

Geostrophic Turbulence and the General Circulation

Geoffrey K. Vallis

Lectures in Kyoto

These are some notes on the material I will present in Kyoto. Sections marked with a * are either more advanced or are not central to the main theme, and I will not cover some or all the material in these sections. Also, some material may already be familiar to the audience; if so, I will skip that during the lectures.

The lectures will fall into two parts:

- (i) The theory of turbulence and geostrophic turbulence (chapters 1 and 2).
- (ii) Applications of this theory to the general circulation, including the midlatitude atmosphere and parts of the ocean (chapter 3).

We will not try to cover all this material in the lectures. Rather, the lectures will draw from these notes as appropriate.

Contents

1	Turbulence, Basic Theory	<i>page</i> 1
1.1	The Fundamental Problem of Turbulence	1
1.1.1	The Closure Problem	1
1.1.2	Triad Interactions in turbulence	3
1.2	The Kolmogorov Theory	5
1.2.1	The physical picture	5
1.2.2	Inertial range theory	6
1.2.3	* An alternative scaling argument for inertial ranges	11
1.2.4	A final note on our assumptions	12
1.3	Two-Dimensional Turbulence	13
1.3.1	Energy and Enstrophy Transfer in Two-Dimensional Turbulence	14
1.3.2	Inertial ranges in 2D turbulence	17
1.3.3	* More about the phenomenology	21
1.3.4	Numerical illustrations	22
1.4	* Predictability of Turbulence	24
1.5	* Spectrum of a Passive Tracer	25
1.5.1	Examples of tracer spectra	27
2	Geostrophic Turbulence and Baroclinic Eddies	31
2.1	Effects of Differential Rotation	31
2.1.1	Organization of turbulence into zonal flow	32
2.2	Stratified Geostrophic Turbulence	36
2.2.1	Quasi-geostrophic flow as an analogue to two-dimensional flow	36
2.2.2	Two-layer geostrophic turbulence	39
2.2.3	Triad interactions	41

2.3	* Phenomenology of Baroclinic Eddies in the Atmosphere and Ocean	44
2.3.1	The Magnitude and Scale of Baroclinic Eddies	45
2.3.2	Baroclinic Eddies in the Atmosphere	47
2.3.3	Baroclinic Eddies in the Ocean	49
3	Eddies and the General Circulation	53
3.1	* The Eliassen-Palm Flux	53
3.1.1	The Eliassen-Palm relation	54
3.1.2	The group velocity property	55
3.2	* The Transformed Eulerian Mean	56
3.2.1	Quasi-geostrophic form	56
3.2.2	The TEM in isentropic coordinates	58
3.3	The Maintenance of Jets in the Atmosphere	60
3.3.1	Observations and Motivation	60
3.3.2	The mechanism of jet production	60
3.3.3	A numerical example	69
3.4	The Ferrel Cell	71
3.4.1	Equations of motion	71
3.4.2	Dynamics of the model	75
3.5	The Antartic Circumpolar Channel	80
3.5.1	Steady and eddy flow	81
3.5.2	Vertically integrated momentum balance	82
3.5.3	Form drag and baroclinic eddies	83

Sections marked with a * are either more advanced or are not central to the main theme, and I will not cover some or all the material in these sections.

Basic Theory of Incompressible Turbulence

Turbulence is high Reynolds number fluid flow, dominated by nonlinearity, with both spatial and temporal disorder. A turbulent flow has eddies with a spectrum of sizes between some upper and lower bounds, the former usually determined by the forcing scale or the domain scale, and the latter usually by viscosity.

1.1 THE FUNDAMENTAL PROBLEM OF TURBULENCE

1.1.1 The Closure Problem

Although in a turbulent flow it may be virtually impossible to predict the detailed motion of each eddy, the statistical properties — time averages for example — might not be changing and we might like to predict such averages. Thus, we might accept we can't predict the weather but we can try to predict the climate. Even though we know which equations determine the system, this task proves to be very difficult because the equations are nonlinear, and we come up against the *closure problem*. To see what this is, let us decompose the velocity field into mean and fluctuating components,

$$\mathbf{v} = \bar{\mathbf{v}} + \mathbf{v}'. \quad (1.1)$$

Here $\bar{\mathbf{v}}$ is the mean velocity field, and \mathbf{v}' is the deviation from that mean. The mean might be a time average, in which case $\bar{\mathbf{v}}$ is a function only of space and not time, or it might be a time mean over a finite period (e.g., a season if we are dealing with the weather), or it might be some form of ensemble mean. The average of the deviation is, by definition, zero; that is $\overline{\mathbf{v}'} = 0$. The idea is to substitute (1.1) into the momentum equation and try to obtain a closed equation for the mean quantity $\bar{\mathbf{v}}$. Rather than

dealing with the full Navier-Stokes equations, let us carry out this program for a model nonlinear system which obeys

$$\frac{du}{dt} + uu + ru = 0 \quad (1.2)$$

where r is a constant. The average of this equation is:

$$\frac{d\bar{u}}{dt} + \bar{u}\bar{u} + r\bar{u} = 0 \quad (1.3)$$

The value of the term $\bar{u}\bar{u}$ is not deducible simply by knowing \bar{u} , since it involves correlations between eddy quantities $\overline{u'u'}$. That is, $\overline{uu} = \bar{u}\bar{u} + \overline{u'u'} \neq \bar{u}\bar{u}$. We can go to next order to try (vainly!) to obtain an equation for \overline{uu} . First multiply (1.2) by u to obtain an equation for u^2 , and then average it to yield:

$$\frac{1}{2} \frac{d\overline{u^2}}{dt} + \overline{uuu} + r\overline{u^2} = 0 \quad (1.4)$$

This equation contains the undetermined cubic term \overline{uuu} . An equation determining this would contain a quartic term, and so on in an unclosed hierarchy. Many methods of ‘closing the hierarchy’ make assumptions about the relationship of $(n + 1)$ ’th order terms to n ’th order terms, for example by supposing that:

$$\overline{uuuu} = \alpha \bar{u}\bar{u}\bar{u}\bar{u} + \beta \overline{uuuu} \quad (1.5)$$

where α and β are some parameters, and closures set in physical space or in spectral space (i.e., acting on the Fourier transformed variables) both exist. If we know that the variables are distributed normally then such closures can be made exact, but this is not generally the case in fluid turbulence.

This same closure problem arises in the Navier-Stokes equations. If density is constant (say $\rho = 1$) the x -momentum equation for an averaged flow is

$$\frac{\partial \bar{u}}{\partial t} + (\bar{\mathbf{v}} \cdot \nabla) \bar{u} = -\frac{\partial \bar{p}}{\partial x} - \nabla \cdot \overline{\mathbf{v}'\mathbf{u}'}. \quad (1.6)$$

Written out in full in Cartesian coordinates, the last term is

$$\nabla \cdot \overline{\mathbf{v}'\mathbf{u}'} = \frac{\partial}{\partial x} \overline{u'u'} + \frac{\partial}{\partial y} \overline{u'v'} + \frac{\partial}{\partial z} \overline{u'w'} \quad (1.7)$$

These terms, and the similar ones in the y - and z - momentum equations, represent the effects of eddies on the mean flow and are known as *Reynolds stress* terms. One way of expressing the ‘problem of turbulence’ is to find a representation of such Reynolds stress terms in terms of mean flow quantities. Nobody has been able to close the system, in any useful way, without introducing physical assumptions not directly deducible from the equations of motion themselves. Indeed, not only has the problem not been solved, it is not clear that a useful closed-form solution generally exists.

1.1.2 Triad Interactions in turbulence

The nonlinear term in the equations of motion not only leads to difficulties in closing the equations, but it leads to interactions among different length scales, and in this section we write the equations of motion in a form that makes this explicit. Purely for algebraic simplicity we will restrict attention two-dimensional flows.

The equation of motion for an incompressible fluid in two-dimensions may be written as

$$\frac{Dz}{Dt} \text{eta} = \frac{\partial \zeta}{\partial t} + J(\psi, \zeta) = F + \nu \nabla^2 \psi, \quad \zeta = \nabla^2 \psi. \quad (1.8)$$

where we include a forcing and viscous term but no Coriolis term. Let us suppose that the fluid is contained in a square, doubly-periodic domain of side L , and let us expand the streamfunction and vorticity in Fourier series so that, with a tilde denoting a Fourier coefficient,

$$\psi(x, y, t) = \sum_{\mathbf{k}} \tilde{\psi}(\mathbf{k}, t) e^{i\mathbf{k} \cdot \mathbf{x}}, \quad \zeta(x, y, t) = \sum_{\mathbf{k}} \tilde{\zeta}(\mathbf{k}, t) e^{i\mathbf{k} \cdot \mathbf{x}}, \quad (1.9)$$

where $\mathbf{k} = \mathbf{i}k^x + \mathbf{j}k^y$, $\tilde{\zeta} = -k^2 \tilde{\psi}$ where $k^2 = k^{x^2} + k^{y^2}$ and, to ensure that ψ is real, $\tilde{\psi}(k^x, k^y, t) = \tilde{\psi}^*(-k^x, -k^y, t)$, a property known as conjugate symmetry. The summations are over all positive and negative x - and y -wavenumbers, and $\tilde{\psi}(\mathbf{k}, t)$ is shorthand for $\tilde{\psi}(k^x, k^y, t)$. Substituting (1.9) in (1.8) gives, with (for the moment) F and ν both zero,

$$\begin{aligned} \frac{\partial}{\partial t} \sum_{\mathbf{k}} \tilde{\zeta}(\mathbf{k}, t) e^{i\mathbf{k} \cdot \mathbf{x}} &= - \sum_{\mathbf{p}} p^x \tilde{\psi}(\mathbf{p}, t) e^{i\mathbf{p} \cdot \mathbf{x}} \times \sum_{\mathbf{q}} q^y \tilde{\zeta}(\mathbf{q}, t) e^{i\mathbf{q} \cdot \mathbf{x}} \\ &+ \sum_{\mathbf{p}} p^y \tilde{\psi}(\mathbf{p}, t) e^{i\mathbf{p} \cdot \mathbf{x}} \times \sum_{\mathbf{q}} q^x \tilde{\zeta}(\mathbf{q}, t) e^{i\mathbf{q} \cdot \mathbf{x}}. \end{aligned} \quad (1.10)$$

where \mathbf{p} and \mathbf{q} are, like \mathbf{k} , horizontal wave vectors. We may obtain an evolution equation for the wavevector \mathbf{k} by multiplying (1.10) by $\exp(-i\mathbf{k} \cdot \mathbf{x})$ and integrating over the domain, and using the fact that the Fourier modes are orthogonal; that is

$$\int e^{i\mathbf{p} \cdot \mathbf{x}} e^{i\mathbf{q} \cdot \mathbf{x}} dA = \frac{1}{L^2} \delta(\mathbf{p} + \mathbf{q}). \quad (1.11)$$

where $\delta(\mathbf{p} + \mathbf{q})$ equals unity if $\mathbf{p} = -\mathbf{q}$ and is zero otherwise. Using this, (1.10) becomes, restoring the forcing and dissipation terms,

$$\frac{\partial}{\partial t} \tilde{\psi}(\mathbf{k}, t) = \sum_{\mathbf{p}, \mathbf{q}} A(\mathbf{k}, \mathbf{p}, \mathbf{q}) \tilde{\psi}(\mathbf{p}, t) \tilde{\psi}(\mathbf{q}, t) + \tilde{F}(\mathbf{k}) - \nu k^4 \tilde{\psi}(\mathbf{k}, t), \quad (1.12)$$

where $A(\mathbf{k}, \mathbf{p}, \mathbf{q}) = (q^2/k^2)(p^x q^y - p^y q^x) \delta(\mathbf{p} + \mathbf{q} - \mathbf{k})$ is an ‘interaction coefficient’, and the summation is over all \mathbf{p} and \mathbf{q} ; however, note that only those wavevector triads with $\mathbf{p} + \mathbf{q} = \mathbf{k}$ make a nonzero contribution, because of presence of the delta function.

Consider, then, a fluid a fluid in which just two Fourier modes are initially excited, with wavevectors \mathbf{p} and \mathbf{q} say (along with their conjugate-symmetric partners at $-\mathbf{p}$

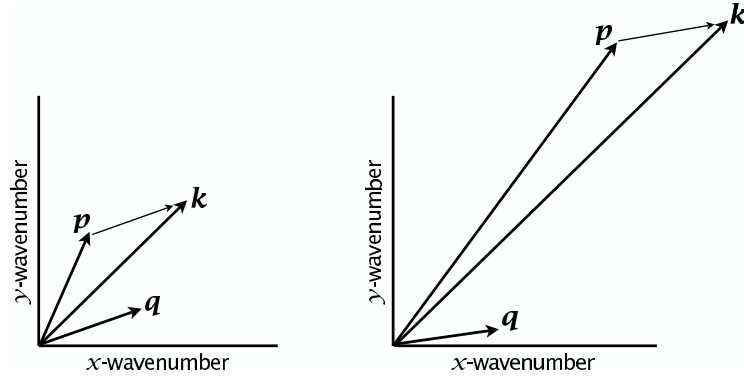


Fig. 1.1 Two interacting triads, each with $\mathbf{k} = \mathbf{p} + \mathbf{q}$. On the left, a local triad with $k \sim p \sim q$. On the right, a nonlocal triad with $k \sim p \gg q$.

and $-\mathbf{q}$). These modes interact [obeying (1.12)] to generate third and fourth wavenumbers, $\mathbf{k} = \mathbf{p} + \mathbf{q}$ and $\mathbf{m} = \mathbf{p} - \mathbf{q}$ (again along with their conjugate-symmetric partners). These four wavenumbers can interact among themselves to generate several additional wavenumbers, $\mathbf{k} + \mathbf{p}$, $\mathbf{k} + \mathbf{m}$ etc, and these in turn lead to still more interactions so potentially filling out the entire spectrum of wavenumbers. The individual interactions are called *triad interactions*, and it is by way of such interactions that energy is transferred between scales in turbulent flows, in both two and three dimensions. The dissipation term does not lead to interactions between modes with different wavevectors; rather, it acts like a drag on each Fourier mode, with a coefficient that increases with wavenumber and therefore that preferentially affects small scales.

The selection rule for triad interactions — that $\mathbf{k} = \mathbf{p} + \mathbf{q}$ — does not restrict the scales of these interacting wavevectors, and the types of triad interactions fall between two extremes:

- (i) Local interactions, in which $k \sim p \sim q$;
- (ii) Nonlocal interactions, in which $k \sim p \gg q$.

These two kinds of triads are schematically illustrated in Fig. 1.1. Without very detailed analysis of the solutions of the equations of motion — an analysis that is impossible for fully-developed turbulence — it is impossible to say whether one particular kind of triad interaction dominates. The theory of Kolmogorov considered below, and its two-dimensional analog, assume that it is the *local* triads that are most important in transferring energy; this is a reasonable assumption because from the perspective of a small eddy, large eddies appear as a nearly-uniform flow, and so simply translate the small eddies around without distorting them and thus without transferring energy between scales.

1.2 THE KOLMOGOROV THEORY

The foundation of many theories of turbulence is the spectral theory of Kolmogorov. This theory does not close the equations in quite as explicit a manner as (1.5), but it does provide a prediction for the energy spectrum of a turbulent flow (loosely speaking, how much energy is present at a particular spatial scale) and it does this by suggesting a relationship between the energy spectrum (a second order quantity in velocity) and the spectral energy flux (a third order quantity).

1.2.1 The physical picture

Consider high Reynolds number (Re) incompressible flow that is being maintained by some external force. Then the evolution of the system is governed by

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v} = -\nabla p + \mathbf{F} + \nu \nabla^2 \mathbf{v} \quad (1.13)$$

and

$$\nabla \cdot \mathbf{v} = 0 \quad (1.14)$$

Here, \mathbf{F} is some force we apply to maintain fluid motion — for example, we stir the fluid with a spoon. (A pedant might argue that such stirring is not a force like gravity but a continuous changing of the boundary conditions. Having noted this, we treat it as a force.) A simple scale analysis of these equations seems to indicate that the relative sizes of the inertial terms on the left-hand side to the viscous term is the Reynolds number VL/ν . To be explicit let us consider the ocean, and take $V = 0.1 \text{ m s}^{-1}$, $L = 1000 \text{ km}$ and $\nu = 10^{-6} \text{ m}^2 \text{ s}^{-1}$. Then $Re = VL/\nu \approx 10^{11}$, and it seems that we can neglect the viscous term on the right hand side of (1.13). But this can lead to a paradox. The fluid is being forced, and this forcing is likely to put energy into the fluid. We obtain the energy budget for (1.13) by multiplying by \mathbf{v} and integrating over a domain. If there is no flow into or out of our domain, the inertial terms in the momentum equation conserve energy and, recalling the results of section ??, the energy equation is

$$\frac{d\hat{E}}{dt} = \frac{d}{dt} \int \frac{1}{2} \mathbf{v}^2 dV = \int (\mathbf{F} \cdot \mathbf{v} + \nu \mathbf{v} \cdot \nabla^2 \mathbf{v}) dV = \int (\mathbf{F} \cdot \mathbf{v} - \nu \omega^2) dV \quad (1.15)$$

where \hat{E} is the total energy. If we neglect the viscous term we are led to an inconsistency, since the forcing term is a source of energy: $\overline{\mathbf{F} \cdot \mathbf{v}} > 0$, because a force will normally, on average, produce a velocity that is correlated with the force itself. Without viscosity, energy keeps on increasing.

What is amiss? It is true that for motion with a 1000 km length scale and a velocity of a few centimetres per second we can neglect viscosity when considering the balance of forces in the momentum equation. But this does not mean that there is no motion at much smaller length scales — indeed we seem to be led to the inescapable conclusion that there must be some motion at smaller scales in order to remove energy. Scale analysis of the momentum equation suggests that viscous terms will be comparable with

the inertial terms at a scale L_ν where the Reynolds number based on that scale is of order unity, giving

$$L_\nu \sim \frac{\nu}{V}. \quad (1.16)$$

This is a very small scale for geophysical flows, of order millimetres or less. Where and how are such small scales generated? Boundaries are one important region. If there is high Reynolds number flow above a solid boundary, for example the wind above the ground, then viscosity *must* become important in bringing the velocity to zero in order that it can satisfy the no-slip condition at the surface, as illustrated in Fig. ??.

Motion on very small scales may also be generated in the fluid interior. How might this happen? Suppose the forcing acts only at large scales, and its direct action is to set up some correspondingly large scale flow, composed of eddies and shear flows and such-like. Then typically there will be an instability in the flow, and a smaller eddy will grow: initially, the large scale flow may be treated as an unchanging shear flow, and the disturbance while small will obey linear equations of motion similar to those applicable in idealized Kelvin-Helmholtz instability. This instability clearly must draw from the large scale quasi-stationary flow, and it will eventually saturate at some finite amplitude. Although it has grown in intensity, it is still typically smaller than the large scale flow which fostered it (remember how the growth rate of the shear instability gets larger as wavelength of the perturbation decreased). As it reaches finite amplitude, the perturbation itself may become unstable, and smaller eddies will feed off its energy and grow, and so on. The picture that emerges is of a large scale flow that is unstable to eddies somewhat smaller in scale. These eddies grow, and develop still smaller eddies. Energy is transferred to smaller and smaller scales in a cascade-like process, sketched in Fig. 1.2. Finally, eddies are generated which are sufficiently small that they feel the effects of viscosity, and energy is drained away. Thus, there is a flux of energy from the large scales to the small scales, where it becomes dissipated.

1.2.2 Inertial range theory

Given the above picture it becomes possible to predict what the energy spectrum is. Let us suppose that the flow is statistically isotropic (i.e., the same in all directions) and homogeneous (i.e., the same everywhere; note that all isotropic flows are homogeneous, but not vice versa). Homogeneity precludes the presence of solid boundaries but can be achieved in a periodic domain. This puts an upper limit, sometimes called the outer scale, on the size of eddies.

If we decompose the velocity field into Fourier components, then in a finite domain we may write

$$u(x, y, z, t) = \sum_{k^x, k^y, k^z} \tilde{u}(k^x, k^y, k^z, t) e^{i(k^x x + k^y y + k^z z)} \quad (1.17)$$

where \tilde{u} is the Fourier transformed field of u , with similar identities for v and w . The sum is over all wavenumbers and in finite domain the wavenumbers are quantized, so that, for example, $k^x = 2\pi n/L$, where n is an integer and L the domain size. Finally,

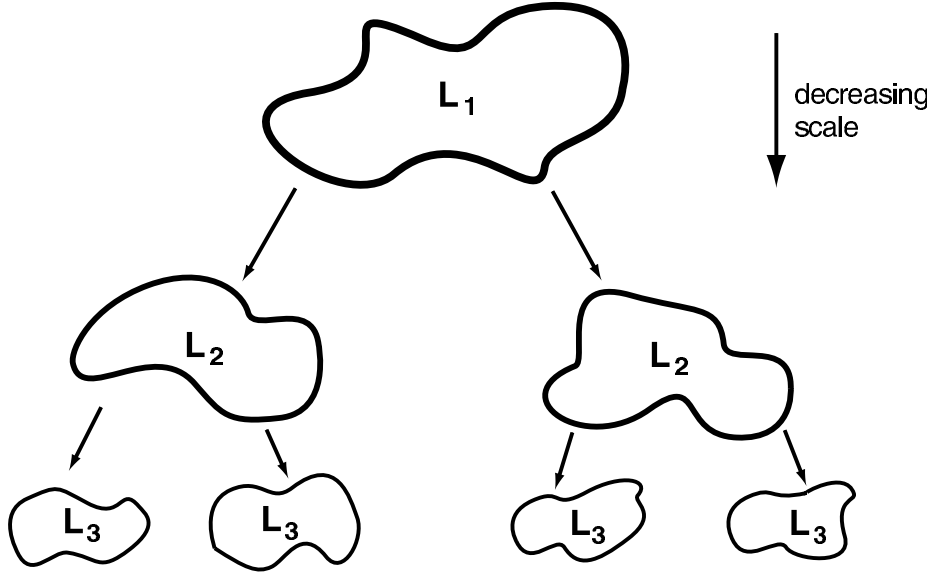


Fig. 1.2 Schema of a 'cascade' of energy to smaller scales: eddies at a large scale break up into smaller scale eddies, thereby transferring energy to smaller scales. If the transfer occurs between eddies of similar sizes (i.e., if it is spectrally local) the transfer is said to be a cascade. The eddies in reality are embedded within each other.

to ensure that u is real we require that $\tilde{u}(-k^x, -k^y, -k^z) = \tilde{u}^*(k^x, k^y, k^z)$, where the asterisk denotes the complex conjugate. The energy in the fluid is given by (assuming density is unity)

$$\begin{aligned}\hat{E} &= \int E dV = \frac{1}{2} \int (u^2 + v^2 + w^2) dV \\ &= \frac{1}{2} \sum (|\tilde{u}|^2 + |\tilde{v}|^2 + |\tilde{w}|^2) d\mathbf{k}\end{aligned}\quad (1.18)$$

using Parseval's theorem, where \hat{E} is the total energy and E is the energy per unit volume. We will now suppose that the turbulence is homogeneous and isotropic, and furthermore we will suppose that the domain is sufficiently large that the sums in the above equations may be replaced by integrals. We then write (1.18) as

$$\hat{E} \equiv \int \mathcal{E}(k) dk \quad (1.19)$$

where $\mathcal{E}(k)$ is the energy spectral density, or the energy spectrum, (so that $\mathcal{E}(k)\delta k$ is the energy in the small wavenumber interval δk) and because of the assumed isotropy, the energy is a function only of the scalar wavenumber k , where $k^2 = k^{x^2} + k^{y^2} + k^{z^2}$.

We now suppose that the fluid is stirred at large scales and, via the nonlinear terms in the momentum equation, that this energy is transferred to small scales where it is dissipated by viscosity. The key assumption is to suppose that, if forcing scale is sufficiently

Dimensions and the Kolmogorov Spectrum

Quantity	Dimension
Wavenumber, k	$1/L$
Energy per unit mass, E	$U^2 = L^2/T^2$
Energy spectrum, $\mathcal{E}(k)$	$EL = L^3/T^2$
Energy Flux, ε	$E/T = L^2/T^3$

If $\mathcal{E} = f(\varepsilon, k)$ then the only dimensionally consistent relation for the energy spectrum is

$$\mathcal{E} = \mathcal{K} \varepsilon^{2/3} k^{-5/3}$$

where \mathcal{K} is a dimensionless constant.

larger than the dissipation scale, there exists a range of scales intermediate between the large scale and the dissipation scale where neither forcing nor dissipation are explicitly important to the dynamics. This assumption, known as the *locality hypothesis*, depends on the nonlinear transfer of energy being sufficiently local (in spectral space). This intermediate range is known as the *inertial range*, because the inertial terms and not forcing or dissipation must dominate in the momentum balance. If the rate of energy input per unit volume by stirring is equal to ε , then if we are in a steady state there must be a flux of energy from large scales to small also equal to ε , and an energy dissipation rate, also ε .

