The smallest eddies are almost numberless, and large things are rotated only by large eddies and not by small ones, and small things are turned by small eddies and large.

Leonardo da Vinci, describing *turbolenza* in a sketch book, c. 1500.

CHAPTER 13

Turbulent Diffusion and Eddy Transport

THE TRANSPORT OF FLUID PROPERTIES BY UNSTEADY MOTION — that is, the way in which the properties of a fluid may be carried from one location to another by waves and turbulence — is one of the most important topics in geophysical fluid dynamics. It may be the dominant transport in a fluid, greatly exceeding that of the mean flow — in the atmosphere, for example, heat is transferred polewards primarily by the action of unsteady weather systems, not by the much weaker time-mean flow. However, we are often not interested in the details of the turbulent eddies and hence we might seek to *parameterize* the turbulent transport in terms of the mean flow; unfortunately, no general theory exists for such transport, for indeed such a theory would amount to a theory of turbulence. In the absence of this, we focus our attention in this chapter on the theory (such as it is) and practice of *turbulent diffusion*. In models of turbulent diffusion, the turbulent transport is generally related to the gradient of the mean flow, and it is the simplicity of the resulting expressions that has led to their wide adoption in areas as different as turbulent pipe flow, atmospheric boundary layer transport and large-scale ocean modelling. Diffusive models are, or aim to be, rational, simple and tractable — a blend of heuristic reasoning and elementary mathematics, the latter needed to ensure that certain basic requirements (conservation laws, for example) of a physical process are captured by a parameterization. However, just like other turbulent closures, they rely on physical assumptions that cannot be rigorously justified. In the first part of the chapter we consider turbulent diffusion from a general standpoint, and then specialize our discussion to geofluids, and in particular to large-scale transport by baroclinic eddies. Those readers with some prior knowledge of turbulent diffusion may choose to skip ahead to Section 13.6.

13.1 DIFFUSIVE TRANSPORT

We begin with a brief discussion of the diffusion equation itself, to wit

$$\frac{\partial \varphi}{\partial t} = \kappa \nabla^2 \varphi, \tag{13.1}$$

where κ is a constant, positive, scalar diffusivity and the tracer φ is a scalar field. We expect that an initially concentrated blob of tracer would spread out — it would diffuse — and thus small parcels of tracer are transported. How quickly does this occur, or, put another way, is there an effective diffusive transport velocity?

If the rate of spreading becomes independent of the initial conditions then, purely from dimensional considerations, the spreading can depend only on the diffusivity and time itself and we can write

$$\overline{X^2} = \alpha\kappa t, \tag{13.2}$$

where $\overline{X^2}$ is the mean-square displacement, t is time and α is a nondimensional constant. Let us quantify this with an explicit calculation. If φ is interpreted as the density of markers of fluid parcels, then the mean-square displacement of the markers is given by (in three dimensions)

$$\overline{X^2} = \frac{\int_0^\infty r^2 \varphi r^2 \,\mathrm{d}r}{\int_0^\infty \varphi r^2 \,\mathrm{d}r}, \tag{13.3}$$

where the denominator, the total amount of tracer present, is a constant and we have assumed a spherically symmetric distribution of tracer. Using (13.1) we find

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_0^\infty r^2\varphi r^2 \,\mathrm{d}r = \kappa \int_0^\infty r^2 \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial \varphi}{\partial r}\right) r^2 \,\mathrm{d}r = 6\kappa \int_0^\infty \varphi r^2 \,\mathrm{d}r, \tag{13.4}$$

after a couple of integrations by parts. Thus

$$\frac{\mathrm{d}}{\mathrm{d}t}\overline{X^2} = 6\kappa, \tag{13.5}$$

and because κ is a constant we have the important result that

$$\boxed{\overline{X^2} = 6\kappa t.} \tag{13.6}$$

In two dimensions the equivalent calculation begins with

$$\overline{X^2} = \frac{\int_0^\infty r^2 \varphi r \,\mathrm{d}r}{\int_0^\infty \varphi r \,\mathrm{d}r} \tag{13.7}$$

and using the diffusion equation we find

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_0^\infty r^2\varphi r \,\mathrm{d}r = \kappa \int_0^\infty r^2 \frac{1}{r}\frac{\partial}{\partial r}\left(r \frac{\partial \varphi}{\partial r}\right) r \,\mathrm{d}r = 4\kappa \int_0^\infty \varphi r \,\mathrm{d}r. \tag{13.8}$$

Thus we obtain

$$\overline{X^2} = 4\kappa t. \tag{13.9}$$

Finally, in one dimension (i.e., spreading along a line) it is easy to show that

$$\overline{X^2} = 2\kappa t. \tag{13.10}$$

Thus, in both three and two dimensions, *the spread of a diffused scalar increases with the half power of time.*

13.1.1 An Explicit Example

We gain a little more intuition about what the above calculations mean by considering the case in which the initial tracer distribution is a delta function at the origin. If the total amount of tracer is unity, then in three dimensions at subsequent times the tracer is given by the distribution

$$\varphi(r, t) = \frac{1}{8(\pi\kappa t)^{3/2}} \exp(-r^2/4\kappa t), \tag{13.11}$$

as may be checked by substitution back into the equation of motion. The distribution clearly broadens with time, and the mean-square distance from the origin is given by

$$\overline{X^2} = \int_0^\infty \frac{4\pi r^2}{8(\pi\kappa t)^{3/2}} \exp(-r^2/4\kappa t)\,\mathrm{d}r = 6\kappa t, \tag{13.12}$$

as in (13.6). The important point is that the mean distance travelled by a particle during a time interval t is proportional to the square root of that time interval. This is, of course, redolent of a random walk (see the shaded box on the following page), which brings us to the subject of turbulent diffusion.

13.2 TURBULENT DIFFUSION

Fluids differ from solids in that they can transport properties by advection — thus, heat is primarily transferred polewards in the atmosphere by means of air movement and not by molecular diffusion. Turbulent fluid motion differs from laminar fluid motion in that such advective transport may be greatly enhanced by the seemingly random motion of the fluid, the net transport being much larger than that which would be effected by the time-mean fluid motion alone. Indeed, to continue the atmospheric example, away from the tropics the poleward transport of heat in the atmosphere is largely effected by way of the (large-scale) turbulent transfer of heat in mid-latitude weather systems. Of course, such transfer *is* simply by advection, and if we could explicitly calculate the motion of the fluid parcels we could explicitly calculate the transport. However, turbulent transport is both very complicated and very sensitive to the initial conditions, so that any hope of performing such a calculation exactly in a real situation is often a forlorn one.

Turbulent transport is most important in *inhomogeneous* situations, because it is the divergence of the transport that is important and the mean divergence is non-zero only if there is inhomogeneity. The theories of Chapters 11 and 12 do not lend themselves to an easy extension to inhomogeneous flow, and we turn to a slightly more empirical approach.[1]

13.2.1 Simple Theory

Let us consider how fluid markers are transported in a statistically steady, homogeneous, turbulent flow. The markers are introduced at the origin $x = y = z = 0$ at $t = 0$; we may create an ensemble of such markers by performing many such tracer release experiments on different realizations of the turbulent flow, but with each flow having the same statistical properties. The question is, what is the average rate of dispersion of a single particle of fluid?

The displacement of a marker at a time t is given by

$$\boldsymbol{X}(t) = \int_0^t \boldsymbol{V}(t')\,\mathrm{d}t', \tag{13.13}$$

where $\boldsymbol{V}$ is the velocity of the fluid parcel — a material velocity. (We will use uppercase variables to denote material ('Lagrangian') quantities.) The mean-square displacement is

$$\overline{X^2(t)} = \int_0^t \mathrm{d}t_1 \int_0^t \overline{\boldsymbol{V}(t_1)\cdot\boldsymbol{V}(t_2)}\,\mathrm{d}t_2, \tag{13.14}$$

where the overbar denotes an ensemble average, and thus $\overline{\boldsymbol{V}(t_1)\cdot\boldsymbol{V}(t_2)}$ is a measure of the velocity correlation between the velocities of the fluid parcels at times t_1 and t_2. That is,

$$\overline{\boldsymbol{V}(t_1)\cdot\boldsymbol{V}(t_2)} = \overline{v^2}R(t_2 - t_1) = \overline{v^2}R(\tau), \tag{13.15}$$

A Random Walk

Here we give an elementary derivation of the most basic result in random walk theory, the relationship of the mean-square displacement to the number of steps taken.[2] A loose analogy is that of drunkards staggering randomly from here to there, with no correlation between their successive steps. After any number of steps, the *mean* displacement of the drunkards is zero, but we expect their *root-mean-square* displacement to increase: this is because drunkards independently thrown out of the same bar will generally wander off in different directions (which is why the mean displacement is zero), but after some time most of them will indeed end up some distance away.

For simplicity consider steps, $\boldsymbol{s}_n$, each with random orientation but equal magnitude, s. The displacement after n steps is related to the displacement after $n-1$ steps by

$$\boldsymbol{D}_n = \boldsymbol{D}_{n-1} + \boldsymbol{s}_n, \tag{R.1}$$

so that the amplitude of $\boldsymbol{D}_n$, namely D_n, is given by

$$\begin{aligned} D_n^2 &= (\boldsymbol{D}_{n-1} + \boldsymbol{s}_n) \cdot (\boldsymbol{D}_{n-1} + \boldsymbol{s}_n) \\ &= D_{n-1}^2 + s^2 + 2\boldsymbol{D}_{n-1} \cdot \boldsymbol{s}_n. \end{aligned} \tag{R.2}$$

Taking an ensemble average over many realizations gives

$$\overline{D_n^2} = \overline{D_{n-1}^2} + s^2, \tag{R.3}$$

having used $\overline{\boldsymbol{D}_{n-1} \cdot \boldsymbol{s}_n} = 0$, because each step is random.

Now, $D_0 = 0$, so that $\overline{D_1^2} = s^2$, $\overline{D_2^2} = 2s^2$ and so on. Thus, using (R.3) to proceed inductively, we have

$$\overline{D_n^2} = ns^2, \tag{R.4}$$

or

$$\overline{D_n^2}^{1/2} = \sqrt{n}s. \tag{R.5}$$

Thus, in a random walk the *root-mean-square displacement increases with the half-power of the number of steps taken.* More work is required to calculate the distribution of the random walkers, but it may be shown that in the limit of infinitesimally small steps the random walk becomes a Wiener process and the distribution becomes Gaussian, as in (13.11) (with the exact form depending on the dimensionality of the problem), indicating a diffusive process. Diffusion may be thought of as a continuous random walk, with the displacement proportional to the square root of time.

where $R(t_2 - t_1)$ is the velocity correlation function and, because the turbulence is statistically steady, this depends only on the time difference $\tau = t_2 - t_1$. Furthermore, $R(-\tau) = R(\tau)$. Thus,

$$\overline{X^2(t)} = \int_0^t \mathrm{d}t_1\, \overline{v^2} \int_0^t R(t_2 - t_1)\, \mathrm{d}t_2 = \int_0^t \mathrm{d}\hat{t}\, \overline{v^2} \int_{-\hat{t}}^{t-\hat{t}} R(\tau)\, \mathrm{d}\tau, \tag{13.16}$$

changing variables to τ and $\hat{t} = t_1$ (Fig. 13.1). We expect the velocity correlation function to fall monotonically from its initial value of unity to a value approaching zero as $\tau \to \infty$, as in Fig. 13.2,

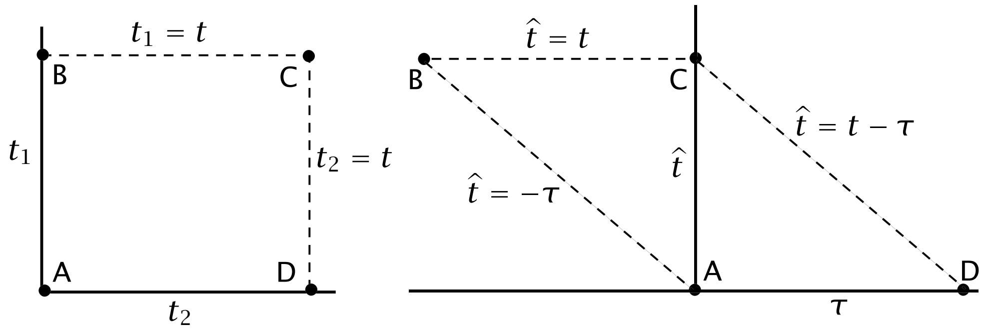

Fig. 13.1 Changes of time variables involved in (13.16) and (13.32). The original two-dimensional integral is over the rectangle ABCD. Defining $\tau = t_2 - t_1$ and $\hat{t} = t_1$, then the area is spanned by $[\hat{t} = (0, t),\ \tau = (-\hat{t}, t - \hat{t})]$ as in (13.16), or by $[\tau = (0, t),\ \hat{t} = (0, t - \tau)]$ (i.e., ACD) plus $[\tau = (-t, 0),\ \hat{t} = (-\tau, t)]$ (i.e., ABC) in (13.32).

and typically, there will be some characteristic time τ_{corr} that parameterizes the behaviour of the function. In general, we cannot obtain explicit general solutions without detailed knowledge of this correlation function, but there are two interesting limits:

(i) The short-time limit. For small times, i.e., for $t \ll \tau_{corr}$ (and so $t_1, t_2 \ll \tau_{corr}$) the correlation function will be approximately unity and so (13.16) becomes

$$\overline{X^2(t)} \approx \int_0^t \mathrm{d}\hat{t}\, \overline{v^2} \int_0^t \mathrm{d}\hat{t} = \overline{v^2} t^2. \tag{13.17}$$

Thus, the root-mean-square displacement increases linearly with time, and linearly with the root-mean-square velocity of the flow. For small times, the fluid parcel's behaviour is well correlated with that at the initial time, and so the displacement increases linearly in the direction it was initially going. Indeed, directly from (13.13) we have, for small times, $\boldsymbol{X}(t) \approx \boldsymbol{V}t$, which leads directly to (13.17).

(ii) The long-time limit. We are now concerned with the case $t \gg \tau_{corr}$. Because the correlation function falls with time, most of the contributions to the second integrand (involving $R(\tau)$) in (13.16) are from $\tau \le \tau_{corr}$. Thus, without much loss in accuracy, we can replace the limits of integration by $-\infty$ and $+\infty$; that is

$$\overline{X^2(t)} \approx \int_0^t \mathrm{d}\hat{t}\, \overline{v^2} \int_{-\infty}^{\infty} R(\tau)\, \mathrm{d}\tau. \tag{13.18}$$

Assuming the second integral converges, it is just a number; in fact, noting that $R(\tau) = R(-\tau)$, we may use it to *define* the correlation time τ_{corr} by

$$\tau_{corr} \equiv \int_0^{\infty} R(\tau)\, \mathrm{d}\tau. \tag{13.19}$$

We then have the important result that

$$\boxed{\overline{X^2(t)} \approx 2\overline{v^2} t \int_0^{\infty} R(\tau)\, \mathrm{d}\tau = 2\overline{v^2} \tau_{corr} t.} \tag{13.20}$$

That is, for times that are long compared with the turbulence correlation time, the distance travelled by a fluid parcel in some time interval is proportional to the square-root of that time interval, just as for a diffusive process; this is because the fluid parcels are essentially undergoing random walks. Equation (13.20) connects two quite different fluid properties: the left-hand side tells us how tracers are dispersed in a turbulent flow, a material property, whereas the right-hand side can be evaluated from the Eulerian velocity field at different times. Both the left- and right-hand sides can be directly measured, by looking at the dispersion of a dye and by measuring the velocity at successive times.

