critically examined by Bennett (1996). Davis (1987), making a different assumption about the persistence in time of statistical dependence, infers an alternative closure to (C11) which exhibits a "memory" of past concentrations. The derivation of (C12) furnishes a version of the conventional Fickian flux-gradient closure. Whatever the shortcomings of the approximations underlying it, they are at least rationally stated. Some of the elements of the sub-grid diffusivity tensor have been estimated from observations of meso-scale eddy motions (Freeland *et al.,* 1975; Davis, 1985). Others, particularly from the third row and column, which involve dientropic flow, are difficult. In the case of such elements, Eq. (C12) may suggest estimates of their order of magnitude.

Substituting (C11) into (C7), and using (C4), one obtains

$$\frac{D\hat{C}}{Dt} = \frac{\partial \hat{C}}{\partial x_0} + \hat{u}_i \frac{\partial \hat{C}}{\partial x_i} = \frac{1}{\langle h \rangle} \frac{\partial}{\partial x_i} \left(\langle h \rangle (E_{ik} + K_{ik}) \frac{\partial \hat{C}}{\partial x_k} \right), \tag{C13}$$

where the effect of the K_{ik} tensor comes from the averaged infra-grid flux divergence $\hat{q}^C$ [eq. (5.1)]. This may also be written

$$\frac{D\hat{C}}{Dt} = \partial_i^T (D_{ik} \partial_k \hat{C}) = \text{div} \, (\mathbf{D} : \text{grad} \, \hat{C}), \tag{C14}$$

where

$$D_{ik} = \Delta_{im}(E_{mn} + K_{mn})\Delta_{kn}, \quad \Delta_{im} = \text{diag}\,(1, 1, Z_\theta). \tag{C15}$$

The elements of D_{ik} all have the same units: $\text{m}^2\,\text{s}^{-1}$; "div" and all three elements of the "grad" operator have units of m^{-1} (Appendix E). The infra-grid diffusivity tensor K_{mn} contributes significantly only to the (3,3) element of D_{ik}, for which it gives K_D.

APPENDIX D

Transformation to quasi-Cartesian pressure coordinates

The main text describes how to sub-grid-average the equations of motion at fixed $\mathbf{x}$, t, i.e., with respect to isentropic coordinates. It is then useful to consider these averaged equations transformed back to quasi-Cartesian pressure coordinates X_i, $i = 0, 1, 2, 3$ (see appendix E), where $X_3 = P(x_0, x_1, x_2, \theta)$ is the mean pressure height P of isentrope θ. The inverse of this function is $\theta = \theta(X_0, X_1, X_2, X_3)$, potential temperature's dependence on quasi-Cartesian coordinates. The transformations of coordinates may be written

$$\frac{\partial}{\partial X_i} = \bar{Q}_{ik}\frac{\partial}{\partial x_k}, \quad \frac{\partial}{\partial x_k} = \tilde{\bar{Q}}_{ki}\frac{\partial}{\partial X_i}, \qquad i, k = 0, 1, 2, 3 \tag{D1}$$

(cf. (B1), (B4)), where $\bar{\mathbf{Q}}$, $\tilde{\bar{\mathbf{Q}}}$ are the counterparts of (B2), (B3) with $P = X_3$ substituted for $z = x_3^*$. The reverse of the transformations (B5), (B6) shows that

$$\frac{\partial\langle M\rangle}{\partial x_3} - \Pi = \check{V}\frac{\partial P}{\partial x_3} + \Pi\frac{\partial\theta}{\partial x_3} - \Pi + g\frac{\partial Z}{\partial x_3} = \left(\check{V} + g\frac{\partial Z}{\partial X_3}\right)\frac{\partial P}{\partial x_3} = \left(g\frac{\partial Z}{\partial X_3} + \check{V}\right)\frac{1}{\theta_P}, \tag{D2}$$

and

$$\frac{\partial\langle M\rangle}{\partial x_\alpha} = \check{V}\frac{\partial X_3}{\partial x_\alpha} + g\left(\frac{\partial Z}{\partial X_\alpha} + \frac{\partial X_3}{\partial x_\alpha}\frac{\partial Z}{\partial X_3}\right) = g\frac{\partial Z}{\partial X_\alpha} - \frac{\theta_{,\alpha}}{\theta_P}\left(g\frac{\partial Z}{\partial X_3} + \check{V}\right). \tag{D3}$$

