FROM CLASSICAL TO STATISTICAL OCEAN DYNAMICS

GREG HOLLOWAY
Institute of Ocean Sciences, 9860W Saarich Road, Sidney, British Columbia, V8L 4B2, Canada
E-mail: g@planetwater.ca

(Received 9 January 2004; Accepted 4 March 2004)

Abstract. Traditional ocean modeling treats fields resolved on the model grid according to the classical dynamics of continua. Variability on smaller scales is included through sundry "eddy viscosities", "mixing coefficients" and other schemes. In this paper we develop an alternative approach based on statistical dynamics. First, we recognize that we treat probabilities of flows, not the flows themselves. Modeled dependent variables are the moments (expectations) of the probabilities of possible flows. Second, we address the challenge to obtain the equations of motion for the moments of probable flows rather than the (traditional) equations for explicit flows. For linear terms and on larger resolved scales, the statistical equations agree with classical dynamics where those of traditional modeling works well. Differences arise where traditional modeling would relegate unresolved motion to "eddy viscosity", etc.. Instead, changes of entropy ($<-\log P>$ over the probability distribution of possible flows) with respect to the modeled moments act as forcings upon those moments. In this way we obtain a consistent framework for specifying the terms which, traditionally, represent subgridscale effects. Although these statistical equations are close to the classical equations in many ways, important differences are also evident; here, two phenomena are described where the results differ. We consider eddies interacting with bottom topography. It is seen that traditional "eddy viscosity" and/or "topographic drag", which would reduce large scale flows toward rest, are wrong. The second law of thermodynamics is violated; the "arrow of time" is running backwards! From statistical dynamics, approximate corrections are obtained, yielding a practical improvement to the fidelity of ocean models. Another phenomenon occurs at much smaller scales in the turbulent mixing of heat and salt. Even when both heat and salt are stably stratifying, their rates of turbulent transfer should differ. This suggests a further model improvement.

Keywords: eddies, entropy, mixing, ocean models, probability, statistical dynamics

1. Oceans, lakes and (most) duck ponds are too big

Advances in computing capacity along with advances in numerical methods provide ever greater power to solve for flows in oceans, lakes and duck ponds. In principle we suppose that we know a good approximation to the equations of motion on some scale, e.g., the Navier–Stokes equations coupled with heat and salt balances under gravity and rotation. In practice we cannot solve for oceans, lakes or most duck ponds on the scales for which

Surveys in Geophysics **25:** 203–219, 2004.

these equations apply. For example, the length scales over which ocean salinity varies are often shorter than 1 mm. We try to solve for fields represented on grids (or other bases) that are far larger than the scales to which "known" equations apply. Then we are compelled to guess the equations of motion.

Guessing equations is uncomfortable, often causing us to assume without question the equations used by some previous author. When we are brave, we realize that this too is uncomfortable. It is natural to wish that, as computers grow ever more powerful, we guess less and less. What would be needed from the computer? In the oceans there are about 1.36×10^{18} m^3 of water. If we felt that variability was unimportant within volumes of $O(10^{-9}\ m^3)$ then the computer should track $O(10^{27})$ volumes, each described by several degrees of freedom. Clearly one can fiddle these numbers. Today's "big computer models", e.g., for weather forecasting or turbulence research, may advance 10^7, 10^8 or 10^9 variables. Over time we are assured that computers will become bigger yet. Even if we imagine computer models advancing 10^{12} variables (not on my desk in my lifetime!), we still face the situation that, for each one variable we track, we must guess how that variable interacts with 10^{15} variables about which we are uninformed. Limiting ourselves to coastal oceans or lakes, the mismatch in degrees of freedom might reduce to 10^8 or less. Might computations for a suitably modest duck pond – someday – be possible? Maybe.

This enterprise is like seeking to reinvent the steam engine from molecular dynamics' simulation of water vapour. What a brave, but bizarre, thing to attempt! For oceans, lakes and ponds the circumstance is even worse than the dismal numbers above suggest. As computing power opens more explicit degrees of freedom, solutions exhibit chaos in the sense that arbitrarily nearby solutions at some time become entirely uncorrelated at a later time. This invites ensemble forecasting to allow "meaningful" results such as means, variances, most probable results, etc.. We are led to ask what is an appropriate size of ensemble. 3? No. Perhaps 10^3? Maybe.

The point of scary big numbers is: for the numerical ocean, lake or duck pond to be soluble, huge numbers of degrees of freedom must be somehow "slaved" to a much smaller number. Indeed, we suppose that this is true. But the keyword is "somehow", and that is the motivation for this paper.

2. Geophysical fluid dynamics acquires an "arrow of time"

Our goal is to represent subgridscale variability, often seen in eddy viscosity or mixing parameterizations. A wealth of experience in such matters has shown, arguably, greater and lesser success. Numerical ocean models run, and outputs are not conspicuously absurd (sometimes). Successes have depended in parts on choices of eddy viscosities and mixing schemes, drawn from history of trial-and-error tuning. Our purpose here is not to re-tune traditional eddy viscosities, but rather to focus upon underlying fundamentals which can lead to results strikingly different from traditional schemes.

Ocean modeling rests upon geophysical fluid dynamics (GFD), a branch of classical mechanics applied to continua under gravity and rotation. Apart from the role of viscosity or mixing (which are "added on" to GFD without derivation), GFD inherits from classical mechanics a marvelous geometric quality about space and time. Past and future in t are not fundamentally different from left or right in x. An evolution from past to future is no more usual than from future to past (bearing in mind that the Earth's rotation reverses for a time reversal experiment). But real flows do not evolve "backwards" from future to past. This is true in global oceans, lakes, duck ponds and tea cups. (Milk stirred into tea turns brown. Reversing the stirring does not reseparate the milk from the tea.)

How should GFD learn about time's arrow from past to future? The question is at the heart of moving from classical to statistical mechanics, e.g., Coveney and Highfield (1992). Two key ideas come to mind: probability and entropy.