Now, we have no general theory for the energy spectrum of a turbulent fluid but we might write it in the general form

$$\mathcal{E}(k) = g(\varepsilon, k, k_0, k_\nu) \quad (1.20)$$

where the right-hand side denotes a function of the energy flux ε , the wavenumber k , the forcing wavenumber k_0 and the wavenumber at which dissipation acts, k_ν (and $k_\nu \sim L_\nu^{-1}$). The function f will of course depend on the particular nature of the forcing. Now, the locality hypothesis essentially says that at some scale within the inertial range the flux of energy to smaller scales depends only on processes occurring at or near that scale. That is to say, the energy flux is only a function of \mathcal{E} and k , or equivalently that the energy spectrum can be a function *only* of the energy flux ε and the wavenumber itself. From a physical point of view, as energy cascades to smaller scales the details of the forcing are forgotten but the effects of viscosity are not yet apparent, and the energy spectrum takes the form,

$$\mathcal{E}(k) = g(\varepsilon, k). \quad (1.21)$$

The function g is, within this theory, *universal*, the same for every turbulent flow.

Let us now use dimensional analysis to give us the form of the function $f(\varepsilon, k)$ (see

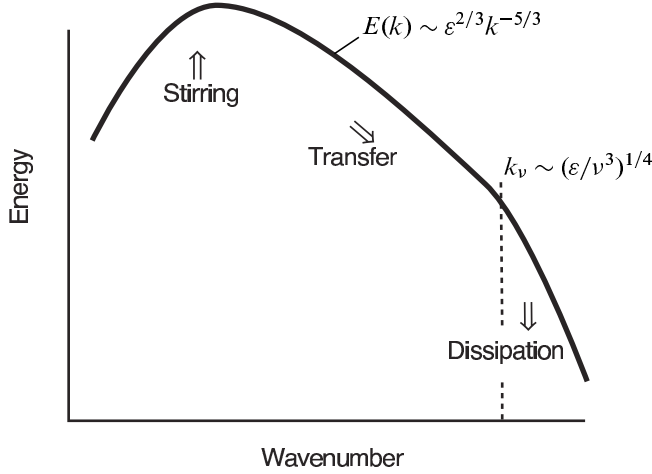


Figure 1.3 Schema of energy spectrum in three-dimensional turbulence, in the theory of Kolmogorov. Energy is supplied at some rate ϵ ; it is transferred ('cascaded') to small scales, where it is ultimately dissipated by viscosity. There is no systematic energy transfer to scales larger than the forcing scale, so here the energy falls off.

the shaded box). In (1.21), the left hand side has dimensionality L^3/T^2 ; the dimension T^{-2} on the left-hand side can only be balanced by $\epsilon^{2/3}$ because k has no time dependence; that is,

$$\begin{aligned} \mathcal{E}(k) &\sim \epsilon^{2/3} g(k) \\ \frac{L^3}{T^2} &\sim \frac{L^{4/3}}{T^2} g(k). \end{aligned} \quad (1.22)$$

where $g(k)$ is some function; this function $g(k)$ must have dimensions $L^{5/3}$ and the functional relationship we must have, if the physical assumptions are right, is

$$\boxed{\mathcal{E}(k) = \mathcal{K} \epsilon^{2/3} k^{-5/3}}. \quad (1.23)$$

This is the famous 'Kolmogorov -5/3 spectrum', enshrined as one of the cornerstones of turbulence theory, and sketched in Fig. 1.3, and some experimental results are shown in Fig. 1.4. The parameter \mathcal{K} is a dimensionless constant, undetermined by the theory. It is known as Kolmogorov's constant and experimentally it is found to be approximately 1.5.

An equivalent, perhaps slightly more intuitive, way to derive this is to first define an eddy turnover time τ_k , which is the time taken for a parcel with velocity v_k to move a distance $1/k$, v_k being the velocity associated with the (inverse) scale k . On dimensional considerations $v_k = (\mathcal{E}(k)k)^{1/2}$ so that

$$\tau_k = (k^3 \mathcal{E}(k))^{-1/2}. \quad (1.24)$$

Kolmogorov's assumptions are then equivalent to setting

$$\epsilon \sim \frac{v_k^2}{\tau_k} = \frac{k \mathcal{E}(k)}{\tau_k}. \quad (1.25)$$

If we demand that ϵ be constant then (1.24) and (1.25) yield (1.23).

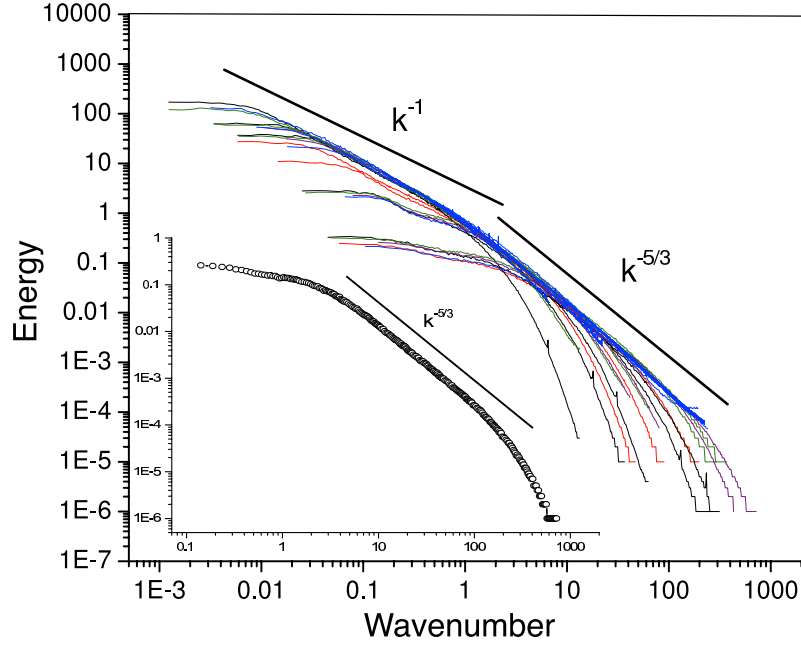


Fig. 1.4 The energy spectrum of 3D turbulence measured in some experiments at the Princeton Superpipe facility. The outer plot shows the spectra from a large number of experiments at different Reynolds numbers, with the magnitude of their spectra appropriately rescaled. Smaller scales show a good $-5/3$ spectrum, whereas at larger scales the eddies feel the effects of the pipe wall and the spectra are a little shallower. The inner plot shows the spectrum in the centre of the pipe in a single experiment at $Re \approx 10^6$.

The viscous scale and energy dissipation

At some small length-scale we should expect viscosity to become important and the scaling theory we have just set up will fail. What is that scale? In the inertial range friction is unimportant because the timescales on which it acts are too long for it to be important and dynamical effects dominate. In the momentum equation the viscous term is $\nu \nabla^2 u$ so that a viscous or dissipation timescale at a scale k^{-1} , τ_k^v , is

$$\tau_k^v \sim \frac{1}{k^2 \nu}, \quad (1.26)$$

so that the viscous timescale decreases with scale. The eddy turnover time, τ_k — that is, the inertial timescale — in the Kolmogorov spectrum is

$$\tau_k = \varepsilon^{-1/3} k^{-2/3}. \quad (1.27)$$

The wavenumber at which dissipation becomes important is then given by equating these two timescales, yielding the dissipation wavenumber, k_v , and the associated length-scale, L_v ,

$$k_v \sim \left(\frac{\varepsilon}{\nu^3} \right)^{1/4}, \quad L_v \sim \left(\frac{\nu^3}{\varepsilon} \right)^{1/4}. \quad (1.28a,b)$$

L_ν is called the *Kolmogorov scale*. It is the *only* quantity which can be created from the quantities ν and ε that has the dimensions of length. (It is the same as the scale given by provided that in that expression V is the velocity magnitude at the Kolmogorov scale.) Thus, for $L \gg L_\nu$, $\tau_k \ll \tau_k^\nu$ and inertial effects dominate. For $L \ll L_\nu$, $\tau_k^\nu \ll \tau_k$ and frictional effects dominate. In fact for length-scales smaller than the dissipation scale, (1.27) is inaccurate; the energy spectrum falls off more rapidly than $k^{-5/3}$ and the inertial timescale falls off less rapidly than (1.27) implies, and dissipation dominates even more.

Given the dissipation scale, let us estimate the energy dissipation rate. This is given by (section ??)

$$\hat{E} = \int \nu \mathbf{v} \cdot \nabla^2 \mathbf{v} \, dV. \quad (1.29)$$

The length at which dissipation acts is the Kolmogorov scale and, noting that $v_k^2 \sim \varepsilon^{2/3} k^{-2/3}$ and using (1.28a), the energy dissipation rate scales as (for a box of unit size)

$$\hat{E} \sim \nu k_\nu^2 v_{k_\nu}^2 \sim \nu k_\nu^2 \frac{\varepsilon^{2/3}}{k_\nu^{2/3}} \sim \varepsilon. \quad (1.30)$$

That is, the energy dissipation rate is equal to the energy cascade rate. On the one hand this seems sensible, but on the other hand it is *independent of the viscosity*. In particular, in the limit of viscosity tending to zero, L_ν tends to zero, but the energy dissipation does not! Surely the energy dissipation rate must go to zero if viscosity goes to zero? To see that this is not the case, consider that energy is input at some large scales, and the magnitude of the stirring largely determines the energy input and cascade rate. The scale at which viscous effects then become important is determined by the viscous scale, L_ν , given by (1.28b). *As viscosity tends to zero L_ν^{-1} becomes smaller in just such a way as to preserve the constancy of the energy dissipation.* This is one of the most important results in three-dimensional turbulence. Now, we established in section ?? that the Euler equations (i.e., the fluid equations with the viscous term omitted from the outset) do conserve energy. This means that the Euler equations are a *singular limit* of the Navier-Stokes equations: the behaviour of the Navier-Stokes equations as viscosity tends to zero is different from the behaviour resulting from ‘simply omitting the viscous term from the equations *ab initio*’.

How big is L_ν in the atmosphere? A crude estimate, perhaps wrong by an order of magnitude, comes from noting that ε has units of U^3/L , and that at length-scales of order 100 m in the atmospheric boundary layer (where there might be a three-dimensional energy cascade to small scales) velocity fluctuations are of order 1 cm s^{-1} , giving $\varepsilon \approx 10^{-8} \text{ m}^2 \text{ s}^{-3}$. Using (1.28b) we then find the dissipation scale to be of order a millimetre or so. In ocean the dissipation scale is also of order millimetres.

1.2.3 * An alternative scaling argument for inertial ranges

Kolmogorov’s spectrum, as well as some other useful scaling relationships, can be obtained in a slightly different way as follows. If we for the moment ignore viscosity, the

Euler equations are invariant under the following scaling transformation:

$$x \rightarrow x\lambda \quad v \rightarrow v\lambda^r \quad t \rightarrow t\lambda^{1-r}, \quad (1.31)$$

where r is an arbitrary scaling exponent. So far there is minimal physics. Now make the following *physical* assumptions about the behaviour of a turbulent fluid, *with viscosity*:

- (i) That the flux of energy from large to small scales (i.e., ε) is finite and constant.
- (ii) That the scale invariance (1.31) holds, on a time-average, in the intermediate scales between the forcing scales and dissipation scales.

The second assumption plays the role of the locality hypothesis. Dimensional analysis then tells us that the energy flux at some wavenumber k scales as

$$\varepsilon_k \sim \frac{v_k^3}{l_k} \sim \lambda^{3r-1}. \quad (1.32)$$

where v_k and l_k are the velocity and length scales at wavenumber k . Invoking assumption (i), that ε is independent of scale, gives $r = 1/3$. The velocity then scales as

$$v_k \sim \varepsilon^{1/3} k^{-1/3}, \quad (1.33)$$

and the velocity gradient (and so vorticity) scales as $kv_k \sim \varepsilon^{1/3} k^{2/3}$. (This becomes infinite at very small scales, but this behaviour is avoided in any real physical situation by the presence of viscosity.) We can now recover (1.23) because, on dimensional grounds,

$$\mathcal{E}(k) \sim v_k^2 k^{-1} \sim \varepsilon^{2/3} k^{-2/3} k^{-1} \sim \varepsilon^{2/3} k^{-5/3}. \quad (1.34)$$

In general, the slope of the energy spectrum, k^n , is related to the scaling exponent by $n = -(2r + 1)$. The ‘structure functions’ S_m , which are the average of the m ’th power of the velocity difference over distances $l \sim 1/k$, scale as $S_m \sim (v_k)^m \sim \varepsilon^{m/3} k^{-m/3}$. In particular the second-order structure function, which is the Fourier transform of the energy spectra, scales as $S_2 \sim \varepsilon^{2/3} k^{-2/3}$. Other results of the Kolmogorov theory follow similarly.

1.2.4 A final note on our assumptions

The essential *physical* assumptions are: (i) that there exists an inertial range in which the energy flux is constant, and (ii) that the energy is cascaded from large to small scales in a series of small steps, for then the energy spectra will be determined by spectrally local quantities. The second assumption is the locality assumption and without it we could have

$$\mathcal{E}(k) = C\varepsilon^{2/3} k^{-5/3} g(k/k_0) h(k/k_\nu), \quad (1.35)$$

where g and h are unknown functions; this is just as dimensionally consistent as (1.23). Kolmogorov essentially postulated that there exists a range of intermediate wavenumbers over which the energy spectrum has no functional dependence of the energy spectra on the forcing or dissipation scale, and $g(k/k_0) = h(k/k_\nu) = 1$.

The first assumption might be called the *non-intermittency* assumption, and it demands that rare events (in time or space) with large amplitudes do not dominate the energy flux or the dissipation rate. If they were to do so, then the flux would fluctuate strongly, the turbulent statistics would not be completely characterized by ε and Kolmogorov's theory would not be exactly right. (Note that in the theory ε is the mean energy cascade rate.) In fact, in high Reynolds turbulence the $-5/3$ spectra is often observed to a fairly high degree of accuracy (e.g., as in Fig. 1.4), although the higher-order statistics (e.g., higher-order structure functions) predicted by the theory are often found to be in error, and it is generally believed that Kolmogorov's theory is not exact.

1.3 TWO-DIMENSIONAL TURBULENCE

Two-dimensional turbulence behaves in a profoundly different way from three-dimensional turbulence, largely because of the presence of another quadratic invariant, the enstrophy (see also section ??). In two dimensions, the vorticity equation for incompressible flow is:

$$\frac{\partial \zeta}{\partial t} + \mathbf{u} \cdot \nabla \zeta = F + \nu \nabla^2 \zeta \quad (1.36)$$

where $\mathbf{u} = u\mathbf{i} + v\mathbf{j}$ and $\zeta = \mathbf{k} \cdot \nabla \times \mathbf{u}$ and F is a stirring term. In terms of a streamfunction, $u = -\partial\psi/\partial y$, $v = \partial\psi/\partial x$, and $\zeta = \nabla^2\psi$, and (1.36) may be written:

$$\frac{\partial \nabla^2 \psi}{\partial t} + J(\psi, \nabla^2 \psi) = F + \nu \nabla^4 \psi. \quad (1.37)$$

We obtain an energy equation by multiplying by $-\psi$ and integrating over the domain, and an enstrophy equation by multiplying by ζ and integrating. When $F = \nu = 0$ we find:

$$\hat{E} = \frac{1}{2} \int_A (u^2 + v^2) dA = \frac{1}{2} \int_A (\nabla \psi)^2 dA, \quad \frac{d\hat{E}}{dt} = 0, \quad (1.38a)$$

$$\hat{Z} = \frac{1}{2} \int_A \zeta^2 dA = \frac{1}{2} \int_A (\nabla^2 \psi)^2 dA, \quad \frac{d\hat{Z}}{dt} = 0. \quad (1.38b)$$

where the integral is over a finite area with either no-normal flow or periodic boundary conditions. The quantity \hat{E} is the energy, and \hat{Z} is known as the *enstrophy*. This enstrophy invariant arises because the vortex stretching term, so important in three-dimensional turbulence, vanishes identically in two dimensions. In fact, because vorticity is conserved on parcels it is clear that the integral of *any* function of vorticity is zero when integrated over A ; that is, from (1.36)

$$\frac{Dg(\zeta)}{Dt} = 0 \quad \text{and} \quad \frac{d}{dt} \int_A g(\zeta) dA = 0. \quad (1.39)$$

where $g(\zeta)$ is an arbitrary function. Of this infinity of conservation properties, enstrophy conservation (with $g(\zeta) = \zeta^2$) in particular has been found to have enormous consequences to the flow of energy between scales, as we soon discover.

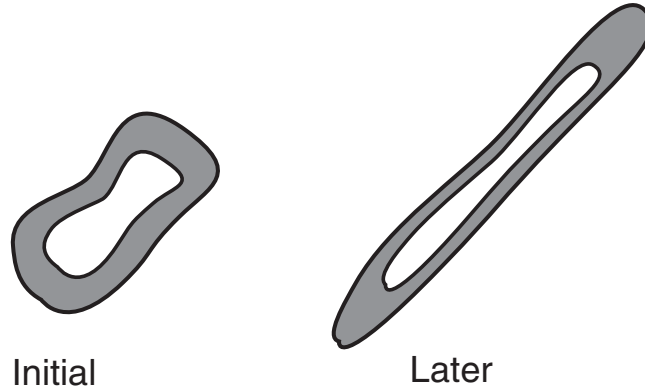


Fig. 1.5 In incompressible two-dimensional flow, a band of fluid will generally be elongated, but its area will be preserved. Since vorticity is tied to fluid parcels, the values of the vorticity in the hatched area (and in the hole in the middle) are maintained; thus, vorticity gradients will increase and the enstrophy is thereby, on average, moved to smaller scales.

1.3.1 Energy and Enstrophy Transfer in Two-Dimensional Turbulence

In three dimensional turbulence we posited that energy is cascaded to small scales via vortex stretching. In two dimensions that mechanism is absent, and it turns out that it is more reasonably to expect energy to be transferred to *larger scales*. This counter-intuitive behaviour arises from the twin integral constraints of energy and enstrophy conservation, and the following three arguments illustrate why this should be so.

I Vorticity elongation

Consider a band or a patch of vorticity, as in Fig. 1.5, in a nearly inviscid fluid. The vorticity of each element of fluid is conserved as the fluid moves. Now, we should expect that quasi-random motion of the fluid will act to elongate the band but, as its area must be preserved, the band narrows and so vorticity gradients will increase. This is equivalent to the enstrophy moving to smaller scales. Now, the energy in the fluid is

$$\hat{E} = -\frac{1}{2} \int \psi \zeta \, dA, \quad (1.40)$$

where the streamfunction is obtained by solving the Poisson equation $\nabla^2 \psi = \zeta$. If the vorticity is locally elongated primarily only in one direction (as it must be to preserve area), the integration involved in solving the Poisson equation will lead to the scale of the streamfunction becoming larger in the direction of stretching, but virtually no smaller in the perpendicular direction. Because stretching occurs, on average, in all directions, the overall scale of the streamfunction will increase in all directions, and the cascade of enstrophy to small scales will be accompanied by a transfer of energy to large scales.

II An energy-enstrophy conservation argument

A moments thought will reveal that the distribution of energy and enstrophy in wavenumber space are respectively analogous to the distribution of mass and moment of inertia of a lever, with wavenumber playing the role of distance from the fulcrum. Any rearrangement of mass such that its distribution also becomes wider must be such that the centre of mass moves toward the fulcrum. Thus, analogously, any rearrangement of a flow that preserves both energy and enstrophy, and that causes the distribution to spread out in wavenumber space, will tend to move energy to small wavenumbers and enstrophy to large. To prove this we begin with expressions for the total energy and enstrophy:

$$\hat{E} = \int \mathcal{E}(k) dk, \quad \hat{Z} = \int \mathcal{Z}(k) dk = \int k^2 \mathcal{E}(k) dk, \quad (1.41)$$

where $\mathcal{E}(k)$ and $\mathcal{Z}(k)$ are the energy and enstrophy spectra. A wavenumber characterizing the spectral location of the energy is the centroid,

$$k_e = \frac{\int k \mathcal{E}(k) dk}{\int \mathcal{E}(k) dk} \quad (1.42)$$

and, for simplicity, we normalize units so that the denominator is unity. The spreading out of the energy distribution is formalized by setting

$$I \equiv \int (k - k_e)^2 \mathcal{E}(k) dk, \quad \frac{dI}{dt} > 0. \quad (1.43)$$

Here, I measures the width of the energy distribution, and this is assumed to increase. Expanding out the integral gives

$$\begin{aligned} I &= \int k^2 \mathcal{E}(k) dk - 2k_e \int k \mathcal{E}(k) dk + k_e^2 \int \mathcal{E}(k) dk \\ &= \int k^2 \mathcal{E}(k) dk - k_e^2 \int \mathcal{E}(k) dk, \end{aligned} \quad (1.44)$$

where the last equation follows because $k_e = \int k \mathcal{E}(k) dk$ is, from (1.42), the energy-weighted centroid. Because both energy and enstrophy are conserved, (1.44) gives

$$\frac{dk_e^2}{dt} = -\frac{1}{\hat{E}} \frac{dI}{dt} < 0. \quad (1.45)$$

Thus, the centroid of the distribution moves to smaller wavenumber and to larger scale (see Fig. 1.6).

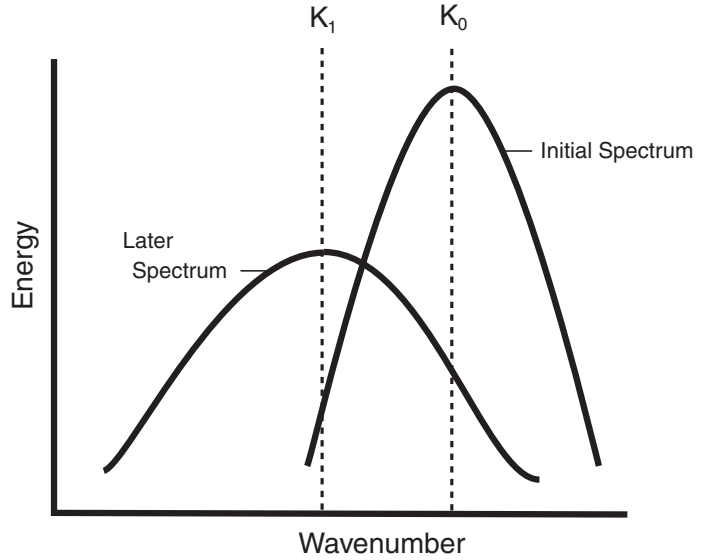
An appropriately defined measure of the centre of the enstrophy distribution, on the other hand, moves to higher wavenumber. The demonstration follows easily if we work with the inverse wavenumber, which is a direct measure of length. Let $q = 1/k$ and assume that the enstrophy distribution spreads out by nonlinear interactions, so that, analogously to (1.43),

$$J = \int (q - q_e)^2 \mathcal{Z}(q) dq, \quad \frac{dJ}{dt} > 0, \quad (1.46)$$

where

$$q_e = \frac{\int q \mathcal{Z}(q) dq}{\int \mathcal{Z}(q) dq}. \quad (1.47)$$

Figure 1.6 In two-dimensional flow, the centroid of the energy spectrum will move to large scales (smaller wavenumber) provided that the width of the distribution increases, which can be expected in a nonlinear, eddying flow



Expanding the integrand in (1.46) and using (1.47) gives

$$J = \int q^2 Z(q) dq - q_e^2 \int Z(q) dq, \quad (1.48)$$

But $\int q^2 Z(q) dq$ is conserved, because this is the energy. Thus,

$$\frac{dJ}{dt} = -\frac{d}{dt} q_e^2 \int Z(q) dq \quad (1.49)$$

whence

$$\frac{dq_e^2}{dt} = -\frac{1}{Z} \frac{dJ}{dt} < 0 \quad (1.50)$$

Thus, the length scale characterizing the enstrophy distribution gets smaller, and the corresponding wavenumber gets larger.