The substantial rate of change of $\hat{C}$ following the mean motion $\hat{u}_i$ is invariant,

$$\frac{\hat{D}\hat{C}}{Dt} \equiv \frac{\partial\hat{C}}{\partial x_0} + \hat{u}_j\frac{\partial\hat{C}}{\partial x_j} = \frac{\partial\hat{C}}{\partial X_0} + U_j\frac{\partial\hat{C}}{\partial X_j}, \tag{D4}$$

where $U_1 = \hat{u}_1$, $U_2 = \hat{u}_2$, $U_3 \equiv P_{,0} + \hat{u}_1 P_{,1} + \hat{u}_2 P_{,2} + \hat{\theta}P_{,3} \equiv \dot{P} \equiv \hat{D}P/Dt$, the last being the apparent vertical motion of an isentrope θ. The divergence of mean motion transforms as

$$\frac{\partial\hat{u}_i}{\partial x_i} = \frac{\partial U_i}{\partial X_i} + \frac{1}{\theta_P}\frac{\hat{D}\theta_P}{Dt} = \frac{\partial U_i}{\partial X_i} - \frac{1}{\langle h\rangle}\frac{\hat{D}\langle h\rangle}{Dt}, \tag{D5}$$

where recall that $1/\theta_P = P_\theta = -g\langle h\rangle$. The counterpart of (B10) is

$$\frac{1}{\langle h\rangle}\left\{\frac{\hat{D}\langle h\rangle}{Dt} + \langle h\rangle\frac{\partial\hat{u}_i}{\partial x_i}\right\} = \frac{\partial U_i}{\partial X_i} = 0.$$ (D6)

A vector other than velocity is taken to transform as

$$F_k = \bar{Q}_{ik}\bar{F}_i, \quad \bar{F}_k = \tilde{\bar{Q}}_{ik}F_i,$$ (D7)

where $\bar{F}_k$ are the components in the quasi-Cartesian system. The divergence transforms as

$$\frac{\partial F_i}{\partial x_i} = \frac{1}{\theta_P}\frac{\partial(\theta_P\bar{F}_i)}{\partial X_i}.$$ (D8)

Application: the flux of scalar $\hat{C}$,

$$F_i = \langle h\rangle E_{ij}\frac{\partial\hat{C}}{\partial x_j}$$ (D9)

[cf. (B14)], transforms into

$$\bar{F}_i = \langle h\rangle\bar{E}_{ij}\frac{\partial\hat{C}}{\partial X_j}, \quad \text{where } \bar{E}_{ij} = \tilde{\bar{Q}}_{ki}E_{kl}\tilde{\bar{Q}}_{lj};$$ (D10)

and its divergence becomes

$$\frac{1}{\langle h\rangle}\frac{\partial}{\partial x_i}\left(\langle h\rangle E_{ij}\frac{\partial\hat{C}}{\partial x_j}\right) = \frac{\partial}{\partial X_i}\left(\bar{E}_{ij}\frac{\partial\hat{C}}{\partial X_j}\right).$$ (D11)

The physical units of the elements of the quasi-Cartesian diffusion tensor $\bar{E}_{ij}$ are not all the same, although the elements of the re-scaled tensor $\bar{\Delta}_{mi}\bar{E}_{ij}\bar{\Delta}_{jn}$, where $\bar{\Delta}_{mi} = \text{diag}(1, 1, Z_P = -g^{-1}\check{V})$, all have units of m²s⁻¹.

APPENDIX E

Glossary

A list of most symbols used in the text follows; some symbols that are very common or are only used locally have been omitted. The symbols are listed alphabetically, Latin first, Greek second. A list of operators is given at the end. A brief description, and the number of the equation in which the symbol or operator is first used, are given.

a Thermal expansion coefficient; $= \check{V}^{-1}\partial\check{V}/\partial\theta$. (2.7)

$\mathbf{A}(\mathbf{a}, s|t)$ The Lagrangian "position" (specified by two horizontal coordinates and temperature) of a fluid parcel at time t that is at position $\mathbf{a}$ at time s. (C1)

$\mathbf{A}^*(\mathbf{x}^*, t|w)$ Lagrangian trajectory in conventional Cartesian coordinates. (6.10)

B_j^{*C} Eddy transport of scalar C in a Cartesian reference frame by infra-grid fluctuation covariances (similarly $B_j^{*\theta}$); $= -\bar{\rho}\tilde{u}_j\tilde{C}/\rho + u_j^*C = -\overline{\bar{\rho}u_j''C''}/\rho$.[9] (2.1)