Symbolically, let $\mathbf{y}$ denote a vector in phase space (of dimension 10^{27}, or whatever), governed by $\mathrm{d}\mathbf{y}/\mathrm{d}t = \mathbf{F}(\mathbf{y}) + \mathbf{G}$. $\mathbf{y}$ is simply the collection of all variables needed to describe the ocean, lake or pond fully while $\mathbf{F}$ and $\mathbf{G}$ express terms in the forced equations of motion. Depending upon one's interest, $\mathbf{y}$ may include biological and chemical components. Computation of the highly erratic (presumably chaotic) $\mathbf{y}(t)$ in 10^{27} elements is not feasible, or even desirable. We consider instead the probability $p(\mathbf{y})\mathrm{d}y$ for actual "$\mathbf{y}$" within a phase volume $\mathrm{d}y$ of any $\mathbf{y}$. Then we would seek an equation for $\mathrm{d}p(\mathbf{y})/\mathrm{d}t = \ldots$ But this is even less feasible than attempting to solve $\mathrm{d}\mathbf{y}/\mathrm{d}t$.

One sometimes solves "toy" stochastic models, such as a Fokker–Planck equation $\partial_t p = -\partial_y J = -\partial_y A(y,t)p + \frac{1}{2}\partial_y\partial_y B(y,t)p$, where J is a flux of probability and A and B are modeled after some properties estimated from F and G. For realistic ocean modeling, the actual derivations of A and B from F and G are not feasible (to my knowledge), and Fokker–Planck equations are explored as intuition-building models.

Practical progress results from reducing the problem of $p(\mathbf{y})$ to that of seeking a chosen set of moments $\mathbf{Y} = \int \mathbf{y}\mathrm{d}p$ or, generally, $\mathbf{K} = \int \mathbf{k}(\mathbf{y})\mathrm{d}p$ for functions $\mathbf{k}(\mathbf{y})$. An advantage is that one may choose as few $\mathbf{Y}$ as one likes, perhaps 10^8 or perhaps only 10, "integrating out" as much of $\mathbf{y}$ as one wishes. The outstanding problem is that we do not have equations for $\mathrm{d}\mathbf{Y}/\mathrm{d}t = \ldots$ Progress is made by carrying out probability averaging (expectation) over the components of $\mathbf{y}$ retained in $\mathbf{Y}$, yielding

$$\frac{\mathrm{d}\mathbf{Y}}{\mathrm{d}t} = \mathbf{F}'(\mathbf{Y}) + \mathbf{G}' + \cdots \tag{1}$$

where $\mathbf{F}'(\mathbf{Y})$ includes linear dynamics and couples explicitly the retained components of $\mathbf{Y}$ while discarding the myriad couplings among $\mathbf{Y}$ and 10^{27} $\mathbf{y}$. $\mathbf{G}'$ are the expectations of $\mathbf{G}$ for the retained $\mathbf{Y}$. Eq. (1) expresses traditional GFD symbolically with ... being supplied from *ad hoc* mixing. The manifest successes of GFD are due largely to circumstances near forced linear dynamics that are fully captured by $\mathbf{F}'$ and $\mathbf{G}'$.

Suggesting that the "textbook equations of GFD" are wrong often meets with resistance. The books have been in use for so long that the "work" of GFD turns to applications and the properties of these equations, not to questioning the equations. How could such well-used equations be "wrong"? The problem, I suggest, has been that too little attention has been paid to the dependent variables which we mean to describe. Viewing the output of numerical models, it is easy to see maps of "flows", or "temperatures", or whatever. Pausing for a moment, we retreat to something like "grid-cell and time-step averaged flow", etc.. Yet if we try to define space-time filtering operators to pass over underlying equations, it is deeply problematic to obtain traditional GFD. The enterprise is far more troubled if we advance to modern computing with an eddy-active, and presumably chaotic, output. Then what meaning do we attach to "temperature", say, in some cell over some step in some specific (chance!) realization? Perhaps what we should mean is "temperature moment over probabilities of possible states". However, if we then asked: what is the equation of motion of the temperature moment over probabilities of possible states?, we would be far more ready to suppose that a textbook equation for temperature might not be right.

Apart from ..., Eq. (1) expresses the marvelous symmetry between past and future inherited from classical mechanics. Time is without its arrow, bringing us to our second key idea: entropy. Entropy is most familiar in thermodynamic context, often as $\mathrm{d}H = \mathrm{d}Q/T$, where $\mathrm{d}H$ is the increment of entropy resulting from the supply of an increment of heat $\mathrm{d}Q$ to a system with constant volume at temperature T. Entropy has also entered common language as an allusion to disorder. But we should be careful to be precise; for a process defined by a probability function p, entropy is

$$H = -\int \mathrm{d}p \log(p) \tag{2}$$

to within an arbitrary multiplicative factor and an arbitrary additive constant. Describing probabilistically the field of molecular chaos, H from Eq.

(2) is the thermodynamic entropy $dH = dQ/T$. In another context, describing probabilistically the likelihoods of transmitted messages, H from Eq. (2) is the entropy of information theory. While the "log" in Eq. (2) is considered to be the natural $\log_e$, this is sometimes $\log_2$ for the information theory application, rendering units of H in bits. The equivalence of thermodynamic H and information theoretical H is discussed in texts such as Katz (1967). However, the duality of entropy under statistical physics and information theory has not been exploited in the context of oceans, e.g., relating dynamics and state estimation.

Finally, the arrow of time arrives with the Second Law of Thermodynamics. For "internal" or "free" interactions (i.e., interactions not subject to external influence), p evolves such that, for $t_2 > t_1$, $H_2 \geq H_1$. We are left to instruct GFD that ... in Eq. (1) should provide this property – appropriately!

3. Equations of motion for moments of probable flows

The alert reader already asks: do not traditional eddy viscosity or mixing (as ...) break time reversal symmetry? They do. The challenge is that traditional schemes, while irreversible, may not satisfy $H_2 \geq H_1$ for $t_2 > t_1$. In many cases, traditional schemes, guided by intuition and experience, are adequate. Milk stirred into tea turns brown and eddy diffusion accomplishes such irreversible browning. H for the milk–tea mix is increased after browning. If intuition and experience worked this reliably, the present paper would be unnecessary and we could instead invest effort refining our estimates of eddy diffusion coefficients. But intuition, and GFD, go very wrong as two illustrations (eddy–topography interaction and stably stratified mixing) will show. For now we continue developing the equations of motion.