III A similarity argument

Consider an initial value problem, in which a fluid with some initial distribution of energy is allowed to freely evolve, unencumbered by boundaries. We note two aspects of the problem:

- (i) There is no externally imposed length-scale (because of the way the problem is posed).
- (ii) The energy is conserved (this being an assumption).

It is the second condition that limits the argument to two dimensions, for in three dimensions energy is quickly cascaded to small scales and dissipated, but let us here posit that this does not occur. These two assumptions are then sufficient to infer the general direction of transfer of energy, using a rather general similarity argument. To begin,

write the total energy (per unit mass) of the fluid as

$$\hat{E} = U^2 = \int \mathcal{E}(k, t) dk, \quad (1.51)$$

where $\mathcal{E}(k, t)$ is the energy spectrum and U is measure of the total energy, with units of velocity. Now, solely on dimensional considerations we can write

$$\mathcal{E}(k, t) = U^2 L \hat{\mathcal{E}}(\hat{k}, \hat{t}), \quad (1.52)$$

where $\hat{\mathcal{E}}$, and its arguments, are nondimensional quantities, and L is some length-scale. However, on physical considerations, the only parameters available to determine the energy spectrum are U , t and k , the wavenumber. A little thought reveals that the most general form for the energy spectrum with no L dependence is

$$\mathcal{E}(k, t) = U^3 t \hat{\mathcal{E}} = U^3 t g(Ukt), \quad (1.53)$$

where g is an arbitrary function of its arguments. The argument of g is the only non-dimensional grouping of U , t and k , and $U^3 t$ provides the proper dimensions for \mathcal{E} . Conservation of energy now implies that the integral

$$I = \int_0^\infty t g(Ukt) dk \quad (1.54)$$

not be a function of time. Defining $\vartheta = Ukt$, this requirement is met if

$$\int_0^\infty g(\vartheta) d\vartheta = \text{constant}. \quad (1.55)$$

Now, the spectrum is a function of k only through the combination $\vartheta = Ukt$. Thus, as time proceeds features in the spectrum move to smaller k . Suppose, for example, that the energy is initially peaked at some wavenumber k_p ; the product tk_p is preserved, so k_p must diminish with time and the energy must move to larger scales. Similarly, the energy weighted mean wavenumber, k_e , moves to smaller wavenumber, or larger scale. To see this explicitly, we have

$$k_e = \frac{\int k \mathcal{E} dk}{\int \mathcal{E} dk} = \frac{\int k \mathcal{E} dk}{U^2} = \int k U t g(Ukt) dk = \int \frac{\vartheta g(\vartheta)}{U t} d\vartheta = \frac{C}{U t} \quad (1.56)$$

where all the integrals are over the interval $(0, \infty)$ and $C = \int \vartheta g(\vartheta) d\vartheta$ is a constant. Thus, the wavenumber centroid of the energy distribution decreases with time, and the characteristic scale of the flow, $1/k_e$, increases with time. Interestingly, the enstrophy does not explicitly enter this argument, and in general it is not conserved; rather, it is the requirement that energy be conserved that limits the argument to two dimensions. If we accept *ab initio* that energy is conserved, it must be transferred to larger scales.

1.3.2 Inertial ranges in 2D turbulence

If, unlike the case in three dimensions, energy is transferred to larger scales in inviscid, nonlinear, two-dimensional flow then we might expect two-dimensional turbulence, and any associated inertial ranges, to be quite different from their three-dimensional counterparts. Before looking in detail at the inertial ranges themselves, we establish a couple of general properties of forced-dissipative flow in two dimensions.

Some properties forced-dissipative flow

We will first show that, unlike the case in three dimensions, energy dissipation goes to zero as Reynolds number rises. In the absence of forcing terms, the total dissipation of energy is, from (1.36)),

$$\frac{d\hat{E}}{dt} = -\nu \int \zeta^2 dA \quad (1.57)$$

Energy dissipation can only remain finite as $\nu \rightarrow 0$ if vorticity becomes infinite. However, this cannot happen because vorticity is conserved on parcels except for the action of viscosity, meaning that $D\zeta/Dt = \nu \nabla^2 \zeta$. However, the viscous term can only *reduce* the value of vorticity on a parcel, and so vorticity can never become infinite if it is not so initially, and therefore using (1.57) energy dissipation goes to zero with ν . (In three dimensions vorticity becomes infinite as viscosity goes to zero because of the effect of vortex stretching.) This conservation of energy is related to the fact that energy is trapped at large scales, even in forced-dissipative flow. On the other hand, enstrophy is transferred to small scales and therefore we expect it to be dissipated at large wavenumbers, even as the Reynolds number becomes very large.

We can show that energy is trapped at large scales in forced-dissipative two-dimensional flow (in a sense that will be made explicit) by the following argument.

Suppose that the forcing of the fluid is confined to a particular scale, characterized by the wavenumber k_f , and that dissipation is effected by a linear drag and a small viscosity. The equation of motion is

$$\frac{\partial \zeta}{\partial t} + J(\psi, \zeta) = F - r\zeta + \nu \nabla^2 \zeta. \quad (1.58)$$

where F is the stirring and r and ν are positive constants. This leads to the following energy and enstrophy equations:

$$\frac{d\hat{E}}{dt} = -2r\hat{E} - \int \psi F d\mathbf{x} + \int \nu \zeta^2 dA \approx -2r\hat{E} - \int \psi F dA, \quad (1.59a)$$

$$\frac{d\hat{Z}}{dt} = -2r\hat{Z} + \int \zeta F dA - D_Z \approx -2r\hat{Z} - k_f^2 \int \psi F dA - D_Z, \quad (1.59b)$$

where $D_Z = \int \nu (\nabla \zeta)^2 dA$ is the enstrophy dissipation. To obtain the right-most expressions, in (1.59a) we assume there is no dissipation of energy by the viscous term, and in (1.59b) we assume that the forcing is confined to wavenumbers near k_f . In a statistically steady state, and writing $\hat{E} = \int E(k) dk$ and $\hat{Z} = \int k^2 E(k) dk$, then eliminating the integral involving ψF between (1.59a) and (1.59b) gives

$$\int k^2 E(k) dk + \frac{D_Z}{2r} = \int k_f^2 E(k) dk, \quad (1.60)$$

where the integrations are over all wavenumbers. Now, from the obvious inequality $\int (k - k_e)^2 E(k) dk \geq 0$, where k_e is the energy centroid defined in (1.42), we obtain

$$\int (k^2 - k_e^2) E(k) dk \geq 0. \quad (1.61)$$

Combining (1.60) and (1.61) gives

$$\int (k_f^2 - k_e^2) E(k) dk \geq \frac{D_Z}{2r} > 0. \quad (1.62)$$

Thus, in a statistically steady state, the energy containing scale, as characterized by k_e^{-1} , must be larger than the forcing scale k_f^{-1} . This demonstration (rather like argument II in section 1.3.1) relies both on the conservation of energy and enstrophy by the nonlinear terms and on the particular relationship between the energy and enstrophy spectra.

This result, and (especially) the arguments of section 1.3.1, suggest that in a forced-dissipative two-dimensional fluid, energy is transferred to larger scales and enstrophy is transferred to small scales. To obtain a statistically steady state friction (such as the Rayleigh drag of (1.58)) is necessary to remove energy at large scales, and enstrophy must be removed at small scales, but if the forcing scale is sufficiently well separated in spectral space from such frictional effects then two inertial ranges may form — an *energy inertial range* carrying energy to larger scales, and an *enstrophy inertial range* carrying enstrophy to small scales (Fig. 1.7). These ranges are analogous to the three-dimensional inertial range of section 1.2, and similar conditions must apply if the ranges are to be truly inertial — in particular we must assume spectral locality of the energy or enstrophy transfer. But given that, we can calculate their properties, as follows.

The enstrophy inertial range

In the enstrophy inertial range the enstrophy cascade rate η , equal to the rate at which enstrophy is supplied by stirring, is assumed constant. By analogy with (1.25) we may assume that this rate is given by

$$\eta \sim \frac{k^3 \mathcal{E}(k)}{\tau_k}. \quad (1.63)$$

With τ_k (still) given by (1.24) we obtain

$$\boxed{\mathcal{E}(k) = \mathcal{K}_\eta \eta^{2/3} k^{-3}}, \quad (1.64)$$

where \mathcal{K}_η is, we presume, a universal constant, analogous to the Kolmogorov constant of (1.23).

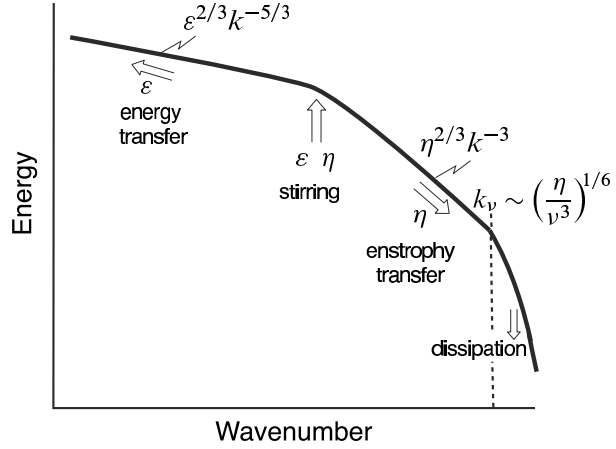
It is also possible to obtain (1.64) from scaling arguments similar to those in section 1.2.3. The scaling transformation (1.31) still holds, but now instead of (1.32) we assume that the enstrophy flux is constant with wavenumber. Dimensionally, and analogously to (1.32), we have

$$\eta \sim \frac{v_k^3}{l_k^3} \sim \lambda^{3r-3}, \quad (1.65)$$

and the constancy of η gives $r = 1$ for the scaling exponent. The exponent n determining the slope of the inertial range is given, as before, by $n = -(2r+1)$ yielding the -3 spectra of (1.64). The velocity at a particular wavenumber then scales as

$$v_k \sim \eta^{1/3} k^{-1}, \quad (1.66)$$

Figure 1.7 The energy spectrum of two-dimensional turbulence. (Compare with Fig. 1.3.) Energy supplied at some rate ε is transferred to large scales, whereas enstrophy supplied at some rate η is transferred to small scales, where it may be dissipated by viscosity. If the forcing is localized at a scale k_f^{-1} then $\eta \approx k_f^2 \varepsilon$.



and the time scales as

$$t_k \sim l_k / v_k \sim \eta^{-1/3}. \quad (1.67)$$

We may also obtain (1.67) by substituting (1.64) into (1.24). Thus, *the eddy turnover time in the enstrophy range of two-dimensional turbulence is length-scale invariant*. The appropriate viscous scale is given by equating the inertial and viscous terms in (1.36). Using (1.66) we obtain, analogously to (1.28a),

$$k_v \sim \left(\frac{\eta^{1/3}}{\nu} \right)^{1/2}. \quad (1.68)$$

The enstrophy dissipation, analogously to (1.30) goes to a finite limit given by

$$\dot{Z} = \nu \int_A \zeta \nabla^2 \zeta \, dA \sim \nu k_v^4 v_{k_v}^2 \sim \eta, \quad (1.69)$$

using (1.66) and (1.68). Thus, the enstrophy dissipation in two-dimensional turbulence is (at least according to this theory) independent of the viscosity.

Energy inertial range

The energy inertial range of two-dimensional turbulence is quite similar to that of three-dimensional turbulence, except in one major respect: the energy flows from smaller to larger scales! Because the atmosphere and ocean both behave in some ways as two-dimensional fluids, this has profound consequences on their behaviour, and is something we return to in the next chapter. The upscale energy flow is known as the *inverse cascade*, and the associated energy spectrum is, as in the three-dimensional case,

$$\mathcal{E}(k) = \mathcal{K}_\varepsilon \varepsilon^{2/3} k^{-5/3}, \quad (1.70)$$

where \mathcal{K}_ε is a nondimensional constant [not necessarily equal to \mathcal{K} in (1.23)], and ε is the rate of energy transfer to larger scales. Of course we now need a mechanism to remove energy at large scales, else it will pile up at the scale of the domain and a

statistical steady state will not be achieved. Introducing a linear drag, $-r\zeta$, into the vorticity equation, as in (1.58), is one means to achieve this, and such a term may be physically justified by appeal to Ekman layer theory (section ??). Although such a term appears to be scale invariant, its effects will be felt only at large scales because at smaller scales the timescale of the turbulence is much shorter than that of the friction. We may estimate the scale at which the drag becomes important by equating the drag timescale to the inertial timescale. The latter is given by (1.27), and equating this to the frictional timescale r^{-1} gives

$$r^{-1} = \varepsilon^{-1/3} k_r^{-2/3} \quad \rightarrow \quad k_r = \left(\frac{r^3}{\varepsilon} \right)^{1/2}, \quad (1.71)$$

where k_r is the frictional wavenumber. Frictional effects are important at scales *larger* than k_r^{-1} .

1.3.3 * More about the phenomenology

The phenomenology of two-dimensional turbulence is not quite as settled as the above arguments imply. Note, for example, that timescale (1.67) is independent of length scale, whereas in three dimensional turbulence the timescale decreases with length scale, which seems more physical and more conducive to spectrally local interactions. A useful measure of this locality is given by estimating the contributions to the straining rate, $S(k)$, from motions at all scales.

The strain rate scales like the shear, so that an estimate of the total strain rate is given by

$$S(k) = \left[\int_{k_0}^k \mathcal{E}(p) p^2 dp \right]^{1/2}, \quad (1.72)$$

where k_0 is the wavenumber of the largest scale present. The contributions to the integrand from a given wavenumber octave are given by

$$\int_p^{2p} \mathcal{E}(p') p'^3 d \log p' \sim \mathcal{E}(p) p^3. \quad (1.73)$$

In three dimensions, use of the $-5/3$ spectrum indicates that the contributions from each octave below a given wavenumber k increase with wavenumber, being a maximum close to k , and this is *a posteriori* consistent with the locality hypothesis. However, in two-dimensional turbulence with a -3 spectrum each octave makes the same contribution. That is to say, the contributions to the strain rate at a given wavenumber, as defined by (1.72), are not spectrally local. This does not prove that the enstrophy transfer is spectrally non-local, but nor does it build confidence in the theory.

Dimensionally the strain rate is the inverse of a time, and if this is a spectrally non-local quantity then, instead of (1.24), we might use the inverse of the strain rate as an eddy turnover time giving

$$\tau_k = \left[\int_{k_0}^k p^2 \mathcal{E}(p) dp \right]^{-1/2}. \quad (1.74)$$

This has the advantage over (1.24) in that it is a non-increasing function of wavenumber, whereas if the spectrum is steeper than k^{-3} (1.24) implies a timescale increasing with wavenumber. Using this in (1.63) gives a prediction for the enstrophy inertial range, namely

$$\mathcal{E}(k) = \mathcal{K}_\eta \eta^{2/3} [\log(k/k_0)]^{-1/3} k^{-3}, \quad (1.75)$$

which is similar to (1.64) except for a logarithmic correction. This expression is, of course, spectrally non-local, in contradiction to our original assumption: this new prediction has arisen by noting the spectral locality inherent in (1.72), and proposing a reasonable, although *ad hoc*, solution.

The discussion above suggests that phenomenology of the forward enstrophy cascade is on the verge of being internally inconsistent, and that the k^{-3} spectral slope might be the shallowest limit that is likely to be actually achieved in nature or in any particular computer simulation rather than a robust, universal slope. To see this argument, suppose the detailed fluid dynamics attempts in some way to produce a slope shallower than k^{-3} ; then, using (1.73), the strain is local and the shallow slope is forbidden by the Kolmogorovian scaling results. However, if the dynamics organizes itself into structures with a slope steeper than k^{-3} the strain is quite nonlocal. The fundamental assumption of Kolmogorov scaling is not satisfied, and there is no internal inconsistency — the theory simply doesn't apply. The k^{-3} slope itself is at the margin.

There are two other potential problems with the theory of two-dimensional turbulence that we have described. One is that enstrophy is only one of an infinity of invariants of inviscid two-dimensional flow, and the theory takes no account of the presence of others. The second is that, as in three-dimensional turbulence, if there is strong intermittency the flow cannot be fully characterized by single enstrophy and energy cascade rates. In spite of all this, the notions of a forward transfer of enstrophy and an inverse transfer of energy are quite robust, and have considerable numerical support.

1.3.4 Numerical illustrations

Numerical simulations nicely illustrate both the classical phenomenology and its shortcomings. In the simulations shown in Fig. 1.8 and Fig. 1.9 the vorticity field is initialized 'randomly', meaning that there is no structure in the initial field, but with only a few non-zero Fourier components, and the flow is allowed to freely evolve, save for the effects of a weak viscosity. Vortices soon form, and between them enstrophy is cascaded to small scales where it is dissipated, producing a flat and nearly featureless landscape. The energy cascade to larger scales is reflected in the streamfunction field, the length-scale of which slowly grows larger with time. The vortices themselves form through a roll-up mechanism, similar to that illustrated in Fig. ??, and their presence provides problems to the phenomenology. Because circular vortices are nearly exact, stable solutions of the inviscid equations they can 'store' enstrophy, disrupting the relationship between enstrophy flux and enstrophy itself that is assumed in the Kolmogorov-Kraichnan phenomenology and providing a form of intermittency.

Nevertheless, some forced-dissipative numerical simulations suggest that the pres-

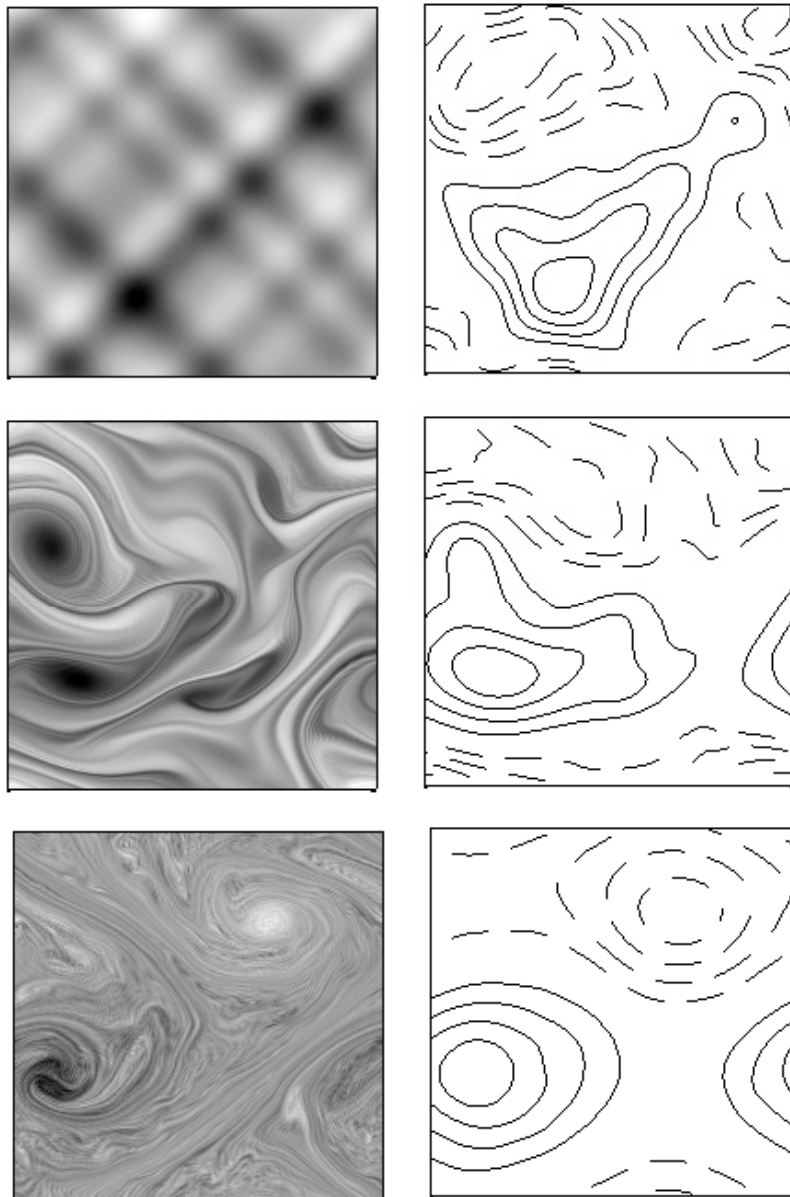


Fig. 1.8 Nearly-free evolution of vorticity (left column) and streamfunction (right column) in a doubly-periodic square domain (of length 2π) at times (from the top, and in units of inverse vorticity) $t = 0$, $t = 50$ and $t = 260$, obeying the two-dimensional vorticity equation with no forcing but with a weak viscous term. The initial conditions have just a few non-zero Fourier modes with randomly generated phases, producing a maximum value of vorticity of about 3. Kelvin-Helmholtz instability leads to vortex formation and roll-up (as in Fig. ??), and like-signed vortices merge, ultimately leading to a state of just two oppositely-signed vortices. Between the vortices enstrophy cascades to smaller scales. The scale of the stream function grows larger, reflecting the transfer of energy to larger scales.

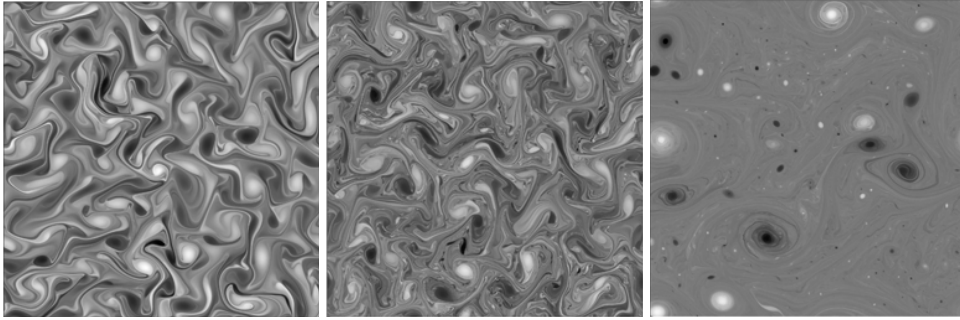


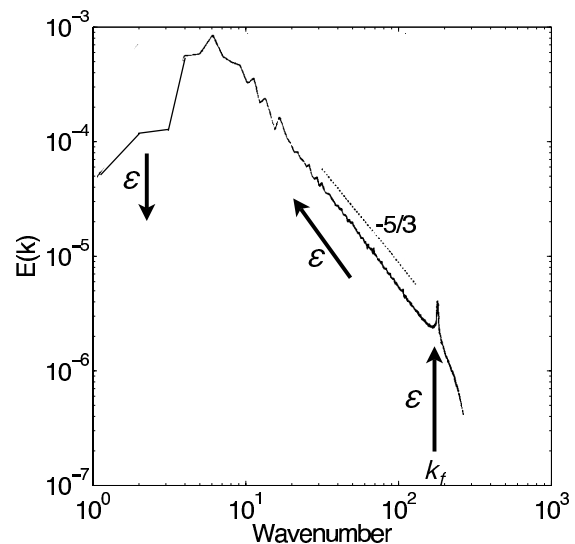
Fig. 1.9 Snapshots of the vorticity field in decaying two-dimensional turbulence, similar to Fig. 1.8, with time increasing left to right. The flow ultimately consists of a small number of vortices whose trajectories are similar to that of interacting point vortices, with occasional close encounters leading to vortex merger.

ence of vortices may be confined to scales close to that of the forcing, and if the resolution is sufficiently high then the $-5/3$ inverse cascade and -3 forward enstrophy cascade may appear. Certainly, if the forcing is spectrally localized, then a well-defined $-5/3$ spectrum robustly forms, as illustrated in Fig. 1.10. Typically, however, the forward k^{-3} spectrum is more delicate, being influenced by the presence of coherent vortices.

1.4 * PREDICTABILITY OF TURBULENCE

A turbulent flow contains multiple scales of motion; the error is typically initially largely confined to small scales, but the ‘predictability time’ of the atmosphere may be taken as the time taken to contaminate all scales of motion. Let us suppose that errors on a small scale will mostly contaminate the motion on the next larger scale (in a logarithmic

Figure 1.10 The energy spectrum in a numerical simulation of forced-dissipative two-dimensional turbulence. The fluid is stirred at wavenumber k_f and dissipated at large scales with a linear drag, and there is an $k^{-5/3}$ spectrum at intermediate scales. The arrows schematically indicate the direction of the energy flow.



sense), at that this contamination occurs on the local eddy turnover time. Eddies on this larger scale then grow and affect the next larger scale, and the error field is so cascaded upscale via local local triad interactions finally reaching the largest scales of the fluid.