B_j^{C} Eddy transport of scalar C in an isentropic reference frame by infra-grid fluctuation covariances (similarly B_j^{θ}); $B_j^{C} = B_j^{*C}$, $j = 1, 2$; $B_3^{C} = (B_3^{*C} - z_{,\beta}B_\beta^{*C})/z_\theta$ (summation over $\beta = 1,2$). (3.3), (3.6)

$\tilde{C}$ Total scalar concentration (similarly potential temperature $\tilde{\theta}$); $= C + C''$. (A4)

C'' Infra-grid scalar concentration fluctuation (similarly potential temperature θ''); $\overline{\bar{\rho}C''} = 0$.[10] (A4)

C Infra-grid averaged, density-weighted scalar concentration (similarly potential temperature θ); $= \overline{\bar{\rho}\tilde{C}}/\rho = \hat{C} + C'$. (2.1)

$\hat{C}$ Sub-grid averaged, thickness-weighted, scalar concentration; $= \langle hC \rangle/\langle h \rangle$. (4.2)

C' Sub-grid scalar concentration fluctuation; $\langle hC' \rangle = 0$. (4.2)

c Sound speed; $= -\check{V}_P/\check{V}^2$. (4.17)

Cx   Cox number; $= (Z_\theta)^2 \dfrac{\langle h\chi^\theta \rangle}{\kappa^\theta \langle h \rangle} \approx \dfrac{K_D}{\kappa^\theta}$. (5.7)

D   Total eddy diffusion tensor. (6.1)

D_{ij} The total diffusivity tensor, expressed relative to isentropic coordinates, generated by sub-grid and infra-grid eddy motions; elements of **D**. (6.2)

D_H Horizontal (along-isentrope) sub-grid eddy diffusivity. (6.2)

D_V Vertical (across-isentrope) sub-grid eddy diffusivity. (6.2)

$\bar{D}_{ij}$ The total diffusivity tensor, expressed relative to quasi-Cartesian coordinates, generated by sub-grid and infra-grid eddy motions. (7.2)

E_{ik} The diffusivity tensor, expressed relative to isentropic coordinates, generated by sub-grid eddy motion covariances. (C11)

$\bar{E}_{ij}$ The diffusivity tensor, expressed relative to quasi-Cartesian coordinates, generated by sub-grid eddy motion covariances. (7.2)

G_0^{C} Infra-grid averaged scalar concentration variance (similarly G_0^θ); $= \frac{1}{2}\overline{\bar{\rho}\tilde{C}^2}/\rho - \frac{1}{2}C^2 = \frac{1}{2}\overline{\bar{\rho}C''^2}/\rho$.[9] (2.6)

G_j^{*C} Infra-grid eddy transport of scalar concentration variance relative to a Cartesian reference frame (similarly $G_j^{*\theta}$); $= -\frac{1}{2}\overline{\bar{\rho}\tilde{u}_j\tilde{C}^2}/\rho + \frac{1}{2}u_j^*C^2 - B_j^{*C}C + u_j^*G_0^{C} = -\frac{1}{2}\overline{\bar{\rho}u_j''C''^2}/\rho$.[9] (2.6)

G_j^{C} Infra-grid eddy transport of scalar concentration variance relative to an isentropic reference frame; related to G_j^{*C} as B_j^{C} is related to B_j^{*C} (similarly G_j^θ). (3.7)

h Specific thickness of infra-grid averaged isentropic layers; $= \rho z_\theta = -p_\theta/g = \langle h \rangle + h'$; units: kg m^{-2} K^{-1}. (3.1)

---

9. There are extra terms after the second equality caused by cross-scale interaction terms; they are given in appendix A.

10. More generally, the practical average of this fluctuation need not vanish (appendix A).

$\langle h \rangle$ Sub-grid averaged specific thickness of isentropic layers; $= \hat{V}^{-1} Z_\theta = -P_\theta / g$. (4.1)

h' Sub-grid fluctuation of isentropic specific thickness; $\langle h' \rangle = 0$. (4.8)

H Heat content function, $= \int_0^P \check{V}(p, \theta) dp$. (B5)

$\langle H \rangle$ Sub-grid averaged heat function. (4.9)

K_V^* Vertical diffusion coefficient due to infra-grid eddy fluctuations. (5.2)

K_D Dientropic diffusion coefficient due to infra-grid eddy fluctuations. (5.1)

K_H^* Horizontal diffusion coefficient due to infra-grid eddy fluctuations. (5.1)

