Greg Holloway

g@planetwater.ca

DIFFERENTIAL MIXING, theoretically

Inst Ocean Sciences, Canada

www.planetwater.ca

Preamble in a teacup

Light cream stirred into dark tea produces a uniformly brown mix. Why? Because stirring causes a cascade of tracer (cream) variance from large to small scales, yielding at small scales to molecular diffusion. OK? Were ocean mixing so simple, there would be no need for this poster (nor, perhaps, for this meeting).

Ocean mixing remains profoundly challenging. Neither is it clear that application of classical mechanics (the basis of GFD) plus fudge factors (like mixing coefficients) plus ever-bigger computers will see us through.

Consider the humble teacup. What we know from classical mechanics (or include quantum and relativity too) has a time-symmetry. Re-inserting the spoon with reversed stirring should restore the original segregated cream and tea (other than for effects of molecular diffusion, an empirical fudge, and 'eddy diffusion', a made-up fudge). Classical mechanics fails, rescued only by empirical and/or made-up fudges -- not a safe way for doing ocean mixing.

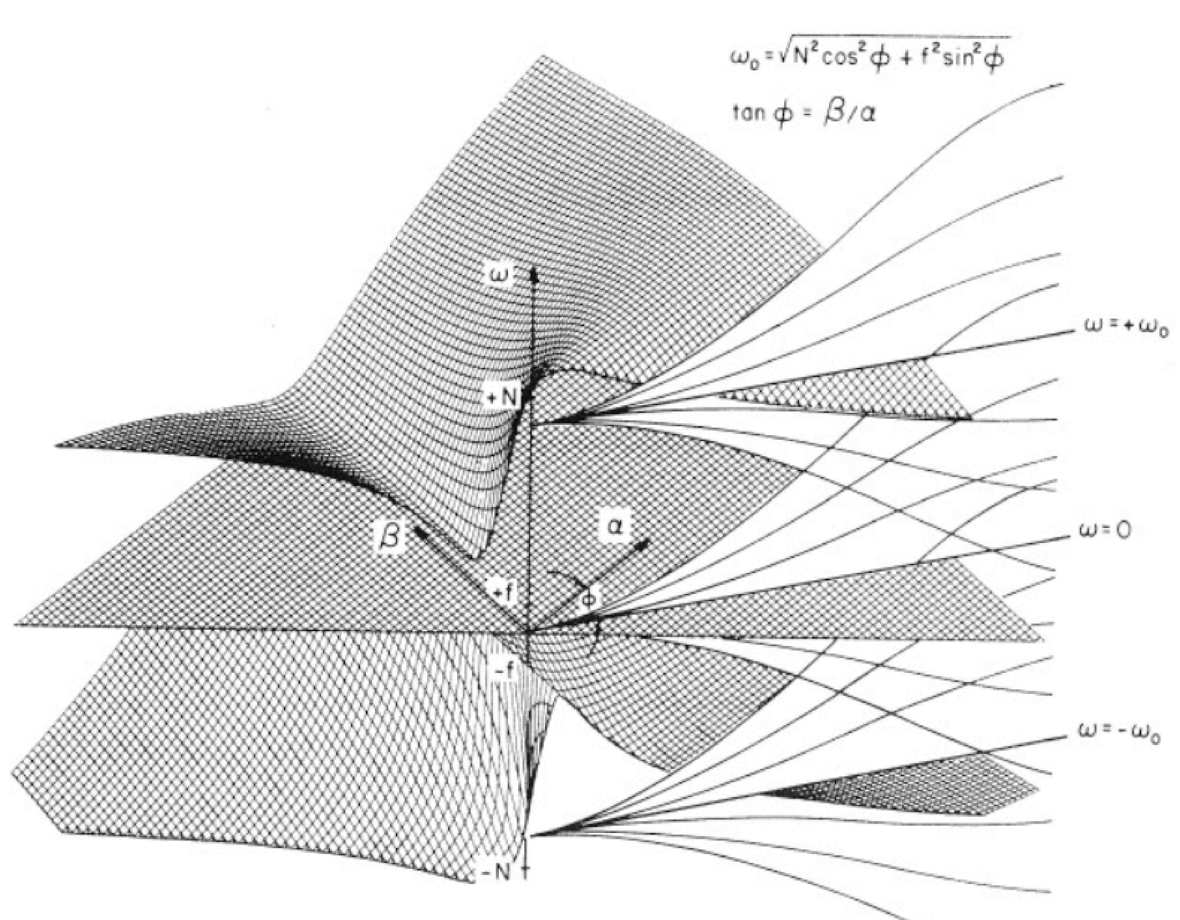
Consider again the humble teacup, here including statistical mechanics. A probability function p(C;**x**,t) might describe the likelihood of finding cream of concentration C near any x at any t. Entropy H = - ∫ log(p)dp is a measure of p. Stirring opens a pathway for p(C;**x**,t1) with H1 to evolve to p(C;**x**,t2) with H2 > H1. The Second Law ("arrow of time") enters, finally driving p to equipartition at molecular scales ("brown mix").