We may define a *coefficient of turbulent diffusivity* by

$$\boxed{K_{turb} = \tfrac{1}{3}\overline{\upsilon^2}\tau_{corr}}\,, \tag{13.21}$$

and then we have the result that

$$\overline{X^2(t)} = 6K_{turb}t. \tag{13.22}$$

Comparison of (13.22) with (13.6) indicates that the transport of turbulent flow, under these conditions, is like a diffusive transport, with a coefficient of diffusivity given by (13.21). [Sometimes, the numerical factors are neglected, and a diffusivity is defined by the expressions

$$K_{turb} \equiv \frac{\mathrm{d}\overline{X^2(t)}}{\mathrm{d}t} \qquad \text{or} \qquad K_{turb} \equiv \frac{1}{2}\frac{\mathrm{d}\overline{X^2(t)}}{\mathrm{d}t}. \tag{13.23}$$

These lose the exact connection with a true diffusion coefficient, but usually the turbulent diffusivity can only be estimated, anyway.]

We may define a correlation length scale to be the approximate distance that a parcel moves, on average, in a material (i.e., 'Lagrangian') correlation time. Thus

$$l_{corr} \equiv \upsilon_{rms}\tau_{corr}, \tag{13.24}$$

where $\upsilon_{rms} = (\overline{\upsilon^2})^{1/2}$, whence

$$\boxed{K_{turb} = \tfrac{1}{3}\upsilon_{rms}l_{corr}}\,. \tag{13.25}$$

In most situations, the numerical coefficient (1/3 here) cannot be trusted because a real turbulent flow is unlikely to satisfy the restrictions of stationarity and homogeneity that we have imposed. Nevertheless, a relationship similar to (13.25) — that a turbulent diffusivity is proportional to an r.m.s. turbulent velocity and a correlation length scale, is the foundation for semi-empirical *mixing length* theories that we discuss in Section 13.4.

The simple relationships between the mean-square displacement, the Lagrangian time scale, the mean-square velocity and the eddy diffusivity allow the diffusivity to be computed from the statistics of particle trajectories. Thus, suppose that a cluster of floats is released into the ocean, or some balloons are released in the atmosphere. If neutrally buoyant, these instruments then essentially become labelled fluid particles, and one may compute K_{turb} directly from the dispersion of the cluster using (13.22). If it is possible to measure their root-mean-square velocity, then one may use (13.21) to estimate the diffusivity from this and the material correlation time scale.

13.2.2 ✦ An Anisotropic Generalization

We now consider the correlation between the different components of the displacement in anisotropic, but still homogeneous, flow.[3] The displacement of a fluid particle is given by (13.13), and this is a

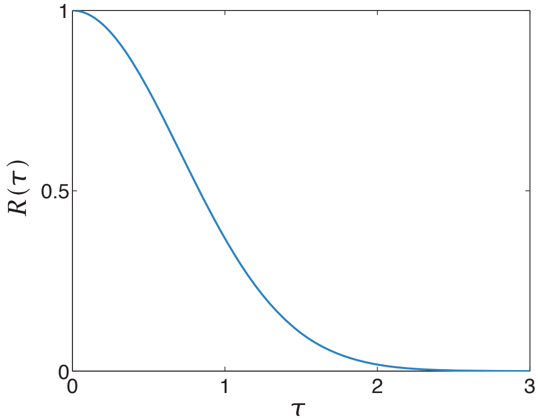

Fig. 13.2 Idealized velocity correlation function in turbulent flow with correlation time $\tau_{corr} = \mathcal{O}(1)$. For small times, $\tau \ll \tau_{corr}$, $R(\tau) \approx 1$. For large times, $\tau \gg \tau_{corr}$, $R(\tau) \ll 1$. We may define the correlation time by $\tau_{corr} = \int_0^\infty R(\tau)\,d\tau$.

random vector. Thus, generalizing (13.14), we may define the fluid particle displacement covariance tensor by

$$D_{ij}(t) = \overline{X_i(t)X_j(t)} = \int_0^t \int_0^t \overline{V_i(t_1)V_j(t_2)}\,dt_1\,dt_2, \tag{13.26}$$

where the velocity denoted by V_i is the ith component of the velocity of a fluid element. For small times, $X_i(t) \approx v_i(\boldsymbol{a}, 0)t$, where $v_i(\boldsymbol{a}, 0)$ is the fluid velocity at the parcel's initial position, $\boldsymbol{a}$, and we obtain

$$D_{ij}(t) \approx \overline{v_i(\boldsymbol{a}, 0)v_j(\boldsymbol{a}, 0)}t^2. \tag{13.27}$$

If the flow is statistically steady and homogeneous the average of any quantity has no spatial or temporal dependence and so

$$D_{ij}(t) = A_{ij}t^2, \tag{13.28}$$

where the tensor $A_{ij} = \overline{v_i v_j}$ has constant entries, and this is a slight generalization of (13.17).

The velocity covariance of a fluid parcel at times t_1 and t_2 is, as before, a function only of the time difference $t_1 - t_2$ and so it must have the form

$$\overline{V_i(t_1)V_j(t_2)} = \left(\overline{v_i^2}\,\overline{v_j^2}\right)^{1/2} R_{ij}(t_2 - t_1). \tag{13.29}$$

Except in the case of isotropic flow $R_{ij}(\tau) \neq R_{ij}(-\tau)$, but we do have, in general,

$$R_{ij}(\tau) = R_{ji}(-\tau). \tag{13.30}$$

Now, to obtain a generalization of (13.20), we first use (13.29) in (13.26) to obtain

$$D_{ij}(t) = \left(\overline{v_i^2}\,\overline{v_j^2}\right)^{1/2} \int_0^t \int_0^t R_{ij}(t_2 - t_1)\,dt_1\,dt_2. \tag{13.31}$$

If we change variables to $\tau = t_2 - t_1$ and $\hat{t} = t_1$ (see Fig. 13.1) we obtain[4]

$$D_{ij}(t) = \left(\overline{v_i^2}\,\overline{v_j^2}\right)^{1/2} \left(\int_0^t d\tau \int_0^{t-\tau} d\hat{t}\,R_{ij}(\tau) + \int_{-t}^0 d\tau \int_{-\tau}^t d\hat{t}\,R_{ij}(\tau) \right), \tag{13.32}$$

and using (13.30) this becomes

$$D_{ij}(t) = 2\left(\overline{v_i^2}\,\overline{v_j^2}\right)^{1/2} \int_0^t d\tau \int_0^{t-\tau} d\hat{t}\,\widehat{R}_{ij}(\tau), \tag{13.33}$$

where $\widehat{R}_{ij} = (R_{ij} + R_{ji})/2$. This order of integration enables us to perform the integration over $\hat{t}$, giving

$$D_{ij}(t) = 2\left(\overline{v_i^2}\,\overline{v_j^2}\right)^{1/2} \int_0^t (t-\tau)\widehat{R}_{ij}(\tau)\,\mathrm{d}\tau. \tag{13.34}$$

For long times, i.e., for $t \gg \tau_{corr}$, the upper limit of the integration may be taken to be infinity, again because the contributions to the integrand from $\widehat{R}_{ij}(\tau)$ all come from small τ. Furthermore, we expect that for large t,

$$\int_0^\infty t\widehat{R}_{ij}(\tau)\,\mathrm{d}\tau \gg \int_0^\infty \tau\widehat{R}_{ij}(\tau)\,\mathrm{d}\tau, \tag{13.35}$$

because $R_{ij}(\tau)$ is only non-negligible for small τ, and $t \gg \tau$ in this range. Thus, we finally obtain a generalization of (13.20) for the displacement covariance of two components of the displacement, namely

$$D_{ij} = 2\left(\overline{v_i^2}\,\overline{v_j^2}\right)^{1/2} t \int_0^\infty \widehat{R}_{ij}(\tau)\,\mathrm{d}\tau. \tag{13.36}$$

The integral is a tensor with constant entries, analogous to the turbulent decorrelation time scale of (13.19). Then, with $\tau_{ij} \equiv \int_0^\infty \widehat{R}_{ij}(\tau)\,\mathrm{d}\tau$, the corresponding turbulent diffusivity is

$$K_{ij} = \tfrac{1}{3}\left(\overline{v_i^2}\,\overline{v_j^2}\right)^{1/2} \tau_{ij}. \tag{13.37}$$

13.2.3 Discussion

We have shown that, for sufficiently long times, the distance travelled by a fluid parcel in some time is proportional to the square root of that time, just as for a diffusive process and just as for a random walk. The motion of our fluid parcel is analogous to that of a dust particle undergoing Brownian motion — both are continually buffeted and undergo random walks as a result. Still, it may appear that the usefulness of our results is limited by the assumptions of stationarity and homogeneity — it is well-nigh impossible in nature to produce a statistically stationary, homogeneous turbulent flow, because statistical stationarity implies there must be an energy source and this, as well as the presence of boundaries, militates against homogeneity. However, we should not be so pessimistic, on two counts:

(i) Similar ideas may be directly applied to flows that are homogeneous in one direction, which is more easily achievable in nature.

(ii) Often, a flow will *not* be homogeneous in any direction. However, if the *statistics* of the eddy motion vary on a space scale that is longer than $v_{rms}\tau_{corr}$, then the eddy transport properties may be determined by a local theory. For example, the size of the eddy diffusivity is then determined by $D_t \sim v_{\mathrm{rms}} l_{corr}$ where the parameters, and hence the diffusion coefficient, vary, but only on a scale longer than the energy-containing scale.

The essential results of this section thus lie in (13.20), (13.21) and (13.25): that the dispersion of a fluid particle in a turbulent flow is *diffusive* in nature, and that the turbulent diffusivity is proportional to the product of the root-mean-square velocity and the correlation length.

13.3 TWO-PARTICLE DIFFUSIVITY

Let us now consider the problem of determining the mutual separation of two fluid parcels; the problem is relevant to geofluids because by tracking the separation of floats in the ocean, or balloons in the atmosphere, we can learn much about the nature of large-scale turbulence in those systems. The problem differs from the one-particle problem, because the separation of the particles itself will affect the rate of increase of the separation. In the one-particle problem in homogeneous flow, the position of the particular tagged fluid particle plays no direct role in determining its rate of spreading from its initial condition — any one position is the same as any other. But if two particles are close together, they may be swept away together by some large eddy, without affecting their mutual separation whereas two particles that are widely separated will undergo largely uncorrelated motion. Thus, we identify two regimes:

(i) A regime in which the separation of the particles is greater than the scale of the largest eddies. In this case, each particle is undergoing a random walk that is effectively uncorrelated with that of the other particle.

(ii) A regime in which the separation of the particles is less than the energy-containing scale of the motion. In this case, the eddies that contribute most to the two-particle separation are those that are comparable in scale to the separation itself.

If we attempt to apply the Taylor analysis ab initio we evidently have, by analogy to (13.13)

$$\boldsymbol{Y}(t) = \boldsymbol{X}_1(0) - \boldsymbol{X}_2(0) + \int_0^t [\boldsymbol{V}_1(t') - \boldsymbol{V}_2(t')]\,\mathrm{d}t', \tag{13.38}$$

and a mean-square separation of

$$\overline{\boldsymbol{Y}^2(t)} = \overline{(\boldsymbol{X}_1(0) - \boldsymbol{X}_2(0))^2} + \int_0^t \mathrm{d}t_1 \int_0^t [\boldsymbol{W}_1(t_1)\cdot\boldsymbol{W}_2(t_2)]\,\mathrm{d}t_2, \tag{13.39}$$

where $\boldsymbol{W}(t) = \boldsymbol{V}_1(t) - \boldsymbol{V}_2(t)$. However, it is now difficult to proceed much further. The problem is that we cannot write

$$\overline{\boldsymbol{W}(t_1)\cdot\boldsymbol{W}(t_2)} = \overline{w^2}R(t_2 - t_1), \tag{13.40}$$

because the correlation will depend on the initial separation of the particles as well as the time since then. Thus, the diffusivity itself will depend on both time and the initial particle separation, and the results analogous to those of the single-particle diffusivity cannot easily be recovered. However, we can make some progress by separately considering the two above-mentioned regimes.

13.3.1 Large Particle Separation

This case is analogous to the single-particle case. The particle separation is given by

$$\boldsymbol{Y}(t) = \boldsymbol{X}_1(t) - \boldsymbol{X}_2(t), \tag{13.41}$$

so the mean-square separation is

$$\overline{Y^2(t)} = \overline{X_1(t)^2} + \overline{X_2(t)^2} - 2\overline{\boldsymbol{X}_1(t)\cdot\boldsymbol{X}_2(t)}. \tag{13.42}$$

For long times, the last term is zero because the motion of the two particles is uncorrelated. Furthermore, each of the first two terms is given by (13.20) or (13.22), so that the mean separation varies as

$$\overline{Y^2(t)} = 4\overline{v^2}\tau_{corr}t \tag{13.43}$$

and the rate of separation, for large t, is given by

$$\frac{\mathrm{d}\overline{Y^2(t)}}{\mathrm{d}t} = 4\overline{v^2}\tau_{corr} = 12K_{turb}. \tag{13.44}$$

Thus, the relative diffusion is twice that of the single-particle process, in the limit that the particles are separated by an amount larger than the largest eddies.