Let us suppose that the error is initially confined to some small scale characterized by the (inverse of) the wavenumber k_1 , as determined by the resolution of our observing network. For modes at that scale the error may be considered finite rather than infinitesimal, and it will saturate and contaminate the next largest scale in a timescale comparable to the eddy turnover time at that scale. Thus, in general, errors initially confined to a scale k will contaminate the scale $2k$ after a time τ_k , with τ_k given by (1.24). The total time taken for errors to propagate from the small scale k_1 to the largest scale k_0 is then given by

$$T = \int_{k_0}^{k_1} \tau_k d(\ln k) = \int_{k_0}^{k_1} [k^3 \mathcal{E}(k)]^{-1/2} d(\ln k), \quad (1.76)$$

treating the wavenumber spectrum as continuous. The logarithmic integral arises because the cascade proceeds logarithmically — error cascades from k to $2k$ in a time τ_k . For an energy spectrum of the form $E = Ak^{-n}$ this becomes

$$T = \frac{2}{A^{1/2}(n-3)} \left[k^{(n-3)/2} \right]_{k_0}^{k_1}. \quad (1.77)$$

for $n \neq 3$, and $T = A^{-1/2} \ln(k_1/k_0)$ for $n = 3$. If in two dimensional turbulence we have $n = 3$ and $A = \eta^{2/3}$, and if in three-dimensional turbulence we have $n = 5/3$ and $A = \varepsilon^{2/3}$, then the respective predictability times are given by:

$$\boxed{\begin{aligned} T_{2d} &\sim \eta^{-1/3} \ln(k_1/k_0), \\ T_{3d} &\sim \varepsilon^{-1/3} k_0^{-2/3} \end{aligned}}. \quad (1.78a,b)$$

As $k_1 \rightarrow \infty$, that is as the initial error is confined to smaller and smaller scales, predictability time grows larger for two dimensional turbulence (and for $n \geq 3$ in general), but remains finite for three dimensional turbulence.

1.5 * SPECTRUM OF A PASSIVE TRACER

Let us now consider, heuristically, the spectrum of a passive tracer that obeys

$$\frac{D\phi}{Dt} = F[\phi] + \kappa \nabla^2 \phi, \quad (1.79)$$

where $F[\phi]$ is the stirring of the dye, and κ is its diffusivity, which in general differs from the kinematic molecular viscosity ν . If ϕ is temperature, the ratio of viscosity to diffusivity is called the *Prandtl number* and denoted σ , so that $\sigma \equiv \nu/\kappa$. If ϕ is a passive tracer, the ratio is sometimes called the *Schmidt number*, but we shall call it the Prandtl number in all cases. We assume that the tracer variance is created at some well-defined scale k_0 , and that κ is sufficiently small that dissipation only occurs at very

small scales. (Note that dissipation only reduces the tracer *variance*, not the amount of tracer itself.) The turbulent flow will generically tend to stretch patches of dye into elongated filaments, in much the same way as vorticity in two-dimensional turbulence is filamented — note that Fig. 1.5 applies just as well to a passive tracer in either two or three dimensions as it does to vorticity in two dimensions. Thus we expect a transfer of tracer variance from large-scales to small. If the dye is stirred at a rate χ then, by analogy with our treatment of the cascade of energy, we posit that

$$\mathcal{K}_\chi \chi \propto \frac{\mathcal{P}(k)k}{\tau_k}, \quad (1.80)$$

where $\mathcal{P}(k)$ is the spectrum of the tracer, k is the wavenumber, τ_k is an eddy timescale and \mathcal{K}_χ is a constant, not necessarily the same constant in all cases. (In the rest of the section, Kolmogorov-like constants will be denoted \mathcal{K} , differentiated with miscellaneous superscripts or subscripts.) Let us first assume that τ_k is given by

$$\tau_k = [k^3 \mathcal{E}(k)]^{-1/2}. \quad (1.81)$$

Suppose that the turbulent spectrum is given by $\mathcal{E}(k) = Ak^{-n}$, then using (1.81), (1.80) becomes

$$\mathcal{K}_\chi \chi = \frac{\mathcal{P}(k)k}{[Ak^{3-n}]^{-1/2}}, \quad (1.82)$$

and

$$\boxed{\mathcal{P}(k) = \mathcal{K}_\chi A^{-1/2} \chi k^{(n-5)/2}}. \quad (1.83)$$

Note that the steeper the energy spectrum the shallower the tracer spectrum. If the energy spectrum is steeper than -3 then (1.81) may not be a good estimate of the eddy turnover time, and we use instead

$$\tau_k = \left[\int_{k_0}^k p^2 \mathcal{E}(p) dp \right]^{-1/2}, \quad (1.84)$$

where k_0 is the low-wavenumber limit of the spectrum. If the energy spectrum is shallower than -3 , then the integrand is dominated by the contributions from high wavenumbers and (1.84) effectively reduces to (1.81). If the energy spectrum is steeper than -3 , then the integrand is dominated by contributions from low wavenumbers. For $k \gg k_0$ we can approximate the integral by $[k_0^3 \mathcal{E}(k_0)]^{-1/2}$, that is the eddy-turnover time at large scales, τ_{k_0} , given by (1.81). The tracer spectrum then becomes

$$\boxed{\mathcal{P}(k) = \mathcal{K}'_\chi \chi \tau_{k_0} k^{-1}}, \quad (1.85)$$

where \mathcal{K}'_χ is a constant. In all these cases the tracer cascade is to smaller scales even if, as may happen in two-dimensional turbulence, energy is cascading to larger scales.

The scale at which diffusion becomes important is given by equating the turbulent time-scale τ_k to the diffusive time-scale $(\kappa k^2)^{-1}$. This is independent of the flux of tracer, χ , essentially because the equation for the tracer is linear. Determination of expressions for these scales in two and three dimensions are left as problems for the reader.

1.5.1 Examples of tracer spectra

Energy inertial range flow in three dimensions

Consider a range of wavenumbers over which neither viscosity nor diffusivity directly influence the turbulent motion and the tracer. Then in (1.83) $A = \mathcal{K}\varepsilon^{2/3}$ where ε is the rate of energy transfer to small scales, \mathcal{K} is the Kolmogorov constant, and $n = 5/3$. The tracer spectrum becomes

$$\mathcal{P}(k) = \mathcal{K}_\chi^{3d} \varepsilon^{-1/3} \chi k^{-5/3}. \quad (1.86)$$

where \mathcal{K}_χ^{3d} is a (putatively universal) constant. It is interesting that the $-5/3$ exponent appears in both the energy spectrum and the passive tracer spectrum. Using (1.81), this is the only spectral slope for which this occurs. Experiments show that this range does, at least approximately, exist with a value of \mathcal{K}_χ^{3d} of about 0.5–0.6 in three dimensions.

Inverse energy-cascade range in two-dimensional turbulence

Suppose that the energy injection occurs at a smaller scale than the tracer injection, so that there exists a range of wavenumbers over which energy is cascading to larger scales while tracer variance is simultaneously cascading to smaller scales. The tracer spectrum is then

$$\mathcal{P}(k) = \mathcal{K}_\chi^{2d} \varepsilon^{-1/3} \chi k^{-5/3}, \quad (1.87)$$

the same as (1.86), although ε is now the energy cascade rate to larger scales and the constant \mathcal{K}_χ^{2d} does not necessarily equal \mathcal{K}_χ^{3d} .

Enstrophy inertial range in two-dimensional turbulence

In the forward enstrophy inertial range the eddy timescale is $\tau_k = \eta^{-1/3}$ (assuming of course that the classical phenomenology holds). Directly from (1.80) the corresponding tracer spectrum is then

$$\mathcal{P}(k) = \mathcal{K}_\chi^{2d*} \eta^{-1/3} \chi k^{-1}. \quad (1.88)$$

The passive tracer spectrum now has the same slope as the spectrum of vorticity variance (i.e., the enstrophy spectrum), which is perhaps comforting since the tracer and vorticity obey similar equations in two dimensions.

The viscous-advective range of large Prandtl number flow

If $\sigma = \nu/\kappa \gg 1$ (and in seawater $\sigma \approx 7$) then there may exist a range of wavenumbers in which viscosity is important but not tracer diffusion. The energy spectrum is then very steep, and (1.85) will apply. The straining then comes from wavenumbers near the viscous scale, so that for three dimensional flow the appropriate k_0 to use in (1.85) is the viscous wavenumber, and $k_0 = k_\nu = (\varepsilon/\nu^3)^{1/4}$. The dynamical timescale at this wavenumber is given by

$$\tau_{k_\nu} = \left(\frac{\nu}{\varepsilon}\right)^{1/2}, \quad (1.89)$$

and using this and (1.85) the tracer spectrum in this viscous-advective range becomes

$$\mathcal{P}(k) = \mathcal{K}'_B \left(\frac{\nu}{\varepsilon}\right)^{1/2} \chi k^{-1}. \quad (1.90)$$

This spectral form applies for $k_\nu < k < k_\kappa$, where k_κ is the wavenumber at which diffusion becomes important, found by equating the eddy turnover time given by (1.89) with the diffusive timescale $(\kappa k^2)^{-1}$. This gives

$$k_\kappa = \left(\frac{\varepsilon}{\nu \kappa^2} \right)^{1/4}, \quad (1.91)$$

and k_κ is known as the Batchelor wavenumber (and its inverse is the Batchelor scale). Beyond k_κ , the diffusive flux is not constant and the tracer spectrum can be expected to decay as wavenumber increases. A heuristic way to calculate the spectrum in this range is to first note that in the diffusive range the flux of the tracer is no longer constant but diminishes according to

$$\frac{d\chi'(k)}{dk} = -2\kappa k^2 \mathcal{P}(k). \quad (1.92)$$

where χ' is the wavenumber-dependent rate of tracer transfer. Let us nevertheless assume that χ' and $\mathcal{P}(k)$ are related by (1.80), except that now we take the eddy turnover time to be a constant, given by (1.89). Thus,

$$\mathcal{K}_B \chi' = \frac{\mathcal{P}(k)k}{\tau_{k_\kappa}} = \frac{\mathcal{P}(k)k}{(\nu/\varepsilon)^{1/2}} \quad (1.93)$$

where \mathcal{K}_B is a constant. Using (1.92) and (1.93) we obtain

$$\frac{d\chi'}{dk} = -2\mathcal{K}_B \kappa k \left(\frac{\nu}{\varepsilon} \right)^{1/2} \chi. \quad (1.94)$$

Solving this, using $\chi' = \chi$ for small k , gives

$$\mathcal{P}(k) = \mathcal{K}_B \left(\frac{\nu}{\varepsilon} \right)^{1/2} \chi k^{-1} \exp[-\mathcal{K}_B (k/k_\kappa)^2]. \quad (1.95)$$

This reduces to (1.90) if $k \ll k_\kappa$, and is known as the Batchelor spectrum.

In two dimensions the viscous-advective range occurs for wavenumbers greater than $k_\nu = (\eta/\nu^3)^{1/6}$. The appropriate timescale within this subrange by $\eta^{-1/3}$, and therefore gives a spectrum with the precisely the same form as (1.88). At sufficiently high wavenumbers tracer diffusion becomes important, with the diffusive scale now given by equating the eddy turnover time $\eta^{-1/3}$ with the viscous timescale $(\kappa k^2)^{-1}$. This gives the diffusive wavenumber, analogous to (1.91), of $k_\kappa = (\eta/\kappa^3)^{1/6}$. Using (1.94) and the procedure above we then obtain an expression for the spectrum in the region $k > k_\nu$, that is a two-dimensional analog of (1.95), namely

$$\mathcal{P}(k) = \mathcal{K}'_B \eta^{-1/3} \chi k^{-1} \exp[-\mathcal{K}'_B (k/k_\kappa)^2]. \quad (1.96)$$

For $k \ll k_\kappa$ this reduces to (1.88), possibly with a different value of the Kolmogorov-like constant.

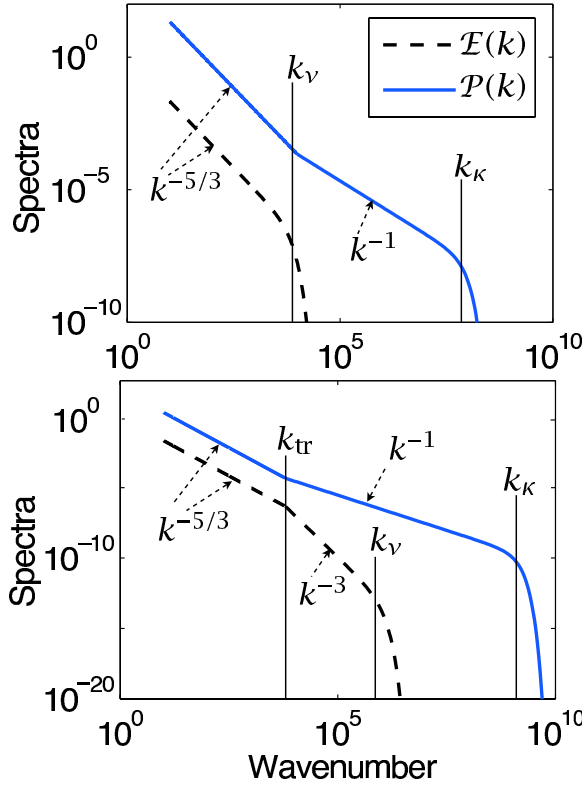


Figure 1.11 The energy spectra, $\mathcal{E}(k)$ and passive tracer spectra $\mathcal{P}(k)$ in large Prandtl number three-dimensional turbulence (top) and two-dimensional turbulence (bottom). In three dimensions $\mathcal{P}(k)$ is given by (1.86) for $k < k_v$ and by (1.95) for $k > k_v$. In two dimensions, if k_{tr} marks the transition between a $k^{-5/3}$ inverse energy cascade and a k^{-1} forward enstrophy cascade, then $\mathcal{P}(k)$ is given by (1.87) for $k < k_{tr}$ and by (1.96) for $k > k_{tr}$. In both two and three dimensions the tracer spectra falls off rapidly for $k > k_\kappa$.

* *The inertial-diffusive range of small Prandtl number flow*

For small Prandtl number ($\nu/\kappa \ll 1$) the energy inertial range may co-exist with a range over which tracer variance is being dissipated, giving us the so-called inertial-diffusive range. The tracer will begin to be dissipated at a wavenumber obtained by equating a dynamical eddy turnover time with a diffusive time, and this gives a diffusive wavenumber

$$k'_\kappa = \begin{cases} (\varepsilon/\kappa^3)^{1/4} & \text{in three dimensions,} \\ (\eta/\kappa^3)^{1/6} & \text{in two dimensions.} \end{cases} \quad (1.97)$$

Beyond the diffusive wavenumber the flux of the tracer is no longer constant but diminishes according to (1.92).

Given a non-constant flux and an eddy-turnover time that varies with wavenumber there is no self-evidently correct way to proceed. One way is to assume that χ and $\mathcal{P}(k)$ are related by $\mathcal{K}'_\chi \chi = \mathcal{P}(k)k/\tau_k$ [as in (1.93), but with a potentially different proportionality constant] and with τ_k given by (1.81); that is, $\tau_k = \varepsilon^{-1/3}k^{-2/3}$ in three dimensional turbulence. Using this in (1.92) leads to

$$\mathcal{P}(k) = \mathcal{K}'_\chi \chi \varepsilon^{-1/3} k^{-5/3} \exp[-(\mathcal{K}'_\chi 3/2)(k/k'_\kappa)^{4/3}]. \quad (1.98)$$

where χ is the tracer flux at the beginning of the tracer dissipation range. (A similar

expression emerges in two dimensional turbulence.) However, given such a steep spectrum an argument based on spectral locality is sometimes thought to be suspect. Another argument posits a particular relationship between the tracer spectrum and energy spectrum in the inertial-diffusive range, and this leads to

$$\mathcal{P}(k) = \frac{\mathcal{K}_B''}{3} \chi_0 \varepsilon^{2/3} \kappa^{-3} k^{-17/3} = \mathcal{K}_B'' \chi_0 \varepsilon^{-1/3} k^{-5/3} g(k/k_\kappa), \quad (1.99)$$

where $g(\alpha) = \alpha^{-4}/3$ and \mathcal{K}_B'' is a constant.

Geostrophic Turbulence and Baroclinic Eddies

Geostrophic turbulence may be defined as turbulence in stably-stratified flow that is in near-geostrophic balance. The constraining effects of rotation and stratification that are so important are captured in a simple and direct way by the quasi-geostrophic equations and these will be our main tool. Let us consider the effects of rotation first, then stratification.

2.1 EFFECTS OF DIFFERENTIAL ROTATION IN TWO-DIMENSIONAL TURBULENCE

One of the effects of rapid rotation on a fluid is its two-dimensionalization, a manifestation of the Taylor-Proudman effect. In the limit of motion of a scale much shorter than the deformation radius, and with no topography, the quasi-geostrophic potential vorticity equation reduces to the two-dimensional equation,

$$\frac{Dq}{Dt} = 0 \tag{2.1}$$

where $q = \zeta + f$. This is perhaps the simplest equation with which to study the effects of rotation on turbulence. The effects of rotation are of course already playing a role in enabling us to reduce a complex three-dimensional flow to two-dimensional flow. Further, suppose that the Coriolis parameter is constant. Then (2.1) becomes simply the two-dimensional vorticity equation

$$\frac{D\zeta}{Dt} = 0. \tag{2.2}$$

Thus constant rotation has *no* effect on purely two-dimensional motion. Flow that is already two-dimensional — flow on a soap film, for example — is unaffected by rotation.

Suppose, though, that the Coriolis parameter is variable, as in $f = f_0 + \beta y$. Then we have

$$\frac{D}{Dt}(\zeta + \beta y) = 0 \quad \text{or} \quad \frac{D\zeta}{Dt} + \beta v = 0. \quad (2.3a,b)$$

If the dominant term in these equations is the one involving β , then we obtain $\beta v = 0$. That is, there is no flow in the meridional direction and any flow is purely *zonal*. This constraint may be interpreted as a consequence of angular momentum and energy conservation. A ring of fluid encircling the earth at a velocity u has an angular momentum per unit mass $a \cos \theta (u + \Omega a \cos \theta)$ where θ is the latitude and a is the radius of the earth. Moving this ring of air polewards (i.e., giving it a meridional velocity) while conserving its angular momentum requires its velocity and hence energy to increase. Unless there is a source for that energy the flow is constrained to remain zonal.

2.1.1 Organization of turbulence into zonal flow

Scaling

Let us now consider how flow can become organized into zonal bands, from the perspective of two-dimensional turbulence. Re-write (2.1) in full as

$$\frac{\partial \zeta}{\partial t} + \mathbf{u} \cdot \nabla \zeta + \beta v = 0. \quad (2.4)$$

If $\zeta \sim U/L$ and if $t \sim T$ then the respective terms in this equation scale as

$$\frac{U}{LT} \quad \frac{U^2}{L^2} \quad \beta U \quad (2.5)$$

How time scales (i.e., advectively or with a Rossby wave frequency scaling) is determined by which of the other two terms dominates, and this in turn is scale dependent. For large scales the β -term will be dominant, and at smaller scales the advective term is dominant. The cross-over scale, or the ' β -scale' or 'Rhines scale' L_β , is given by

$$L_\beta \sim \sqrt{\frac{U}{\beta}}. \quad (2.6)$$

This is not a unique definition of the cross-over scale, since we have chosen the same length scale that connects vorticity to velocity and to be the β -scale, and it is by no means *a priori* clear that this should be so. If the scale is different, the three terms in (2.4) scale as

$$\frac{Z}{T} : \frac{UZ}{L} : \beta U \quad (2.7)$$

where Z is the scaling for vorticity (i.e., $\zeta = \mathcal{O}Z$). Equating the second and third terms gives the scale

$$L_{\beta Z} = \frac{Z}{\beta}. \quad (2.8)$$

In any case, (2.6) and (2.8) both indicate that at some *large* scale Rossby waves are likely to dominate whereas at small scales advection, and turbulence, dominates.

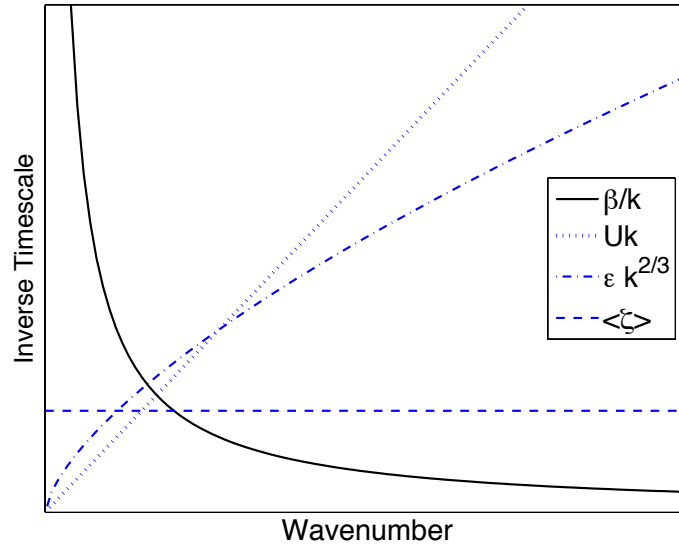


Fig. 2.1 Three estimates of the wave-turbulence cross-over, in wavenumber space. The solid curve is the frequency of Rossby waves, proportional to β/k . The other three curves are various estimates of the inverse turbulence timescale, or ‘turbulence frequency’. These are the turbulent eddy transfer rate, proportional to $\varepsilon k^{2/3}$ in a $k^{-5/3}$ spectrum; the simple estimate Uk where U is an rms velocity; and the mean vorticity, which is constant. Where the Rossby wave frequency is larger (smaller) than the turbulent frequency, i.e., at large (small) scales, Rossby waves (turbulence) dominate the dynamics.

Another heuristic way to derive (2.6) is by a direct consideration of timescales. The Rossby wave frequency is β/k and an inverse advective timescale is Uk , where k is the wavenumber. Equating these two gives an equation for the β -wavenumber

$$k_{\beta} \sim \sqrt{\frac{\beta}{U}}. \quad (2.9)$$

This equation is the inverse of (2.6), but note that factors of order unity cannot be revealed by simple scaling arguments such as these. The cross-over between waves and turbulence is reasonably sharp, as indicated in Fig. 2.1.

Phenomenology

Can we be more precise about the scaling, using the phenomenology of turbulence? Let us suppose that the fluid is stirred at some well-defined scale k_f , producing an energy input ε . Then (assuming no energy is lost to smaller scales) energy cascades to large scales at that same rate. At some scale, the β term in the vorticity equation will start to make its presence felt. By analogy with the procedure for finding the dissipation scale in turbulence, we can find the scale at which linear Rossby waves dominate by equating

the inverse of the turbulent eddy turnover time to the Rossby wave frequency. The eddy-turnover time is

$$\tau_k = \varepsilon^{-1/3} k^{-2/3}, \quad (2.10)$$

and equating this to the inverse Rossby wave frequency k/β gives the β -scale

$$k_\beta \sim \left(\frac{\beta^3}{\varepsilon} \right)^{1/5}. \quad (2.11)$$

From a practical perspective this is less useful than (2.9), since it is generally much easier to measure velocities than energy transfer rates, or even vorticity. Nonetheless, it is a little more fundamental from the point of view of turbulence since one can often imagine that ε is determined by processes largely independent of β , whereas the magnitude of the eddies (i.e. U) at the energy containing scales is likely to be a function of β .

Generation of anisotropy

None of the measures discussed so far take into account the anisotropy inherent in Rossby waves, nor do they suggest how the flow might organize itself into zonal structures. To understand that, let us note that energy transfer will be relatively inefficient at those scales where linear Rossby waves dominate. But the wave-turbulence boundary is not isotropic; the Rossby wave frequency is quite anisotropic, being given by

$$\omega_\beta = -\frac{\beta k^x}{k^{x^2} + k^{y^2}}. \quad (2.12)$$

If, as a first approximation, we suppose that the turbulent part of the flow remains isotropic, the wave turbulence boundary is then given from the solution of

$$\varepsilon^{1/3} k^{2/3} = \frac{\beta k^x}{k^2} \quad (2.13)$$

where k is the isotropic wavenumber. Solving this gives expressions for the x- and y-wavenumber components of the wave-turbulence boundary, namely

$$k_\beta^x = \left(\frac{\beta^3}{\varepsilon} \right)^{1/5} \cos^{8/5} \theta, \quad k_\beta^y = \left(\frac{\beta^3}{\varepsilon} \right)^{1/5} \sin \theta \cos^{3/5} \theta, \quad (2.14)$$

where the polar coordinate is parameterized by the angle $\theta = \tan^{-1}(k^y/k^x)$. This rather uninformative-looking formula is illustrated in Fig. 2.3.