Is this funny way of talking practically useful? Yes. (Sort of.)

Vertical mixing from outer stratosphere to ocean microscale

At first consider only stably stratified, rotating flow with buoyancy given by a single scalar field. Later we extend to two scalars.

Supposing incompressible flow (*i.e.*, omitting acoustic modes), the velocity and buoyancy fields may be decomposed onto three eigenmodes of the linearized equations. These modes are two branches ("upward" and "downward") of internal inertio-gravity waves plus one geostrophic (or "vortical") mode. Let's also omit external modes. For motion of very small amplitude, a frequency spectra for modal coefficients will peak close to natural frequencies. Shorter scales of motion, more readily advected, will see their frequencies "Doppler smeared" about their natural frequencies. At still shorter scales, frequency smearing becomes so great that distributions of the three modal coefficients quite overlap. This is sketched in the following figure.



The scales for which differential mixing (heat vs. salt) occur are so short that the three branches quite overlap and a "turbulence" jargon might seem natural. However, the goal is to reat the entirety from largest to smallest scales in a single consistent dynamic framework.

Comment: I suggest it's a really bad idea to separate fluids in "turbulent" plus "wavy" parts. There's just one fluid consisting of coupled waves.

Atmospheric and oceanic modes tend to be forced at large scales while molecular dissipation removes macroscale motion at small scales, tending to produce "red" spectra. Wave and vortical modes also are forced differently. Myriad interactions then transfer variance across scales and among modes, yielding vertical buoyancy transport (among other things).

The *Arrow of Time* enters

How to anticipate the outcome of more interactions than we can ever keep track of? Even in a teacup? When we ask about anything, say "X", already that tends to average over vastly many things we don't see. X is a moment of probability, X = ∫ x dp. Evolution of X may depend upon a lot of other stuff, some of which we know about, say "Y", and much more we don't know (can't observe, can't calculate, wouldn't care to if we could).

So what is dX/dt? $\quad dX/dT = f(X, Y, ...) + A \bullet \partial H / \partial(X, Y, ...)$

where f(X, Y, ...) are the usual (classical) mechanics which would respect time-reversal symmetry except for fudge factors while "A" project the gradient of entropy, H, back onto dX/dt where H is understood to depend upon macroscopic observables X, Y, ...

It is at " $A \bullet \partial H$ " that the "arrow of time" enters, propelling us beyond classical mechanics (*i.e.*, ocean GFD) to incorporate nonequilibirum statistical mechanics. The challenge remains to see if this can be practically useful.

Practical next steps

We need $A \bullet \partial H$. Only we don't have H, hence no ∂H, and no A. But be brave!

There have been two approaches. First we realize $A \bullet \partial H$ tries to drive the system X, Y, .. toward higher entropy. This sometimes motivates a maximum entropy production ("MEP") principle, requiring an expression for H-production which can be optimised relative to X, Y, .. subject to appropriate "subject to"s. To my knowledge no one's done this in present context.