13.3.2 Separation Within the Inertial Range

How do fluid parcels whose separation is at inertial scales behave relative to each other?[5] Suppose that two particles are tagged, and that their separation is greater than the viscous scale but smaller than the scales of the largest eddies — that is, the separation lies within the inertial range of the flow. Then, the rate of separation of the two particles can depend only on two quantities, the separation itself and properties of the inertial range, meaning (in three dimensions) the energy flux, ε, through the system. It cannot depend on the time, because this would imply that the subsequent rate of particle separation depends on the history of how the particles came to their current positions. Thus we can write

$$\frac{\mathrm{d}\overline{L}^2}{\mathrm{d}t} = g(\overline{L}, \varepsilon), \tag{13.45}$$

where $\overline{L} \equiv \overline{Y(t)^2}^{1/2}$. Dimensional analysis then gives

$$\frac{\mathrm{d}\overline{L}^2}{\mathrm{d}t} = A\varepsilon^{1/3}\overline{L}^{4/3}, \tag{13.46}$$

where A is a nondimensional constant, and this is known as 'Richardson's four-thirds law'. We can integrate (13.46) to give

$$\overline{L}^2 \sim \varepsilon t^3. \tag{13.47}$$

Another way of deriving (13.46) is to suppose that the separation obeys the diffusive law

$$\frac{\mathrm{d}\overline{L}^2}{\mathrm{d}t} = K_{turb}, \tag{13.48}$$

where K_{turb} is a turbulent diffusivity that is *a function of the separation itself*. This is because the farther apart the eddies are, the larger the scale of the eddies that can move the two particles independently, rather than just sweeping them along together. An estimate of the diffusivity is then

$$\mathcal{K}_{turb} \sim vl, \tag{13.49}$$

where v is the characteristic velocity of an eddy of scale l, and $l \sim \overline{L} = \overline{Y^2}^{1/2}$. Using the inertial range scaling $v \sim (l\varepsilon)^{1/3}$ this is

$$\mathcal{K}_{turb} \sim \varepsilon^{1/3}\overline{L}^{4/3}, \tag{13.50}$$

and so (13.48) becomes

$$\frac{\mathrm{d}\overline{L}^2}{\mathrm{d}t} \sim \varepsilon^{1/3}\overline{L}^{4/3}, \tag{13.51}$$

as before. Of course, dimensional consistency demands that we obtain the same result, but the derivation is intuitive and the estimate of the two-particle diffusivity (i.e., (13.50), that the eddy diffusivity governing the separation of two fluid parcels goes as the 4/3 power of their root-mean-square separation) is useful. If the particle separation is greater than the scale of the largest eddies in the system, l_{max}, then

$$K_{turb} \sim v(l_{max})l_{max} \sim \varepsilon^{1/3}l_{max}^{4/3} = \text{constant}. \tag{13.52}$$

The two-particle separation then proceeds as a conventional random walk or diffusive process, with the mean-square separation increasing linearly with time.

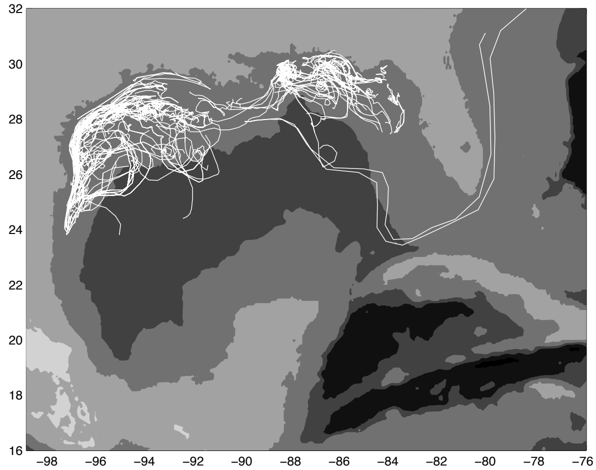

Fig. 13.3 Trajectories of surface drifters in the Gulf of Mexico, each truncated to produce paths of just 25 days. The drifters were released as part of 'SCULP' — the Surface CUrrent and Lagrangian drift Program.[6]

Diffusion in two-dimensional flow

In two dimensions the turbulent diffusivity will differ depending on whether the two-particle separation is in the energy inertial range or in the enstrophy inertial range. In the energy inertial range the scaling is the same as in the three-dimensional case, but in the enstrophy inertial range the rate of separation will depend on the enstrophy cascade rate, η, and the separation itself. Dimensional analysis then leads to

$$\frac{\mathrm{d}\overline{L}^2}{\mathrm{d}t} = B\eta^{1/3}\overline{L}^2, \tag{13.53}$$

where B is a nondimensional constant. This integrates to

$$\overline{L}^2 = \overline{L(0)}^2 \exp(B\eta^{1/3}t), \tag{13.54}$$

or $\overline{Y^2(t)} = \overline{Y^2(0)}\exp(B\eta^{1/3}t)$. Thus, the rate of separation is *exponential* in the enstrophy inertial range, a result unique to two-dimensional turbulence. Similarly, using $v \sim \eta^{1/3}l$, the turbulent diffusivity is given by

$$K_{turb} \sim \eta^{1/3}\overline{L}^2. \tag{13.55}$$

A geophysical example

The above ideas are well illustrated by analysing the trajectories of surface drifters in the Gulf of Mexico. The drifters are free-moving buoys which float about a half metre below the surface and which thus act as imperfect fluid markers — imperfect because they cannot follow the full three-dimensional motion of water parcels. Nevertheless, the motion at these scales can be expected to be quasi-geostrophic and nearly horizontal, so the associated error will be small. The drifters are tracked by satellite and their trajectories, proxies for the motion of fluid parcels, are shown in Fig. 13.3. The two-particle, or two-drifter, separation is illustrated in Fig. 13.4 and two regimes may be discerned. In the first, the pair separations grow approximately exponentially in time, with an e-folding time of 2 days, consistent with motion within an enstrophy inertial range using (13.54). The second regime is characterized by a power-law growth, proportional to $t^{2.2}$, somewhat slower than the t^3 separation expected for an energy inverse cascade using (13.47).[7] The boundary for the two regimes occurs at about 75 km, which is similar to the first deformation radius. No late-time diffusive regime (where the dispersion goes like t) is observed, suggesting that there exist long-time drifter correlations; these correlations arise because the separation of the drifters is never significantly larger than the energy-containing scale of the eddies themselves.

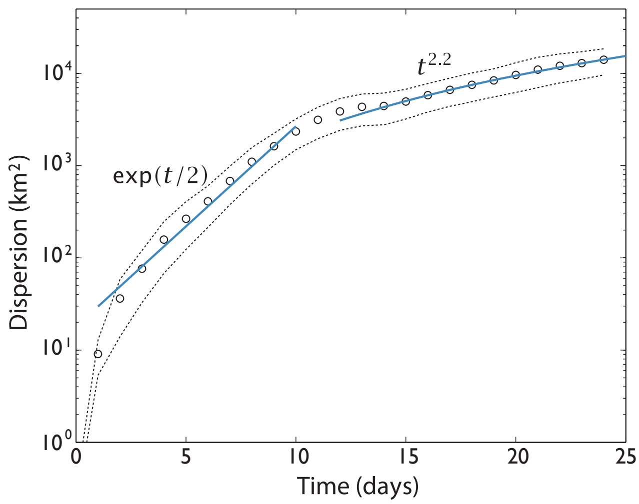

Fig. 13.4 Relative dispersion (the mean-square separation) for 140 drifter pairs as a function of time. The analysis utilizes all drifter pairs which come within 1 km of each other during their lifetimes.

In the atmosphere similar exponential separation of pairs of drifting balloons in the stratosphere at scales of less than 1000 km has been seen, consistent with an enstrophy inertial range. Evidence of a t^3 separation at larger scales, consistent with an energy inverse cascade, has been less forthcoming.[8]

13.4 MIXING LENGTH THEORY

The discussion of the previous two sections deals with the dispersion of marked fluid parcels. However, both for practical and fundamental reasons, we would like to be able to represent the turbulent transport of a fluid property in an Eulerian form. Thus, consider the equation for a conserved quantity φ in an incompressible turbulent flow:

$$\frac{\mathrm{D}\overline{\varphi}}{\mathrm{D}t} = -\nabla \cdot (\overline{\boldsymbol{v}'\varphi'}) + \kappa\nabla^2\overline{\varphi}, \tag{13.56}$$

where κ is the molecular diffusivity and the overbar denotes some kind of averaging or a filtering, so that $\overline{\varphi}$ represents only large scales. (We also adopt the convention that, unless noted, whenever the material derivative written as $\mathrm{D}/\mathrm{D}t$ is applied to an averaged field, the advection is by the averaged velocity only.) We expect the transport of φ to be enhanced by the turbulent flow and, as we saw in the previous sections, in some circumstances this transport will have a diffusive nature, completely overwhelming the molecular diffusivity. Let us consider this from an Eulerian angle, and by analogy with molecular mixing.

Given the mean distribution $\overline{\varphi}(x, y, z)$, let a fluid parcel be displaced from its mean position by a turbulent fluctuation. Suppose that the displaced parcel of fluid is able to carry its initial properties a distance l' before mixing with its surroundings. Then just prior to mixing with the environment the fluctuation of φ is given by, in the one-dimensional case,

$$\varphi' = -l'\frac{\partial\overline{\varphi}}{\partial x} - \frac{1}{2}l'^2\frac{\partial^2\overline{\varphi}}{\partial x^2} + \mathcal{O}(l'^3). \tag{13.57}$$

If the mean gradient is varying on a space-scale, L, that is larger than the mixing length l', that is if

$$L \equiv \frac{|\partial\overline{\varphi}/\partial x|}{|\partial^2\overline{\varphi}/\partial x^2|} \gg l', \tag{13.58}$$

then we can neglect terms in l'^2 and higher. The turbulent flux of φ-stuff is then given by

$$F = \overline{u'\varphi'} = -\overline{u'l'}\frac{\partial\overline{\varphi}}{\partial x}. \tag{13.59}$$

In more than one dimension, we have

$$\boldsymbol{F} = F_i = -\overline{v_i'l_j'}\partial_j\overline{\varphi} = -K_{ij}\partial_j\overline{\varphi}, \tag{13.60}$$

with summation over repeated indices, where $K_{ij} \equiv \overline{v_i'l_j'}$. The quantity K_{ij} (which we also write as $\boldsymbol{K}$) is known as the eddy (or turbulent) diffusivity tensor. At high Reynolds number it is a property of the flow rather than the fluid itself but, supposing that it can somehow be determined, the equation for the mean value of φ becomes

$$\frac{\mathrm{D}\overline{\varphi}}{\mathrm{D}t} = \nabla\cdot(\boldsymbol{K}\nabla\overline{\varphi}) = \partial_i(K_{ij}\partial_j\overline{\varphi}), \tag{13.61}$$

neglecting molecular diffusion.

Suppose that there exists a coordinate system in which the displacements in one direction ($l'^{\hat{x}}$, displacements in the $\hat{x}$-direction) are not correlated with the fluctuating velocity in another, orthogonal, direction (v', the velocity in the $\hat{y}$-direction) and for simplicity we restrict ourselves to two dimensions. Then, in that coordinate system $\boldsymbol{K}$ is symmetric and

$$\boldsymbol{K} = \begin{pmatrix} \overline{u'l'^{\hat{x}}} & \overline{u'l'^{\hat{y}}} \\ \overline{v'l'^{\hat{x}}} & \overline{v'l'^{\hat{y}}} \end{pmatrix} = \begin{pmatrix} \overline{u'l'^{\hat{x}}} & 0 \\ 0 & \overline{v'l'^{\hat{y}}} \end{pmatrix}. \tag{13.62}$$

The tensor may then, if needs be, be rotated to some other Cartesian coordinate system, but it will remain a symmetric tensor. In isotropic flow the two diagonal entries are equal and the equation of motion is,

$$\frac{\mathrm{D}\overline{\varphi}}{\mathrm{D}t} = \nabla\cdot(K\nabla\overline{\varphi}), \tag{13.63}$$

which is identical to the equation with molecular diffusion, save that the eddy diffusivity scalar, K, is different from the molecular diffusivity. To the extent, then, that K_{ij} is a symmetric tensor with constant entries, the turbulence acts like an enhanced diffusion. If the flow is homogeneous, then K does not vary spatially.

13.4.1 Requirements for Turbulent Diffusion

Turbulent diffusion evidently is a tractable and rational approach for parameterizing the effects of turbulent transport.[9] However, the premises required for the derivations above are not always satisfied and the derivation itself is rather heuristic, and turbulent diffusion is in no way a fundamental solution to the turbulence closure problem. Nonetheless, it can be an extremely useful parameterization in the appropriate circumstances, these being:

(i) There should be a scale separation between the mean gradient and the maximum mixing length, and the mixing length and decorrelation time scale should be well-defined.

(ii) The diffused property φ should be a materially conserved quantity, except for the effects of molecular diffusion.

(iii) The diffused property φ should be able to *mix* with its environment.

These are all largely self-evident from the derivation, but let us discuss items *(ii)* and *(iii)* a little more.

(ii) Material conservation of tracer

We assumed that a parcel of fluid carries its value of φ a distance, on average, equal to its mixing length before irreversibly mixing with its environment; this assumption is necessary in order that one may write $\varphi' = -l'\partial\overline{\varphi}/\partial x$. If φ is not materially-conserved over this scale other terms enter this formula. In particular, momentum is affected by the pressure force, and so is not normally a good candidate for turbulent diffusion. Potential vorticity is a better candidate, because it is a true material invariant, save for dissipative terms, and in large-scale geophysical flows potential vorticity also contains a great deal of the information about the flow. There is no ab initio requirement that the tracer be passive, and if it is not then its turbulent transport will affect the flow itself.

(iii) Tracer mixing and turbulent cascades

If a parcel cannot mix with its surroundings, then turbulent mixing cannot take place at all. Instead, we have what might be called turbulent stirring and if φ were, say, a dye then it would merely become threaded through the environment, producing streaks and swirls of colour rather than a truly mixed fluid. As another example, let φ be temperature and suppose that it has a mean gradient, so that temperature falls in the direction of increasing y. If a displaced parcel of fluid does not mix with or assume the value of its new environment at some stage, then there will be no correlation between the velocity producing the displacement and the value of the fluctuating quantity φ'. Suppose, for example, that an eddy causes parcels to be displaced from their mean positions. If a displaced parcel mixes with its surroundings, then a correlation will develop between v' and φ', and we would have $\overline{v'\varphi'} \neq 0$. However, if no mixing occurs, then the eddy simply recirculates with eddies retaining their initial values, and $\overline{v'\varphi'}$ is zero because of a lack of correlation between the two quantities. Thus, it is essential that there be a degree of *irreversibility* to the flow in order for turbulent diffusion to be appropriate.

Molecular diffusion is not the only process that enables an eddy to assume the value of its surroundings — a Newtonian or other relaxation back to a specified temperature may have much the same effect. Indeed, in the atmosphere a displaced parcel will be subject to a radiation field that acts qualitatively in this way. For example, suppose that the temperature equation is

$$\frac{\mathrm{D}T}{\mathrm{D}t} = -\lambda(T - T^*(y,z)), \tag{13.64}$$

where the right-hand side crudely represents radiative effects via a relaxation back to a specified profile. Then a displaced parcel will be subject to a radiative damping that is different from that at its initial position, and this will allow the parcel to take on the value of its surroundings, and so potentially enable turbulent diffusion to occur (provided λ is small so that T is approximately materially conserved).

For molecular diffusion to be the mechanism whereby a parcel mixes with its surroundings, the turbulence must create scales that are small enough for diffusion to act. This means that turbulence must create a cascade of φ-stuff to small scales. This is quite consistent with the notion that φ is a materially conserved quantity, because a scalar field φ that satisfies

$$\frac{\mathrm{D}\varphi}{\mathrm{D}t} = F + \kappa\nabla^2\varphi, \tag{13.65}$$

where F might represent a spectrally local source of variance of φ, is certainly cascaded to smaller scales. The presence of a molecular diffusion does not substantially affect the requirement that φ be conserved on parcels, because on scales comparable to the eddy mixing length the effect of molecular diffusion is negligible. (And if it were not, perhaps because κ was extremely large or because the turbulence was anaemic, we would not be particularly interested in the turbulent transport.)