There have been two approaches which we consider briefly. In these approaches the aim is not to complete Eq. (1) itself but rather to characterize solutions of Eq. (1). The approach by Paltridge (1975, 1978) hypothesized that the mean state (**Y**) of the atmosphere is such as to maximize the production of (thermodynamic) entropy. This has been further advanced by Ozawa and Ohmura (1997) in the context of the Earth's atmosphere, and shown to be a plausible descriptor for other planetary atmospheres by Lorenz et al. (2001). Of particular note for ourselves are (1) the role of thermodynamic entropy, absent in other work below, and (2) a principle of maximum entropy production (MEP) to which we shall return.

The second approach to solutions of Eq. (1) resulting from ... is seen in Salmon et al. (1976) and in a host of papers since, some referenced by Holloway (1986) and Salmon (1998). In these cases the mean states of idealized oceans are described by the maximizing of entropy due to uncertain macroscale eddies. These studies have focused upon expectations for potential vorticity $q = (\Omega + \nabla \times \mathbf{u}) \bullet \nabla\rho$, where Ω is the Earth's angular frequency of rotation, $\nabla \times \mathbf{u}$ is vorticity, and $\nabla\rho$ is density gradient. Calculations assume no external forcing and no internal dissipation. Thus the macroscale flow field does not communicate with the field of molecular chaos, and thermodynamic entropy is not included. For such unforced, non-dissipative, ideal dynamics, H is maximized subject to conserved integrals of the motion. In early studies, conservation constraints included domain-integrated energy, potential vorticity (q) and enstrophy (q^2). Subsequent studies by Miller (1990), Robert and Sommeria (1991) or Robert and Rosier (1997) have considered the roles of further invariants derived from the advection of potential vorticity. Despite such idealizations, many aspects of these solutions seemed to capture realistic features. However, the construction of such solutions does not yet provide the missing ... in Eq. (1). For later reference, we denote these maximum entropy (ME) solutions as **Y***.

In many cases, our concern is not only to characterize stationary, or equilibrium, solutions to Eq. (1), such as addressed by MEP or ME, but rather to ask how the solutions to Eq. (1) evolve from the assigned initial conditions and under assigned external forcing. For this we need to complete the equations of motion, i.e., to represent ... in Eq. (1).

Two paths to ... have been opened up. One path is seen in the recent research by Chavanis and Sommeria (1997), Kazantsev et al. (1998) and Polyakov (2001) based upon MEP for subgridscale potential vorticity. The promise and challenges along this path are not yet clear. For the present I refer the interested reader to the cited sources while we here turn to another path for which there is more experience to date. It will be seen that the two approaches have much in common. Indeed the recent work of Polyakov (2001) also finds quantitative results with much in common.

The other path, sketched below and described by Holloway (2003), follows the "generalized thermodynamic force" (GTF) after Onsager (1931a,b) or Onsager and Machlup (1953). If entropy depends upon some set of macroscopic parameters, **X** (e.g., the expectations **Y**), then a forcing acts upon **X** due to the gradient of H with respect to **X**, i.e., $\mathbf{C} \bullet \nabla_{\mathbf{X}} H$ where **C** is a coefficient tensor projecting $\nabla_{\mathbf{X}} H$ onto $\mathrm{d}\mathbf{X}/\mathrm{d}t$. As an illustration consider a gas having different concentrations in two chambers separated by a membrane, with X as the fraction of total gas in one chamber. The derivatives of H with respect to available volume (per molecule) give thermodynamic pressures in the two chambers, yielding nonzero $\mathrm{d}H/\mathrm{d}X$ when the pressures are not equal. $\mathrm{d}H/\mathrm{d}X$ can force $\mathrm{d}X/\mathrm{d}t$ depending upon the extent of perforations in the membrane, etc., described by C. Importantly, there are two parts to the GTF – $\nabla_{\mathbf{X}} H$ and **C** – which, together provide the ... in Eq. (1).

$$\frac{\mathrm{d}\mathbf{Y}}{\mathrm{d}t} = \mathbf{F}'(\mathbf{Y}) + \mathbf{G}' + \mathbf{C} \bullet \nabla_{\mathbf{Y}} H \,. \tag{3}$$

Clearly, GTF and MEP have very much in common. Each seeks to drive **Y** as rapidly as possible toward higher H. The physics challenge is focussed upon "as rapidly as possible", considering both the availability and skill of different methods of estimation. We return to Eq. (3) in Section 3.3.

3.1. Illustration: eddies and topography, the wrong way

Consider the interaction among eddies and mean flows in basins with complex topography. First let us consider a wrong answer, the answer provided by nearly every major ocean model run at every major oceanographic research institution on Earth. Because it can be difficult to assess "wrong" in realistic circumstances of many complicated, poorly known inputs while outputs are compared with limited data, let us instead pose a thought experiment. Suppose that we only observe the ocean on the larger scales amenable to numerical modeling, perhaps some tens km to 100 km. Suppose that on these scales we find our ocean to be at rest, motionless with flat density surfaces. Although we cannot see them, we are aware that on smaller scales the ocean is filled with ubiquitous eddies. Maybe the eddies arise from previous episodes of forcing, or maybe they are driven by smaller scale forcing which we cannot see, perhaps by enthusiastic goldfish. Suppose that on the large scale we observe that there is no imposed forcing. We wish to predict the future ocean on the scales which we can see. We poll the major ocean models worldwide, asking: if the ocean is stably at rest and no forcing acts, what is the future? With extraordinary unanimity across different models, the answer is: nothing. That unanimous answer is absolutely wrong.

We can test the answer within the same models. For a given model, we obtain a bigger computer, allowing the resolution on the eddy-active scales not previously seen. From our awareness that small scale eddies exist, we randomly excite the newly realized small scales. The test remains the same as before: on larger scales the ocean is at rest, and no large scale forcing acts. We run our newer, higher resolution model to compare with the previous prediction that no large scale flow will occur. We find instead (as the reader with computer access – or a friend – can check) that large scale flows emerge with a definite sense (moving with shallow water to the right in the northern hemisphere).