Alternatively we begin expanding, *e.g.*,

$$A \bullet \partial H / \partial X \approx A \bullet \partial^2 H/\partial X \partial X|^* \bullet (X-X^*) \equiv C \bullet (X-X^*)$$

where $\partial^2 H/\partial X\partial X|^*$ is evaluated at X* such that $\partial H/\partial X|^*=0$ and $C \equiv A \bullet \partial^2 H/\partial X\partial X|^*$ Of course one may continue, considering cross and higher derivatives. But practicality counts, and it is helpful not to proceed blindly.

Importantly, $A \bullet \partial H$ has two parts: A and ∂H. There are directions in configuration space for which A is so small that ∂H cannot matter. Such configurations are effectively inaccessible. For example, one would not consider rapid conversion of all mechanical energy to thermal energy. Thus, we seek $A \bullet \partial H \approx C \bullet (X-X^*)$ taking account of those X* which are "accessible". In present context, this means that interactions (particularly at larger scales) are nearly ideal (advective, adiabatic). Although our ultimate concern may focus on shorter scales supporting differential mixing, we first address the larger scales which set up the variance transfers to smaller scales.

"Oh, fudge!"

Our conceptual tools are now in place to move ahead. But the savvy reader appreciates that, while we began with lofty words -- seeking to avoid "fudging" -- already we feel the fudge thickening. It is. That's the bad news. The good news is:

1. the conceptual framework organizes, and may limit, where and how we fudge;
2. at each stage, fudging can be reduced or eliminated at a cost of greater effort.

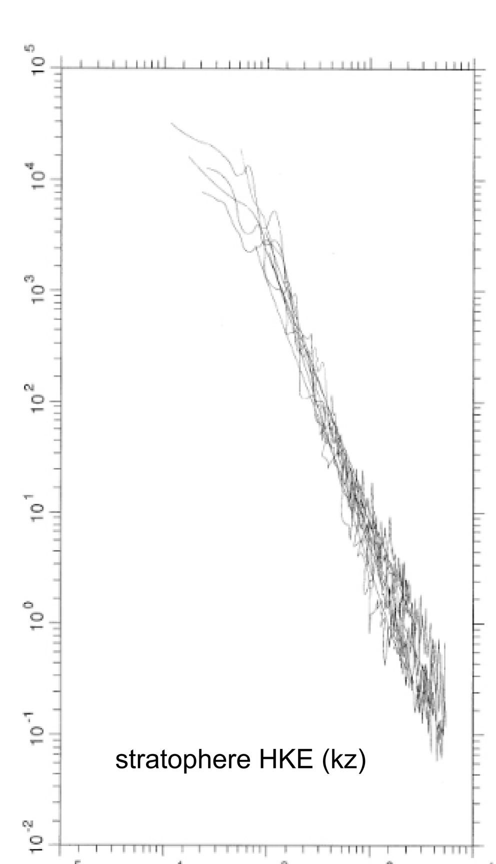
Energetics from large- to micro-scale

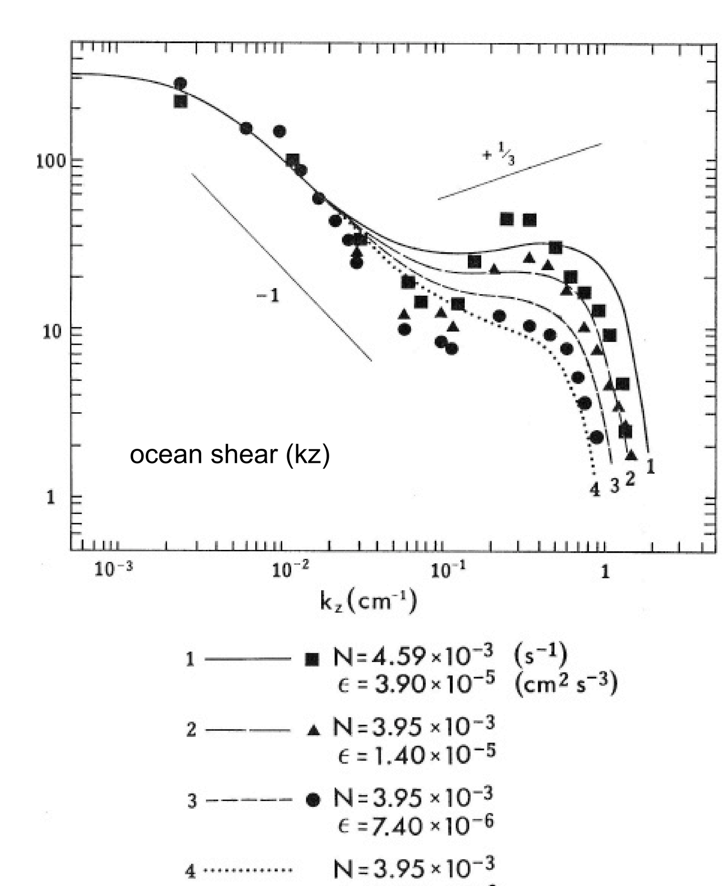
Consider fluctuations along the "z" (upward) axis, with wavenumber k referring to this one axis. Although fields are far from isotropy (certainly at larger scales), anisotropy needn't be our immediate concern. Let $U(k)$ be total velocity variance and $B(k)$ be total buoyancy variance (expressed as potential energy) at k.