13.4.2 A Macroscopic Perspective

Consider turbulent diffusion from a more macroscopic point of view, and in particular consider the transport of a nearly materially conserved tracer obeying

$$\frac{\mathrm{D}\varphi}{\mathrm{D}t} = D, \tag{13.66}$$

where the advecting flow is incompressible and D is a dissipative process such that $\overline{D\varphi} \leq 0$ (a conventional harmonic diffusion has this property). By decomposing the fields into mean and eddy components in the usual way an equation for the evolution of the tracer variance can be straightforwardly derived, namely

$$\frac{1}{2}\frac{\partial}{\partial t}\overline{\varphi'^2} + \overline{\boldsymbol{v}'\varphi'} \cdot \nabla\overline{\varphi} + \frac{1}{2}\overline{\boldsymbol{v}} \cdot \nabla\overline{\varphi'^2} + \frac{1}{2}\nabla \cdot \overline{\boldsymbol{v}'\varphi'^2} = \overline{D'\varphi'}, \tag{13.67}$$

and where we may assume $\overline{D'\varphi'} < 0$. If the mean flow is small and if the third-order term may be neglected then in a statistically steady state we have

$$\overline{\boldsymbol{v}'\varphi'} \cdot \nabla\overline{\varphi} \approx \overline{D'\varphi'} < 0. \tag{13.68}$$

Therefore, on average, the flux of φ is downgradient in regions of dissipation, implying a positive average eddy diffusivity, and a balance is maintained between the downgradient flux of φ (which increases the variance) and dissipation. However, it should also be clear from (13.67) that if the turbulence is not statistically stationary, or if there is a mean flow, then downgradient transport cannot necessarily be expected. Indeed, the transport may be upgradient in regions where the eddy variance is falling, for then we may have the balance

$$\overline{\boldsymbol{v}'\varphi'} \cdot \nabla\overline{\varphi} \approx -\frac{1}{2}\frac{\partial}{\partial t}\overline{\varphi'^2} > 0. \tag{13.69}$$

13.5 HOMOGENIZATION OF A SCALAR THAT IS ADVECTED AND DIFFUSED

Let us now assume that the effects of turbulence on a tracer are indeed diffusive. An important consequence of this is that, in the absence of additional forcing, there can be no extreme values of the tracer in the interior of the fluid and, in some circumstances, the diffusion will *homogenize* values of the tracer in broad regions. In this section we demonstrate and explore these properties.

13.5.1 Non-existence of Extrema

Consider a tracer that obeys the equation

$$\frac{\mathrm{D}\varphi}{\mathrm{D}t} = \nabla \cdot (\kappa\nabla\varphi) + S, \tag{13.70}$$

where $\kappa > 0$ and the advecting velocity is divergence-free. We now show that in regions where the source term, S, is zero there can be no interior extrema of φ if the flow is steady. The proof is in the form of a *reductio ad absurdum* argument — we first suppose there *is* an extrema of φ in the fluid, and show a contradiction.

Given an extremum, there will then be a surrounding surface (in three dimensions), or a surrounding contour (in two), connecting constant values of φ. For definiteness consider two-dimensional incompressible flow for which the steady flow satisfies

$$\nabla \cdot (\boldsymbol{u}\varphi) = \nabla \cdot (\kappa\nabla\varphi). \tag{13.71}$$

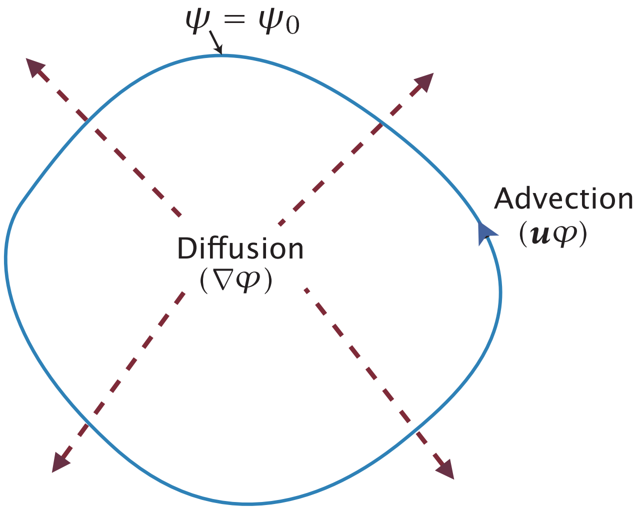

Fig. 13.5 If an extremum of a tracer φ exists in the fluid interior, then diffusion will provide a downgradient tracer flux. But over an area bounded by a streamline, or by an isoline of φ, the net advective flux is zero. Thus, the diffusion cannot be balanced by advection and so in a steady state no extrema can exist.

Integrating the left-hand side over the area, A, enclosed by the iso-line of φ, and applying the divergence theorem, gives

$$\iint_A \nabla\cdot(\boldsymbol{u}\varphi)\,\mathrm{d}A = \oint(\boldsymbol{u}\varphi)\cdot\boldsymbol{n}\,\mathrm{d}l = \varphi\oint\boldsymbol{u}\cdot\boldsymbol{n}\,\mathrm{d}l = \varphi\iint_A \nabla\cdot\boldsymbol{u}\,\mathrm{d}A = 0, \tag{13.72}$$

where $\boldsymbol{n}$ is a unit vector normal to the contour. (If we were to integrate over an area bounded by a velocity contour, then $\boldsymbol{u}\cdot\boldsymbol{n} = 0$ and the integral would similarly vanish.) But the integral of the right-hand side of (13.71) over the same area is non-zero; that is

$$\iint_A \nabla\cdot(\kappa\nabla\varphi)\,\mathrm{d}A = \oint\kappa\nabla\varphi\cdot\boldsymbol{n}\,\mathrm{d}l \neq 0, \tag{13.73}$$

if the integral surrounds an extremum. This is a contradiction for steady flow. Hence, there can be no isolated extrema of a conserved quantity in the interior of a fluid, if there is any diffusion at all. The result (which applies in two or three dimensions) is kinematic, in that φ can be any tracer at all, active or passive. The physical essence of the result is that the integrated effects of diffusion are non-zero surrounding an extremum, and cannot be balanced by advection. Thus, if the initial conditions contain an extremum, diffusion will smooth away the extremum until it no longer exists. This process, and the homogenization discussed below, are illustrated in Fig. 13.5.

13.5.2 Homogenization in Two-dimensional Flow

For two-dimensional flow we can obtain a still stronger result if we allow ourselves to make more assumptions about the strength and nature of the diffusion. The steady distribution of a scalar quantity being advected by an incompressible flow is governed by

$$J(\psi,\varphi) = \nabla\cdot(\kappa\nabla\varphi) + S, \tag{13.74}$$

where the terms on the right-hand side represent diffusion and source terms. Suppose that these terms are small, in the sense that the individual terms on the left-hand side nearly balance each other, so that

$$|J(\psi,\varphi)| \ll \frac{U|\varphi|}{L}. \tag{13.75}$$

This means we are in the high Peclet number limit ($P = UL/\kappa \gg 1$), and the dominance of advection suggests that any steady solution to (13.74) is of the form

$$\varphi = G(\psi) + \mathcal{O}(P^{-1}), \tag{13.76}$$

where G is (for the moment) any function of its argument. Thus, isolines of φ are nearly coincident with streamlines, and

$$\nabla\varphi \approx \nabla\psi \frac{\mathrm{d}\varphi}{\mathrm{d}\psi}. \tag{13.77}$$

On integrating (13.74) over the area, A, bounded by some closed streamline, $\psi = \psi_0$ say, the left-hand side vanishes and we obtain

$$0 = \iint_A S\,\mathrm{d}A + \oint_{\psi_0} \kappa\nabla\varphi \cdot \boldsymbol{n}\,\mathrm{d}l. \tag{13.78}$$

Using (13.77) then gives

$$\iint_A S\,\mathrm{d}A = -\oint_{\psi_0} \kappa \frac{\mathrm{d}\varphi}{\mathrm{d}\psi} \nabla\psi \cdot \boldsymbol{n}\,\mathrm{d}l. \tag{13.79}$$

Since $\mathrm{d}\varphi/\mathrm{d}\psi$ is constant along streamlines, and using $\boldsymbol{u} = \nabla^{\perp}\psi$, we have

$$\frac{\mathrm{d}\varphi}{\mathrm{d}\psi} = -\frac{\iint S\,\mathrm{d}A}{\oint_{\psi_0} \kappa\boldsymbol{u} \cdot \mathrm{d}\boldsymbol{l}}. \tag{13.80}$$

This relationship determines φ as a function of ψ — that is, it determines $G(\psi)$ — in terms of the forcing and dissipation acting on the fluid. If the fluid is both unforced and inviscid, then a steady solution obtains when φ is an *arbitrary* function of ψ. If the source term S is zero, but dissipation is non-zero then the denominator of (13.80) is non-zero and therefore φ must be uniform: φ has been *homogenized*. The homogenization result also follows if we choose to integrate over an area surrounded by an isoline of φ, φ_0 say, but we leave this as a problem for the reader.

Interpretation

The homogenization result applies to a statistically steady flow in which the eddy transport of φ-stuff by the eddying motion may be parameterized diffusively, and in which there is an approximate functional relationship between mean φ and mean ψ. The first of these assumptions we have discussed at length in previous sections. The second requires that the diffusion must not be too strong, so that locally the tracer is conserved on fluid parcels. In the steady state the tracer is then a function of the streamfunction, the same function everywhere within the closed region.

Given these assumptions, the dynamics giving rise to homogenization is transparent: integrating round a contour of ψ or φ the effect of the advective terms vanishes; the source (S) and the diffusion must balance each other, and if there is no source term there can be no tracer gradient. Put another way, the flow will circulate endlessly and steadily around the contours of ψ, which nearly coincide with contours of φ. Advection cannot alter the mean value of φ, so diffusion smooths out gradients within the closed contours, effectively *expelling* gradients of φ to the boundaries and forming a plateau of φ-values. Because extrema of φ are forbidden, the value of φ on the plateau cannot be a maximum or minimum: at the edge of the plateau the values of φ must fall somewhere, and rise somewhere else. The plateau can be a flat region etched out of a hillside, but a plateau on top of a butte is forbidden, for in that case diffusion would erode the butte down to the level of the surrounding land. Our derivation makes no distinction between a passive scalar like a dye and an active scalar, like potential vorticity. In reality, in the latter case the dynamics will further constrain the flow because the scalar distribution must be consistent with the velocity field that advects it, and this is particularly important in the dynamics of ocean gyres.

13.6 ✦ DIFFUSIVE FLUXES AND SKEW FLUXES

Thus far we have considered diffusion using a scalar diffusivity, and this is what is usually meant by diffusion. However, if only from a mathematical point of view, we may allow the diffusivity to be a tensor and this turns out to be very useful when considering the effects of baroclinic eddies. Having a tensor diffusivity allows for the possibility of *skew fluxes,* which are perpendicular to the gradient of the diffused quantity, unlike conventional diffusive fluxes, which are downgradient.

13.6.1 Symmetric and Antisymmetric Diffusivity Tensors

A tracer evolving freely save for the effects of molecular diffusivity, κ_m, obeys the equation

$$\frac{\mathrm{D}\varphi}{\mathrm{D}t} = \nabla\cdot(\kappa_m\nabla\varphi), \tag{13.81}$$

where κ_m is a positive scalar quantity. In the more general case we might have

$$\frac{\mathrm{D}\varphi}{\mathrm{D}t} = -\nabla\cdot\boldsymbol{F} = \nabla\cdot\mathbf{K}\nabla\varphi, \tag{13.82}$$

where $\mathbf{K}$ is (if φ is a scalar) a second-rank tensor and $\boldsymbol{F} = -\mathbf{K}\nabla\varphi$ is the diffusive flux of φ. The flux has a component across the isosurfaces of φ, called the diffusive flux, and a component along the iso-surfaces, called the skew flux We will see that these fluxes are associated with the symmetric and antisymmetric components of the diffusivity tensor, respectively, where

$$\mathbf{K} = \mathbf{S} + \mathbf{A}, \tag{13.83}$$

and, using component notation,

$$S_{mn} = \frac{1}{2}(K_{mn} + K_{nm}), \qquad A_{mn} = \frac{1}{2}(K_{mn} - K_{nm}). \tag{13.84}$$

The diagonal elements of the antisymmetric tensor are zero. The transport that is effected by these two tensors has different physical characteristics, as we now discuss.

Diffusion with the symmetric tensor

In the simplest case of all, with an isotropic medium, $\mathbf{K}$ is diagonal with equal entries,

$$\mathbf{K} = \mathbf{S} = \begin{pmatrix} \kappa & 0 & 0 \\ 0 & \kappa & 0 \\ 0 & 0 & \kappa \end{pmatrix}, \tag{13.85}$$

and we have the familiar $\boldsymbol{F} = -\kappa\nabla\varphi$, and (13.82) has the same form as (13.81). If κ is positive, then the flux is *downgradient*, meaning that

$$\boldsymbol{F}\cdot\nabla\varphi < 0, \tag{13.86}$$

even if κ is spatially non-uniform. Furthermore, such a diffusion is variance-dissipating; to see this, suppose we have the equation of motion

$$\frac{\mathrm{D}\varphi}{\mathrm{D}t} = \nabla\cdot(\kappa\nabla\varphi). \tag{13.87}$$

Multiplying by φ and integrating over the domain V gives

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_V \varphi^2\,\mathrm{d}V = \int_V \boldsymbol{F}\cdot\nabla\varphi\,\mathrm{d}V = -\int_V \kappa(\nabla\varphi)^2\,\mathrm{d}V \le 0, \tag{13.88}$$

after an integration by parts and assuming that the normal derivative of φ vanishes at the boundaries; that is, there is no flux of φ-stuff through the boundary. However, diffusion does preserve the first moment of the field; that is

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_V \varphi \, \mathrm{d}V = \int_V \nabla\cdot(\kappa\nabla\varphi)\,\mathrm{d}V = 0, \tag{13.89}$$

again assuming no flux through the boundaries.