The region inside the dumbbell shapes in Fig. 2.3 is dominated by Rossby waves, where the natural frequency of the oscillation is *higher* than the turbulent frequency. If the flow is stirred at a wavenumber higher than this the energy will cascade to larger scales, but because of the frequency mismatch the turbulent flow will be unable to efficiently excite modes within the dumbbell. Nevertheless, there is still a natural tendency of the energy to seek the gravest mode, and it may do this by cascading toward the $k^x = 0$ axis — that is, toward zonal flow. In this way zonally elongated structures are produced.

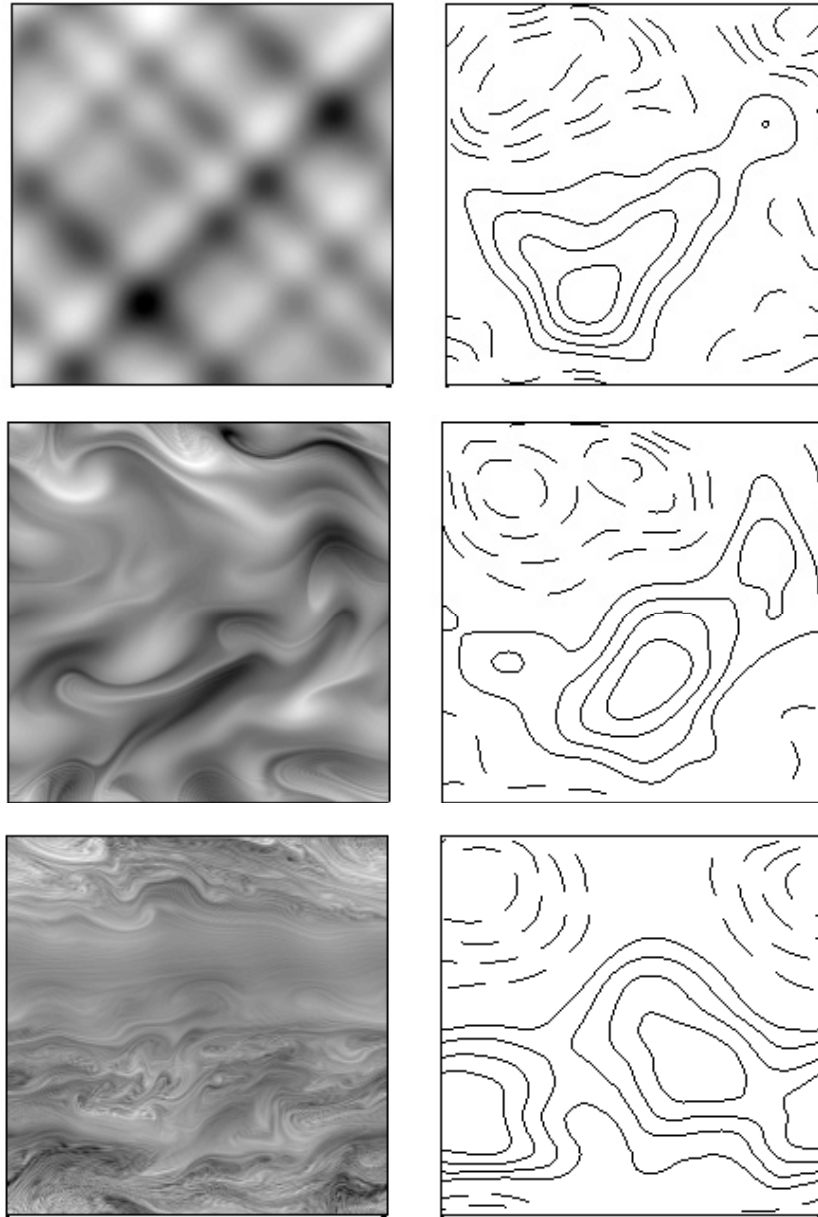
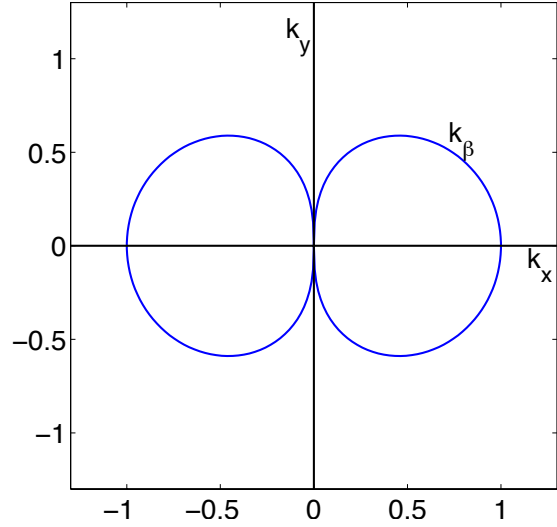


Fig. 2.2 Evolution of vorticity (greyscale, left column) and streamfunction (contour plots, right column) in a doubly-periodic square domain (of length 2π) at times $t = 0$, $t = 50$ and $t = 260$ (in units of inverse vorticity), obeying (2.4) with the addition of a weak viscous term on the right-hand side. The initial conditions are the same as for Fig. 1.8, with maximum value of vorticity about 3. As $\beta = 3$, the β -Rossby number, $|\zeta|/\beta L$ is about unity. Compared to Fig. 1.8, vortex formation is inhibited and there is tendency toward zonal flow.

Figure 2.3 The anisotropic wave-turbulence boundary k_β , in wave-vector space calculated by equating the turbulent eddy transfer rate, proportional to $k^{2/3}$ in a $k^{-5/3}$ spectrum, to the Rossby wave frequency $\beta k^x/k^2$, as in (2.14). Within the dumbbell Rossby waves dominate and energy transfer is inhibited. The inverse cascade plus Rossby waves thus leads to a generation of zonal flow.



Slight variations on this theme are produced by using different expressions for the ‘turbulence frequency’. For example, if we use the simple expression Uk then the wave turbulence boundary is given from

$$Uk = \frac{\beta k^x}{k^2}, \quad (2.15)$$

which has solutions that may be written as

$$k_\beta^x = \left(\frac{\beta}{U}\right)^{1/2} \cos^{3/2} \theta, \quad k_\beta^y = \left(\frac{\beta}{U}\right)^{1/2} \sin \theta \cos^{1/2} \theta. \quad (2.16a,b)$$

A plot of this is very similar to Fig. 2.3.

Does this putative mechanism actually work? Fig. 2.4 shows the freely evolving (unforced, inviscid) energy spectrum in a simulation on a β -plane, with an initially isotropic spectrum. The energy implodes, cascading to larger scales but avoiding the region inside the dumbbell and piling up at $k^x = 0$. A similar picture emerges in a forced-dissipative simulations, and with zonally-periodic boundary conditions these show a robust tendency to produce zonally-elongated structures and jets (Fig. 2.5). In closed domains, such as occur in the earth’s ocean, the production of such jets is interrupted by the meridional boundaries.

2.2 STRATIFIED GEOSTROPHIC TURBULENCE

2.2.1 Quasi-geostrophic flow as an analogue to two-dimensional flow

Now let us consider stratified effects in a simple setting, namely the quasi-geostrophic equations with constant Coriolis parameter and constant stratification. The (dimensional) unforced and inviscid governing equation may then be written

$$\frac{Dq}{Dt} = 0, \quad q = \nabla^2 \psi + Pr^2 \frac{\partial^2 \psi}{\partial z^2}, \quad (2.17a)$$

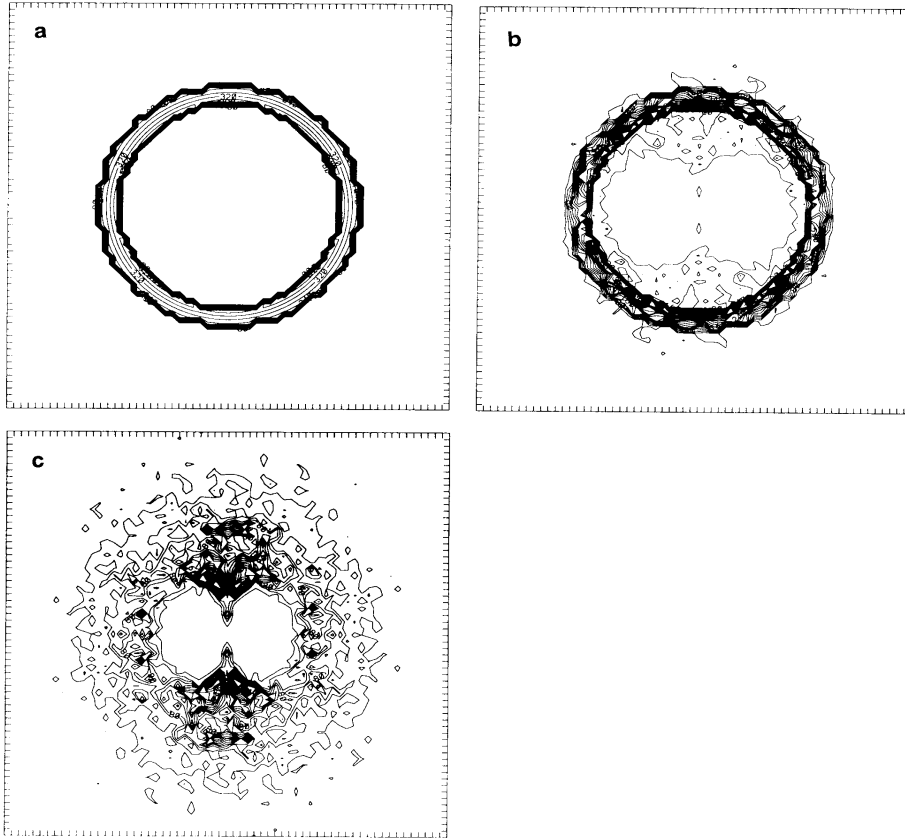


Fig. 2.4 Evolution of the energy spectrum in a freely-evolving two-dimensional simulation on the β -plane. The panels show contours of energy in wavenumber (k, l) space. The initial spectrum (a) is isotropic. The energy ‘implodes’, but its passage to large scales is impeded by the β -effect, and panels (b) and (c) show the spectrum at later times, illustrating the dumbbell predicted by (2.14) and Fig. 2.3.

where $Pr = f_0/N$ is the *Prandtl ratio* (and Pr/H is the inverse of the deformation radius) and $D/Dt = \partial/\partial t + \mathbf{u} \cdot \nabla$ is the two-dimensional material derivative. The vertical boundary conditions are

$$\frac{D}{Dt} \left(\frac{\partial \psi}{\partial z} \right) = 0, \quad \text{at } z = 0, H. \quad (2.17b)$$

These equations are analogous to the equations of motion for purely two-dimensional flow. In particular, with periodic lateral boundary conditions, or conditions of no-normal flow, there are two quadratic invariants of the motion, the energy and the enstrophy, which are obtained by multiplying (2.17a) by ψ and q and integrating over the domain.

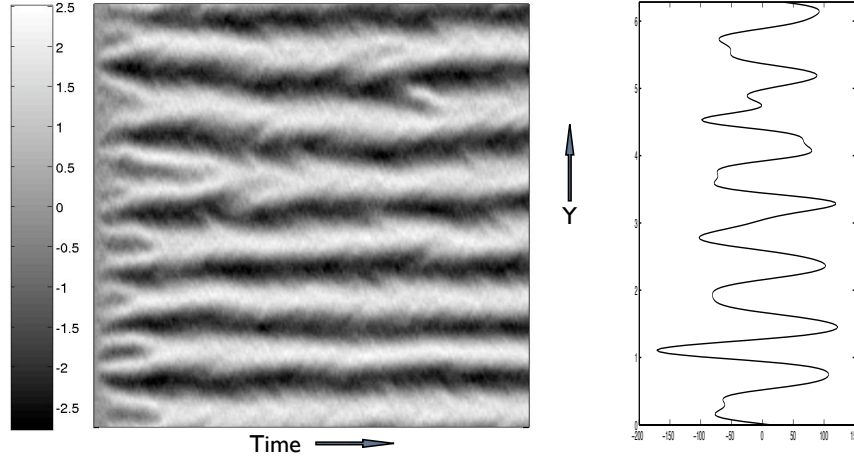


Fig. 2.5 Left: Gray-scale image of zonally average zonal velocity (\bar{u}) as a function of time and latitude (Y), produced in a simulation forced around wavenumber 80 and with $k_\beta = \sqrt{\beta/\bar{U}} \approx 10$ (in a domain of size 2π). Right: Values of $\partial^2 \bar{u} / \partial y^2$ as a function of latitude, late in the integration. Jets form very quickly from the random initial conditions, and are subsequently quite steady.

The conserved energy is

$$\frac{d\hat{E}}{dt} = 0, \quad \hat{E} = \int_V \left[(\nabla\psi)^2 + Pr^2 \left(\frac{\partial\psi}{\partial z} \right)^2 \right] dV, \quad (2.18)$$

where the integral is over a *three-dimensional* domain. The enstrophy is conserved at each vertical level, and of course the volume integral is also conserved, namely

$$\frac{d\hat{Z}}{dt} = 0, \quad \hat{Z} = \int_V q^2 dV = \int_V \left[\nabla^2 \psi + Pr^2 \left(\frac{\partial^2 \psi}{\partial z^2} \right) \right]^2 dV. \quad (2.19)$$

The analogy with two-dimensional flow is even more transparent if we further rescale the vertical coordinate by $1/Pr$, and so let $z' = z/Pr$. Then the energy and enstrophy invariants are:

$$\hat{E} = \int (\nabla_3 \psi)^2 dV, \quad \hat{Z} = \int q^2 dV = \int (\nabla_3^2 \psi)^2 dV \quad (2.20)$$

where $\nabla_3 = \mathbf{i} \partial / \partial x + \mathbf{j} \partial / \partial y + \mathbf{k} \partial / \partial z'$. The invariants then have almost same form as the two-dimensional invariants, but with a three-dimensional Laplacian operator instead of a two-dimensional one.

Given these invariants, we should expect that any dynamical behaviour that occurs in the two-dimensional equations *that depends solely on the energy/enstrophy constraints* should have an analogy in quasi-geostrophic flow. In particular, the transfer of energy to large-scales and enstrophy to small scales will also occur in quasi-geostrophic flow with, in so far as these transfers are effected by a local cascade, corresponding spectra of $k^{-5/3}$ and a k^{-3} . However, the wavenumber is the now *three-dimensional* wavenumber,

appropriately scaled by the Prandtl ratio in the vertical. Interestingly, then, the energy cascade to larger horizontal scales is accompanied by a cascade to larger vertical scales — a *barotropization* of the flow. This is an important and robust process in geostrophic turbulence and we come back to it later. However, the analogy should not be taken too far, because in quasi-geostrophic flow the potential vorticity is advected only by the horizontal flow. Thus, the dynamics of quasi-geostrophic turbulence will *not* in general be isotropic in three-dimensional wavenumber. To examine the detailed dynamical behaviour of quasi-geostrophic turbulence, we turn to a simpler model, that of two-layer flow

2.2.2 Two-layer geostrophic turbulence

Let us now consider flow in two layers of equal depth, governed by the quasi-geostrophic equations with (for now) $\beta = 0$, namely

$$\frac{\partial q_i}{\partial t} + J(\psi_i, q_i) = 0, \quad i = 1, 2, \quad (2.21)$$

where

$$q_1 = \nabla^2 \psi_1 + \frac{1}{2} k_d^2 (\psi_2 - \psi_1), \quad q_2 = \nabla^2 \psi_2 + \frac{1}{2} k_d^2 (\psi_1 - \psi_2), \quad (2.22a)$$

$$J(a, b) = \frac{\partial a}{\partial x} \frac{\partial b}{\partial y} - \frac{\partial b}{\partial y} \frac{\partial a}{\partial x}, \quad \frac{1}{2} k_d^2 = \frac{2f_0^2}{g'H} \equiv \frac{4f_0^2}{N^2 H^2}. \quad (2.22b)$$

The wavenumber k_d is inversely proportional to the baroclinic radius of deformation, and the two equivalent expressions given are appropriate in a layered model and a level model, respectively. The equations conserve the total energy,

$$\frac{d\hat{E}}{dt} = 0, \quad \hat{E} = \frac{1}{2} \int \left[(\nabla \psi_1)^2 + (\nabla \psi_2)^2 + \frac{1}{2} k_d^2 (\psi_1 - \psi_2)^2 \right] dA, \quad (2.23)$$

and the enstrophy in each layer

$$\frac{d\hat{Z}_1}{dt} = 0, \quad \hat{Z}_1 = \int_A q_1^2 dA, \quad (2.24)$$

$$\frac{d\hat{Z}_2}{dt} = 0, \quad \hat{Z}_2 = \int_A q_2^2 dA. \quad (2.25)$$

The first two terms in the energy expression represent the kinetic energy, and the last term is the available potential energy, proportional to the variance of temperature.

Baroclinic and barotropic decomposition

Define the barotropic and barotropic streamfunctions by

$$\psi \equiv \frac{1}{2} (\psi_1 + \psi_2), \quad \tau \equiv \frac{1}{2} (\psi_1 - \psi_2). \quad (2.26)$$

Then the potential vorticities for each layer may be written:

$$q_1 = \nabla^2 \psi + (\nabla^2 - k_d^2) \tau \quad (2.27a)$$

$$q_2 = \nabla^2 \psi - (\nabla^2 - k_d^2) \tau \quad (2.27b)$$

and the equations of motion may be rewritten as evolution equations for ψ and τ as follows:

$$\frac{\partial}{\partial t} \nabla^2 \psi + J(\psi, \nabla^2 \psi) + J(\tau, (\nabla^2 - k_d^2) \tau) = 0 \quad (2.28a)$$

$$\frac{\partial}{\partial t} (\nabla^2 - k_d^2) \tau + J(\tau, \nabla^2 \psi) + J(\psi, (\nabla^2 - k_d^2) \tau) = 0 \quad (2.28b)$$

We note the following:

- (i) ψ and τ are like vertical modes. That is, ψ is the barotropic mode with a ‘vertical wavenumber’, k^z , of zero, and τ is a baroclinic mode with a vertical wavenumber of one.
- (ii) Just as purely two dimensional turbulence can be considered to be a plethora of interacting triads, whose two-dimensional vector wavenumbers sum to zero, it is clear from (2.28b) geostrophic turbulence may be considered to be similarly comprised of a sum of interacting triads. The types of triad interaction are:

$$(\psi, \psi) \rightarrow \psi, \quad (\tau, \tau) \rightarrow \psi, \quad (\psi, \tau) \rightarrow \tau. \quad (2.29)$$

The first kind is a *barotropic triad*, for it involves only the barotropic mode. The other two are examples of a *baroclinic triad*. If a barotropic mode has a vertical wavenumber of zero, and a baroclinic mode has a vertical wavenumber of plus or minus one, then the vertical wavenumbers of the triad interactions must sum to zero. There is no triad that involves only the baroclinic mode, as we may see from the form of (2.28). (If the layers are of unequal depths, then purely baroclinic triads do exist.)

- (iii) Wherever the Laplacian operator acts on τ , it is accompanied by $-k_d^2$. That is, it is as if the effective horizontal wavenumber (squared) of τ is shifted, so that $k^2 \rightarrow k^2 + k_d^2$.

Conservation properties

Multiplying (2.28a) by ψ and (2.28b) by τ and horizontally integrating over the domain, assuming once again that the domain is either periodic or has solid walls, gives

$$\hat{T} = \int_A (\nabla \psi)^2 dA, \quad \frac{d\hat{T}}{dt} = \int_A \psi J(\tau, (\nabla^2 - k_d^2) \tau) dA \quad (2.30a)$$

$$\hat{C} = \int_A [(\nabla \tau)^2 + k_d^2 \tau^2] dA, \quad \frac{d\hat{C}}{dt} = \int_A \tau J(\psi, (\nabla^2 - k_d^2) \tau) dA. \quad (2.30b)$$

Here, \hat{T} is the energy associated with the barotropic flow and \hat{C} is the energy of the baroclinic flow. An integration by parts shows that

$$\int_A \psi J(\tau, (\nabla^2 - k_d^2)\tau) dA = - \int_A \tau J(\psi, (\nabla^2 - k_d^2)\tau) dA, \quad (2.31)$$

and therefore

$$\frac{d\hat{E}}{dt} = \frac{d}{dt}(\hat{T} + \hat{C}) = 0. \quad (2.32)$$

That is, total energy is conserved.

An enstrophy invariant is obtained by multiplying (2.28a) by $\nabla^2\psi$ and (2.28b) by $(\nabla^2 - k_d^2)\tau$ and integrating over the domain and adding the two expressions. The result is

$$\frac{d\hat{Z}}{dt} = 0, \quad \hat{Z} = \int_A (\nabla^2\psi)^2 + [(\nabla^2 - k_d^2)\tau]^2 dA. \quad (2.33)$$

This also follows from (2.24).

Just as for two-dimensional turbulence, we may define the spectra of the energy and enstrophy. Then, with obvious notation, for the energy we have

$$\hat{T} = \int \mathcal{T}(k) dk \quad \text{and} \quad \hat{C} = \int C(k) dk, \quad (2.34)$$

and the enstrophy spectrum $Z(k)$ is related to the energy spectra by

$$\hat{Z} = \int Z(k) dk = \int [k^2\mathcal{T}(k) + (k^2 + k_d^2)C(k)] dk. \quad (2.35)$$

which is analogous to the relationship between energy and enstrophy in two-dimensional flow. We thus begin to suspect that the phenomenology to two-layer turbulence is closely related to, but perhaps richer than, that of two-dimensional turbulence.

2.2.3 Triad interactions

Two types of triad interactions are possible:

Barotropic triads: An interaction that is purely barotropic (i.e., as if $\tau = 0$) conserves \hat{T} , the barotropic energy, and the associated enstrophy $\int k^2\mathcal{T}(k)dk$, and a barotropic triad behaves as purely two-dimensional flow. Explicitly, the conserved quantities are

$$\text{Energy:} \quad \frac{d}{dt}(\mathcal{T}(k) + \mathcal{T}(p) + \mathcal{T}(q)) = 0, \quad (2.36)$$

$$\text{Enstrophy:} \quad \frac{d}{dt}(k^2\mathcal{T}(k) + p^2\mathcal{T}(p) + q^2\mathcal{T}(q)) = 0. \quad (2.37)$$

Baroclinic triads: Baroclinic triads involve two baroclinic wavenumbers (say p, q) interacting with a barotropic wavenumber (say k). The energy and enstrophy conservation laws for this triad are

$$\text{Energy:} \quad \frac{d}{dt} (\mathcal{T}(k) + C(p) + C(q)) = 0, \quad (2.38a)$$

$$\text{Enstrophy:} \quad \frac{d}{dt} (k^2 \mathcal{T}(k) + (p^2 + k_d^2)C(p) + (q^2 + k_d^2)C(q)) = 0. \quad (2.38b)$$

Consider the following four cases of baroclinic triad:

- i. $(p, q) \gg k_d$. Then neglect k_d^2 in (2.38a) and (2.38b), and a baroclinic triad behaves like a barotropic triad, for (2.38b) is similar to (2.37). Alternatively, but equivalently, reconsider the layer form of the equations,

$$\frac{\partial q_i}{\partial t} + J(\psi_i, q_i) = 0 \quad (2.39)$$

where

$$q_i = \nabla^2 \psi_i + k_d^2 (\psi_j - \psi_i) \approx \nabla^2 \psi_i \quad i = 1, 2, j = 3 - i \quad (2.40)$$

In this case, each layer is decoupled from the other. Enstrophy is cascaded to small scales and, were there to be an energy source at small scales, energy would be transferred upscales until it reached a scale comparable with the deformation scale. Note that the transfer of enstrophy to small scales in a purely two-dimensional fashion depends on the two-layer nature of the flow. In reality, the small scales of a continuously stratified flow may not be representable by a two-layer model: remember that in a continuously stratified quasi-geostrophic model the enstrophy cascade occurs in *three-dimensional* wavenumber. Thus, as the horizontal scales become smaller, so does the vertical scale and higher deformation radii will start to play a role.