Energy balance has $\partial_t U(k) = -\partial_k \varepsilon(k) + F(k)$ and $\partial_t B(k) = -\partial_k \chi(k) - F(k)$ where $\varepsilon(k)$ and $\chi(k)$ are rates of variance transfer from wavenumbers less than k to greater than k. At this point we have no prejudice about the signs of ε and χ, *i.e.* the sense of transfers with respect to scale. F is vertical buoyancy transport, converting B to U.

At first let's set aside F. This is a personal prejudice: I just don't think buoyancy flux is very important to energetics of stably stratified turbulence. *(I may reconsider that?!)* ε and χ arise because gradients of U and B with respect to k set up large ∂H which the flow can readily address by energy "cascades" across k, *i.e.*, large A in the conceptual frame. Specifically, if U and B are equipartitioned (constant) across wavevectors, ∂H (with respect to U and B distribution) vanishes. Following the previous conceptual guide, we seek $-\partial_k \varepsilon(k) \approx C \bullet (U - U^*)$ where C is an operator details of which are accessible via spectral closure theory. For present purpose we can fudge this to $C = \partial_k D_U \partial_k$ hence $\varepsilon \approx -D_U \partial_k U$, a "classic" turbulence phenomenology. Then it's a very short step to observe that, if U and D_U depend only "locally" on k and ε (presumed a weak function of k per my prejudice against F), one has $U \approx \varepsilon^{2/3}k^{-5/3}$ and $D_U \approx \varepsilon^{1/3}k^{-4/3} \approx \tau^{-1}k^2$ where $\tau \approx \varepsilon^{-1/3}k^{-2/3}$.

While the preceding is just a quick fudge to "classic" turbulence (inertial, Kolmogorov), and far from the ocean interior, there's a further easy step. Study of nonlinear waves leads to energy exchange on wavevector triads weighted by nearness to frequency resonance. As energy is farther "smeared" (first figure) at higher k, more triads efficiently participate in energy transfer. Details are again accessible via spectral closure theory but can again be fudged (for present purpose), replacing D_U above as $D_U \approx \theta\tau^{-2}k^2$ with $\theta = \tau^{-1}/(\tau^{-2} + N^2)$ where N^2 is background stability frequency. The result is $U \approx N^2k^{-3} + \varepsilon^{2/3}k^{-5/3}$ compared (below) with stratospheric and oceanic observations.





Implications for tracer equations

Thus far we've got velocity variance. The preceding figure included also viscous decay for $k > k_K = (\varepsilon/\nu^3)^{1/4}$. The other transitional scale was $k_b = (N^3/\varepsilon)^{1/2}$. A similar development for buoyancy yields (in energy units) $B \approx N^2k^{-3} + \varepsilon^{2/3}k^{-5/3}$ for $k < k_K$.