The transport that is effected by the symmetric diffusion tensor is the diffusive flux, $\boldsymbol{F}_d$, where

$$\boldsymbol{F}_d = -\mathbf{S}\nabla\varphi = -S_{mn}\partial_n\varphi, \tag{13.90}$$

where we employ the common convention that repeated indices are summed. In general, the flux has a component that is parallel to the tracer gradient; that is, $\boldsymbol{F}_d\cdot\nabla\varphi \neq 0$. Suppose we have the simple equation of motion

$$\frac{\partial\varphi}{\partial t} = -\nabla\cdot\boldsymbol{F}_d = \nabla\cdot(\mathbf{S}\nabla\varphi). \tag{13.91}$$

This equation preserves the first moment of φ, provided there is no flux through the boundary. Tracer variance evolves according to

$$\frac{1}{2}\frac{\partial}{\partial t}\int_V \varphi^2\,\mathrm{d}V = \int_V \varphi\nabla\cdot(\mathbf{S}\nabla\varphi) = -\int_V(\mathbf{S}\nabla\varphi)\cdot\nabla\varphi\,\mathrm{d}V. \tag{13.92}$$

This can be shown to be negative or zero, provided that $\mathbf{S}$ is positive semi-definite, meaning that

$$\nabla\varphi\,\mathbf{S}\,\nabla\varphi = \partial_m\varphi S_{mn}\partial_n\varphi \geq 0. \tag{13.93}$$

The flux effected by such a diffusivity is then downgradient in the sense that

$$\boldsymbol{F}_d\cdot\nabla\varphi = -\mathbf{S}\nabla\varphi\cdot\nabla\varphi \leq 0. \tag{13.94}$$

The skew flux

The transport associated with the antisymmetric transport tensor is perpendicular to the gradient of φ, and so is neither upgradient nor downgradient. The flux is

$$\boldsymbol{F}_{sk} = -\mathbf{A}\nabla\varphi = -A_{mn}\partial_n\varphi, \tag{13.95}$$

and thus

$$\boldsymbol{F}_{sk}\cdot\nabla\varphi = -\mathbf{A}\nabla\varphi\cdot\nabla\varphi = -A_{mn}\partial_n\varphi\partial_m\varphi = 0, \tag{13.96}$$

where the final result follows because of the antisymmetry of $\mathbf{A}$ — the contraction of a symmetric tensor and an antisymmetric tensor is zero.[10] For this reason, the associated transport is known as a *skew flux* (a term applying in general to fluxes that are perpendicular to the tracer gradient) or a *skew diffusion* (when those fluxes are parameterized using an antisymmetric diffusivity). It follows from this that if a tracer obeys

$$\frac{\partial\varphi}{\partial t} = \nabla\cdot(\mathbf{A}\nabla\varphi), \tag{13.97}$$

then the tracer variance is conserved. This may be verified by multiplying this equation by φ and integrating by parts, assuming that the flux vanishes at the boundaries. That is, *a skew diffusion has no effect on the variance of the skew diffused variable.* One other familiar physical process shares these properties, and that is advection by a divergence-free flow. A skew diffusion is physically equivalent to such an advection in that the divergence of a skew diffusive flux is the same as the divergence of an appropriately chosen advective flux.

To explore this more, define an advective flux of a tracer φ to be a flux of the form

$$\boldsymbol{F}_{ad} \equiv \tilde{\boldsymbol{v}}\varphi, \tag{13.98}$$

where $\tilde{\boldsymbol{v}}$ is a divergence-free vector field. The divergence of the flux is just

$$\nabla \cdot \boldsymbol{F}_{ad} = \nabla \cdot (\tilde{\boldsymbol{v}}\varphi) = \tilde{\boldsymbol{v}} \cdot \nabla\varphi. \tag{13.99}$$

The field $\tilde{\boldsymbol{v}}$ might be called a *pseudovelocity* or a *quasi-velocity* — it acts like a velocity but is not necessarily the velocity of any fluid particle. Because $\tilde{\boldsymbol{v}}$ is divergence-free, we may define a vector streamfunction $\boldsymbol{\psi}$ such that

$$\tilde{\boldsymbol{v}} = \nabla \times \boldsymbol{\psi} \qquad \text{or} \qquad \tilde{v}_n = \epsilon_{lmn}\partial_l \psi_m. \tag{13.100}$$

The Levi–Civita symbol ϵ_{lmn} is such that $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1$, $\epsilon_{132} = \epsilon_{321} = \epsilon_{213} = -1$, and $\epsilon_{lmn} = 0$ for other combinations. The equivalence of the two expressions may be verified by expansion in Cartesian coordinates. The field $\boldsymbol{\psi}$ is not unique: the gradient of an arbitrary function may be added to it, this gradient vanishing upon taking the curl, resulting in the same velocity field. That is, if $\boldsymbol{\psi}' = \boldsymbol{\psi} + \nabla\gamma$, then $\tilde{\boldsymbol{v}} = \nabla \times \boldsymbol{\psi} = \nabla \times \boldsymbol{\psi}'$. The scalar field γ is known as the *gauge,* and the freedom to choose it is the gauge freedom.

The advective flux $\boldsymbol{F}_{ad}$ is related to the skew flux $\boldsymbol{F}_{sk}$ by

$$\varphi\tilde{\boldsymbol{v}} = \varphi\nabla \times \boldsymbol{\psi} = \nabla \times (\varphi\boldsymbol{\psi}) - \nabla\varphi \times \boldsymbol{\psi}, \tag{13.101}$$

or

$$\boldsymbol{F}_{ad} = \boldsymbol{F}_r + \boldsymbol{F}_{sk}, \tag{13.102}$$

where $\boldsymbol{F}_r = \nabla \times (\varphi\boldsymbol{\psi})$ is a rotational flux with no divergence, and

$$\boldsymbol{F}_{sk} = -\nabla\varphi \times \boldsymbol{\psi} \tag{13.103}$$

is a skew flux — 'skewed' because it is manifestly orthogonal to the gradient of φ, i.e., $\nabla\varphi \cdot \boldsymbol{F}_{sk} = 0$. Because $\nabla \cdot \boldsymbol{F}_{\mathrm{r}} = 0$ the divergence of the skew flux and advective flux are equal:

$$\boxed{\nabla \cdot \boldsymbol{F}_{ad} = \nabla \cdot \boldsymbol{F}_{sk}.} \tag{13.104}$$

However, the skew flux, $-\nabla\varphi \times \boldsymbol{\psi}$, and the advective flux, $\varphi\nabla \times \boldsymbol{\psi}$, may have, and in general do have, different magnitudes and directions; only their divergences are equal. If the divergences of the skew fluxes given by (13.95) and (13.103) are to be the same then $\boldsymbol{\psi}$ must be related to the antisymmetric tensor A. Using (13.95) we have

$$\begin{aligned} \nabla \cdot \boldsymbol{F}_{sk} &= -\partial_m(A_{mn}\partial_n\varphi) \\ &= -(\partial_n\varphi)(\partial_m A_{mn}) - [A_{mn}\partial_n\partial_m\varphi] \\ &= -\partial_n(\varphi\partial_m A_{mn}) + [\varphi\partial_n\partial_m A_{mn}], \end{aligned} \tag{13.105}$$

where the quantities in square brackets are zero as a consequence of the antisymmetry of A — a symmetric operator acting on an antisymmetric tensor is zero. But the skew flux divergence is equal to the advective flux divergence

$$\nabla \cdot \boldsymbol{F}_{sk} = \nabla \cdot \boldsymbol{F}_{ad} = \partial_n(\varphi\tilde{v}_n), \tag{13.106}$$

so the associated skew velocity is related to the antisymmetric tensor

$$\boxed{\tilde{v}_n = -\partial_m A_{mn},} \tag{13.107}$$

and this is divergence-free because $\partial_n \partial_m A_{mn} = 0$. The streamfunction and the antisymmetric tensor are thus related.

Using (13.107) and (13.100), and just a little algebra, gives

$$A_{mn} = \epsilon_{mnp}\psi_p = \begin{pmatrix} 0 & \psi_3 & -\psi_2 \\ -\psi_3 & 0 & \psi_1 \\ \psi_2 & -\psi_1 & 0 \end{pmatrix}, \tag{13.108}$$

which provides an explicit connection between the antisymmetric tensor A_{mn} and the streamfunction for the skew velocity $\tilde{\boldsymbol{v}}$. Thus, to summarize:

- Any flux can be decomposed into a component across iso-surfaces of a scalar (the along-gradient or diffusive flux) and a component along isosurfaces and so perpendicular to the gradient (the skew flux).
- The along-gradient (usually downgradient) flux is effected by a diffusion using a symmetric diffusivity tensor.
- The skew flux is effected by a diffusion using an antisymmetric diffusivity tensor, and this is equivalent to an advection by some divergence-free velocity.
- The diffusive flux reduces tracer variance if the diffusivity is positive (in which case the diffusion is downgradient), whereas the skew diffusion has no effect on variance.

Let us now consider how all of this is relevant to the large-scale flow in the atmosphere and the ocean.

13.7 † EDDY DIFFUSION IN THE ATMOSPHERE AND OCEAN

We now, rather heuristically, discuss the transport of fluid properties by large-scale eddies typically generated by baroclinic instability — mesoscale eddies in the ocean, and weather systems in the atmosphere. The practical motivation is perhaps more oceanographic than atmospheric. Specifically, mesoscale and submesoscale eddies in the ocean cannot be easily resolved in numerical models of its large-scale circulation, especially those used for climate simulations involving integrations of the global ocean over decades and centuries. In such models, the effects of eddies must be *parameterized* in terms of properties of the mean flow.

In the end this problem will be solved for us by the increasing power of computers, as it largely has in atmospheric flows since numerical models of the general circulation already resolve most of the effects of baroclinic eddies. However, we then have to deal with the hardly less difficult problem of understanding those massive, turbulent, numerical integrations, and for that task a theory of turbulent transport is a *sine qua non.*

13.7.1 Preliminaries

Consider a tracer that obeys the advective-diffusive equation

$$\frac{\mathrm{D}\varphi}{\mathrm{D}t} = \nabla \cdot (\kappa_m \nabla\varphi). \tag{13.109}$$

If the advecting flow is divergence-free then the ensemble average or filtered flow obeys, neglecting the molecular diffusion,

$$\frac{\mathrm{D}\overline{\varphi}}{\mathrm{D}t} = -\nabla \cdot \overline{\boldsymbol{v}'\varphi'}, \tag{13.110}$$

where the right-hand side is the eddy transport (akin to Reynolds stresses). If we parameterize this transport by a diffusion then

$$\overline{\boldsymbol{v}'\varphi'} = -\boldsymbol{K}\nabla\overline{\varphi}, \tag{13.111}$$

where $\boldsymbol{K}$ is, in general, a second-rank tensor. If, say, the average is a zonal average then

$$\frac{\mathrm{D}\overline{\varphi}}{\mathrm{D}t} = -\frac{\partial\overline{v'\varphi'}}{\partial y} - \frac{\partial\overline{w'\varphi'}}{\partial z}. \tag{13.112}$$

If we are to employ a diffusive parameterization for the eddy terms in these equations, the issues that then arise fall into two general camps:

(i) the overall *magnitude* of the eddy diffusivity, possibly as a function of the mean flow;

(ii) the *structure* of the diffusivity tensor, and in particular the separate structure of its symmetric and antisymmetric parts.

13.7.2 Magnitude of the Eddy Diffusivity

The magnitude and scale of eddies were considered in Sections 12.3 and 12.4. Here we see how these give rise to corresponding estimates of the magnitude of an eddy diffusivity. If we restrict attention for the moment to the meridional transfer of tracer properties, then we might write

$$\overline{v'\varphi'} = -\kappa^{vy}\frac{\partial\overline{\varphi}}{\partial y} - \kappa^{vz}\frac{\partial\overline{\varphi}}{\partial z}, \tag{13.113}$$

where κ^{vy} and κ^{vz} are components of the eddy diffusivity tensor with obvious notation. These components have the dimensions of a length times a velocity and, to the extent that the diffusion represents the eddying motion we expect that κ^{vy} has an approximate magnitude of

$$\kappa^{vy} \sim v'l', \tag{13.114}$$

where v' is a typical magnitude of the horizontal eddy velocity, and l' is the *mixing length* of the eddies, generally taken to be a typical length scale of the eddies. Larger and more energetic eddies thus have a larger effect on the mean flow. We can estimate v' and l' in a number of reasonable ways depending on the flow conditions, and we consider a few such below.[11] The magnitude of the component κ^{vz} may then be estimated by making choices about the plane of parcel displacements, and this is considered in the next section.

Perhaps the simplest assumption to make follows from the fact that the eddies are a consequence of baroclinic instability, and so one might suppose that the eddy length scale is the scale of the instability — the first deformation radius. One might also suppose that the eddy velocity is of the same approximate magnitude as the mean flow, $\overline{u}$, thus giving

$$\kappa^{vy} \sim L_d\overline{u} = \frac{NH\overline{u}}{f}. \tag{13.115}$$

Another way of deriving this result is by noting that $\kappa^{vy} \sim l'^2/T_e$, where T_e is a characteristic eddy time scale. If T_e is the Eady time scale, $L_d/\overline{u}$, and if $l' \sim L_d$, we reproduce (13.115). Equation (13.115) may be written as

$$\kappa^{vy} \sim L_d\overline{u} \sim \frac{L_d^2 f}{\sqrt{Ri}} \sim L_d^2 Fr f, \tag{13.116}$$

where $Ri \equiv N^2/\Lambda^2 = N^2H^2/\overline{u}^2$ and $Fr \equiv U/(NH)$ are the Richardson and Froude numbers for this problem, respectively.

A little more generally, if there is a cascade to larger scales then the eddy scale, L_e say, may be larger than the deformation scale. Depending on circumstances, L_e might be the domain scale (if eddies grow to the size of the domain), the β-scale (if the β-effect halts the cascade), or some scale determined by frictional effects (possibly in conjunction with β). However, the arguments of

Section 12.3 suggest that the eddy time scale is the Eady time scale in all cases. We therefore have:

$$\text{eddy length scale} \sim L_e, \tag{13.117a}$$
$$\text{eddy time scale, } T_e \sim L_d/\overline{u}, \tag{13.117b}$$
$$\text{eddy velocity scale, } U_e \sim \overline{u}(L_e/L_d). \tag{13.117c}$$

These give the general estimate for the horizontal diffusivity of

$$\boxed{\kappa^{vy} \sim \overline{u}\left(\frac{L_e^2}{L_d}\right).} \tag{13.118}$$

The estimate (13.115) is a special case of this, with $L_e = L_d$; the two estimates will thus differ if the eddy scale is much larger than the deformation radius.

In the case in which the inverse cascade is modified by the Rossby waves we might (and neglecting friction effects; see Section 12.1.2) suppose that the eddy scale is the β-scale, (12.6), and we have

$$L_e \sim L_\beta = \left(\frac{U_e}{\beta}\right)^{1/2} = \frac{\overline{u}}{\beta L_d}, \tag{13.119}$$

using (13.117c). The eddy velocity scale is, using (13.117b),

$$U_e \sim \overline{u}\frac{L_e}{L_d} = \frac{\overline{u}^2}{\beta L_d^2}, \tag{13.120}$$

and combining (13.119) and (13.120) gives the estimate for the eddy diffusivity,

$$\kappa^{vy} \sim \frac{\overline{u}^3}{\beta^2 L_d^3}. \tag{13.121}$$

A similar estimate can be written in terms of the inverse energy cascade rate, ε, giving

$$\kappa^{vy} \sim \left(\frac{\varepsilon^3}{\beta^4}\right)^{1/5}. \tag{13.122}$$

This expression may be obtained purely by dimensional analysis, if it is assumed that the only factors determining κ are ε and β. The estimate may be useful if ϵ is known independently, for example by calculating the energy throughput in the system.