- ii. $(p, q, k) \ll k_d$. The energy and enstrophy conservation laws collapse to:

$$\frac{d}{dt} (C(p) + C(q)) = 0. \quad (2.41)$$

That is to say, energy is conserved among the baroclinic modes alone, with the barotropic mode k mediating the interaction. There is no constraint preventing the transfer of baroclinic energy to smaller scales, and no production of barotropic energy at $k \ll k_d$.

- iii. $(p, q, k) \sim k_d$. In this case both baroclinic and barotropic modes are important. Suppose that we define the pseudo-wavenumber k' by $k'^2 \equiv k^2 + k_d^2$ for a baroclinic mode and $k'^2 = k^2$ for a barotropic mode, and similarly for p' and q' . Then energy and enstrophy conservation can be written

$$\frac{d}{dt} (\mathcal{E}(k) + \mathcal{E}(p) + \mathcal{E}(q)) = 0, \quad (2.42a)$$

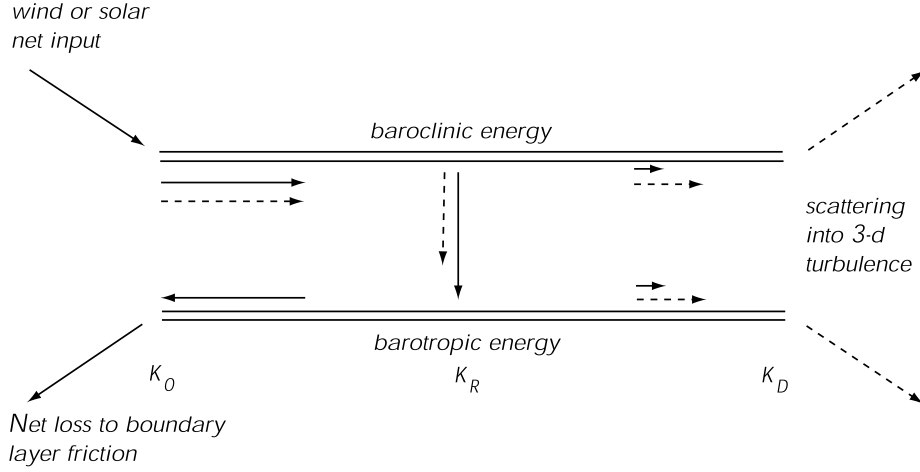


Fig. 2.6 Schema of idealized two-layer baroclinic turbulence. The horizontal axis represents horizontal wavenumber, and the vertical variation is decomposed into two vertical modes — the barotropic and first baroclinic. Large-scale forcing maintains the available potential energy, and so provides energy to the baroclinic mode at very large scales. At these large scales, the equation for the baroclinic streamfunction is approximately that of a passive tracer, and so energy is transferred to smaller scales. It is also transferred to barotropic energy, at horizontal scales comparable to and larger than the deformation radius (this is baroclinic instability) and thence to larger barotropic scales. The entire process of energy transfer may be thought of as a generalized inverse cascade in which the energy passes to smaller pseudo-wavenumber $k'^2 \equiv k^2 + k_d^2$. At scales smaller than the first deformation radius the layers are decoupled and enstrophy in each layer cascades to smaller scales. The two-layer model may become less accurate for such small scales, because of the influence of higher baroclinic modes not present in a two-layer model.

$$\frac{d}{dt} (k'^2 \mathcal{E}(k) + p'^2 \mathcal{E}(p) + q'^2 \mathcal{E}(q)) = 0 \quad (2.42b)$$

where $\mathcal{E}(k)$ is the total energy (barotropic plus baroclinic) of the particular mode. These are formally identical with the conservation laws for purely two-dimensional flow and so we expect energy to seek the gravest (smallest pseudo-wavenumber) mode. Since the gravest mode has $k_d = 0$ this implies a *barotropization* of the flow.

- IV. *Baroclinic Instability.* Baroclinic instability in the classic two-layer problem concerns the instability of a flow with vertical but no horizontal shear. This is like a triad interaction for which $p \ll (k, q, k_d)$. The conservation laws are,

$$\begin{aligned} \frac{d}{dt} (\mathcal{T}(k) + C(p) + C(q)) &= 0, \\ \frac{d}{dt} (k^2 \mathcal{T}(k) + k_d^2 C(p) + (q^2 + k_d^2) C(q)) &= 0. \end{aligned} \quad (2.43)$$

From these, and with $k^2 \approx q^2$, we derive

$$k^2 \dot{C}(q) = (k_d^2 - k^2) \dot{\mathcal{T}}(k). \quad (2.44)$$

Baroclinic instability requires that both $\dot{C}(q)$ and $\dot{T}(k)$ be positive. This can only occur if

$$k^2 < k_d^2. \quad (2.45)$$

Thus, there is a *high-wavenumber cut-off* for baroclinic instability. This cut-off arises solely from considerations of energy and enstrophy conservation, and is not dependent on linearizing the equations and looking for exponentially growing normal mode instabilities.

For small scales, i.e., $k^2 \gg k_d^2$, the potential vorticity in each layer is, with $\beta = 0$,

$$q_1 = \nabla^2 \psi_1 + \frac{1}{2} k_d^2 (\psi_2 - \psi_1) \approx \nabla^2 \psi_1, \quad (2.46a)$$

$$q_2 = \nabla^2 \psi_2 + \frac{1}{2} k_d^2 (\psi_1 - \psi_2) \approx \nabla^2 \psi_2. \quad (2.46b)$$

Thus, each layer is decoupled from the other. Thus, enstrophy will cascade to smaller scales and, should there be an energy source at scales smaller than the deformation scale it will cascade to larger scales. However, baroclinic instability (of the mean flow) occurs at scales *larger* than the deformation radius. Thus, energy extracted from the mean flow is essentially trapped at scales larger than the deformation scale.

Summary of phenomenology

Putting together the considerations above leads to the following picture of geostrophic turbulence in a two-layer system (and see Fig. 2.6). At large horizontal scales we imagine some source of baroclinic energy, which in the atmosphere might be the differential heating between pole and equator, or in the ocean might be the wind and surface heat fluxes. Baroclinic instability effects a nonlocal transfer of energy to the deformation scale, where both baroclinic and barotropic modes are excited. From here there is an enstrophy cascade in each layer to smaller and smaller scales, until eventually the scale is small enough so that non-geostrophic effects become important and enstrophy is scattered by three-dimensional effects. At scales larger than the deformation radius, there is an inverse barotropic cascade of energy to larger scales. The energy of the large-scale barotropic modes is eventually dissipated by boundary layer effects such as Ekman drag. These ideas do not precisely apply to either atmosphere or ocean. In the latter, the turbulence is quite inhomogeneous except perhaps in the Antarctic Circumpolar Current. In the atmosphere, the deformation radius is almost as large as the Rhines scale, leaving little room for an inverse cascade. However, the atmosphere does display k^{-3} spectra at scales similar to and somewhat smaller than the deformation radius, as in Fig. 2.7, and analysis of this indicates that it may indeed be associated with a forward cascade of enstrophy.

2.3 * PHENOMENOLOGY OF BAROCLINIC EDDIES IN THE ATMOSPHERE AND OCEAN

In the remaining sections of this chapter we take a phenomenological approach, illustrated by numerical experiments and observations, to the problem of baroclinic eddies in

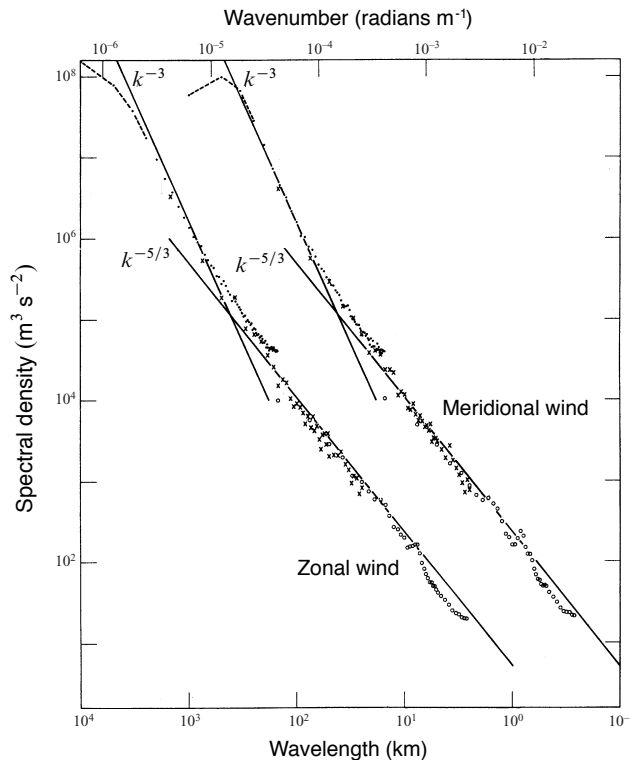


Figure 2.7 Energy spectra of the zonal and meridional wind near the tropopause, from thousands of commercial aircraft measurements between 1975 and 1979. The meridional spectrum is shifted one decade to the right. (From Gage and Nastrom 1986)

the atmosphere and ocean. We draw from our treatment of geostrophic turbulence but by being a little less precise we are able to travel farther, for we spend less time looking at the map (but with a concomitant danger that we lose our way).

2.3.1 The Magnitude and Scale of Baroclinic Eddies

How big, in both amplitude and scale, do baroclinic eddies become? Suppose that the time-mean flow is given, and that it is baroclinically unstable. Eddies will grow, initially according to linear baroclinic instability theory, but they cannot and do not continue to amplify: they ultimately equilibrate, and this by way of nonlinear mechanisms. The eddies will extract energy from the mean flow, but at the same time the available energy of the mean flow is being replenished by external forcing (i.e., the maintenance of an equator-pole temperature gradient by radiative forcing in the atmosphere, and wind and buoyancy forcing at the surface in the ocean). Thus, we cannot a priori determine the amplitude of baroclinic eddies by simply assuming that all of the available potential energy in the mean flow is converted to eddying motion. To close the problem we find we need to make three, not necessarily independent, assumptions:

- (i) An assumption about the magnitude of the baroclinic eddies;
- (ii) An assumption relating eddy kinetic energy to eddy available potential energy;
- (iii) An assumption about the horizontal scale of the eddies.

Baroclinic eddies extract available potential energy (APE) from the mean flow, and it is reasonable to suppose that an eddy of horizontal scale L_e can extract, as an upper bound, the APE of the mean flow contained within that scale. The APE is proportional to the variation of the buoyancy field so that

$$(\Delta b')^2 \sim |\Delta \bar{b}|^2 \sim L_e^2 |\nabla \bar{b}|^2 \quad (2.47)$$

where $\Delta \bar{b}$ is the variation in the buoyancy over the horizontal scale L_e . (For simplicity we stay with the Boussinesq equations, and $b = -g\delta\rho/\rho_0$. However, we might easily apply this to an ideal-gas atmosphere with $b = g\delta\theta/\theta_0$.) Equivalently, we might simply write

$$\boxed{b' \sim L_e |\nabla \bar{b}|}, \quad (2.48)$$

which arises from a mixing-length approach. Supposing that the temperature gradient is mainly in the y -direction then, using thermal wind, we have

$$b' \sim L_e f \frac{\partial \bar{u}}{\partial z} \quad \text{and} \quad v'_\tau \sim \bar{u}, \quad (2.49a,b)$$

where v'_τ is an estimate of the shear (multiplied by the depth scale) of the eddying flow. [These estimates are the same as (??), with \bar{u} replacing U .]

Our second assumption is to relate the barotropic eddy kinetic energy to the eddy available potential energy, and the most straightforward one to make is that there is a rough equipartition between the two. This assumption is reasonable because in the baroclinic lifecycle (or baroclinic inverse cascade) energy is continuously transferred from eddy available potential energy to eddy kinetic energy, and the assumption is then equivalent to supposing that the relevant eddy magnitude is always proportional to this rate of transfer. Thus we assume $v'_\psi^2 \sim (b'/N)^2$ or

$$\boxed{v'_\psi \sim \frac{b'}{N}}. \quad (2.50)$$

Finally, the scale of the eddies is determined by the extent to which the eddies might grow through nonlinear interactions. As we discussed earlier, possibilities for this scale include the deformation radius itself (if the inverse cascade is weak) or the Rhines scale (if the inverse cascade is slowed by the beta effect), or even the domain scale if neither of these apply.

Some consequences

These simple manipulations have some very interesting consequences. Using (2.49) and (2.50) we find

$$v'_\psi \sim \frac{fL_e}{NH} \bar{u} \approx \frac{L_e}{L_d} \bar{u} \quad (2.51)$$

where $L_d = NH/f_0$ is the deformation radius and \bar{u} is the amplitude of the mean baroclinic velocity, that is the mean shear multiplied by the height scale. This important

relationship relates the magnitude of the eddy kinetic energy to that of the mean. In the atmosphere the scale of the motion not much larger than the deformation radius (which is about 1000 km) the eddy and mean kinetic energies are, consistently, comparable to each other. In the ocean the deformation radius (about 50 km over large areas) is significantly smaller than the scale of mesoscale eddies (which typically might be more like 200 km), and observations consistently reveal that the eddy kinetic energy is an order of magnitude larger than the mean kinetic energy.

One other important and somewhat counter-intuitive result concerns the timescale of eddies. From (2.51) we have

$$T_E \sim \frac{L_e}{v'_\psi} \sim \frac{L_d}{u}, \quad (2.52)$$

and this is simply the Eady timescale. That is, the eddy timescale (at the scale of the largest eddies) is independent of the process that ultimately determines the spatial scale of those eddies; if the eddy length scale increases somehow, perhaps because friction or β are decreased, the velocity scale increases in proportion.

Let us now consider various aspects of baroclinic eddies in the atmosphere and the ocean.

2.3.2 Baroclinic Eddies in the Atmosphere

Amplitude and Scale

We saw in section ?? that baroclinic instability in the atmosphere occurs predominantly in the troposphere, i.e., in the lowest 10 km or so of the atmosphere, with the higher stratification of the eponymous stratosphere inhibiting instability. In the mid-latitude troposphere the vertical shear and the stratification are also relatively uniform which is why fairly simple models, such as the two-layer model or the Eady model (with the addition of the β -effect) are reasonable first-order models.

The mean pole-equator temperature gradient is about 40 K and the deformation radius NH/f is about 1000 km. The Rhines scale, $\sqrt{u_{\text{rms}}/\beta}$ is a little larger than this, perhaps 2000 km, and is similar to the width of the main mid-latitude baroclinic zone which lies between about 40° and 65°, in either hemisphere. Given these, and especially given that the maximum wavelength for instability occurs at scales somewhat larger than the deformation radius, there is little prospect of an extended upscale cascade, and for this reason the earth's atmosphere has comparable eddy kinetic and mean kinetic energies.

The baroclinic lifecycle

The baroclinic lifecycle of geostrophic turbulence, sketched schematically in Fig. 2.6, can be nicely illustrated by way of numerical initial value problems, and we describe two such. The first is very idealized: take a doubly-period quasi-geostrophic model on the f -plane, initialize it with baroclinic energy at large horizontal scales, and then let the flow freely evolve. Fig. 2.8 shows the results. The flow, initially concentrated in high vertical wavenumbers to best illustrate the energy transfer, is baroclinically unstable, and energy is transferred to barotropic flow at wavenumbers close to the first radius of

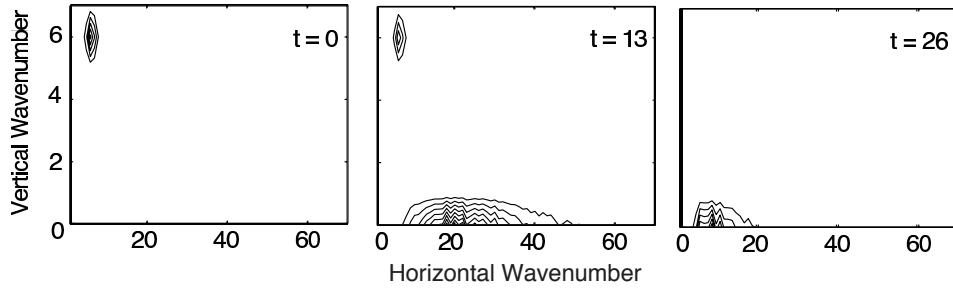


Fig. 2.8 A numerical simulation of a very idealized baroclinic lifecycle, showing contours of energy in spectral space at successive times. Initially, there is baroclinic energy at low horizontal wavenumber, as in a large-scale shear. Baroclinic instability transfers this energy to barotropic flow at the scale of the deformation radius, and this is followed by a barotropic inverse cascade to large scales. Most of the transfer to the barotropic mode in fact occurs quite quickly, between times 11 and 14, but the ensuing barotropic inverse cascade is slower. The entire process may be thought of as a generalized inverse cascade. The stratification (N^2) is uniform, and the first deformation radius is at about wavenumber 15. There is no friction in the simulation, except for a small hyperviscosity to remove small scale noise. Times are in units of eddy turnover time.

deformation, here at about wavenumber 15. Energy then slowly cascades back to large scales in a predominantly barotropic inverse cascade, piling up at the largest scales much as in decaying, two-dimensional turbulence. Nearly all of the initial baroclinic energy is converted to barotropic, eddy kinetic energy and, even without any surface friction, the flow evolves to a baroclinically stable state. Couched in these terms, it is easy to see the baroclinic lifecycle as a form of baroclinic inverse cascade, with an energy transfer to large total wavenumber, K_{tot} , that is made up of contributions from both horizontal and vertical wavenumbers:

$$K_{\text{tot}}^2 = K_h^2 + k_d^2 m^2 \quad (2.53)$$

where m is the vertical and K_h the horizontal wavenumber. As we noted earlier, the twin constraints of energy and enstrophy conservation prevent the excitation of horizontal scales with very large horizontal wavenumbers, and so the lifecycle proceeds through wavenumbers at the deformation scale.

The results of second, and more realistic, initial value problem are illustrated in Fig. 2.9. Here, the atmospheric primitive equations on a sphere are integrated forward, beginning from a baroclinically unstable zonal flow, plus a small-amplitude disturbance at zonal wavenumber six. The disturbance grows rapidly through baroclinic instability, accompanied by a conversion of energy initial from the zonal mean potential energy to eddy available potential energy (EAPE), and then from EAPE to eddy kinetic energy (EKE), and finally from EKE to zonal kinetic energy (ZKE). The last stage of this corresponds to the barotropic inverse cascade of quasi-geostrophic theory, and because of the presence of a β -effect the flow becomes organized into a zonal jet. The parameters in the earth's atmosphere are such that there is only one such jet, and in the lower panel of

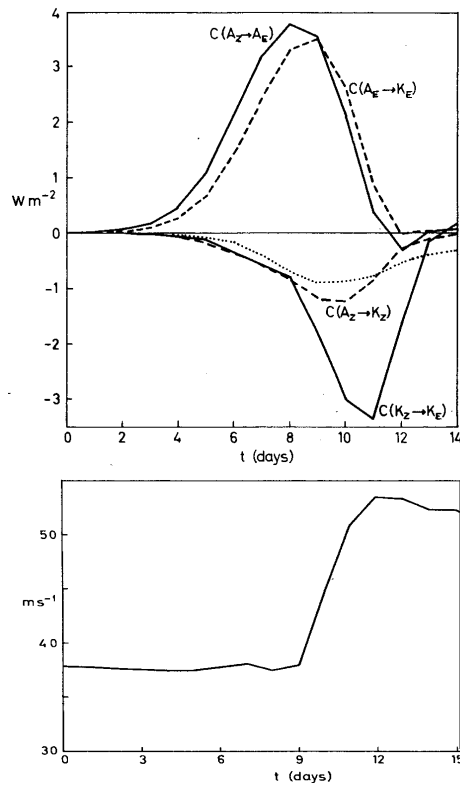


Figure 2.9 Top: Energy conversion and dissipation processes in a numerical simulation of an idealized atmospheric baroclinic lifecycle, simulated with a GCM Bottom: Evolution of the maximum zonal-mean velocity. A_Z and A_E are zonal and eddy available potential energies, and K_Z and K_E the corresponding kinetic energies. Initially baroclinic processes dominate, with conversions from zonal to eddy kinetic energy and then eddy kinetic to eddy available potential energy, followed by the barotropic conversion of eddy kinetic to zonal kinetic energy. The latter process is reflected in the *increase* of the maximum zonal-mean velocity at about day 10.

Fig. 2.9 we see its amplitude increase quickly from days 10 though 12, associated with the conversion of EKE to ZKE.

Of course, the atmosphere is never in a zonally uniform state as used in our baroclinic instability studies or the lifecycle study. At any given time, finite amplitude eddies exist and these provide a finite amplitude perturbation to the baroclinically unstable zonal flow, and thus we will rarely, if ever, see an exponentially growing normal mode. Furthermore, given any instantaneous atmospheric state, zonally symmetric or otherwise, the fastest growing (linear) instability is not necessarily exponential but may be ‘non-modal’, with a secular or linear growth that, over some finite time period and in some given norm, is much more rapid than exponential. A baroclinically turbulent atmosphere is of course maintained because of the underlying presence of baroclinic instability, and the classic baroclinic instability problems and nonlinear lifecycles illustrate, in an idealized way, the continuous growth, maturation and decay of eddies embedded in that flow.

2.3.3 Baroclinic Eddies in the Ocean

Basic ideas

Baroclinic instability was first developed as a theory for midlatitude synoptic-scale instabilities in the atmosphere and the original, now classic, problems are accordingly set in a zonally re-entrant channel. The ocean, apart from the Antarctic Circumpolar Current (ACC), is not zonally re-entrant. However, the ocean is driven by buoyancy and wind-forcing at the surface, and these combine to produce a region of enhanced stratification and associated shear in the ocean in the upper 500–1000 m or so, in the ‘thermocline’, as discussed more fully in chapter ?? (e.g., Fig. ??). The sloping isopycnals indicate that there is a pool of available potential energy that might be converted to kinetic energy, and so that the ocean is potentially baroclinically unstable. Satellite observations indicated that baroclinic eddies are almost ubiquitous in the mid- and high-latitude oceans, two particularly eddy-rich regions being the areas in and surrounding intense western boundary currents, such as the Gulf Stream, and in the ACC.

In addition to the geometry, the main differences between the oceanic and atmospheric problems are:

- (i) In the ocean, the shear and the stratification are not uniform between two rigid lids, nor even uniform between one rigid lid and a structure like the tropopause. Instead, both stratification and shear are largest in the upper ocean, decaying into a quiescent and nearly unstratified abyss.
- (ii) The first radius of deformation is much smaller than the scale of the large-scale flow — that is, of the gyres or the large-scale overturning circulation.

A consequence of the first item is that the amplitude of the growing waves is also largely concentrated in the upper ocean, as we saw in Fig. ?. Regarding item (ii), we can estimate the oceanic deformation radius as

$$L_d = \frac{NH}{f} \approx \frac{10^{-2} \times 500}{10^{-4}} \approx 50 \text{ km.} \quad (2.54)$$

More precisely, in quasi-geostrophic theory we may define the deformation radii by solution of the eigenvalue problem

$$\frac{\partial}{\partial z} \frac{f_0^2}{N^2} \frac{\partial \phi_n}{\partial z} + E_n \phi_n = 0. \quad (2.55)$$

The successive eigenvalues, E_n , are related to the successive deformation radii by $L_{dn}^2 = 1/E_n$, and the results of a similar calculation are given in Fig. 2.10. Note that in uniform stratification the deformation radius as defined by (2.55) and displayed in Fig. 2.10 is a factor of π smaller than the simple estimate NH/f , so that the most baroclinically unstable waves have a wavelength several times L_{d1} . Nevertheless, we may expect baroclinic instability to occur on a scale much smaller than that in the atmosphere, and much smaller than the scale of the domain.