U and B look the "same". But they aren't quite and the difference is hidden in the fudge. Hard numbers in lieu of "≈" are accessible with (a lot of!) effort from spectral closure theory or from laboratory or numerical experiments. Nonetheless it's worth pausing to think where we are for two reasons:

1. Theory given here differs radically from previous accounts. Especially the N^2k^{-3} range differs (a) from turbulent-theoretic accounts of "buoyancy" subrange governed by F and strong divergences $\partial_k\varepsilon$, $\partial_k\chi$ and (b) from "saturation" accounts with internal waves limited by "overtopping" all along N^2k^{-3}, hence again governed by F. The present kinder-gentler theory simply has U and B transferred from larger to smaller scales, with no significant role for F, while N^2k^{-3} and $\varepsilon^{2/3}k^{-5/3}$ ranges only differ on account of the role of off-resonant transfers among nonlinear waves. Someone's got it wrong. Who me?

2. Before coming to differential mixing, it is important to consider larger scale issues. Otherwise a concern is that we study differential mixing in either laboratory or numerical experiments for which "turbulence" is imposed by shaken rods or grids or by numerical injection whereas, if differential mixing in the environment sits at the end of long cascades of buoyancy and velocity from much larger scales, results may depend upon how the buoyancy and velocity were delivered to the differential mixing.

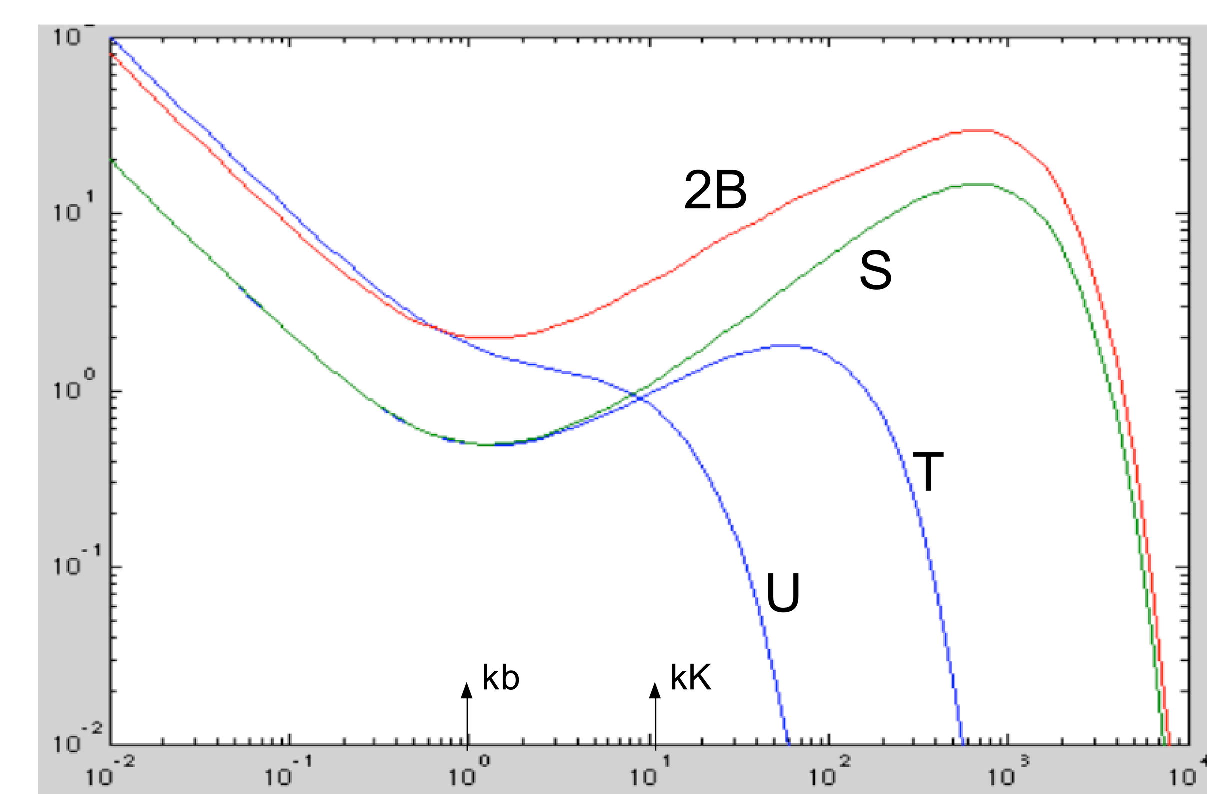
A last remark is important at larger scales. Spectra of U and B are not the same. It is long known from theory and experiments concerning turbulent advection of passive tracers that tracer cascades are more "efficient" than velocity cascades. This is because the vector velocity field preserves near-incompressibility ($\nabla \bullet \mathbf{u} = 0$), imposed via pressure gradient forces, impeding the rates of variance transfer from large to small scales.