K_{ik}^g Sub-grid eddy diffusivity tensor obtained from averaging with respect to fixed coordinates. (6.9)

$\mathbf{K}^g$ Sub-grid eddy diffusivity matrix from averaging with respect to fixed coordinates; $= K_{ik}^g$. (6.10)′

$\mathbf{K}^I$ Sub-grid eddy diffusivity matrix, calculated wholly with respect to fixed coordinates, and rotated to local isentropic coordinates. (6.11)

K_I Principal along-isentropic eddy diffusivity coefficient calculated with respect to fixed coordinates. (6.11)

K_{ij} Diffusivity tensor due to infra-grid eddy fluctuations, relative to isentropic coordinates. (5.1)

K_{kl}^* Diffusivity tensor due to infra-grid eddy fluctuations, relative to Cartesian coordinates. (6.9)

L_H Horizontal (along-isentrope) sub-grid mixing length. (6.2)

L_V Vertical (across-isentrope) sub-grid mixing length. (6.2)

M Montgomery function, $= H + gz = \langle M \rangle + M'$. (3.4)

$\langle M \rangle$ Sub-grid averaged Montgomery function; $= \langle H \rangle + gZ$. (4.6)

M' Sub-grid Montgomery function fluctuation; $\langle M' \rangle = 0$. (4.8)

M_{jk}^* Eddy transport of momentum in a Cartesian reference frame by infra-grid velocity fluctuation covariances; $= -\tilde{\rho} \tilde{u}_j \tilde{u}_k / \rho + u_j^* u_k^* = -\tilde{\rho} \overline{u_j'' u_k''} / \rho$.[9] (2.3)

M_{jk} Eddy transport of momentum in an isentropic reference frame by infra-grid velocity fluctuation covariances; $M_{j\alpha} = M_{j\alpha}^*, j = 1, 2$; $M_{3\alpha} = (M_{3\alpha}^* - z_{,\beta} M_{\beta\alpha}^*)/z_\theta$. (3.4)

$N(p)$ Typical buoyancy frequency profile, as a function of pressure. (4.12)

p Infra-grid averaged pressure (*not* density weighted); $= P + p'$. (2.3)

P Sub-grid averaged pressure, or pressure depth (in decibars, say) of isentrope θ; $= \langle p \rangle$ (*not* thickness weighted). (4.10)

P_θ $= \dfrac{\partial P}{\partial \theta}$. (4.1)

p' Sub-grid pressure fluctuation; $\langle p' \rangle = 0$. (4.10)

$\langle p'^2 \rangle$ Sub-grid pressure variance on an isentrope θ. (4.10)

p_0 Mid-depth rms pressure fluctuation of an isentrope θ; ≈ 100 dbar. (4.12)

P   Sub-grid averaged production of infra-grid eddy kinetic energy; $= \left\langle h M_{ji} \dfrac{\partial u_i}{\partial x_j} \right\rangle.$ (5.9)

p_θ $= p_{,3} = \dfrac{\partial p}{\partial \theta}.$ (3.2)

q^C Infra-grid average creation rate of scalar concentration by eddy transport divergence (similarly q^θ); $= \dfrac{1}{\rho} \dfrac{\partial}{\partial x_j^*}(\rho B_j^{*C}) = \dfrac{1}{h} \dfrac{\partial}{\partial x_j}(h B_j^C)$ (see formula (B12)). (2.1)

q_α^M Infra-grid average creation rate of horizontal momentum ($\alpha = 1, 2$) by eddy transport divergence; $= \dfrac{1}{\rho} \dfrac{\partial}{\partial x_j^*}(\rho M_{j\alpha}^*) = \dfrac{1}{h} \dfrac{\partial}{\partial x_j}(h M_{j\alpha}).$ (2.3)

$\hat{q}^C$ Thickness-weighted sub-grid average of q^C (similarly $\hat{q}^\theta$); $= \dfrac{1}{\langle h \rangle} \dfrac{\partial}{\partial x_j} \langle h B_j^C \rangle.$ (4.4)

$\hat{q}_\alpha^M$ Thickness-weighted sub-grid average of q_α^M; $= \dfrac{1}{\langle h \rangle} \dfrac{\partial}{\partial x_j} \langle h M_{j\alpha} \rangle.$ (4.6)

Q_{ik} Transformation matrix between Cartesian (x_i^*) and isentropic (x_k) coordinates; $= \dfrac{\partial x_k}{\partial x_i^*}.$ (B1)