Amplitude and Scale

The consequences of this small deformation radius on the lifecycle and finite-amplitude equilibration of oceanic baroclinic eddies are potentially far-reaching, the most important of which is that there is more scope for an inverse cascade than in the atmosphere,

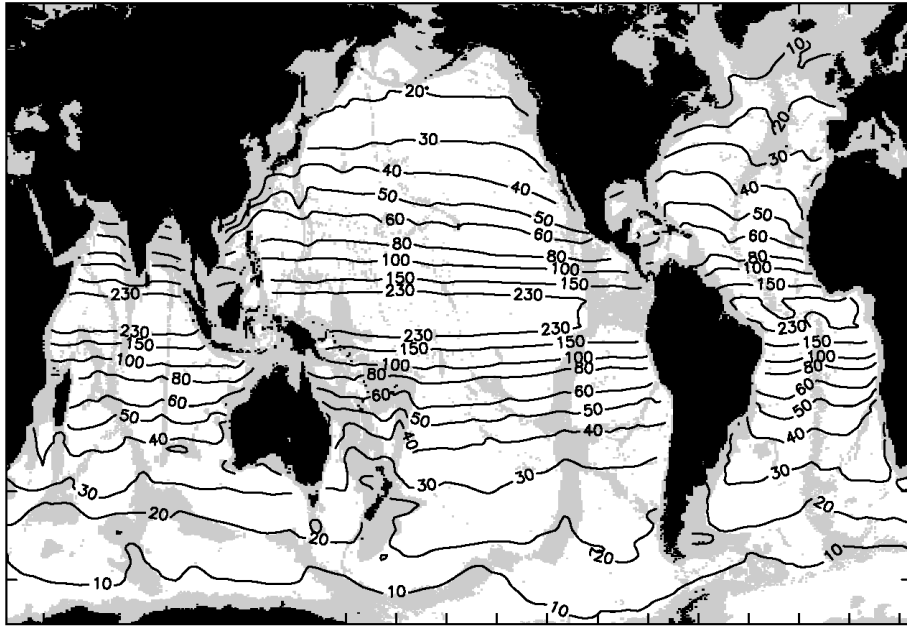


Fig. 2.10 The oceanic first deformation radius L_d , calculated by using the observed stratification from the eigenproblem:

$\partial^2 \phi / \partial z^2 + (N^2(z)/c^2) \phi = 0$ with $\phi = 0$ at $z = 0$ and $z = -H$, where H is the ocean depth and N is the observed buoyancy frequency. The deformation radius is given by $L_d = c/f$ where c is the first eigenvalue and f is the latitudinally varying Coriolis parameter. Near equatorial regions are excluded, and regions of ocean shallower than 3500 m are shaded. Variations in Coriolis parameter are responsible for much of large-scale variability, although weak stratification also reduces the deformation radius at high latitudes.

and indeed observations indicate that the horizontal scale of the eddies is typically a few to several times larger than the local deformation radius itself. The situation is not clear cut, however: the eddy scale does seem to be somewhat larger than the deformation scale, but some observations indicate that the eddy size nevertheless scales with the local deformation radius, suggesting that the eddy scale may be set by the instability scale and not an inverse cascade. In any case, suppose then that an ocean eddy is of horizontal scale 200 km, and that it sits in the subtropical gyre where the mean temperature gradient is 10^{-5} K m^{-1} , that the mean shear and ensuing baroclinic activity is mainly confined to the upper 1000 m of the ocean, and that the deformation radius is 50 km. The temperature gradient corresponds to a temperature difference of about 20 K across 2000 km, a horizontal buoyancy gradient of about $2 \times 10^{-9} \text{ s}^{-2}$ (using the simple equation of state $\rho = \rho(1 - \beta_T \Delta T)$ where $\beta_T = 2 \times 10^{-4} \text{ K}^{-1}$) and a shear of about 2 cm s^{-1} over the upper 1 km of ocean. Then, using (2.51), we can estimate a typical eddy velocity scale as

$$v'_\psi \sim \frac{L_e}{L_d} \bar{u} \approx 4\bar{u} \approx 8 \text{ cm s}^{-1}, \quad (2.56)$$

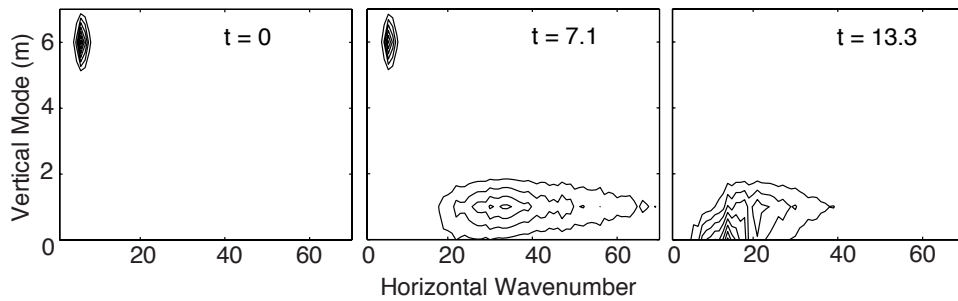


Fig. 2.11 Idealized baroclinic lifecycle, similar to that in Fig. 2.8, but with enhanced stratification of the basic state in the upper domain, representing the oceanic thermocline.

implying, as we noted earlier, an EKE that is an order of magnitude larger than the mean kinetic energy. Associated with this are typical temperature perturbations whose magnitude we can estimate using (2.48) or (2.49) as being about 2 K. These estimates are comparable to those observed in mid ocean, with more energetic eddies forming near intense western boundary currents where gradients are large and barotropic instability also provides a source of energy for the eddies. There is at least a factor of a few uncertainty, but it is noteworthy that they are roughly comparable to the values observed.

Lifecycles

The lifecycle of a mid-oceanic baroclinic eddy will differ from its atmospheric counterpart in two main respects:

- (i) Baroclinic eddies may be advected by the mean flow into regions with quite different properties from where they initially formed.
- (ii) The nonuniformity of the stratification affects the passage to barotropic flow.

Both of these can properly only be studied by numerical means. Regarding the first, eddies will often form in or near intense western boundary currents, but then will be advected by that current into the potentially less unstable open ocean before completing their lifecycle. Regarding the second, an oceanic analog of the lifecycle illustrated in Fig. 2.8 is shown in Fig. 2.11.

Eddies and the General Circulation

In this chapter we look at how eddies affect the general circulation of the atmosphere. The first two sections — on the Eliassen-Palm flux and the Transformed Eulerian Mean — are a little technical. In the lectures we may start with section 3.3 and refer back to these sections as needed. (On the other hand...)

3.1 * THE ELIASSEN-PALM FLUX

The eddy flux of potential may be expressed in terms of vorticity and buoyancy fluxes as

$$v'q' = v'\zeta' + f_0v' \frac{\partial}{\partial z} \left(\frac{b'}{N^2} \right) \quad (3.1)$$

The second term on the right-hand side can be written

$$\begin{aligned} f_0v' \frac{\partial}{\partial z} \left(\frac{b'}{N^2} \right) &= f_0 \frac{\partial}{\partial z} \left(\frac{v'b'}{N^2} \right) - f_0 \frac{\partial v'}{\partial z} \frac{b'}{N^2} \\ &= f_0 \frac{\partial}{\partial z} \left(\frac{v'b'}{N^2} \right) - f_0 \frac{\partial}{\partial x} \left(\frac{\partial \psi'}{\partial z} \right) \frac{b'}{N^2} \\ &= f_0 \frac{\partial}{\partial z} \left(\frac{v'b'}{N^2} \right) - \frac{f_0^2}{N^2} \frac{\partial}{\partial x} \left(\frac{1}{2} \frac{\partial \psi'}{\partial z} \right)^2 \end{aligned} \quad (3.2)$$

using $b' = f_0 \partial \psi' / \partial z$.

Similarly, the flux of relative vorticity can be written

$$v'\zeta' = -\frac{\partial}{\partial y} u'v' + \frac{1}{2} \frac{\partial}{\partial x} (v'^2 - u'^2) \quad (3.3)$$

Using (3.2) and (3.3), (3.1) becomes

$$\boxed{v'q' = -\frac{\partial}{\partial y}(u'v') + \frac{\partial}{\partial z}\left(\frac{f_0}{N^2}v'b'\right) + \frac{\partial}{\partial x}\left(\frac{1}{2}(v'^2 - u'^2) - \frac{b'^2}{N^2}\right)} \quad (3.4)$$

Thus the potential vorticity flux, in the quasi-geostrophic approximation, can be written as the divergence of a vector: $v'q' = \nabla \cdot \mathbf{E}$ where

$$\mathbf{E} \equiv \left(\frac{1}{2}(v'^2 - u'^2) - \frac{b'^2}{N^2}\right) \mathbf{i} - (u'v') \mathbf{j} + \left(\frac{f_0}{N^2}v'b'\right) \mathbf{k}. \quad (3.5)$$

A particularly useful form of this arises after zonally averaging, after which (3.4) becomes

$$\boxed{\overline{v'q'} = -\frac{\partial}{\partial y}\overline{u'v'} + \frac{\partial}{\partial z}\left(\frac{f_0}{N^2}\overline{v'b'}\right)}. \quad (3.6)$$

The vector defined by

$$\mathbf{F} \equiv -\overline{u'v'} \mathbf{j} + \frac{f_0}{N^2}\overline{v'b'} \mathbf{k} \quad (3.7)$$

is called the *Eliassen-Palm flux*,¹ and its divergence, given by (3.30), gives the polewards flux of potential vorticity:

$$\overline{v'q'} = \nabla_x \cdot \mathbf{F}, \quad (3.8)$$

where $\nabla_x \cdot \equiv (\partial/\partial y, \partial/\partial z)$ is the divergence in the meridional plane. Unless the meaning is unclear, the subscript x on the meridional divergence will be dropped.

3.1.1 The Eliassen-Palm relation

On dividing by $\partial\bar{q}/\partial y$ and using (3.8), the enstrophy equation (??) becomes

$$\boxed{\frac{\partial \mathcal{A}}{\partial t} + \nabla \cdot \mathbf{F} = \mathcal{D}}, \quad (3.9a)$$

where

$$\mathcal{A} = \frac{\overline{q'^2}}{2\partial\bar{q}/\partial y}, \quad \mathcal{D} = \frac{\overline{D'q'}}{\partial\bar{q}/\partial y} \quad (3.9b)$$

Eq. (3.9a) is known as the *Eliassen-Palm relation*, and it is a conservation law for the the *wave activity density* \mathcal{A} , for if we integrate this expression over a meridional area A bounded by walls where the eddy activity vanishes, and if $\mathcal{D} = 0$, we obtain

$$\frac{d}{dt} \int_A \mathcal{A} dA = 0. \quad (3.10)$$

In general, a wave activity is a quantity that is quadratic in the amplitude of the perturbation and that is conserved in the absence of forcing and dissipation. More specifically, \mathcal{A} is the negative of the *pseudomomentum*, for reasons we will encounter later. Note

that neither perturbation energy nor perturbation enstrophy are wave activities of the linearized equations, because there can be an exchange of energy or enstrophy between mean and perturbation — indeed, this is how a perturbation grows in baroclinic or barotropic instability! This is already evident from (??), or in general take (??) with $D' = 0$ and multiply by q' to give the enstrophy equation

$$\frac{1}{2} \frac{\partial q'^2}{\partial t} + \frac{1}{2} \bar{\mathbf{u}} \cdot \nabla q'^2 + \mathbf{u}' q' \cdot \nabla \bar{q} = 0 \quad (3.11)$$

where here the overbar is an average (although it need not be a zonal average). Integrating this over a volume V gives

$$\frac{d\hat{Z}'}{dt} \equiv \frac{d}{dt} \int_V \frac{1}{2} q'^2 dV = - \int_V \mathbf{u}' q' \cdot \nabla \bar{q} dV. \quad (3.12)$$

The right-hand side does not in general vanish and so \hat{Z}' is not in general conserved. The ave activity \mathcal{A} is thus both a measure of the amplitude of a wave and a conserved quantity, in the sense of (3.9a).

3.1.2 The group velocity property

The vector \mathcal{F} describes how the wave activity propagates. In general, we cannot express it simply in terms of \mathcal{A} , but in the case in which the disturbance is composed of plane or almost plane waves that satisfy a dispersion relation, then $\mathcal{F} = \mathbf{c}_g \mathcal{A}$, where \mathbf{c}_g is the group velocity and (3.9a) becomes

$$\frac{\partial \mathcal{A}}{\partial t} + \nabla \cdot (\mathcal{A} \mathbf{c}_g) = 0. \quad (3.13)$$

We shall demonstrate this when the waves in question are plane Rossby waves.

Group velocity property for Rossby waves

The Boussinesq quasi-geostrophic equation on the β -plane, linearized around a uniform zonal flow and with constant static stability, is

$$\frac{\partial q'}{\partial t} + \bar{u} \frac{\partial q'}{\partial x} + v' \frac{\partial \bar{q}}{\partial y} = 0 \quad (3.14)$$

where $q' = [\nabla^2 + (f_0^2/N^2)\partial^2/\partial z^2]\psi'$ and, if \bar{u} is constant, $\partial \bar{q}/\partial y = \beta$. Thus we have

$$\left(\frac{\partial}{\partial t} + \bar{u} \frac{\partial}{\partial x} \right) \left[\nabla^2 \psi' + \frac{\partial}{\partial z} \left(\frac{f_0^2}{N^2} \frac{\partial \psi'}{\partial z} \right) \right] + \beta \frac{\partial \psi'}{\partial x} = 0. \quad (3.15)$$

Seeking solutions of the form

$$\psi' = \text{Re } \tilde{\psi} e^{i(kx+ly+mz-\omega t)}, \quad (3.16)$$

we find the dispersion relation,

$$\omega = \bar{u}k - \frac{\beta k}{\kappa^2}. \quad (3.17)$$

with group velocities,

$$c_g^y = \frac{2\beta kl}{\kappa^2}, \quad c_g^z = \frac{2\beta km f_0^2/N^2}{\kappa^2}, \quad (3.18)$$

where $\kappa^2 = (k^2 + l^2 + m^2 f_0^2/N^2)$. Also, if $u' = \text{Re } \tilde{u} \exp[i(kx + ly + mz - \omega t)]$, and similarly for the other fields, then

$$\begin{aligned} \tilde{u} &= -\text{Re } ik\tilde{\psi}, & \tilde{v} &= \text{Re } il\tilde{\psi}, \\ \tilde{b} &= \text{Re } imf_0\tilde{\psi}, & \tilde{q} &= -\text{Re } \kappa^2\tilde{\psi}. \end{aligned} \quad (3.19)$$

The wave activity is then

$$\mathcal{A} = \frac{1}{2} \frac{\overline{q'^2}}{\beta} = \frac{\kappa^4}{4\beta} |\tilde{\psi}^2| \quad (3.20)$$

where the additional factor of 2 in the denominator arises from the averaging. Using (3.19) the EP flux, (3.7), is

$$\begin{aligned} \mathcal{F}^y &= -\overline{u'v'} = \frac{1}{2} kl |\tilde{\psi}^2| \\ \mathcal{F}^z &= \frac{f_0}{N^2} \overline{v'b'} = \frac{f_0^2}{2N^2} km |\tilde{\psi}^2|. \end{aligned} \quad (3.21)$$

Using this (3.18) and (3.20) gives

$$\boxed{\mathcal{F} = (\mathcal{F}^y, \mathcal{F}^z) = \mathbf{c}_g \mathcal{A}}. \quad (3.22)$$

If the properties of the medium are varying, but only on scales larger than the scale of the waves and we can still define a group velocity, then this is a useful expression to estimate how the wave activity propagates in the atmosphere and in numerical simulations.

3.2 * THE TRANSFORMED EULERIAN MEAN

3.2.1 Quasi-geostrophic form

For simplicity we will use the Boussinesq equations on the beta-plane, and the zonally-averaged Eulerian mean equations for the zonally-averaged zonal velocity and the buoyancy may then be written, approximately

$$\frac{\partial \bar{u}}{\partial t} = f_0 \bar{v} - \frac{\partial}{\partial y} \overline{u'v'} + \bar{F}, \quad (3.23a)$$

$$\frac{\partial \bar{b}}{\partial t} = -N^2 \bar{w} - \frac{\partial}{\partial y} \overline{v'b'} + \bar{J}, \quad (3.23b)$$

where \bar{b} is in thermal wind balance with \bar{u} ($f_0 \partial \bar{u} / \partial z = -\partial \bar{b} / \partial y$). One less-than-ideal aspect of these equations is that in the extratropics the dominant balance is usually between the first two terms on the right-hand sides of each equation, even in time-dependent cases. Thus, the Coriolis force closely balances the divergence of the eddy

momentum fluxes, and the advection of the mean stratification ($N^2 w$, or ‘adiabatic cooling’) often balances the convergence of eddy heat flux, with heating being a small residual. This may lead to an underestimation of the importance of diabatic heating, for this is ultimately responsible for the mean meridional circulation. Thus, in the thermodynamic equation we might seek to combine the terms $N^2 w$ and the eddy flux into a single total (or ‘residual’) heat transport term that in a steady state is balanced by the diabatic term \bar{J} . The TEM provides this reformulation, and in doing so the eddy terms in the momentum equation also take a different form.

To begin, note that because \bar{v} and \bar{w} are related by mass conservation, we can define a mean meridional streamfunction ψ_m such that

$$(\bar{v}, \bar{w}) = \left(-\frac{\partial \psi_m}{\partial z}, \frac{\partial \psi_m}{\partial y} \right). \quad (3.24)$$

Given this, the velocities satisfy $\partial \bar{v} / \partial y + \partial \bar{w} / \partial z = 0$. Then, if we define a ‘residual’ streamfunction by

$$\psi^* \equiv \psi_m + \frac{1}{N^2} \overline{v' b'} \quad (3.25)$$

the components of the *residual mean meridional circulation* are given by

$$(\bar{v}^*, \bar{w}^*) = \left(-\frac{\partial \psi^*}{\partial z}, \frac{\partial \psi^*}{\partial y} \right), \quad (3.26)$$

and

$$\bar{v}^* = \bar{v} - \frac{\partial}{\partial z} \left(\frac{1}{N^2} \overline{v' b'} \right), \quad \bar{w}^* = \bar{w} + \frac{\partial}{\partial y} \left(\frac{1}{N^2} \overline{v' b'} \right). \quad (3.27)$$

Note that by construction, the residual overturning circulation satisfies

$$\frac{\partial \bar{v}^*}{\partial y} + \frac{\partial \bar{w}^*}{\partial z} = 0. \quad (3.28)$$

Substituting (3.27) into (3.23a) and (3.23b) the zonal momentum and buoyancy equations then take the simple forms

$$\left[\begin{array}{l} \frac{\partial \bar{u}}{\partial t} = f_0 \bar{v}^* + \overline{v' q'} + \bar{F} \\ \frac{\partial \bar{b}}{\partial t} = -N^2 \bar{w}^* + \bar{J} \end{array} \right], \quad (3.29)$$

where

$$\overline{v' q'} = \overline{v' \zeta'} + \frac{\partial}{\partial z} \left(\frac{f_0}{N^2} \overline{v' b'} \right) = -\frac{\partial}{\partial y} \overline{u' v'} + \frac{\partial}{\partial z} \left(\frac{f_0}{N^2} \overline{v' b'} \right). \quad (3.30)$$

These are known as the (quasi-geostrophic) *transformed Eulerian mean equations*.

If we know the potential vorticity flux as well as \bar{F} and \bar{J} , then (3.28) and (3.29), along with thermal wind balance

$$f_0 \frac{\partial \bar{u}}{\partial z} = -\frac{\partial \bar{b}}{\partial y} \quad (3.31)$$

Aspects of the TEM formulation

Properties and features

- ★ The residual mean circulation is equivalent to the total mass-weighted (eddy plus Eulerian mean) circulation, and it is this circulation that is driven by the diabatic forcing.
- ★ There are no explicit eddy fluxes in the buoyancy budget; the only eddy term is the flux of potential vorticity, and this is divergence of the Eliassen-Palm flux; that is $\overline{v'q'} = \nabla_x \cdot \mathbf{E}$.
- ★ The residual circulation, \overline{v}^* , becomes part of the solution, just as \overline{v} is part of the solution in an Eulerian mean formulation.

But note

- ★ The theory and practice are well developed for a zonal average, less so for three-dimensional, non-zonal flow. This is because the geometry enforces simple boundary conditions in the zonal mean case.
- ★ The boundary conditions on the residual circulation are neither necessarily simple nor easily determined; for example, at a horizontal boundary \overline{w}^* is not zero if there are horizontal buoyancy fluxes.

form a complete set. The meridional overturning circulation is obtained by eliminating time derivatives from (3.29) using (3.31) to give

$$f_0^2 \frac{\partial^2 \psi^*}{\partial z^2} + N^2 \frac{\partial^2 \psi^*}{\partial y^2} = f_0 \frac{\partial}{\partial z} \overline{v'q'} + f_0 \frac{\partial \overline{F}}{\partial z} + \frac{\partial \overline{J}}{\partial y}. \quad (3.32)$$

Thus, the residual or net overturning circulation is ‘driven’ by the (vertical derivative of the) potential vorticity fluxes and the diabatic terms — driven in the sense that if we know those terms we can calculate the overturning circulation. Note that this equation applies at every instant, even if the equations are not in a steady state.

3.2.2 The TEM in isentropic coordinates

The residual circulation has an illuminating interpretation if we think of the fluid as comprising multiple layers of shallow water, or equivalently if we cast the problem in isentropic coordinates. The momentum and mass conservation equation can then be written

as

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} - f \mathbf{v} = F \quad (3.33a)$$

$$\frac{\partial h}{\partial t} + \nabla \cdot (h \mathbf{u}) = J \quad (3.33b)$$

In isentropic coordinates h is the thickness between two isentropic surfaces and $H = H(b)$ is its mean thickness, and the layer thickness is a measure of the temperature of the layer. With quasi-geostrophic scaling (so that β and variations in layer thickness are small) zonally averaging in a conventional way gives

$$\frac{\partial \bar{u}}{\partial t} - f_0 \bar{v} = \overline{v' \zeta'} + \bar{F} \quad (3.34a)$$

$$\frac{\partial \bar{h}}{\partial t} + H \frac{\partial \bar{v}}{\partial y} = - \frac{\partial}{\partial y} \overline{v' h'} + \bar{J}[h] \quad (3.34b)$$

The overbars in these equations denote averages taken along isentropes, but are otherwise conventional, and the meridional velocity is purely ageostrophic. We now choose to define the residual circulation by

$$\bar{v}^* = \bar{v} + \frac{1}{H} \overline{v' h'} \quad (3.35)$$

which is analogous to (3.27). Using (3.35) in (3.34) gives

$$\frac{\partial \bar{u}}{\partial t} - f_0 \bar{v}^* = \overline{v' q'} + \bar{F} \quad (3.36a)$$

$$\frac{\partial \bar{h}}{\partial t} + H \frac{\partial \bar{v}^*}{\partial y} = \bar{J}[h]. \quad (3.36b)$$

where

$$\overline{v' q'} = \overline{v' \zeta'} + \frac{f_0}{H} \overline{v' h'}. \quad (3.37)$$

From (3.35) we see that the residual velocity is a measure of the *total meridional mass flux*, eddy plus mean, in an isentropic layer. This is often a more useful quantity than the Eulerian velocity \bar{v} because it is generally the former, not the latter, that is constrained by the external forcing. What we have done, of course, is to effectively use a mass-weighted mean in (3.33b); that is, if we define the mass-weighted mean by

$$\bar{v}_* \equiv \frac{\overline{h v}}{\bar{h}} \quad (3.38)$$

so that

$$\bar{v}_* = \bar{v} + \frac{1}{\bar{h}} \overline{v' h'}, \quad (3.39)$$

then the zonal average of (3.33b) is just

$$\frac{\partial \bar{h}}{\partial t} + \frac{\partial}{\partial y} (\bar{h} \bar{v}_*) = \bar{J}[h], \quad (3.40)$$

which is the same as (3.36b) if $\bar{h} = H$. Similarly, if we use the mass weighted velocity (3.39) in the momentum equation (3.34a) we obtain (3.36a).

Evidently, if the mass-weighted meridional velocity is used in the momentum and thickness equations then the eddy mass flux does not enter the equations explicitly — the only eddy flux in (3.36) is that of potential vorticity. That is, in isentropic coordinates the equations in TEM form are equivalent to the equations that arise from a particular form of averaging — mass weighted averaging — rather than the conventional Eulerian averaging. Does a similar relationship hold in height coordinates? The answer is yes, as we now see.