Batchelor ranges where they oughtn't?

Extending tracer equations to $k > k_K$ we take account of large Prandtl numbers for seawater with $\nu/\kappa \approx 7$ and $\nu/\gamma \approx 700$. The traditional view would put this range at the viscous end of an inertial range, going over to $B \approx \chi(\varepsilon/\nu)^{1/2}k^{-1}$. A difficulty arises in ocean observations where "Batchelor-like" ranges may sometimes be seen with little or no suggestion of their supporting inertial range. From the present view this is not distressing as both ε and χ represent transfer across the larger "buoyancy" range N^2k^{-3}. If "activity", $\alpha \equiv \varepsilon/\nu N^2$, is not very large, then $k_K/k_b = \alpha^{3/4}$ is not large, a velocity "microstructure bump" may be indistinct, but nonetheless we may see a tracer "bump" with "Batchelor-like" range.

This can be summarised (figure below) for a case where we assume that two tracers, T and S, contribute equally to the stable density gradient ($R_\rho = -1$) corresponding to numerical experiments by Merryfield. To improve visualization, the figure shows spectra of derivative variance, *i.e.*, k^2 times energy spectra, and take $\nu/\kappa = 7$ and $\nu/\gamma = 70$. The choice $\nu/\gamma = 70$ rather than $\nu/\gamma = 700$ is consistent with numerical restrictions in Merryfield's experiments as well as aiding visualization. Later we compare $\nu/\gamma = 700$. The turbulence level is "modest", with $k_K/k_b = 10$ or $\alpha = \varepsilon/\nu N^2 \approx 22$. Markings denote

- "T" : buoyancy variance due to tracer T
- "S" : buoyancy variance due to tracer S
- "2B": twice the total buoyancy variance = 2 * ("T" + "S")
- "U" : total velocity variance



To the left of this figure lies the domain of weaker wave-wave interactions where details of resonances on wave triads are important and an earlier simplifying fudge, using $\theta \approx \tau^{-1}/(\tau^{-2} + N^2)$, doesn't apply. On $k < k_b$, T and S spectra are identical (on account of $R_\rho = -1$). U, like T and S, has form N^2k^{-1}. Notably, 2B appears with amplitude just less than U (at small k), this reflecting the higher efficiency of tracer variance transfer relative to velocity variance transfer. Moving to higher k, 2B overtakes U.

For the modest $\alpha \approx 22$ shown, velocity microstructure appears only as a weak "shoulder" above k_b. Yet T microstructure exhibits a clear bump, while S realizes a "Batchelor-like" k^{+1}range. The remaining challenge is: what does this imply for differential vertical transports of T and S?