Q   Matrix of Q_{ik} elements. (B2)

$\tilde{\mathbf{Q}}$ Inverse of **Q**; $= (\tilde{Q}_{ki}) = \left(\dfrac{\partial x_i^*}{\partial x_k} \right).$ (B3)

$\bar{Q}_{ik}$ Transformation matrix between quasi-Cartesian (X_i) and isentropic (x_k) coordinates; $= \dfrac{\partial x_k}{\partial X_i}.$ (D1)

$\bar{\mathbf{Q}}$ Matrix of $\bar{Q}_{ik}$ elements. (D1)

$\tilde{\bar{\mathbf{Q}}}$ Inverse of $\bar{\mathbf{Q}}$; $= (\tilde{\bar{Q}}_{ki}) = \left(\dfrac{\partial X_i}{\partial x_k} \right).$ (D1)

$R^{\phi\psi}$ Residual cross-scale contribution in addition to mass-weighted infra-grid covariance $\overline{\tilde{\rho} \phi'' \psi''}/\rho.$ (A4)

R   Rotation matrix from Cartesian to local isentropic coordinates, sub-grid averaged with respect to fixed-level coordinates. (6.11)

Rf   Flux Richardson number; $= \dfrac{ag \langle h B_3^{*\theta} \rangle}{P}.$ (5.11)

t Time; $= x_0^* = x_0 = X_0.$ (2.1)

T Infra-grid eddy kinetic energy; $= \frac{1}{2} \overline{\tilde{\rho} \tilde{u}_k^2}/\rho - \frac{1}{2} u_k^{*2} = \frac{1}{2} \overline{\tilde{\rho} u_k''^2}/\rho.$[9] (2.7)

T_j^* Infra-grid eddy transport of kinetic energy relative to a Cartesian reference frame; $= -\frac{1}{2} \overline{\tilde{\rho} \tilde{u}_j \tilde{u}_k^2}/\rho + \frac{1}{2} u_j^* u_k^{*2} - M_{jk}^* u_k^* + u_j^* T - \overline{p \tilde{u}_j} + p u_j^* = -\overline{u_j''(\frac{1}{2} u_k''^2 + p - \tilde{p})}/\rho.$[9] (2.7)

T_j Infra-grid eddy transport of kinetic energy relative to an isentropic reference frame; related to T_j^* as B_j^C is related to B_j^{*C}. (3.9)

$\tilde{u}_j$ Total Cartesian velocity components; $= u_j^* + u_j''.$ (2.2)

u_j^* Infra-grid averaged, density-weighted Cartesian velocity components in the directions indicated by the suffix j; $= \overline{\tilde{\rho}\tilde{u}_j}/\rho$. (2.2)

u_j'' Infra-grid velocity fluctuations; $\overline{\tilde{\rho}u_j''} = 0$.[10]

u_j Infra-grid averaged generalized velocity components; $u_\alpha = u_\alpha^*$ ($\alpha = 1$, 2), $u_3 = \dot{\theta} = \dfrac{D\theta}{Dt}$; $u_j = \hat{u}_j + u_j'$. (3.1)

$\hat{u}_j$ Sub-grid averaged, thickness-weighted, generalized velocity; $= \langle hu_j \rangle/\langle h \rangle$. (4.2)

U_j Quasi-Cartesian, sub-grid averaged, velocity; $U_\alpha = \hat{u}_\alpha$ ($\alpha = 1, 2$), $U_3 = \dot{P} = \hat{D}P/Dt$. (7.1)

u_j' Sub-grid generalized velocity fluctuations; $\langle hu_j' \rangle = 0$. (4.2)

u_β^+ Horizontal velocity bias due to thickness-Montgomery function gradient covariance (Gent and McWilliams, 1990); $= f^{-1}\varepsilon_{3\alpha\beta}\dfrac{1}{\langle h \rangle}\left\langle h'\dfrac{\partial M'}{\partial x_\alpha}\right\rangle$. (4.6)

$u_j'^*$ Sub-grid eddy velocity fluctuations with respect to fixed coordinates. (6.10)

$\mathbf{u}'^*$ $= (u_1'^* \quad u_2'^* \quad u_3'^*)$. (6.10)′

$\bar{V}$ Specific volume; $= \tilde{\rho}^{-1} = \check{V}(\tilde{p}, \tilde{\theta})$.