How did (nearly) all the models at all the major institutions (including those advising governments about climate, etc.) fail this test? What the models did, based on intuition, experience and simply getting the models to run, was to replace ... with some manner of eddy viscosity, perhaps even of the fancier, iterated-Laplacian sort. When those eddy viscosities saw fields of no motion, they took no action. If there was some slight motion, it would be damped anyway. That was wrong. While eddy viscosity did serve to break time symmetry in Eq. (1), it broke time symmetry the wrong way, driving H the wrong way (to be shown below). Thus, traditional GFD, with traditional eddy viscosities, violates the Second Law of Thermodynamics, assuring the wrong answers.

Could enough computing power resolve the eddies? The challenge is: how much is enough? While modern models are often termed "eddy-resolving", the more apt term is "eddy-admitting". That is, the resolution is sufficiently fine, allowing explicit damping terms that are sufficiently weak, that model dynamics support internal instabilities and that admit eddies. But eddies are only dynamically "resolved" when a further increase in resolution does not lead to systematic changes. Importantly, it is the feedback of eddies upon a larger scale mean flow which (I speculate) is most difficult to achieve from refined resolution. Ultimately – in principle – we may suppose that we shall have computers capable of approaching the molecular dynamics simulation of steam engines or even duck ponds. Our aim in this paper is to seek another way.

3.2. Illustration: eddies and topography, the hard way

That ocean models would fail the eddy–topography test (above) has been known theoretically for nearly three decades, after a comprehensive theory set out by Herring (1977) following spectral-based statistical closure after Kraichnan (1959). A simpler spectral closure theory by Holloway (1978), after Kraichnan (1971), was consistent with the results from Herring (1977). Despite efforts to simplify, these closure theory calculations are hugely difficult, and the brave reader is referred to the cited references.

Briefly we recall only the relevant aspects from the simpler calculations made by Holloway (1978). To render the problem tractable, barotropic quasigeostrophic (QG) dynamics were considered. Potential vorticity $q = (2\Omega + \nabla \times \mathbf{u}) \bullet \nabla\rho$ is approximated as $q = \zeta + h$, where ζ is the vertical component of QG vorticity $\nabla \times \mathbf{u}$, and $h = f\delta h/h_o$ represents the variation δh of total depth, h_o is a constant reference depth, and f is a constant vertical component of 2Ω. The fluid is considered to be of uniform ρ, with ∇_ρ replaced by the inverse of total depth.

Supposing that fluctuations of ζ and h are spatially statistically homogenous, the evolution of $< \zeta\zeta >$ and $< \zeta h >$ in the spectral domain are predicted for assumed statistics of $<hh>$, where $<>$ denotes probability expectation. Compared with direct numerical simulations, such closure theories showed reasonable skill. Moreover, considering an ensemble of realizations of ζ for a given realization of h, theory easily showed that $< \zeta\zeta >$ confined initially to small scales would readily force nonzero $< \zeta > = < \zeta h > /h$ on all scales for which $h \neq 0$. The thought experiment posed in Section 3.1 was sure to fail.

Eddy–topography closure was taken a little farther in Holloway (1987) to include a nonzero, spatially uniform flow U, and admit the simple $\mathrm{f} = \mathrm{f_o} + \beta \mathrm{y}$. This allowed the calculation of pressure-topography "form drag" with a dynamically responding $\mathrm{d}U/\mathrm{d}t$. If, in addition to external forcing applied to U, there were assumed to be sources of eddy energy (e.g., stochastic wind forcing), then the pressure-topography forces could systematically propel U. Use of the term "form drag" has been largely replaced by "form stress" or "topographic stress" to recognize that this force may not simply retard the mean flow but may also force the mean flow. Numerical experiments confirmed that, with mean wind forcing applied to U of sign $U > 0$, the response tended toward $U < 0$. This counter-intuitive result gave rise to the label "neptune effect" (whimsically suggesting that the phenomenon was otherwise inexplicable).

Although some of the above-mentioned results are satisfying theoretically, they are very limited from a practical perspective. A large amount of tedious calculation is needed to obtain results which are restricted to a statistically homogenous, barotropic QG flow. So what? Clues to a way forward were buried in Holloway (1978), and then made powerfully clear in a key paper by Carnevale et al. (1981) who showed that for entire classes of closure theory after Kraichnan (1959, 1971) the theories strictly assured nonlinear interaction terms yielding $\mathrm{d}H/\mathrm{d}t \geq 0$, driving the system monotonically toward the ME solution $\mathbf{Y}^*$. (Other terms due to external forcing or dissipation could decrease H, preventing $\mathbf{Y}^*$.) Thus the outcome of closure theories was to show in detail how the dynamics drove systems from any $\mathbf{Y}$ toward $\mathbf{Y}^*$. Can we use this property to motivate highly simplified approximations to closure theory?

3.3. Illustration: eddies and topography, the easy way

Apparently the roles of external forces and internal dissipation are to drive realistic $\mathbf{Y}$ away from $\mathbf{Y}^*$, inducing an entropy gradient $\nabla_{\mathbf{Y}} H$ which, if unchecked, would force $\mathbf{Y}$ toward $\mathbf{Y}^*$. In terms of GTF this can be viewed as a Taylor expansion $\mathbf{C} \bullet \nabla_{\mathrm{Y}} H \approx \mathbf{C} \bullet \nabla_{\mathrm{Y}} \nabla_{\mathrm{Y}} H \bullet (\mathbf{Y} - \mathbf{Y}^*)$ about $\mathbf{Y} = \mathbf{Y}^*$, where $\nabla_{\mathrm{Y}} H = 0$. More carefully we would admit that we do not know if typical $\mathbf{Y}$ are sufficiently close to $\mathbf{Y}^*$, and we would be daunted seeking to estimate $\mathbf{C} \bullet \nabla_Y \nabla_Y H$. But a simple scheme emerges. Denoting $\mathbf{K} \equiv \mathbf{C} \bullet \nabla_Y \nabla_Y H$, Eq. (3) becomes

$$\mathrm{d}\mathbf{Y}/\mathrm{d}t = \mathbf{F}'(\mathbf{Y}) + \mathbf{G}' + \mathbf{K} \bullet (\mathbf{Y} - \mathbf{Y}^*) \tag{4}$$