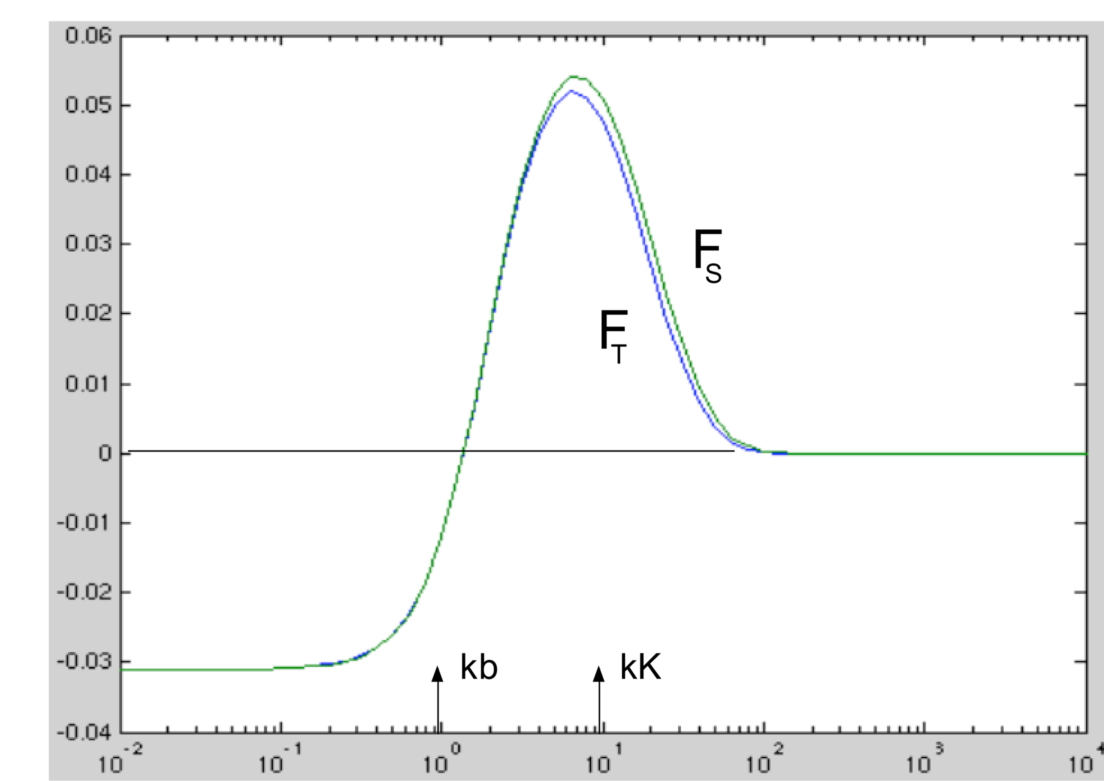
Aside: To improve visualization and to compare with numerical experiments (see Merryfield) which are computationally restricted to $\nu/\gamma = 70$, results on this poster were evaluated at $\nu/\gamma = 70$. It is a question how realistic is $\nu/\gamma = 70$ vs. $\nu/\gamma = 700$. Because theory doesn't suffer the restrictions of numerical experiments, we evaluate δ=.67 when $\alpha = 22$, $R_\rho = -1$ and $\nu/\gamma = 700$ compared with δ=.71 at $\nu/\gamma = 70$. This is encouraging with respect to Merryfield's experiments.

Differential vertical transports

The supposition (mere prejudice to this point) is that vertical buoyancy flux $F = F_T + F_S$, summing the T and S contributions, plays only minor role with respect to U, T and S spectra. However, it is just at F_T and F_S that we evaluate differential mixing $\delta = \dfrac{F_S/\partial_z S}{F_T/\partial_z T}$.

The F arise also in response to entropy gain, ∂H, but there is a competition. On one hand, downgradient F are a source for tracer variance, increasing H. This occurs in turbulent transport of passive tracers, and may be seen in the teacup example. On the other hand, under gravity and rotation (or gravity alone) the modal description in terms of two wave branches plus a vortical branch achieves local maxima of H when energy is equipartioned among the three branches, resulting in $U = 2B$.

For relatively intense turbulence, say $\alpha >> O(10^3)$, simple downgradient F dominates, with high wavenumber F_T suppressed (by $\kappa >> \gamma$) relative to F_S. Then $\delta > 1$. This may be regarded as "usual turbulence", tending toward $\delta = 1$ as $\alpha \Rightarrow \infty$. Through much of the ocean interior (or stratosphere) the "turbulence" is only "marginal" with $\alpha < O(10^2)$. This was the case shown in the previous figure, for which F_T and F_S are shown below resulting in $\delta = .7$.