$\check{V}(\tilde{p}, \tilde{\theta})$ The functional dependence of specific volume on pressure and potential temperature, i.e., the equation of state. (2.5)

V Infra-grid averaged, density-weighted, specific volume; $= \overline{\tilde{\rho}\tilde{V}}/\rho = \rho^{-1}$. (2.5)

$\hat{V}$ Sub-grid averaged, thickness-weighted, specific volume; $= \langle hV \rangle/\langle h \rangle$. (4.1)

ΔV Typical specific volume scale. (4.11)

W A practical space-time filter kernel for an infra-grid average. (A1)

v_i' Isentropic sub-grid velocity fluctuation components, all with units of m s^{-1}; $= \Delta_{il}u_l'$. (6.2)

x_j^* Cartesian space coordinates, horizontal ($j = 1, 2$), vertical ($j = 3$); and time ($j = 0$), i.e., $t = x_0^*$. (2.1)

$\mathbf{x}^*$ $= (x_1^* \quad x_2^* \quad x_3^*)$. (6.10)

x_j Isentropic coordinates; $x_\alpha = x_\alpha^*$ ($\alpha = 1, 2$), $x_3 = \theta$, $x_0 = t$. (3.1)

$\mathbf{x}$ $= (x_1 \quad x_2 \quad x_3)$. (C1)

X_j Quasi-Cartesian coordinates; $X_\alpha = x_\alpha = x_\alpha^*$ ($\alpha = 1, 2$), $X_3 = P$, $X_0 = t$, where P is the depth (in pressure units, decibars) of an infra-grid averaged isentrope θ. (7.1)

z Vertical coordinate, or height of infra-grid averaged isentrope θ; $= x_3^*$. (3.2)

$z_{,j}$ $= \dfrac{\partial z}{\partial x_j}$. (B2)

z_θ $= z_{,3} = \dfrac{\partial z}{\partial \theta}$. (3.2)

$\langle z'^2 \rangle$ Sub-grid height variance of an isentrope θ. (4.17)

Z Sub-grid averaged height of isentrope θ; $= \langle z \rangle$ (*not* thickness weighted). (4.9)

$Z_{,j}$ $= \dfrac{\partial Z}{\partial x_j}$. (5.1)

Z_θ $= Z_{,3} = \dfrac{\partial Z}{\partial \theta}$. (4.1)

ΔZ Error in sub-grid averaged isentrope height. (4.17)

δ $= \sqrt{D_V/D_H}$. (6.2)

Δ_{kj} Diagonal scale matrix; $= \mathrm{diag}\,(1, 1, Z_\theta)$. (C15)

ε $= K_D/D_H$. (6.2)

ε_{3ji} $= 1$ if $ji = 1, 2$; $= -1$ if $ji = 2, 1$; $= 0$ otherwise. (2.3)

κ^θ Molecular diffusion coefficient for temperature (κ^C for scalar C). (5.7)

μ Reference thermobaric parameter; $= \dfrac{\check{V}_{p\theta}}{\check{V}_\theta}\big|_{p=0,\theta=0}$. (4.11)

χ^C Molecular dissipation of scalar concentration variance; $= \overline{\tilde{\rho}\kappa^C(\partial C''/\partial x_j^*)^2}/\rho$. (2.6)

χ^θ Molecular dissipation of temperature variance; $= \overline{\tilde{\rho}\kappa^\theta(\partial\theta''/\partial x_j^*)^2}/\rho$. (3.8)

ρ Infra-grid averaged density; $= \bar{\tilde{\rho}}$. (2.1)

$\tilde{\rho}$ Total density; $= \rho + \rho'' = \check{V}(\tilde{p}, \tilde{\theta})^{-1}$; $\overline{\rho''} = 0$.[11] (2.2)

σ Typical thermocline *e*-folding scale; $\approx (1000\ \mathrm{dbar})^{-1}$. (4.13)

σ_i Rms sub-grid velocities; $= \sqrt{\langle h v_i'^2 \rangle/\langle h \rangle}$. (6.2)

θ Infra-grid averaged potential temperature (entropy); $= \overline{\tilde{\rho}\tilde{\theta}}/\rho$. (2.5)

$\dot{\theta}$ Infra-grid averaged potential temperature rate of change; $= D\theta/Dt$. (3.6)

$\hat{\theta}$ Sub-grid averaged, thickness-weighted, potential temperature rate of change; $= \hat{u}_3 = \hat{D}\theta/Dt$. (4.3)

$\dot{\theta}'$ $= u_3'$. (6.4)

Π Exner function; $= \dfrac{\partial H(p, \theta)}{\partial\theta}$. (3.5)