For application we need **K** and **Y***. **Y*** can be inferred after Salmon et al. (1976) who found at ME a relation between QG streamfunction, ψ, defined by $\nabla^2\psi = \zeta$, and topography, h, viz. $\psi = L^2(\nabla^2\psi + h)$. Here L^2, occurring as a ratio of Lagrange multipliers in the maximization of H, has units of (length)^2 related to coherence scales in the eddy vorticity field. If our scales of interest (model resolved) are significantly larger than L, then the relation for ψ simplifies further to $\psi \approx L^2 h$. However, this is still based upon QG for which $h = f\delta h/h_o$ requires that $|\delta h/h_o| \ll 1$, contrary to the actual ocean whose depth varies by the full depth itself.

For implementation into realistic ocean models, ambiguities involve both h_o and the interpretation of ψ as velocity or transport (integral of **u** over ocean depth D). These questions were considered in Holloway (1992), and then implemented in Alvarez et al. (1994) by taking a ME transport streamfunction to be $\boldsymbol{\Psi}^* = -fL^2 D$, from which the barotropic component of ME velocity **u*** is $\mathbf{u}^* D = z \times \nabla\boldsymbol{\Psi}^*$. The remaining question is how to represent **K**, presumed to be a scale-dependent operator governing the rate at which eddy interactions can force **Y** toward **Y***. Simple choices (with a view towards practicality) include $-K$ (a damping constant) or $A\nabla^2$ (Laplacian diffusion). In practice the choice has been to assign momentum tendency as $A\nabla^2(\mathbf{u}-\mathrm{u}^*)$, with a constant coefficient A.

It must be clear that the choices discussed in the previous paragraph are not strictly derived. They were practical choices made at a time (c. 1992) offering advances over the common practice of eddy viscosity, cf. $\mathrm{A}\nabla^2\mathbf{u}$. Substituting $\mathrm{A}\nabla^2(\mathbf{u}-\mathbf{u}^*)$, with **u*** from $\mathbf{u}^* D = z \times \nabla\boldsymbol{\Psi}^*$, where $\boldsymbol{\Psi}^* = -fL^2 D$ became known as the "neptune parameterization", with L a length scale presumed to take values from some few km to several km.

While the preceding discussion addresses lateral momentum transfers, it is important also to consider vertical transfers. In the case of simplest neptune, u* is independent of depth, and vertical transfer acts to reduce shear similarly to traditional viscosity. However, whereas vertical viscosity in traditional models is assumed to take values perhaps a few tens greater than vertical tracer diffusion, quasigeotrophic neptune scaling implies vertical transfer coefficients of order f^2/N^2 times lateral transfer coefficients, hence very much greater than is traditionally assumed. As well, with baroclinic extension of neptune to depth-dependent **u***, vertical momentum transfers may induce, rather then reduce, vertical shear.

Several papers have explored applications of this for cases ranging from estuarine to global ocean (Alvarez et al., 1994; Eby and Holloway, 1994;

Fyfe and Marinone, 1995; Holloway et al., 1995; Pal and Holloway, 1996; Sou et al., 1996; England and Holloway, 1998; Marinone, 1998; Nazarenko et al., 1998). Two items of particular interest are (1) a study of the impact of neptune parameterization upon global skill (Holloway and Sou, 1996) measured against current meter records, and (2) the Arctic Ocean study by Nazarenko et al. (1998) compared with MEP calculations by Polyakov (2001).

The point of simple (easy!) schemes like neptune is to allow present-day ocean models to produce more skillful results from physics that is closer to statistical dynamics. Such simple schemes are not "right"; they are only "less wrong" than traditional models. This should spur further efforts. Among such efforts, Holloway (1997) considered baroclinic extension, including "thickness" transports in layer models. Merryfield (1998) considered ME QG with continuous stratification. Frederiksen (1999) carefully re-examined closure theory to see how to evaluate terms such as **K** that are only guessed in neptune applications. Merryfield et al. (2001) extended ME without assuming QG, helping to overcome the ambiguity stemming from QG forms of h and ψ, proposing $\Psi^* = fL^2D_o^2/D$ with reference depth D_o. Such efforts, both to establish fundamentals and to devise practical parameterizations, will continue, presumably conjoined by newer work such as MEP.

3.4. Illustration: mixing heat and salt in bi-stably stratified flow

Before closing, we turn to a very different phenomenon on an entirely different scale. In part we seek to test how robust the statistical mechanical approaches are. As well, we could ask: if the statistical mechanical apparatus were brought to bear only to aid the eddy topography problem, maybe that is not so worthwhile. For other circumstances would simpler intuitions suffice?

The stratification of sea water is due to amounts of heat and salt scaled by corresponding density coefficients. At the level of molecular conductivity, sea water is about 100 times as diffusive for heat as for salt. This leads to interesting effects. If the water column is stably stratified with respect to temperature, T, but unstably with respect to salinity, S, while the overall density stratification remains stable one may encounter spontaneous instabilities, called "salt fingers". Contrariwise, if the situation is stable with respect to S but unstable with respect to T, there are instabilities called "layering". Much of the ocean interior is, however, stably stratified with respect both to T and S. Then it is assumed that heat and salt behave similarly and that ambient turbulence, perhaps on account of internal wave breaking, mix the two with a single apparent diffusivity, κ. It has been known for some time that this is not true.

In early experiments, Turner (1968) mechanically agitated a fluid stratified with respect to T and, separately, a fluid stratified with respect to S, taking care that the two stratifications were initially the same. Under the same mechanical agitation, it was found that T was mixed more efficiently than S by an amount greater than could attributed to molecular conduction. Later, Altman and Gargett (1980) performed similar experiments in a tank which was stably stratified with respect to both T and S, arranged so that both made the same initial contribution to stratification. Again it was seen that T mixed more readily than S by amounts exceeding molecular conduction. This is termed "differential diffusion". The observation of differential diffusion in the ocean is technically difficult, with results being reported by Nash and Moum (2002). Although the phenomenon is observed, why does it happen?