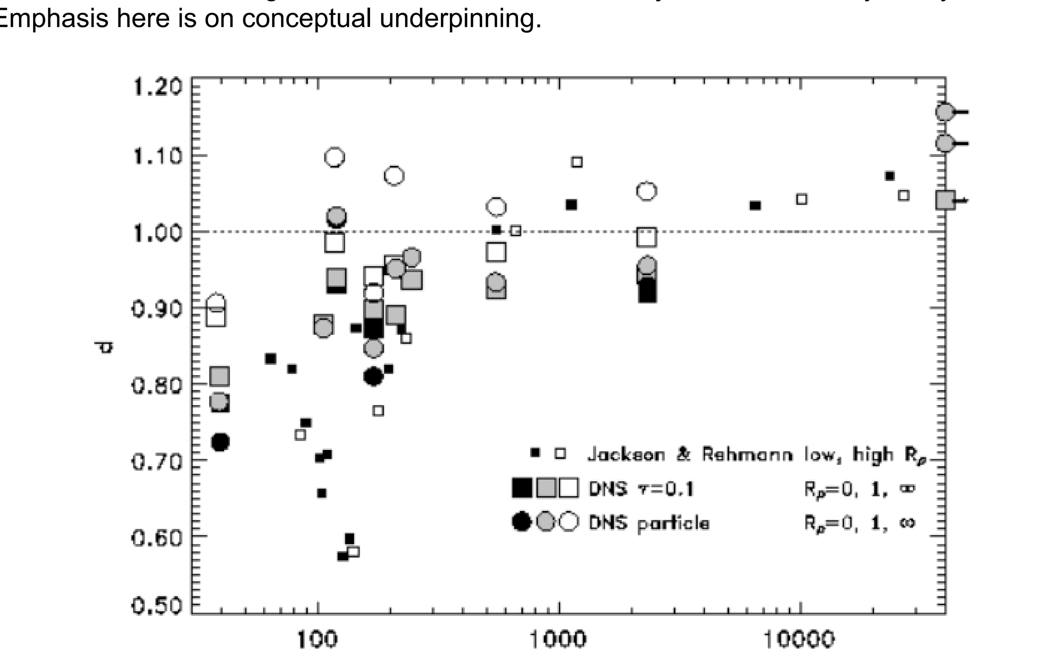
It may seem strange. Such small differences between F_T and F_S produce such large departure from $\delta = 1$. The key is to appreciate that small scale buoyancy fluxes tend to by systematically upwards (restratifying or <u>counter</u>gradient). $\kappa >> \gamma$ tends to limit the restratifying part of F_T, hence $\delta < 1$. When there's significant compensation between downward flux at larger scales and upward flux at smaller scales, the effect on δ is large.

It's worth noting that upwards (<u>counter</u>gradient) F dominate just the scales $k > k_b$ which could be called "turbulent". But this is backwards from "usual turbulence", cautioning us from heuristic application of "turbulence" ideas in ocean mixing and suggesting more fundamental approaches respecting entropy and 2nd Law.

The role of nearly compensating upwards and downwards F further suggests that significant $\delta < 1$ may occur only over a limited range of α. "Intense" (large) α should drive us to $\delta > 1$ while "weak" α has too little velocity variance in short scales to support substantial upward, lessening the tendency for $\delta < 1$. This can be tested comparing the previous case, $\alpha = 22$, with other α:

α:	4.9	8.5	22	35	104	464	2308
δ:	.95	.87	.71	.89	.99	1.01	1.01

These values can be compared, below, with laboratory and numerical experiments. While numerical experiments (Merryfield) suggest decreasing δ with decreasing α, lab results (JR) appear to realize a minimum near $\alpha \approx 100$. Although its authors don't comment on this, I suggest the minimum for JR is real, possibly falling in a parameter "gap" for Merryfield. Also, values of α do not simply map between theory and experiments. In part, given "fudge" I should be first to discount my own numbers. "Tuning" is readily available but would be silly at this point. More fundamentally, we address different problems whether "turbulence" results from cascade from large scale waves/vortices or is injected numerically or by shaken rods. Emphasis here is on conceptual underpinning.



Finally, for the case illustrated above we let T and S make equal contributions, ($R_\rho = -1$), to buoyancy gradient. How does δ vary when $R_\rho \neq -1$? For the case $\alpha = 22$, δ varies from .52 when $|R_\rho| << 1$ (*i.e.*, S dominates buoyancy) to .90 when $|R_\rho| >> 1$ (*i.e.*, T dominates).

This too is easy to explain. But the good news is: the poster has run out of space.