To summarize: the magnitude of any eddy diffusion may be estimated as the product of the velocity scale and the energy-containing length scale of the eddies. If we assume that the time scale is the Eady time scale we obtain (13.118), where L_e is undetermined. If the eddy scale is the β-scale, then (13.118) becomes (13.121). However, in neither the atmosphere nor the ocean is the β-scale significantly (e.g., an order of magnitude) larger than the deformation scale, but nor, complicating matters, does the inverse cascade necessarily halt at the β-scale (see Section 12.1.2). From an observational standpoint, the atmosphere has no $-5/3$ inverse cascade, although there is some evidence for one in some regions of ocean.[12] In other oceanic regions the eddies may be advected away from each other and away from the unstable zone, or dispersed by Rossby waves, before an inverse cascade can be organized, and energy will remain at the deformation scale. These arguments suggest that although we can make sensible estimates, we cannot determine with certainty what the magnitude of an eddy diffusivity should be, in either the atmosphere or ocean.

13.7.3 ✦ Use of the Symmetric Transport Tensor

Along-gradient transport, or diffusion, is transport by a symmetric transport tensor as in (13.90). The structure of the eddy diffusivity will determine, among other things, the surface along which transport occurs; for example, a diffusion of temperature might occur meridionally and/or vertically. To illustrate, let us consider transfer in a re-entrant channel with zonally homogeneous eddy statistics so that the averaging operator is the zonal average; the meridional and upward transport of a tracer φ are then given by

$$\overline{v'\varphi'} = -\kappa^{vy}\frac{\partial\overline{\varphi}}{\partial y} - \kappa^{vz}\frac{\partial\overline{\varphi}}{\partial z}, \tag{13.123}$$

$$\overline{w'\varphi'} = -\kappa^{wy}\frac{\partial\overline{\varphi}}{\partial y} - \kappa^{wz}\frac{\partial\overline{\varphi}}{\partial z}, \tag{13.124}$$

where $\kappa^{wy} = \kappa^{vz}$ by the posited symmetry. The relationship between the various transfer coefficients will be determined by the trajectories of the fluid parcels in the eddying motion. In the Cartesian y–z frame the transport tensor is not necessarily diagonal (i.e., κ^{vz} and κ^{wy} may be nonzero) but locally there is always a natural coordinate system in which the diffusivity tensor is diagonal. (A symmetric matrix may always be diagonalized by a suitable rotation of axes.) In that diagonal frame we can write

$$\mathsf{S}' = \kappa_s\begin{pmatrix}1 & 0\\ 0 & \alpha\end{pmatrix}, \tag{13.125}$$

where κ_s determines the overall size of the transfer coefficients (as estimated in the previous section), and α is the ratio of sizes of the components in the two orthogonal directions. Now, fluid displacements in large-scale baroclinic eddies are nearly, but not exactly, horizontal — they may be along isopycnals, for example, or at an angle between the horizontal and the isopycnals. We may argue that the coordinate system in which the tensor is diagonal is the coordinate system defined by the plane along which fluid displacements occur. This is sensible because the transfers along and orthogonal to the fluid paths are each a consequence of different physical phenomena, and so we may expect the transfer tensor to be diagonal in these coordinates.

Because eddy displacements are predominantly horizontal, the diagonal coordinate system has a small slope, s, at an angle θ with respect to the horizontal, where $s = \tan\theta \approx \theta \ll 1$. Furthermore, we expect the parameter α to be small (i.e., $\alpha \ll 1$), because this represents transfer in a direction orthogonal to the eddy fluid motion. We rotate the tensor S' through an angle θ to move into the usual y-z frame; that is

$$\mathsf{S} = \kappa_s\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & \alpha\end{pmatrix}\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix} \tag{13.126a}$$

$$\approx \kappa_s\begin{pmatrix}1+s^2\alpha & s(1-\alpha)\\ s(1-\alpha) & s^2+\alpha\end{pmatrix} \qquad \text{(for small } s\text{)} \tag{13.126b}$$

$$\approx \kappa_s\begin{pmatrix}1 & s\\ s & s^2+\alpha\end{pmatrix} \qquad \text{(for small } s \text{ and small } \alpha\text{).} \tag{13.126c}$$

We can follow the same procedure in three dimensions. Then, if the eddy transport is isotropic in the plane of eddy displacements, the three-dimensional transport tensor is

$$\mathsf{S}' = \kappa_s\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & \alpha\end{pmatrix}, \tag{13.127}$$

and the slope of the motion is a two-dimensional vector $\boldsymbol{s} = (s^x, s^y)$, with the superscripts denoting components. If we rotate the transport tensor into physical space then we obtain, analogously to

(13.126),

$$\mathsf{S} = \kappa_s \begin{pmatrix} 1 + s^{y2} + \alpha s^{x2} & (\alpha-1)s^x s^y & (1-\alpha)s^x \\ (\alpha-1)s^x s^y & 1 + s^{x2} + \alpha s^{y2} & (1-\alpha)s^y \\ (1-\alpha)s^x & (1-\alpha)s^y & \alpha + s^2 \end{pmatrix} \tag{13.128a}$$

$$\approx \kappa_s \begin{pmatrix} 1 & 0 & s^x \\ 0 & 1 & s^y \\ s^x & s^y & \alpha + s^2 \end{pmatrix}, \tag{13.128b}$$

for small s and small α, where $s^2 = s^{x2} + s^{y2}$.

The plane of eddy displacements

We are now in a position to make heuristic choices about the transfer coefficients, and we will consider two bases for this:

I. *Using linear baroclinic instability theory.*[13] In a simple model of a growing baroclinic (Eady) wave, parcel trajectories that are along half the slope of the mean isopycnals are able to release the most potential energy. We thus suppose that $s = s_\rho/2$, where s_ρ is the isopycnal slope, and that $\alpha = 0$ in (13.126c) or (13.128b). In two dimensions this gives

$$\mathsf{S} = \kappa_s \begin{pmatrix} 1 & s_\rho/2 \\ s_\rho/2 & s_\rho^2/4 \end{pmatrix}, \tag{13.129}$$

and so

$$\overline{v'\varphi'} = -\kappa_s \left(\frac{\partial \overline{\varphi}}{\partial y} + \frac{1}{2} s_\rho \frac{\partial \overline{\varphi}}{\partial z} \right), \tag{13.130a}$$

$$\overline{w'\varphi'} = -\frac{1}{2}\kappa_s s_\rho \left(\frac{\partial \overline{\varphi}}{\partial y} + \frac{1}{2} s_\rho \frac{\partial \overline{\varphi}}{\partial z} \right). \tag{13.130b}$$

If the tracer φ is potential temperature (and not just a passive tracer) then (13.130), along with one of the estimates for the size of κ_s given in Section 13.7.2, constitutes a parameterization for the diffusive poleward and upward heat flux in the atmosphere.

II. *Flow along neutral surfaces.* If the fluid interior is adiabatic and steady, then fluid trajectories are along neutral surfaces; that is, along surfaces of potential density or potential temperature. One might therefore be inclined to assume that the eddy fluxes are aligned along the mean neutral surfaces and choose $s = s_\rho$. However, even in the adiabatic case, this is not always a good choice. From the adiabatic thermodynamic equation $\mathrm{D}b/\mathrm{D}t = 0$ we may derive the equation for the eddy buoyancy variance, namely

$$\frac{1}{2}\frac{\partial \overline{b'^2}}{\partial t} + \frac{1}{2}\overline{\boldsymbol{u}}\cdot\nabla_z \overline{b'^2} + \frac{1}{2}\overline{w}\frac{\partial \overline{b'^2}}{\partial z} + \overline{\boldsymbol{u}'b'}\cdot\nabla_z \overline{b} + \overline{w'b'}\frac{\partial \overline{b}}{\partial z} + \frac{1}{2}\nabla_z\cdot\overline{\boldsymbol{u}'b'^2} + \frac{1}{2}\frac{\partial}{\partial z}\overline{w'b'^2} = 0, \tag{13.131}$$

and specialize to the case of a zonally uniform basic state and small-amplitude wave. In that case

$$\frac{1}{2}\frac{\partial \overline{b'^2}}{\partial t} = -\overline{v'b'}\frac{\partial \overline{b}}{\partial y} - \overline{w'b'}\frac{\partial \overline{b}}{\partial z}. \tag{13.132}$$

If the wave is statistically steady then the left-hand side is zero and

$$\overline{\boldsymbol{v}'b'}\cdot\nabla_x \overline{b} = 0, \tag{13.133}$$

where the subscript x indicates that the vectors are in the meridional plane, with no variation in x. In this case there is indeed no along-gradient flux. However, if the wave is growing then $\overline{\boldsymbol{v}'b'} \cdot \nabla_x \overline{b} < 0$ and if $\partial\overline{b}/\partial y < 0$ and $\overline{v'b'} > 0$, as in the Northern Hemisphere, then

$$\frac{\overline{w'b'}}{\overline{v'b'}} > -\frac{\partial\overline{b}/\partial y}{\partial\overline{b}/\partial z}, \tag{13.134}$$

and so the mixing slope is *less* steep than the mean isopycnal slope, even though the flow may be adiabatic. Similarly, if the wave is decaying the mixing slope is steeper than that of the mean isopycnals. In an inhomogeneous flow, the advection by the mean flow in (13.131) plays a similar role to time dependence: the advection of eddy variance by the mean flow into a region of larger variance will give rise to a mixing slope that is less steep than the isopycnal slope, and conversely for a flow entering a region of less variance. Only for a statistically steady, adiabatic, linear wave field is the mixing slope guaranteed to be along the isopycnals.

Having said all this, let us suppose that the fluid trajectories are indeed along neutral surfaces. If there is no diffusion orthogonal to this then $\alpha = 0$, and the transport tensor is, in two or three dimensions respectively.

$$\mathbf{S} = \kappa_s \begin{pmatrix} 1 & s_\rho \\ s_\rho & s_\rho^2 \end{pmatrix}, \qquad \mathbf{S} = \kappa_s \begin{pmatrix} 1 & 0 & s_\rho^x \\ 0 & 1 & s_\rho^y \\ s_\rho^x & s_\rho^y & |s_\rho|^2 \end{pmatrix}. \tag{13.135}$$

In the two-dimensional case

$$\overline{v'\varphi'} = -\kappa_s \left(\frac{\partial\overline{\varphi}}{\partial y} + s_\rho \frac{\partial\overline{\varphi}}{\partial z} \right), \qquad \overline{w'\varphi'} = -\kappa_s s_\rho \left(\frac{\partial\overline{\varphi}}{\partial y} + s_\rho \frac{\partial\overline{\varphi}}{\partial z} \right). \tag{13.136a,b}$$

Suppose that φ is potential temperature θ, and that surfaces of potential temperature define neutral surfaces. Then plainly eddy motion along potential temperature surfaces does not transfer potential temperature, and the diffusion defined by (13.136) should have no effect. The equations themselves respect this, for then

$$s_\rho = -\frac{\partial_y\overline{\theta}}{\partial_z\overline{\theta}}, \tag{13.137}$$

and using this in (13.136) gives

$$\overline{v'\theta'} = 0, \qquad \overline{w'\theta'} = 0. \tag{13.138a,b}$$

There is no eddy transport at all, as expected.

Application to atmosphere and ocean

In the atmosphere, if we wished to parameterize the heat transporting effects of baroclinic eddies we might choose the mixing slope to be *shallower* than the isothermal slope. Heat may then be transported downgradient, from equator to pole, in a diabatic process. This is a reasonable choice because the eddy transport, in the atmosphere, is diabatic. This choice is *less* appropriate in the ocean, because the ocean interior is almost adiabatic. That is to say, fluid transport is almost along isopycnals, except for some rather small effects involving diapycnal diffusivity. Diffusing buoyancy along isopycnals has no effect at all. In the real ocean the presence of salinity means that the potential temperature, potential density and salinity surfaces are not parallel, and there will be eddy diffusion of θ and S (salinity) along neutral surfaces, but this fact does not help provide a parameterization for the heat flux by baroclinic eddies, because we cannot expect such a flux to depend for its existence on the presence of a second tracer, salinity. However, baroclinic eddies certainly *do* have an effect on the ocean structure, and if our ocean model does not resolve them we must parameterize them. For this, we turn to the antisymmetric transport tensor.

13.7.4 ✦ Use of the Antisymmetric Transport Tensor

The antisymmetric transport tensor gives rise to the skew flux, or the pseudoadvection. In two dimensions (one horizontal, one vertical) we can immediately write down its form, namely

$$\boldsymbol{A} = \begin{pmatrix} 0 & -\kappa_a' \\ \kappa_a' & 0 \end{pmatrix}, \tag{13.139}$$

where the transfer coefficient κ_a', which may vary in space and time depending on the flow itself, determines the overall strength of the transport. In three dimensions we can write, by inspection,

$$\boldsymbol{A} = A_{ij} = \begin{pmatrix} 0 & 0 & -\kappa_a'^x \\ 0 & 0 & -\kappa_a'^y \\ \kappa_a'^x & \kappa_a'^y & 0 \end{pmatrix}, \tag{13.140}$$

where we have used our gauge freedom to choose $A_{21} = -A_{12} = 0$. Equation (13.140) preserves the form of (13.139) if one of the horizontal dimensions is absent — that is, if either row one and column one, or row two and column two, is eliminated. Our remaining choice is to determine the sign and magnitude of the transport coefficients.

An adiabatic, potential-energy diminishing, eddy transport scheme

A very useful parameterization for the transport of tracers in ocean models by baroclinic eddy fluxes, commonly known as the Gent–McWilliams or GM scheme,[14] can be constructed using the antisymmetric transport tensor. The satisfaction of two properties is the foundation of the scheme:

(i) Moments of the tracer should be preserved; in particular, the amount of fluid between two isopycnal surfaces should be preserved. This suggests the scheme should not diffuse buoyancy across constant-buoyancy surfaces.

(ii) The amount of available potential energy in the flow should be reduced, so mimicking the effects of baroclinic instability, which transfers available potential energy to kinetic energy.

The first of these is automatically satisfied by using an antisymmetric diffusivity tensor. The second property can be satisfied by choosing the transfer coefficients to be proportional to the slope of the isopycnals, in which case we may write (13.140) as

$$\boldsymbol{A} = \kappa_a \begin{pmatrix} 0 & 0 & -s^x \\ 0 & 0 & -s^y \\ s^x & s^y & 0 \end{pmatrix}, \tag{13.141}$$

where $\boldsymbol{s} = (s^x, s^y) = \nabla_\rho z = -\nabla_z \rho/(\partial\rho/\partial z)$ is the isopycnal slope (recall that s^x denotes a component of a vector, and s_x a derivative) and κ_a determines the overall magnitude of the diffusivity. In an ocean model separately carrying temperature and salinity fields, then (13.141) would be applied to each of these, with the isopycnal slope being determined using the equation of state. To more easily see what properties are implied by the transport, let us specialize to the salt-free case, with buoyancy, b, the only thermodynamic variable. The isopycnal slope is then $\boldsymbol{s} = -(b_x/b_z, b_y/b_z)$ and the horizontal eddy buoyancy transfer $\boldsymbol{F}_h = (F^x, F^y)$ is given by

$$\boldsymbol{F}_h = -\left(-\kappa_a \boldsymbol{s}\frac{\partial b}{\partial z}\right) = -\kappa_a\left(\frac{\partial b}{\partial x}, \frac{\partial b}{\partial y}\right) = -\kappa_a \nabla_z b, \tag{13.142a}$$

which for positive κ_a is the same as conventional downgradient diffusion.