We try to use our intuition and experience. Omitting gravity at first, turbulent stirring of a (passive) tracer is like the milk-into-tea example. Heat and salt would be stirred down their respective background gradients, with the fluxes in the vertical ($w'T'$ and $w'S'$) being dominated by scales of turbulent energy, with lesser contributions at smaller scales (subject to viscous and diffusive cutoffs). Because molecular diffusion of T is faster than of S, the short scales of $w'T'$ are suppressed more strongly than for $w'S'$. Hence, we might expect the total turbulent salt transport to be greater than the heat transport, the opposite of what is observed.

Including gravity, what changes? Mainly the turbulence is suppressed, exhausting its energy by working against gravity (in traditional thinking). But that is only to say that we expect a weaker version of the stirring-milk-into-tea example, and the previous (wrong) result that salt transport should exceed heat transport is still expected. Where did we go wrong?

Numerical simulations at first in 2-D (motion only in a vertical plane) by Merryfield et al. (1998), and then fully in 3-D by Gargett et al. (2002), reveal what happens. Contributions to $w'T'$ and $w'S'$ reverse sign to become counter-gradient at the shorter scales. When larger molecular diffusion of T preferentially cuts off $w'T'$, the surviving $w'S'$ fluxes are of the counter-gradient sense and, hence, subtract from the overall down-gradient S flux. The result is that the S flux is weaker than the T flux, as observed.

While numerical simulations showed what happens, we are left asking: why? Especially why are counter-gradient (backwards!) fluxes prevalent at shorter scales? I think that such counter-gradient transports are generic to stably stratified turbulence. Closure theory (Holloway, 1988) in 2-D (vertical plane) anticipated very well (quantitatively) the 2-D simulations and (qualitatively) the 3-D simulations. Importantly, that closure theory is of the broad class which, *per* Carnevale et al. (1981), strictly satisfies $dH/dt \geq 0$ (apart from external forcing and dissipation). Then entropy gradient forcing, or GTF, explains each aspect of what happens.

At larger scales (relative to molecular diffusive cutoff scales), a source of turbulent kinetic energy (KE) is assumed. Under gravity, two things happen. In part, KE is redistributed to smaller scales due to entropy gain from such a redistribution (a GTF view of a "turbulent cascade") In part, KE is converted to potential energy (PE) stored in T'^2 and S'^2 (as density variances times gravity), driving the ratio of KE to PE toward higher entropy (*per* GTF). Large scale conversion of KE to PE is by down-gradient w′T′ and w′S′. Large scale T'^2 and S'^2 are also redistributed to shorter scales by another GTF cascade. Recalling that GTF works as $\mathbf{C} \bullet \nabla_Y H$, involving $\mathbf{C}$ as well as $\nabla_\mathbf{Y} H$, the C for redistribution of KE is weaker than for redistribution of PE. This is due to the role of pressure forces maintaining the incompressibility $\nabla \bullet \mathbf{u} = 0$ for a vector field **u**, whereas the scalar fields T and S are unrestricted. A consequence is that PE is more rapidly transferred from large to small scales, causing the ratio of KE to PE on small scales to favor entropy production by converting PE to KE, forced by $\partial H/\partial(\mathrm{KE/PE})$. Conversion of PE to KE are by counter-gradient $\mathrm{w}'T'$ and $\mathrm{w}'S'$, whence stronger diffusion of T' leads to an overall stronger $\mathrm{w}'T'$.

Figure 1 presents idealised wavenumber spectra to indicate what is going on. On the left, larger scale motions are dominantly inertio-gravity waves and vortical modes. In the middle, wave instabilities have led to overturn-

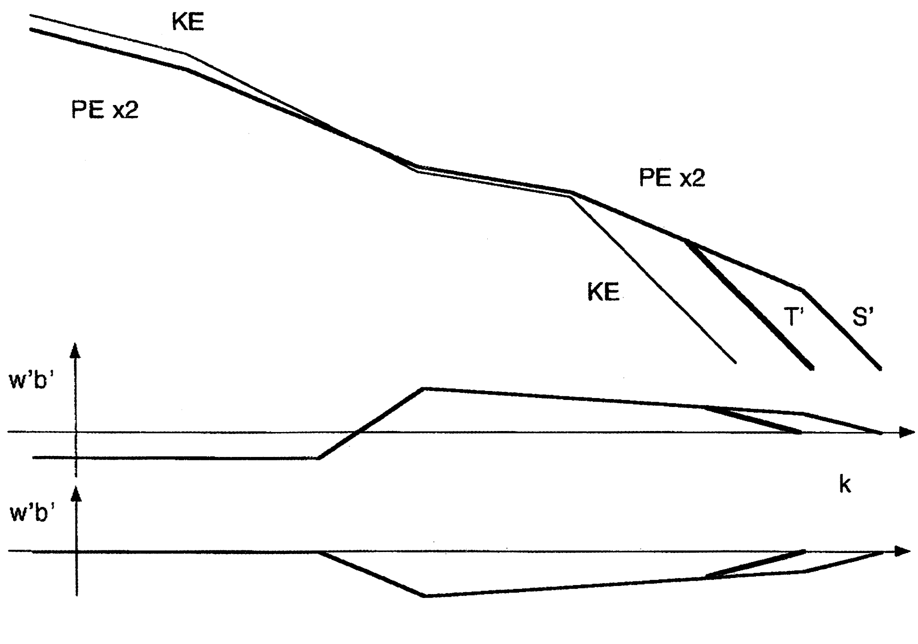

Figure 1. This schematic diagram depicts vertical wavenumber spectra of kinetic energy (KE) and potential energy (PE).

ing and turbulence. On the right, fluctuations decay. The PE spectrum is scaled by a factor of two because, at a maximum of entropy in any isolated mode, KE = 2 PE. Larger scales are characterized by an excess of KE over 2 PE. Then, because PE cascades more effectively, the balance at smaller scales shifts to more than 2 PE over KE.