$\langle\Pi\rangle$ Sub-grid averaged Exner function. (4.7)

Φ Molecular dissipation of kinetic energy; $= \{\frac{1}{2}\overline{\tilde{\rho}v(\partial u_i''/\partial x_j^* + \partial u_i''/\partial x_j^*)^2} - \frac{2}{3}\overline{\tilde{\rho}v(\partial u_i''/\partial x_i^*)^2}\}/\rho$. (2.7)

Ψ Microcanonical probability density function (pdf); $= \hat{\Psi} + \Psi'$. (C1)

$\hat{\Psi}$ Macrocanonical probability density function; $= \langle h\Psi \rangle/\langle h \rangle$. (C4)

Ψ' Probability density function fluctuation; $\langle h\Psi' \rangle = 0$. (C6)

Operations

$\overline{(\dots)}$ Infra-grid average, at fixed x_j^*, t (fixed height z). (A1)

$\langle\dots\rangle$ Sub-grid average, at fixed x_j, t (fixed isentrope θ). (4.1)

$\overline{\tilde{\rho}(\dots)}/\rho$ Density-weighted infra-grid average.

$\langle h(\dots)\rangle/\langle h\rangle$ Thickness-weighted sub-grid average.

div   (6.1)

grad   (6.1)

$\dfrac{DC}{Dt}$ Substantial derivative, following infra-grid averaged motion, of an infra-grid averaged scalar or vector component; $= \dfrac{\partial C}{\partial x_0^*} + u_j^* \dfrac{\partial C}{\partial x_j^*} = \dfrac{\partial C}{\partial x_0}$ $+ u_j \dfrac{\partial C}{\partial x_j}$. (Here, as throughout the main text, repeated indices, such as j, in a monomial term indicate summation over 1,2,3.) (2.1)

$\dfrac{\hat{D}\hat{C}}{Dt}$ Substantial derivative, following sub-grid averaged motion, of a sub-grid averaged scalar or vector component; $= \dfrac{\partial \hat{C}}{\partial x_0} + \hat{u}_j \dfrac{\partial \hat{C}}{\partial x_j} = \dfrac{\partial \hat{C}}{\partial X_0} + U_j \dfrac{\partial \hat{C}}{\partial X_j}$. (4.4)

$\dfrac{\partial}{\partial x_j^*}$ Partial derivative taken with x_k^*, $k \neq j$, fixed. (2.1)

$\dfrac{\partial}{\partial x_j}$ Partial derivative taken with x_k, $k \neq j$, fixed. (3.1)

∂_i $= \dfrac{\partial}{\partial x_i}$, $i = 1, 2$; $= \dfrac{1}{Z_\theta} \dfrac{\partial}{\partial \theta}$, $i = 3$ (grad). (C14)

∂_i^T $= \dfrac{1}{\langle h \rangle} \dfrac{\partial}{\partial x_i} (\langle h \rangle \ \cdot\ )$, $i = 1, 2$; $= \dfrac{\hat{V}}{Z_\theta} \dfrac{\partial}{\partial \theta} (\hat{V}^{-1} \ \cdot\ )$, $i = 3$ (div). (C14)

$\dfrac{\partial}{\partial X_j}$ Partial derivative taken with X_k, $k \neq j$, fixed. (7.1)

$\dfrac{\partial}{\partial \theta}$ $= \dfrac{\partial}{\partial x_3}$. (5.1)

REFERENCES

Bennett, A. F. 1996. Particle displacements in inhomogeneous turbulence, *in* Stochastic Modelling in Physical Oceanography, R. J. Adler, P. Müller, B. L. Rosovskii, eds., Birkhäuser, 1–45.

Davis, R. E. 1985. Drifter observations of coastal surface dynamics during CODE: The statistical and dynamical views. J. Geophys. Res., *90,* 4756–4772.

—— 1987. Modeling eddy transports of passive tracers. J. Mar. Res., *45,* 43–64.

—— 1994a. Diapycnal mixing in the ocean: Equations for large-scale budgets. J. Phys. Oceanogr., *24,* 777–800.

—— 1994b. Diapycnal mixing in the ocean: The Osborn-Cox model. J. Phys. Oceanogr., *24,* 2560–2576.