The vertical transfer is given by

$$F^z = -\kappa_a\left(s^x\frac{\partial b}{\partial x} + s^y\frac{\partial b}{\partial y}\right) = \kappa_a s^2\frac{\partial b}{\partial z}, \tag{13.142b}$$

where $s^2 = \boldsymbol{s} \cdot \boldsymbol{s}$. This flux is *up* the vertical gradient; however, by construction, the total skew flux is neither upgradient nor downgradient.

The combination of the downgradient horizontal flux and the upgradient vertical flux acts to reduce the potential energy of the flow at the same time as preserving the volume of fluid within each density interval. The upgradient flux in the vertical is a consequence of the need to reduce the available potential energy: suppose warm light fluid overlays cold dense fluid in a statically stable configuration, then a downgradient vertical diffusion would raise the centre of gravity of the fluid, increasing its potential energy — just the opposite of the action of baroclinic instability. Thus, the sign on the vertical diffusivity must be negative and this, in combination with the structure of (13.141) (and so a positive horizontal diffusivity) allows both properties *(i)* and *(ii)* above to be satisfied. The parameterization does not preserve total energy; the loss of potential energy is not balanced by a corresponding gain of kinetic energy, rather it is assumed to be lost to dissipation. Finally, to determine the magnitude of the (skew) eddy diffusivity we may turn again to the phenomenological estimates of Section 13.7.2.

The eddy transport velocity

Applying (13.107) to (13.141) gives the eddy transport velocities,

$$\boxed{\widetilde{\boldsymbol{u}} = -\frac{\partial}{\partial z}(\kappa_a \boldsymbol{s}), \qquad \widetilde{w} = \nabla_z \cdot (\kappa_a \boldsymbol{s}).} \tag{13.143}$$

The streamfunction associated with $\boldsymbol{A}$ is found using (13.108) and (13.141) giving

$$\boldsymbol{\psi} = (-\kappa_a s^y, \kappa s^x, 0) = \mathbf{k} \times \kappa_a \boldsymbol{s}. \tag{13.144}$$

Two equivalent ways of implementing the GM parameterization are thus as a skew flux, as in (13.142), or as an advection by the pseudovelocities (13.143). The vanishing of the normal component of the velocity is equivalent to the vanishing of the normal component of the flux at the boundary, and ensures that the scheme conserves tracer moments. The advective flux of buoyancy is just

$$\boldsymbol{F}_{ad} = b\widetilde{\boldsymbol{v}} = b\nabla \times \boldsymbol{\psi} = b\nabla \times (\mathbf{k} \times \kappa_a \boldsymbol{s}), \tag{13.145}$$

whereas using (13.103) the skew flux is given by

$$\boldsymbol{F}_{sk} = -\nabla b \times \boldsymbol{\psi} = -\nabla b \times (\mathbf{k} \times \kappa_a \boldsymbol{s}). \tag{13.146}$$

Vector manipulation readily shows that the divergences of these two fluxes are equal.

13.7.5 Examples

Consider a situation with sloping isotherms (and with the density determined solely by temperature) as illustrated in Fig. 13.6. The vertical flux attempts to tighten the temperature distribution, whereas the horizontal flux, being downgradient, attempts to smooth out horizontal inhomogeneities. Taken together, their net effect is to preserve the amount of fluid between any two isotherms, but at the same time to rotate and flatten the isotherms, so reducing the available potential energy of the flow. This is different from a conventional downgradient diffusion. A purely horizontal diffusion would, in principle, act to equalize values at each level, and a three-dimensional downgradient diffusion would try to equalize all values. Thus, a skew flux behaves quite differently from the usual downgradient diffusion, which merely acts to reduce gradients without caring much about other fluid properties.

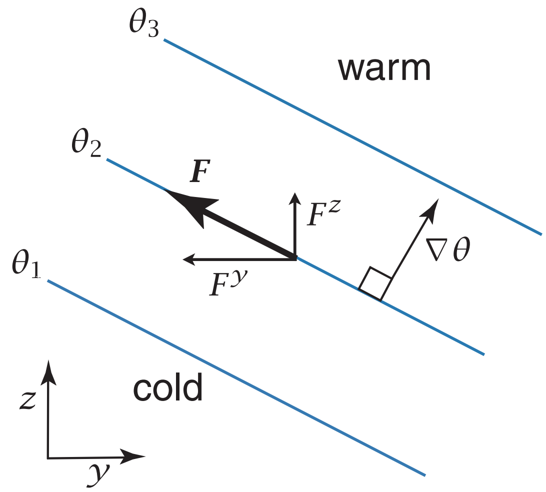

Fig. 13.6 The GM skew fluxes arising from sloping isotherms. The flux itself, $\boldsymbol{F}$, is parallel to the isotherms, with the horizontal flux being directed down the horizontal gradient but the vertical flux being upgradient. The effect of the vertical flux is to lower the centre of gravity of the fluid, and reduce the potential energy. The horizontal flux tries to make the temperature more uniform in the horizontal direction. The net effect of the skew flux is to *flatten* the isotherms.

To illustrate this consider a very simple example, that of a two-dimensional (y–z) fluid in which the initial density field is a 3×3 grid, with initial conditions

$$\rho_{init} = \begin{bmatrix} 2 & 1 & 1 \\ 3 & 2 & 1 \\ 3 & 3 & 2 \end{bmatrix}. \tag{13.147}$$

The isopycnals are sloping, much as in Fig. 13.6, and the flow is statically stable everywhere.

A purely horizontal diffusion would lead to, in the absence of other processes and with zero normal flux at the boundaries, a final state of

$$\rho_{hd} = \begin{bmatrix} 1.33 & 1.33 & 1.33 \\ 2 & 2 & 2 \\ 2.66 & 2.66 & 2.66 \end{bmatrix}, \tag{13.148}$$

and a full (vertical and horizontal) diffusion would give

$$\rho_{hvd} = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix}. \tag{13.149}$$

Neither of the above two final states preserves the density census (i.e., its distribution) and both imply strong diabatic effects — the fluid has been *mixed*, and the density variance has been reduced.

In contrast, a skew diffusion or eddy-transport advection will rotate the density surfaces until the isopycnal slope is zero, at which point the value of the transfer coefficients becomes zero and the process stops. The final state is then

$$\rho_{GM} = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix}. \tag{13.150}$$

This action both preserves the density census and reduces the available potential energy.

We can equally well interpret these effects in terms of eddy-transport velocities, so emphasizing that it is not the eddy flux itself that is important; rather, it is the flux divergence. If the slopes of Fig. 13.6 extended uniformly everywhere, then the associated fluxes would have zero divergence, and the eddy-induced velocities, given by (13.143), would be zero. On the other hand, consider

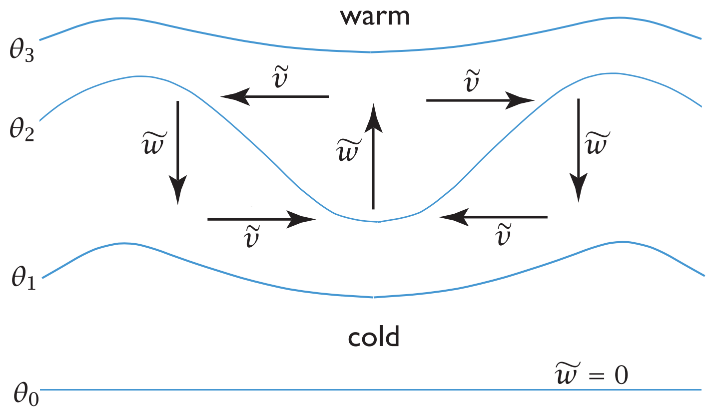

Fig. 13.7 The eddy-induced velocities in the Gent–McWilliams parameterization. The induced circulation attempts to flatten the sloping isopycnals. The induced vertical velocity, $\widetilde{w}$, is zero on flat isopycnals.

the case illustrated in Fig. 13.7, with variously sloping isotherms. For a constant value of the eddy diffusivity κ the slope of the isopycnals, $s = -(\partial\rho/\partial y)/(\partial\rho/\partial z)$, provides the streamfunction for the eddy-induced velocity:

$$\tilde{v} = -\frac{\partial\psi}{\partial z} = -\frac{\partial}{\partial z}(\kappa_a s), \qquad \widetilde{w} = \frac{\partial\psi}{\partial y} = \frac{\partial}{\partial y}(\kappa_a s). \tag{13.151}$$

This induces the velocities illustrated in Fig. 13.7, which evidently serve to flatten the isopycnals.

Downgradient (symmetric) diffusion and skew diffusion would normally be used together, and the transport tensor will then have both symmetric and antisymmetric components. In ocean models the mixing of temperature and salinity is often chosen to be along isopycnal surfaces. On the other hand, in the atmosphere (especially in the troposphere) diabatic effects are quite important and the symmetric tensor may be chosen to represent cross-isothermal transport.

13.8 † THICKNESS AND POTENTIAL VORTICITY DIFFUSION

In the previous section, we considered the structure of the diffusivity tensor, and then chose the entries by physical reasoning to mimic the effects of baroclinic instability. An alternative approach is to choose, a priori, a quantity to be diffused downgradient (i.e., not skew diffused), and then to represent the effect in the equations of motion as commonly used. In this section we first explore the use of thickness as a 'diffusee', and then look at potential vorticity. Thickness is the vertical distance between two isotherms or isopycnals, and it is a candidate for diffusion because a downgradient thickness transfer within an isopycnal layer satisfies the following two conditions:

(i) The total mass contained between two isopycnals is preserved, provided there are no boundary fluxes, so the effect is adiabatic.

(ii) If there is no orography a thickness flux serves to flatten isopycnals, and hence to reduce the available potential energy of the flow, mimicking baroclinic instability.

These are similar to the two properties listed on page 499, and indeed we will find that thickness diffusion is very similar to the GM scheme. However, thickness is *not* a materially conserved quantity; thus, the arguments of Section 13.4 do not apply and turbulent diffusion of thickness is rather ad hoc. Let us explore all these issues further.

13.8.1 Equations for Thickness

Recalling the results of Section 3.10, in a Boussinesq fluid the 'thickness', or the distance between two surfaces of constant buoyancy, is given by

$$\text{thickness} = \int_{z(b_1)}^{z(b_2)} \mathrm{d}z = \int_{b_1}^{b_2} \frac{\partial z}{\partial b}\, \mathrm{d}b, \tag{13.152}$$

and thus we may define the thickness field (strictly a 'thickness density' field), $\sigma \equiv \partial z/\partial b$. The volume of fluid between two isopycnal surfaces is proportional to $\sigma \Delta A$, where ΔA is an infinitesimal area, and in the absence of diabatic processes this is conserved. Thus, we have $\mathrm{D}(\sigma\Delta A)/\mathrm{D}t = 0$ and using $\mathrm{D}\Delta A/\mathrm{D}t = \Delta A \nabla_b \cdot \boldsymbol{u}$ we obtain the equation of motion for thickness (cf. (3.182))

$$\frac{\mathrm{D}\sigma}{\mathrm{D}t} + \sigma\nabla_b \cdot \boldsymbol{u} = D_\sigma \qquad \text{or} \qquad \frac{\partial \sigma}{\partial t} + \nabla_b \cdot (\boldsymbol{u}\sigma) = D_\sigma, \tag{13.153}$$

now including a term D_σ to represent any diabatic terms. From (13.153) we obtain, after a little algebra, the variance equation

$$\frac{1}{2}\frac{\partial}{\partial t}\overline{\sigma'^2} + \overline{\boldsymbol{u}'\sigma'} \cdot \nabla_b \overline{\sigma} + \frac{1}{2}\overline{\boldsymbol{u}} \cdot \nabla_b \overline{\sigma'^2} + \frac{1}{2}\overline{\boldsymbol{u}' \cdot \nabla_b \sigma'^2} = -\overline{w'\sigma'} + \overline{D'_\sigma \sigma'}, \tag{13.154}$$

where we have written $w' \equiv (\sigma \nabla_b \cdot \boldsymbol{u})'$. This equation is to be compared with the corresponding equation for a conserved tracer, (13.67). If the mean flow is small and the third-order correlations may be neglected then (13.154) becomes, in a statistically steady state,

$$\overline{\boldsymbol{u}'\sigma'} \cdot \nabla_b \overline{\sigma} \approx -\overline{w'\sigma'} + \overline{D'_\sigma \sigma'}. \tag{13.155}$$

Unlike the case for a tracer that is materially conserved except for dissipative effects, the transport of thickness is not necessarily downgradient. However, in regions of baroclinic instability, where there is conversion of available potential energy to kinetic energy $\overline{w'\sigma'}$ is positive, thickness may be transferred downgradient, suggesting a diffusive parameterization.

The eddy-induced and residual velocities

Now, let us decompose these variables in the usual manner into a mean component, denoted with an overbar, and an eddy component, denoted with a prime. The averaged thickness equation is

$$\frac{\partial \overline{\sigma}}{\partial t} + \nabla_b \cdot \left(\overline{\sigma}\,\overline{\boldsymbol{u}} + \overline{\sigma'\boldsymbol{u}'}\right) = 0, \tag{13.156}$$

where $\overline{\sigma'\boldsymbol{u}'}$ is the eddy thickness flux. This equation may be written as

$$\frac{\partial \overline{\sigma}}{\partial t} + \nabla_b \cdot [(\overline{\boldsymbol{u}} + \widetilde{\boldsymbol{u}})\overline{\sigma}] = 0, \tag{13.157}$$

where

$$\boxed{\widetilde{\boldsymbol{u}} \equiv \frac{\overline{\sigma'\boldsymbol{u}'}}{\overline{\sigma}}} \tag{13.158}$$

is the 'eddy-induced velocity', sometimes referred to as the 'bolus velocity', so-called because the thickness flux is said to be evocative of a peristaltic transfer along a passage bounded by impermeable but elastic walls. The quantity

$$\overline{\boldsymbol{u}}^* = \overline{\boldsymbol{u}} + \widetilde{\boldsymbol{u}} \tag{13.159}$$

is the residual velocity we encountered in Chapter 10, and it accounts for the total transport of thickness, including both eddy and Eulerian means.