Approaching dissipation scales, KE decreases due to viscosity. At somewhat smaller scales, the thermal contribution (T') to PE decreases due to thermal conductivity, while the salinity contribution (S') persists to even smaller scales until limited by ionic conduction.

These differences between KE and 2 PE set up the entropy gradient GTF which forces the vertical buoyancy flux (correlation of w' with buoyancy b', the fluctuation of density defect due to T' and S'). On larger scales, excess KE forces $w'b' < 0$. This is downward (downgradient) buoyancy transport. From scales of overturning and turbulence through dissipation, the excess of 2 PE over KE forces reversed $w'b' > 0$. This is upwards (upgradient, or restratifying) transport. It seems strange: the sense of transport has turbulence unmixing the ocean! At smaller scales, thermal conductivity first cuts off the T' part of $w'b'$. Because $w'b'$ is upgradient in smaller scales, the effect of cutting off T' is to leave a larger overall downgradient thermal flux while, overall, the downgradient salinity flux is reduced by its more persistent upgradient portion. This is exactly opposite to what would happen if T and S were passive tracers (i.e., in absence of gravity).

At the bottom of the diagram is sketched a more traditional view. In this case, the larger, more wavelike motions are believed to sustain no systematic buoyancy flux. The onset of overturning and turbulence should yield downgradient fluxes cut off at smaller dissipation scales. Were this to be the case, with thermal conductivity cutting off the thermal flux sooner, the overall thermal flux would be less than the salinity flux, contrary to what is observed.

4. Outlook

Broadly, entropy calculus helps to clarify and organize the work that must be done so that traditional ocean models, based on classical mechanics plus *ad hoc* mixing, may acquire a consistent arrow of time. In some cases, traditional mixings and eddy viscosities happen (willy-nilly?) to point the right way. Too often such guessed-at schemes point oppositely to time's arrow, and then go quite wrong. No surprise! The illustrations are drawn from two extremes: (1) the characteristics of ocean currents of scales of tens and hundreds km, and (2) the nature of mixing on scales of cm to mm, and smaller. In each case the approaches from entropy calculus have been sketched, though these are barely sketches. Very little intellectual resource has yet been invested on the statistical dynamical side compared with the investment on the classical (traditional) ocean dynamics side. Among the tasks ahead are: (1) an ongoing effort to build confident fundamentals, and (2) brave efforts to bring statistical dynamics into practical ocean modeling even while further efforts are being made to build and refine the fundamentals. Practical steps will be understood as steps along the way, in place only until they can be superceded. But the sweep of what may be done, with prospects for improvement both in understanding and in practical skill, are powerful motivations for the tasks ahead.

Acknowledgements

I am grateful to Joel Sommeria and Peter Davies for organizing this Grand Combin Summer School and for inviting me to present a lecture. The character of the work, and thoughts expressed above, were inspired from interactions with students and faculty throughout the school period. The research upon which this lecture was based has been supported by the Office of Naval Research.

References

Altman, D.B. and Gargett, A.E.: 1990, 'Differential Property Transport due to Incomplete Mixing in a Stratified Fluid', in E. List and G. Jirka (eds.), Stratified Flows, Amer. Soc. Civil Engr.

Alvarez, A., Tintore, J., Holloway, G., Eby, M. and Beckers, J.M.: 1994, Effect of Topographic Stress on the Circulation in the Western Mediterranean. *J. Geophys. Res.* **99**, 16053–16064.

Carnevale, G.F., Frisch, U., and Salmon, R.: 1981, 'H-theorems in Statistical Fluid Dynamics', *J. Phys. A*. **14**, 1701–1718.

Chavanis, P.H. and Sommeria, J.: 1997, 'Thermodynamical Approach for Small-scale Parameterization in Two-dimensional Turbulence', *Phys. Rev. Lett.* **78**, 3302–3305.

Coveney, P. and Highfield, R.: 1992, '*The Arrow of Time*. Fawcett-Columbine, New York, 378 pp.

Eby, M. and Holloway, G.: 1994, 'Sensitivity of a Large Scale Ocean Model to a Parameterization of Topographic Stress', *J Phys. Oceanogr.* **24**, 2577–2588.

England, M.H. and Holloway, G.: 1998, 'Simulations of CFC-11 Outflow and Seawater Age in the Deep North Atlantic', *J. Geophys. Res.* **103**, 15885–15901.

Frederiksen, J.S.: 1999, 'On Subgrid Scale Parameterization of Eddy-Topographic Force, Eddy Viscosity and Stochastic Backscatter for Flow over Topography', *J. Atmos. Sci.* **56**, 1481–1494.

Fyfe, J. and Marinone, G.: 1995, 'On the Role of Unresolved Eddies in a Model of the Residual Currents in the Central Strait of Georgia, B.C, *Atmos.-Ocean*. **33**, 613–619.

Gargett, A.E., Merryfield, W.J. and Holloway, G.: 2003, 'Direct Numerical Simulation of Differential Scalar Diffusion in Three-Dimensional Stratified Turbulence', *J. Phys. Oceanogr.* **33**, 1758–1782.

Herring, J.R.: 1977, 'Two-dimensional Topographic Turbulence', *J. Atmos. Sci.* **34**, 1731–1750.

Holloway, G.: 1978, 'A Spectral Theory of Nonlinear Barotropic Motion above Irregular Topography', *J. Phys. Oceanogr.* **8**, 414–427.

Holloway, G.: 1986, 'Eddies, Waves, Circulation and Mixing: Statistical Geofluid Mechanics', *Ann. Rev. Fluid Mech.* **18**, 91–147.

Holloway, G.: 1987, 'Systematic Forcing of Large-Scale Geophysical Flows by Eddy–Topography Interaction', *J. Fluid Mech.* **184**, 463–476.

Holloway, G.: 1988, 'The Buoyancy Flux from Internal Gravity Wave Breaking', *Dyn. Atmos. Oceans.* **12**, 107–125.