—— 1996. Sampling turbulent dissipation. J. Phys. Oceanogr., *26,* 341–358.

de Szoeke, R. A. and A. F. Bennett. 1993. Microstructure fluxes across density interfaces. J. Phys. Oceanogr., *23,* 2254–2264.

de Szoeke, R. A. and R. M. Samelson. 2002. The duality between the Boussinesq and non-Boussinesq hydrostatic equations of motion. J. Phys. Oceanogr., *32,* 2194–2203.

Freeland, H. J., P. B. Rhines and T. Rossby. 1975. Statistical observations of the trajectories of neutrally buoyant floats in the North Atlantic. J. Mar. Res., *33,* 383–404.

Fuglister, F. C. 1960. Atlantic Ocean Atlas of Temperature and Salinity Profiles and Data from the

International Geophysical Year 1957–1958. Woods Hole Oceanographic Institution Atlas Series, *1,* 1–209.

Gargett, A. E. and G. Holloway. 1984. Dissipation and diffusion by internal wave breaking. J. Mar. Res., *42,* 15–27.

Gent, P. R. and J. C. McWilliams. 1990. Isopycnal mixing in ocean circulation models. J. Phys. Oceanogr., *20,* 150–155.

Greatbatch, R. J. and T. J. McDougall. 2003. The non-Boussinesq temporal-residual-mean. J. Phys. Oceanogr., *33,* 1231–1239.

Gregg, M. C. 1987. Diapycnal mixing in the thermocline. J. Geophys. Res., *92,* 9686–9698.

—— 1989. Scaling turbulent dissipation in the thermocline: A review. J. Geophys. Res., *94,* 5249–5286.

—— 1999. Uncertainties in measuring ε and χ_t. J. Atmos. Oceanic Tech., *16,* 1483–1490.

Klymak, J. M. and J. D. Nash. 2009. Estimates of mixing, *in* Encyclopedia of Ocean Sciences, J. H. Steele, K. K. Turekian and S. A. Thorpe, eds., Academic Press, 3696–3706.

Ledwell, J. R., A. J. Watson and C. S. Law. 1993. Evidence for slow mixing across the pycnocline from an open-ocean tracer-release experiment. Nature, *364,* 701–703.

Levitus, S. 1982. Climatological Atlas of the World Ocean. NOAA Prof. Paper *13,* U.S. Govt. Printing Office, 173 pp.

Lozier, M. S., M. S. McCartney and W. B. Owens. 1994. Anomalous anomalies in averaged hydrographic data. J. Phys. Oceanogr., *24,* 2624–2638.

Lundgren, T. 1981. Turbulent pair dispersion and scalar diffusion. J. Fluid Mech., *111,* 25–57.

McDougall, T. J. 2003. Potential Enthalpy: A conservative oceanic variable for evaluating heat content and heat fluxes. J. Phys. Oceanogr., *33,* 945–963.

McDougall, T. J. and J. A. Church. 1986. Pitfalls with the numerical representation of isopycnal and diapycnal mixing. J. Phys. Oceanogr., *15,* 1542–1556.

McDougall, T. J. and P. C. McIntosh. 1996. The temporal-residual-mean velocity. Part I: Derivation and the scalar conservation equations. J. Phys. Oceanogr., *26,* 2653–2665.

—— 2001. The temporal-residual-mean velocity. Part II: Isopycnal interpretation and the tracer and momentum equations. J. Phys. Oceanogr., *31,* 1222–1246.

Monin, A. S. and A. M. Yaglom. 1971. Statistical Fluid Mechanics, Vol. 1, M.I.T., 769 pp.

Montgomery, R. B. 1938. Circulation in upper layers of southern North Atlantic deduced with use of isentropic analysis. Paps. in Phys. Oceanogr. and Meteorol., *6,* 55 pp.

Osborn, T. R. and C. S. Cox. 1972. Oceanic fine structure. Geophys. Fluid Dyn., *3,* 321–345.

Polzin, K. L., J. M. Toole. J. R. Ledwell and R. W. Schmitt. 1997. Spatial variability of turbulent mixing in the abyssal ocean. Science, *276,* 93–96.

Redi, M. H. 1982. Oceanic isopycnal mixing by coordinate rotation. J. Phys. Oceanogr., *12,* 1154–1158.

Talley, L. D. 2007. Hydrographic Atlas of the World Ocean Circulation Experiment(WOCE). Vol. 2: Pacific Ocean, M. Sparrow, P. Chapman and J. Gould, eds., International WOCE Project Office, Southampton, U.K.

Veronis, G. 1975. The role of models in tracer studies, *in* Numerical Models of Ocean Circulation, Natl. Acad. Sci., 133–146.