In adiabatic flow, the evolution of a materially conserved tracer τ is given by

$$\frac{\mathrm{D}}{\mathrm{D}t}(\tau\sigma\Delta A) = 0, \tag{13.160}$$

whence, because $\sigma\Delta A$ is a constant,

$$\frac{\partial\tau}{\partial t} + \boldsymbol{u}\cdot\nabla_b\tau = 0. \tag{13.161}$$

Combining this with the thickness equation, (13.153) with $D_\sigma = 0$, gives

$$\frac{\partial}{\partial t}(\sigma\tau) + \nabla_b\cdot(\sigma\boldsymbol{u}\tau) = 0, \tag{13.162}$$

which in turn leads to

$$\frac{\partial}{\partial t}(\overline{\sigma}\,\overline{\tau} + \overline{\sigma'\tau'}) + \nabla_b\cdot(\overline{\sigma}\,\overline{\boldsymbol{u}}\,\overline{\tau}) + \nabla_b\cdot\overline{\sigma'\boldsymbol{u}'}\overline{\tau} + \nabla_b\cdot[\overline{(\sigma\boldsymbol{u})'\tau'}] = 0, \tag{13.163}$$

or, using (13.156),

$$\frac{\partial\overline{\tau}}{\partial t} + \frac{1}{\overline{\sigma}}\frac{\partial}{\partial t}(\overline{\sigma'\tau'}) + \left[\overline{\boldsymbol{u}} + \frac{\overline{\sigma'\boldsymbol{u}'}}{\overline{\sigma}}\right]\cdot\nabla_b\overline{\tau} = -\frac{1}{\overline{\sigma}}\nabla_b\cdot[\overline{(\sigma\boldsymbol{u})'\tau'}]. \tag{13.164}$$

(To derive these, first let $\overline{\sigma\tau\boldsymbol{u}} = \overline{(\sigma\boldsymbol{u})'\tau'} + \overline{\sigma\boldsymbol{u}}\,\overline{\tau}$.) If we neglect the correlation between σ' and τ', then (13.164) has the form

$$\frac{\partial\overline{\tau}}{\partial t} + (\overline{\boldsymbol{u}} + \widetilde{\boldsymbol{u}})\cdot\nabla_b\tau = -\frac{1}{\overline{\sigma}}\nabla_b\cdot[\overline{(\sigma\boldsymbol{u})'\tau'}]. \tag{13.165}$$

Thus, the averaged tracer evolves as if it were advected by two velocity fields: the large-scale field itself, $\overline{\boldsymbol{u}}$, and the eddy-induced velocity $\widetilde{\boldsymbol{u}}$, their sum being the residual velocity. The term on the right-hand side of (13.165) is the divergence of the transport of the tracer along the isopycnals by the eddy transport $(\sigma\boldsymbol{u})'$. These equations are not yet closed because we don't know the eddy-induced velocity, $\widetilde{\boldsymbol{u}}$.

13.8.2 Diffusive Thickness Transport

A downgradient diffusion of thickness parameterizes the eddy transport velocity by

$$\widetilde{\boldsymbol{u}} \equiv \frac{\overline{\sigma'\boldsymbol{u}'}}{\overline{\sigma}} = -\frac{1}{\overline{\sigma}}\kappa\nabla_b\overline{\sigma}, \tag{13.166}$$

where κ is an eddy diffusivity. Similarly, we might parameterize the right-hand side of (13.164) by

$$-\frac{1}{\overline{\sigma}}\nabla_b\cdot[\overline{(\sigma\boldsymbol{u})'\tau'}] = \frac{1}{\overline{\sigma}}\nabla_b\cdot(\kappa\overline{\sigma}\nabla_b\overline{\tau}), \tag{13.167}$$

that is, as a diffusion of the tracer along isopycnals.

In height coordinates the eddy transport velocity will be a three-dimensional field, obtained by appropriately transforming $\widetilde{\boldsymbol{u}}$. We have

$$\widetilde{\boldsymbol{u}} \approx -\frac{1}{\sigma}\kappa\nabla_b\overline{\sigma} = -\kappa\frac{\partial b}{\partial z}\nabla_b\left(\frac{\partial\overline{z}}{\partial b}\right) = -\kappa\frac{\partial b}{\partial z}\frac{\partial\overline{\boldsymbol{s}}}{\partial b} = -\kappa\frac{\partial\overline{\boldsymbol{s}}}{\partial z}, \tag{13.168}$$

where the third equality uses $\nabla_b z = \boldsymbol{s}$, the isopycnal slope. The final result is not quite the same as (13.143), because the diffusivity is now outside the z-derivative. It is a subtle but important distinction, because it means that if κ varies the vertical velocity can no longer be obtained easily as a local function. That is to say, given (13.168), we no longer have $\widetilde{w} = \nabla_z \cdot \kappa \overline{\boldsymbol{s}}$ as in (13.143). Rather, $\widetilde{w}$ must be evaluated by a non-local integration of the mass conservation requirement, so that

$$\widetilde{w} = \int \nabla_z \cdot \left(\kappa \frac{\partial \overline{\boldsymbol{s}}}{\partial z} \right) \mathrm{d}z. \tag{13.169}$$

This result should not be disconcerting from a physical standpoint, because the baroclinic activity of eddies certainly involves vertical communication — recall the tendency toward barotropic flow in baroclinic lifecycles. From a computational standpoint, it is a little less convenient. A less satisfactory feature of thickness diffusion arises when the ocean floor is not flat; a strict thickness diffusion might then increase the slope of the isopycnals and increase the available potential energy, which would be an unwanted effect. Nevertheless, overall the GM scheme is evidently similar to a thickness diffusion.

13.8.3 † Potential Vorticity Diffusion

From a more fundamental perspective, potential vorticity, Q, is a better candidate for diffusion than thickness because it is a materially conserved quantity.[15] It is not the only variable that is materially conserved — potential temperature is also. However, potential temperature is advected by the *three-dimensional* velocity field and the vertical advection complicates matters, since any diffusion tensor certainly cannot be isotropic and is probably not symmetric. On the other hand, in isentropic coordinates the adiabatic potential vorticity advection occurs in the isentropic plane (and in quasi-geostrophic flow the advection is purely horizontal). Thus, only the two-dimensional diffusion need be considered, and the diffusion tensor will be much simplified. Near the upper and lower boundaries buoyancy may still be the appropriate field to diffuse, because $w = 0$ and buoyancy is conserved on parcels when advected by the horizontal flow. Horizontal diffusion of buoyancy is not an adiabatic parameterization, but diabatic effects do occur at the surface. These considerations suggest that downgradient potential vorticity diffusion on isentropic surfaces in the fluid interior, combined with downgradient buoyancy diffusion at the upper and lower boundaries, may be as rational a parameterization of eddy transfer effects as any simple diffusion scheme can be.

Actually implementing a potential vorticity diffusion in the equations of motion is not easily done, because when the equations of motion are written in conventional form the potential vorticity flux does not directly appear. Furthermore, if potential vorticity is diffused with a constant diffusivity, momentum is not conserved. The resolution of issues remains research task, and below we just make a couple of remarks.

Connection to thickness diffusion

Potential vorticity diffusion is closely connected to thickness diffusion, especially in an oceanic setting in which the scales of motion are much larger than the deformation radius and fluxes of relative vorticity are relatively unimportant. To see this, consider the expression for potential vorticity in isentropic coordinates, namely

$$Q = \frac{f + \zeta}{\sigma} \approx \frac{f}{\sigma}, \tag{13.170}$$

where the second expression holds under planetary-geostrophic scaling. The eddy flux of potential vorticity is then

$$\overline{\boldsymbol{u}'Q'} \approx -\frac{f}{\overline{\sigma}^2} \overline{\boldsymbol{u}'\sigma'}, \tag{13.171}$$

which is similar to thickness flux, and using (13.158) and (13.171) gives the bolus velocity

$$\widetilde{\boldsymbol{u}} = \frac{1}{\overline{\sigma}}\overline{\boldsymbol{u}'\sigma'} \approx -\frac{\overline{\sigma}}{f}\overline{\boldsymbol{u}'Q'}. \tag{13.172}$$

Now, the gradient of potential vorticity is approximately given by

$$\nabla_b\overline{Q} = \frac{1}{\overline{\sigma}}\nabla_b f - \frac{f}{\overline{\sigma}^2}\nabla_b\overline{\sigma} = \frac{\beta}{\overline{\sigma}}\mathbf{j} - \frac{f}{\overline{\sigma}^2}\nabla_b\overline{\sigma}, \tag{13.173}$$

so that potential vorticity diffusion is then

$$\overline{\boldsymbol{u}'Q'} = -\boldsymbol{K}\nabla_b\overline{Q} = -\boldsymbol{K}\left(\frac{\beta}{\overline{\sigma}}\mathbf{j} - \frac{f}{\overline{\sigma}^2}\nabla_b\overline{\sigma}\right). \tag{13.174}$$

Using this equation with (13.172) gives the bolus velocity

$$\widetilde{\boldsymbol{u}} = \boldsymbol{K}\left(\frac{\beta}{f}\mathbf{j} - \frac{1}{\overline{\sigma}}\nabla_b\overline{\sigma}\right). \tag{13.175}$$

This differs from (13.166) mainly in the existence of the term involving β on the right-hand side. The expression is singular at the equator, a consequence of ignoring the relative vorticity term in the expression for potential vorticity.

Other recipes for diffusing potential vorticity are possible, but it may fairly be said that, although potential vorticity diffusion has considerable theoretical appeal, no implementation in a comprehensive ocean model has shown practical advantages over simpler GM or thickness-diffusing schemes.

Using the transformed Eulerian mean

A natural framework to discuss how eddy fluxes interact with the mean flow is the transformed Eulerian mean (TEM), discussed in Chapter 10., and that framework may also be useful for the eddy parameterization problem, especially in idealized settings. The connection is not unexpected, given the connection between the residual velocity and the thickness-weighted mean velocity demonstrated in Section 10.3.3. We'll illustrate the use with a simple example.

Recall the TEM form of quasi-geostrophic zonally averaged momentum and thermodynamic equation,

$$\frac{\partial\overline{u}}{\partial t} - f_0\overline{v}^* = \overline{v'q'}, \tag{13.176a}$$

$$\frac{\partial\overline{b}}{\partial t} + N^2\overline{w}^* = 0, \tag{13.176b}$$

with no forcing or dissipation terms for simplicity. The eddy flux terms now explicitly appear *only* in the momentum equations, and the eddy flux on the right-hand side of (13.176a) is the potential vorticity flux. A potential vorticity flux parameterization is thus both natural and adiabatic. Having said that, the advecting velocities are the residual velocities, so that if the Eulerian velocity is required one must pass from $\overline{v}^*$ to $\overline{v}$ using an eddy-flux parameterization. Similar considerations apply to using the TEM in the primitive equations, but once more our reach has exceeded our grasp, and this is where the chapter stops.

Notes

1 A significant fraction of the theory of turbulent diffusion stems from G. I. (Geoffrey Ingram) Taylor (1886–1975), who made important contributions to both fluid and solid mechanics, in the former to meteorology, oceanography and aerodynamics. In addition to his work in turbulence, Taylor is known for his work on the theory of rotating fluids (the 'Taylor–Proudman' effect, for example) and on hydrodynamic stability (analysis of stability of Couette flow, for example), and for his clear and simple laboratory experiments. The main results of this section were first derived by Taylor (1921a). Ludwig Prandtl (1875–1953) was the other great pioneer of turbulent diffusion; he is also famous for his work in boundary-layer theory and aerodynamics.

2 A number of textbooks in both fluid dynamics and stochastic processes give more detail on this topic. Gardiner (1985) is one.

3 See also Monin & Yaglom (1971).

4 The way variables are changed between (13.31) and (13.32) could also have been used to derive (13.20), but in that case a simpler transformation sufficed.

5 This topic was first addressed empirically by Richardson (1926), although it was Obukhov (1941) who first theoretically obtained the '4/3 power law' describing how the eddy diffusivity varies with separation for parcels in the inertial range. Our treatment takes advantage of Kolmogorov scaling.

6 This figure and Fig. 13.4 were kindly provided by Joe LaCasce; see LaCasce & Ohlmann (2003).

7 In the open ocean, Ollitrault *et al.* (2005) do find float separation that increases with t^3, consistent with a $-5/3$ inverse cascade range.

8 Morel & Larcheveque (1974) and Er-El & Peskin (1981). Earlier dispersion calculations were made by Richardson (1926) who measured smoke spreading from chimneys, finding results that are consistent with a three-dimensional energy inertial range at small scales.

9 Turbulent diffusion is both widely used and widely criticized. If there is a scale-separation between a well-defined mean flow and the eddies then turbulent diffusion can be a very useful parameterization. However, this condition is often *not* satisfied, because it is unusual in fluid mechanics for the turbulent eddies to be significantly smaller than the mean flow. Baroclinic turbulence is something of an exception because there is a natural scale of the turbulence — the deformation radius — that is in general different from the scale of the mean flow, although even this scale separation may be lost if there is an inverse cascade or if the deformation scale is sufficiently large, as in the Earth's atmosphere. Furthermore, properly choosing what variable is to be diffused, and ensuring that various fluid conservation properties remain respected by the diffusion, remain difficult problems.

10 If S_{ij} and A_{ij} are symmetric and antisymmetric tensors respectively, then, summing over repeated indices, their contraction is $A_{ij}S_{ij} = -A_{ji}S_{ij} = -A_{ji}S_{ji} = -A_{ij}S_{ij}$, where the last equality follows because the indices are dummy. Thus, the contraction must equal zero.

11 Green (1970) and Stone (1972), in the context of the meridional transport of heat in the Earth's atmosphere, suggested that the magnitude of the turbulent diffusivity coefficients could be obtained by dynamical arguments using such things as baroclinic instability theory and the amount of available potential energy in the atmosphere, although their suggestions differ in such important details as the eddy mixing length. Other efforts have drawn on geostrophic turbulence theory, for example Larichev & Held (1995) and Smith & Vallis (2002).

12 Ollitrault *et al.* (2005).

13 Green (1970).

14 The Gent–McWilliams (GM) scheme originated in Gent & McWilliams (1990) and was much clarified by Gent *et al.* (1995). Previously, Plumb (1979) and Moffatt (1983) had noted the connection between symmetric and antisymmetric diffusivities and diffusive and advective fluxes, and Griffies (1998) explicitly showed how the GM bolus velocities are related to a skew flux and can be calculated using an antisymmetric diffusivity tensor, which led to notable improvements in the scheme. See McDougall (1998) and (Griffies 2004) for reviews and more discussion. Visbeck *et al.* (1997) suggested that the values of eddy diffusivities in the GM scheme might be determined by dynamical

arguments similar to those of Green (1970) and Stone (1972). In spite of the manifest imperfections of the GM scheme, to date (2017) no commonly implemented scheme has proven to be better, in practice, for ocean climate models.

15 Potential vorticity diffusion was suggested by Green (1970) as a parameterization for large-scale eddies in the atmosphere, and further explored and used in ocean contexts by Welander (1973), Marshall (1981), Rhines & Young (1982a), Tréguier *et al.* (1997), Greatbatch (1998) and others. Lee *et al.* (1997), Marshall *et al.* (1999), Drijfhout & Hazeleger (2001), and others, have explored numerically whether the eddy transfer of tracers in the ocean is in fact diffusive, and whether potential vorticity or thickness is a better quantity to diffuse. For other examples and methodologies see, for example, Killworth (1997), Smith & Vallis (2002), Ferrari *et al.* (2010) or Marshall *et al.* (2012).