Holloway, G.: 1992, 'Representing Topographic Stress for Large Scale Ocean Models', *J. Phys. Oceanogr.* **22**, 1033–1046.

Holloway, G., Sou, T., and Eby, M.: 1995, 'Dynamics of Circulation of the Japan Sea', *J. Mar. Res.* **53**, 539–569.

Holloway, G. and Sou, T.: 1996, 'Measuring Skill of a Topographic Stress Parameterization in a Large-scale Ocean Model', *J. Phys. Oceanogr.* **26**, 1088–1092.

Holloway, G.: 1997, 'Eddy Transport of Thickness and Momentum in Layer and Level Models', *J. Phys. Oceanogr.* **27**, 1153–1157.

Holloway, G.: 2003, 'Toward a Statistical Ocean Dynamics', pp. 277–288, in *Statistical Theories and Computational Approaches to Turbulence*, Y. Kaneda and T. Gotoh (eds.), Springer, 409 pp.

Kazantsev, E., Sommeria J. and Verron, J.: 1998, 'Subgridscale Eddy Parameterization by Statistical Mechanics in a Barotropic Ocean Model', *J. Phys. Oceanogr.* **28**, 1017–1042.

Katz, A.: 1967, *Principles of Statistical Mechanics*. W.H. Freeman & Co., San Francisco, 188 pp.

Kraichnan, R.: 1959, 'The Structure of Isotropic Turbulence at Very High Reynolds Numbers, *J. Fluid Mech.* **5**, 497–543.

Kraichnan, R.: 1971, 'An Almost-Markovian Galilean-Invariant Turbulence Model', *J. Fluid Mech.* **47**, 512–524.

Lorenz, R., Lunine, J.I., Withers P.G. and McKay, C.P.: 2001, 'Titan, Mars and Earth: Entropy Production by Latitudinal Heat Transport', *Geophys. Res. Lett.* **98**, 415–418.

Marinone, S.G.: 1998, 'Effect of the Topographic Stress on the Tidal and Wind Induced Residual Currents in the Gulf of California', *J. Geophys. Res.* **103**, 18437–18446.

Merryfield, W.J. and Holloway, G.: 1997, 'Topographic Stress Parameterization in a Quasigeostrophic Barotropic Model', *J. Fluid Mech.* **341**, 1–18.

Merryfield, W.J.: 1998, 'Effects of Stratification on Quasigeostrophic Inviscid Equilibria', *J. Fluid Mech.* **354**, 345–356.

Merryfield, W.J., Holloway G. and Gargett, A.E.: 1998, 'Differential Vertical Transport of Heat and Salt by Weak Stratified Turbulence', *Geophys. Res. Lett.* **25**, 2773–2776.

Merryfield, W.J., Cummins P.F. and Holloway, G.: 2001, 'Equilibrium Statistical Mechanics of Barotropic Flow Over Finite Topography', *J. Phys. Oceanogr.* **31**, 1880–1890.

Miller, J.: 1990, 'Statistical Mechanics of Euler Equations in Two Dimensions', *Phys. Rev. Lett.* **65**, 2137–2140.

Nash, J.D. and Moum, J.N.: 2002, 'Microstructure Estimates of Turbulent Salinity Flux and the Dissipation Spectrum of Salinity', *J. Phys. Oceanogr.* **32**, 2312–2333.

Nazarenko, L., Holloway G., and Tausnev, N.: 1998, 'Dynamics of Transport of 'Atlantic signature' in the Arctic Ocean', *J. Geophys. Res.* **103**, 31003–31015.

Onsager, L.: 1931a, 'Reciprocal Relations in Irreversible Processes', I., *Phys. Rev.* **37**, 405–426.

Onsager, L.: 1931b, 'Reciprocal Relations in Irreversible Processes', II., *Phys. Rev.* **38**, 2265–2279.

Onsager, L. and Machlup, S.: 1953, 'Fluctuations and Irreversible Processes', *Phys. Rev.* **91**, 1505–1515.

Ozawa, H. and Ohmura, A.: 1997, 'Thermodynamics of a Global Mean State of the Atmosphere – A State of Maximum Entropy Increase', *J. Climate* **10**, 441–445.

Pal, B.K. and Holloway, G.: 1996, 'Dynamics of Circulation off the Westcoast of Vancouver Island', *Cont. Shelf Res.* **16**, 1591–1607.

Paltridge, G.W.: 1975, 'Global Dynamics and Climate – A System of Minimum Entropy Exchange', *Quart. J.R. Met. Soc.* **101**, 475–484.

Paltridge, G.W.: 1978, 'The Steady-State Format of Global Climate', *Quart. J.R. Met. Soc.* **104**, 927–945.

Polyakov, I.: 2001, 'An Eddy Parameterization Based on Maximum Entropy Production with Application to Modeling of the Arctic Ocean Circulation', *J. Phys. Oceanogr.* **31**, 2255–2270.

Robert, R. and Sommeria, J.: 1991, 'Statistical Equilibrium State in Two-Dimensional Flows', *J. Fluid Mech.* **229**, 291–310.

Robert, R. and Rosier, C.: 1997, 'The Modelling of Small Scales in 2D Turbulent Flows: A Statistical Mechanics Approach', *J. Stat. Phys.* **86**, 481–515.

Salmon, R., Holloway, G. and Hendershott, M.C.: 1976, 'The Equilibrium Statistical Mechanics of Simple Quasi-Geostrophic Models', *J. Fluid Mech.* **75**, 691–703.

Salmon, R.: 1998, *Lectures on Geophysical Fluid Dynamics*, Oxford University Press, 378 pp.

Sou, T., Holloway, G. and Eby, M.: 1996, 'Effects of Topographic Stress on Caribbean Sea Circulation', *J. Geophys. Res.* **101**, 16449–16453.

Turner, J.S.: 1968, 'The Influence of Molecular Diffusivity on Turbulent Entrainment Across a Density Interface', *J. Fluid Mech.* **33**, 639–656.

Waite, M.L. and Bartello, P. : 2004, 'Stratified Turbulence Dominated by Vertical Motion', *J. Fluid Mech.*, submitted.