3.3 THE MAINTENANCE OF JETS IN THE ATMOSPHERE

3.3.1 Observations and Motivation

The atmosphere above the surface has a generally eastward flow, with a broad maximum about 10 km above the surface at around 40° in either hemisphere. But if we look a little more at the zonally average wind in we see hints there being two jets — one (the subtropical jet) at around 30° , and another somewhat polewards of this, especially apparent in the Southern Hemisphere. Such a jet is particularly noticeable in certain regions of the globe, when a zonal average is not taken, as in Fig. 3.1. The subtropical jet is associated with a strong meridional temperature gradient at the edge of the Hadley Cell, and is quite baroclinic. On the other hand the midlatitude jet (sometimes called the subpolar jet) is fairly barotropic (it has little vertical structure, with less shear than the subtropical jet) and lies above an eastward surface flow. This flow feels the effect of friction and so there must be a momentum *convergence* into this region, and we see this clearly in Fig. ???. We will find that this momentum convergence occurs largely in transient eddies, and the jet is known as the *eddy-driven jet*. Although the eddies themselves are a product of baroclinic instability, the essential mechanism of jet production is present in barotropic dynamics, so let us first consider how an eastward jet can be maintained in a turbulent flow on the surface of a rotating sphere.

In barotropic turbulence, alternating east-west jets can be maintained if β is non-zero, as described in chapter 2. However, that case was homogeneous, with no preferred latitude for a particular direction of jet, whereas in the atmosphere there appears to be but one mid-latitude jet, and although it meanders it certainly has a preferred average location. In the subsections that follow we give four explanations as to how this is achieved. The first of these has a different flavour than the others, but they are all just different perspectives on the same mechanism.²

3.3.2 The mechanism of jet production

1. The vorticity budget

Suppose that the absolute vorticity normal to the surface (i.e., $\zeta + 2\Omega \sin \vartheta$) increases monotonically poleward. (A sufficient condition for this is that the fluid is at rest.) By Stokes' theorem, the circulation around a line of latitude circumscribing the polar cap is

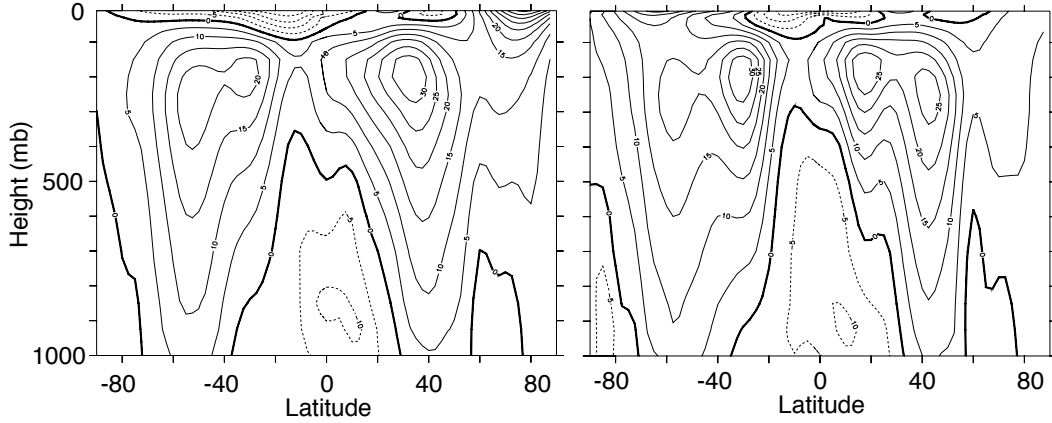


Fig. 3.1 The time-averaged zonal wind at 150°W (in the mid Pacific) in December-January-February (DJF, left), March-April-May (MAM, right). The contour interval is 5 m s^{-1} . Note the double jet in each hemisphere, one in the subtropics and one in midlatitudes. The subtropical jets is associated with strong meridional temperature gradient, whereas the midlatitude jets have a stronger barotropic component and are associated with westerly winds at the surface.

equal to the integral of the absolute vorticity over the cap. That is,

$$I_1 = \int_{\text{cap}} \boldsymbol{\omega}_{ia} \cdot d\mathbf{A} = \oint_C u_{ia} dl = \oint_C (u_i + 2\Omega a \cos \vartheta) dl, \quad (3.41)$$

where $\boldsymbol{\omega}_{ia}$ and u_{ia} are the initial absolute vorticity and velocity, u_i is the initial zonal velocity in the earth's frame of reference, and the line integrals are around the line of latitude. For simplicity let us take $u_i = 0$ and suppose there is a disturbance equatorwards of the polar cap, and that this results in a distortion of the material line around the latitude circle C (Fig. 3.2). Since we are supposing the source of the disturbance is distant from the latitude of interest, then if we neglect viscosity the circulation along the material line is conserved, by Kelvin's circulation theorem. Thus, vorticity with a lower value is brought into the region of the polar cap, and (using Stokes theorem again) the circulation around the latitude circle C must fall. That is, denoting values after the disturbance with a subscript f ,

$$I_f = \int_{\text{cap}} \boldsymbol{\omega}_{fa} \cdot d\mathbf{A} < I_i \quad (3.42)$$

so that

$$\oint_C (u_f + 2\Omega a \cos \vartheta) dl < \oint_C (u_i + 2\Omega a \cos \vartheta) dl \quad (3.43)$$

and

$$\bar{u}_f < \bar{u}_i \quad (3.44)$$

with the overbar indicating a zonal average. Thus, there is a tendency to produce *westward* flow polewards of the disturbance. By a similar argument westward flow is also

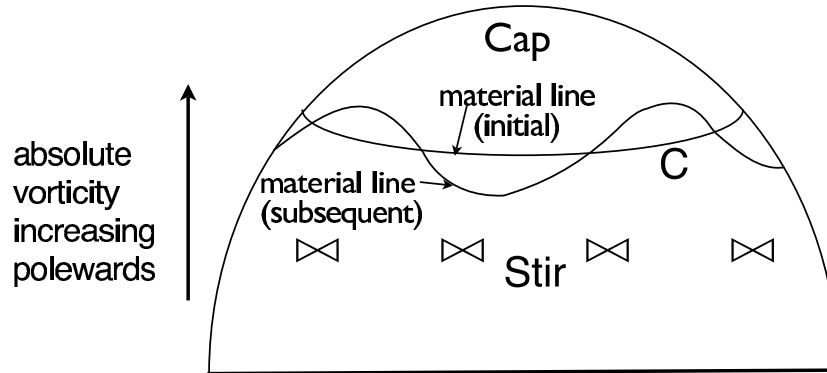


Fig. 3.2 The effects of midlatitude disturbance on the circulation around the latitude line C. If initially the absolute vorticity increases monotonically polewards, then the disturbance will bring fluid with lower absolute vorticity into the cap region. Then, using Stokes theorem, the velocity around the latitude line C will become more westward.

produced equatorward of the disturbance — to see this apply Kelvin's theorem over all of the globe south of the source of the disturbance (taking care to take the dot-product correctly between the direction of the vorticity vector and the direction of normal to the surface). Finally, note that the overall situation is the same in the Southern Hemisphere. Thus, on the surface of a rotating sphere, external stirring will produce westward flow away from the region of the stirring.

Now suppose furthermore that the disturbance imparts no net angular momentum to the fluid. Then the integral of $ua \cos \vartheta$ over the entire hemisphere must be constant. But the fluid is accelerating westward away from the disturbance. Therefore, the fluid in the region of the disturbance must accelerate *eastward*. That is, angular momentum must converge into the stirred region, producing an eastward flow. This simple mechanism is the essence of the production of eastward eddy-driven jets in the atmosphere, and of the eastward surface winds in mid-latitudes. The stirring that here we have externally imposed comes, of course, from baroclinic instability.

If the stirring subsides then the flow may reversibly go back to its initial condition, with a concomitant reversal of the momentum convergence that caused the zonal flow. Thus, we must have some form of dissipation and irreversibility in order to produce permanent changes, and in particular we need to irreversibly mix vorticity. (This result is closely related to the non-acceleration results of chapter ??.) If the fluid is continuously mixed, then of course we also need a source that restores the absolute vorticity gradient, else we will completely homogenize the vorticity over the hemisphere, so let us now set up a simple model that shows how a permanent jet structure can be maintained.

II. The pseudomomentum budget

The kinematic relation between vorticity flux and momentum flux for non-divergent two-dimensional flow is

$$v\zeta = \frac{1}{2} \frac{\partial}{\partial x} (v^2 - u^2) - \frac{\partial}{\partial y} (uv). \quad (3.45)$$

After zonal averaging this gives

$$\overline{v'\zeta'} = -\frac{\partial \overline{u'v'}}{\partial y}, \quad (3.46)$$

noting that $\bar{v} = 0$. For reference, in spherical coordinates this expression becomes

$$\overline{v'\zeta'} \cos \vartheta = -\frac{1}{a \cos \vartheta} \frac{\partial}{\partial \vartheta} (\cos^2 \vartheta \overline{u'v'}). \quad (3.47)$$

If (3.46) [or (3.47)] is integrated with respect to y between two quiescent latitudes then the right-hand side vanishes. That is the zonally-averaged meridional vorticity flux vanishes when integrated over latitude.

Now, the barotropic zonal momentum equation is (for horizontally non-divergent flow)

$$\frac{\partial u}{\partial t} + \frac{\partial u^2}{\partial x} + \frac{\partial uv}{\partial y} - fv = -\frac{\partial \phi}{\partial x} + F_u - D_u, \quad (3.48)$$

where F_u and D_u represent the effects of any forcing and dissipation. Zonal averaging, with $\bar{v} = 0$, gives

$$\frac{\partial \bar{u}}{\partial t} = -\frac{\partial \bar{u}\bar{v}}{\partial y} + \bar{F}_u - \bar{D}_u, \quad (3.49)$$

or, using (3.46),

$$\frac{\partial \bar{u}}{\partial t} = \overline{v'\zeta'} + \bar{F}_u - \bar{D}_u. \quad (3.50)$$

Thus, the zonally averaged wind is maintained by the zonally averaged vorticity flux. On average there is little if any direct forcing of horizontal momentum and we may set $\bar{F}_u = 0$. Also, most of the dissipation comes from the bottom Ekman layer and if this is parameterized by a linear drag (3.50) becomes

$$\frac{\partial \bar{u}}{\partial t} = \overline{v'\zeta'} - r\bar{u}, \quad (3.51)$$

where the constant r is an inverse frictional timescale.

Now consider the maintenance of this vorticity flux. The barotropic vorticity equation is

$$\frac{\partial \zeta}{\partial t} + \mathbf{u} \cdot \nabla \zeta + v\beta = F_\zeta - D_\zeta. \quad (3.52)$$

where F_ζ and D_ζ are forcing and dissipation of vorticity. Linearize about a mean zonal flow to give

$$\frac{\partial \zeta'}{\partial t} + \bar{u} \frac{\partial \zeta'}{\partial x} + yv' = F'_\zeta - D'_\zeta, \quad (3.53)$$

where

$$\gamma = \beta - \frac{\partial^2 \bar{u}}{\partial y^2} \quad (3.54)$$

is the meridional gradient of absolute vorticity. Multiply (3.53) by ζ'/γ and zonally average to form the pseudomomentum equation,

$$\frac{\partial P}{\partial t} + \overline{v'\zeta'} = \frac{1}{\gamma} (\overline{\zeta'F'_\zeta} - \overline{\zeta'D'_\zeta}), \quad (3.55)$$

where

$$P = \frac{1}{2\gamma} \overline{\zeta'^2} \quad (3.56)$$

is the negative of the pseudomomentum for this problem (see also (3.9b) on page 54, and section 3.1 more generally). The parameter γ is positive if the average absolute vorticity increases monotonically northwards, and this is usually the case in both Northern and Southern hemispheres.

In the absence of forcing and dissipation, (3.51) and (3.55) imply an important relationship between the change of the mean flow and the pseudomomentum, namely

$$\frac{\partial \bar{u}}{\partial t} + \frac{\partial P}{\partial t} = 0. \quad (3.57)$$

Now, P is a measure of the wave activity; if for some reason this increases, perhaps because a wave enters an initially quiescent region because of stirring elsewhere, then P increases and the mean flow must decrease. However, because the vorticity flux integrates to zero, the zonal flow cannot decrease everywhere. Thus, if the zonal flow decreases in regions away from the stirring, it must *increase* in the region of the stirring. In the presence of forcing and dissipation this mechanism can lead to the production of a statistically steady jet in the region of the forcing, for (3.51) and (3.55) combine to give

$$\frac{\partial \bar{u}}{\partial t} + \frac{\partial P}{\partial t} = -r\bar{u} + \frac{1}{\gamma} (\overline{\zeta'F'_\zeta} - \overline{\zeta'D'_\zeta}), \quad (3.58)$$

and in a statistically steady state

$$\boxed{r\bar{u} = \frac{1}{\gamma} (\overline{\zeta'F'_\zeta} - \overline{\zeta'D'_\zeta})}. \quad (3.59)$$

The terms on the right-hand side simply represent the stirring and dissipation of vorticity, and integrated over latitude their sum will vanish, or otherwise the pseudomomentum budget cannot be in a steady state. However, let us suppose that forcing is confined to midlatitudes. In that region, the first term on the right-hand side of (3.59) will be larger than the second, and eastward mean flow will be generated. Away from the direct influence of the forcing, the dissipation term will dominate and westward mean flows will be generated, as sketched in Fig. 3.3 *On a β -plane or on the surface of a rotating sphere an eastward mean zonal flow can be maintained by a vorticity stirring that imparts no net momentum to the fluid.* More generally, stirring in the presence

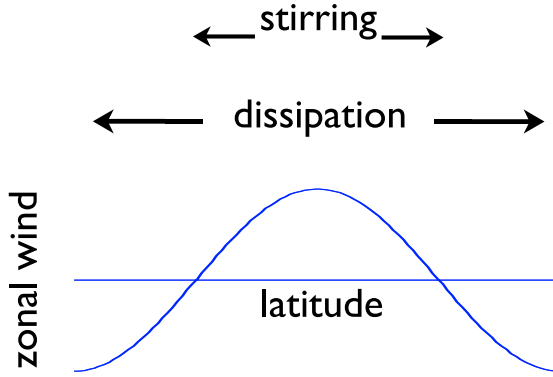


Figure 3.3 Pseudomomentum stirring, which in reality occurs via baroclinic instability, is confined to midlatitudes. Because of Rossby wave propagation away from the source region, the distribution of pseudomomentum dissipation is broader, and the sum of the two leads to the zonal wind distribution shown, with positive (eastward) values in the region of the stirring. See also Fig. 3.8.

of vorticity gradient gives rise to a mean flow, and on a spherical planet that vorticity gradient is provided by rotation.

In above arguments, the vorticity equation (3.52) was linearized about a mean flow whereas the zonal momentum equation (3.50) was not. Is this consistent? If the eddy amplitude is small, then linearization is certainly appropriate in the vorticity equation. However, even in this case we cannot linearize the zonally averaged zonal momentum equation because there is nothing to linearize it about: there is no large term associated with the mean flow that dominates the other terms if the eddy amplitude is small. The reader may also object that we have not *proven* that the forcing and dissipation terms will not locally balance in the region of the forcing, producing no net winds. That can only occur if the dissipation is confined to the region of the forcing, but this is highly unlikely because Rossby waves are generated in the forcing region, and these propagate meridionally before dissipating, as we now discuss.

III. Rossby waves and momentum flux

We have seen that the presence of a mean gradient of vorticity is an essential ingredient in the mechanism whereby a mean flow is generated by stirring. Given such, we expect Rossby waves to be excited, and we now show how Rossby waves are intimately related to the momentum flux maintaining the mean flow.

If a stirring is present in midlatitudes then we expect that Rossby waves will be generated there. To the extent that the waves are quasi-linear and do not interact then away from the immediate source region each wave has the form

$$\psi = \text{Re } C e^{i(kx+ly-\omega t)} = \text{Re } C e^{i(kx+ly-ckt)}, \quad (3.60)$$

where C is a constant, with dispersion relation

$$\omega = ck = \bar{u}k - \frac{\beta k}{k^2 + l^2} \equiv \omega_R, \quad (3.61)$$

provided that there is no meridional shear in the zonal flow. The meridional component of the group velocity is given by

$$c_g^y = \frac{\partial \omega}{\partial l} = \frac{2\beta kl}{(k^2 + l^2)^2}. \quad (3.62)$$

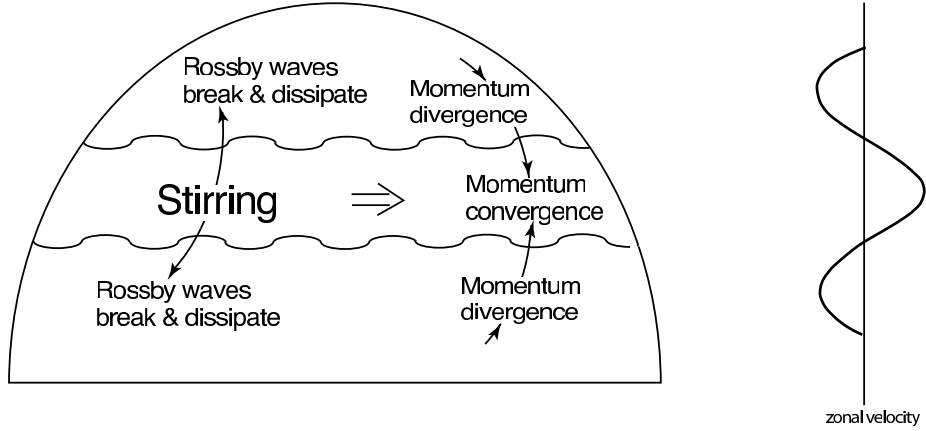


Fig. 3.4 Generation of zonal flow on a β -plane or on a rotating sphere. Stirring in midlatitudes (by baroclinic eddies) generates Rossby waves that propagate away from the disturbance. Momentum converges in the region of stirring, producing eastward flow there and weaker westward flow on its flanks.

Now, the direction of the group velocity must be *away* from the source region; this is a radiation condition (discussed more in the next subsection), demanded by the requirement that Rossby waves transport energy *away* from the disturbance. Thus, northwards of the source kl is positive and southwards of the source kl is negative. That the product kl can be positive or negative arises because for each k there are two possible values of l that satisfy the dispersion relation (3.61), namely

$$l = \pm \left(\frac{\beta}{\bar{u} - c} - k^2 \right)^{1/2}, \quad (3.63)$$

assuming that the quantity in brackets is positive.

The velocity variations associated with the Rossby waves are

$$u' = -\text{Re } C i l e^{i(kx+ly-\omega t)}, \quad v' = \text{Re } C i k e^{i(kx+ly-\omega t)}, \quad (3.64a,b)$$

and the associated momentum flux is

$$\overline{u'v'} = -\frac{1}{2} C^2 k l. \quad (3.65)$$

Thus, given that the sign of kl is determined by the group velocity, northwards of the source the momentum flux associated with the Rossby waves is southward (i.e., $\overline{u'v'}$ is negative), and southwards of the source the momentum flux is northwards (i.e., $\overline{u'v'}$ is positive). That is, the momentum flux associated with the Rossby waves is *toward* the source region. Momentum converges in the region of the stirring, producing net eastward flow there and westward flow to either side.

Another way of seeing this is to note that if kl is positive then lines of constant phase ($kx + ly = \text{constant}$) are tilted north-west/south-east, and the momentum flux

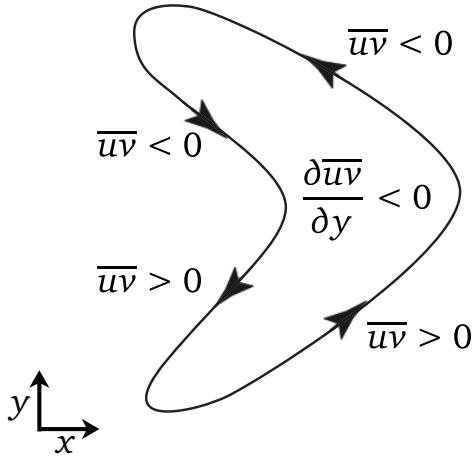


Figure 3.5 The momentum transport in physical space, caused by the propagation of Rossby waves away from a source in midlatitudes. The ensuing bow-shaped eddies are responsible for a convergence of momentum, as indicated in the idealization pictured.

associated with such a disturbance is negative ($\overline{u'v'} < 0$). Similarly, if kl is negative then the constant-phase lines are tilted north-east/south-west and the associated momentum flux is positive ($\overline{u'v'} > 0$). The net result is a convergence of momentum flux into the source region. In physical space this is reflected by having eddies that are ‘bow-shaped’, as in Fig. 3.5.

** The radiation condition and Rayleigh friction*

Why is the group velocity directed away from the source region? It is because the energy flux travels at the group velocity, and the energy flux must be directed away from source region; the reader comfortable with that result may stop here. Another way to determine the direction of the group velocity is to employ a common trick in fluid dynamics, especially in problems of wave propagation, that of adding a small amount of friction to the inviscid problem.³ The solution of the ensuing problem in the limit of small friction will often make clear which solution is physically meaningful in the inviscid problem, and therefore which solution nature chooses. Consider the linear barotropic vorticity equation with linear friction,

$$\frac{\partial \zeta}{\partial t} + \beta \frac{\partial \psi}{\partial x} = -r\zeta \quad (3.66)$$

where r is a small friction coefficient. The dispersion relation is

$$\omega = -\frac{\beta k}{K^2} - ir = \omega_R(k, l) - ir, \quad (3.67)$$

where ω_R is defined by (3.61), and the wave decays with time. Now suppose a wave is generated in some region, and that it propagates meridionally away, decaying as moves away. Then, instead of an imaginary frequency, we may suppose that the frequency is real and the y -wavenumber is imaginary. Specifically, we take $l = l_0 + l'$ where $l_0 = \pm[\beta/(\bar{u} - c) - k^2]^{1/2}$ for some zonal wavenumber k , as in (3.63), and $\omega = \omega_R(k, l_0)$. For small friction, we obtain l' by Taylor-expanding the dispersion relation around its

inviscid value, $\omega_R(k, l_0)$, giving

$$\omega + ir = \omega_R(k, l) \approx \omega_R(k, l_0) + \frac{\partial \omega_R(k, l_0)}{\partial l} l', \quad (3.68)$$

and therefore

$$l' = \frac{ir}{c_g^y} \quad (3.69)$$

where $c_g^y = \partial_l \omega_R(k, l_0)$ is the y -component of the group velocity. The wavenumber is imaginary, so that the wave either grows or decays in the y -direction. The wave solution then obeys

$$\psi \approx C \exp[i(kx - \omega_R t)] \exp(il_0 y - r y / c_g^y). \quad (3.70)$$

We now demand that the solution decay away from the source, because any other choice is manifestly unphysical, even as we let r be as small as we please. Thus, with the source at $y = 0$, c_g^y must be positive for positive y and negative for negative y . In other words, the group velocity must be directed *away* from the source region, and therefore momentum flux converges on the source region.

IV. The Eliassen-Palm flux

The Eliassen-Palm (EP) flux provides a useful framework for determining how waves affect the mean flow, and the barotropic case is a particularly simple and instructive example. The zonally averaged momentum equation may be written, for either a stratified or barotropic model, as

$$\frac{\partial \bar{u}}{\partial t} - f_0 \bar{v}^* = \nabla_x \cdot \mathcal{F} - r \bar{u} \quad (3.71)$$

where \bar{v}^* is the residual meridional velocity and \mathcal{F} is the Eliassen-Palm (EP) flux, and $\nabla_x \cdot$ is the divergence in the meridional plane. In the barotropic case $\bar{v}^* = 0$ and

$$\mathcal{F} = -\mathbf{j} \overline{u'v'}. \quad (3.72)$$

Now, if the momentum flux is primarily the result of interacting nearly-plane Rossby waves, then the EP flux obeys the group velocity property (chapter ??), namely that the flux of wave activity is equal to the group velocity times the wave activity density. Thus,

$$\mathcal{F}^y \equiv \mathbf{j} \cdot \mathcal{F} \approx c_g^y \mathcal{A} \quad (3.73)$$

where \mathcal{A} is the wave activity density, or pseudomomentum,

$$\mathcal{A} = \frac{\overline{\zeta'^2}}{\overline{q}_y} = \frac{\overline{\zeta'^2}}{y} \approx \frac{\overline{\zeta'^2}}{\beta}, \quad (3.74)$$

and, if $y > 0$, \mathcal{A} is a positive definite quantity. Now, the group velocity is directed *away* from the region of disturbance, and furthermore if the vorticity gradient is everywhere positive then the EP flux takes the sign of the group velocity (3.62). Thus, as sketched in Fig. 3.5 and Fig. 3.6, momentum converges in the region of the disturbance and an eastward jet is generated. This argument is largely equivalent to that of the previous subsection, and the result of (3.65) can be regarded as illustrating the group velocity property of the EP flux for barotropic Rossby waves in an argument from first principles. [Using (3.62), (3.73) and (3.74) we can explicitly recover (3.65).]

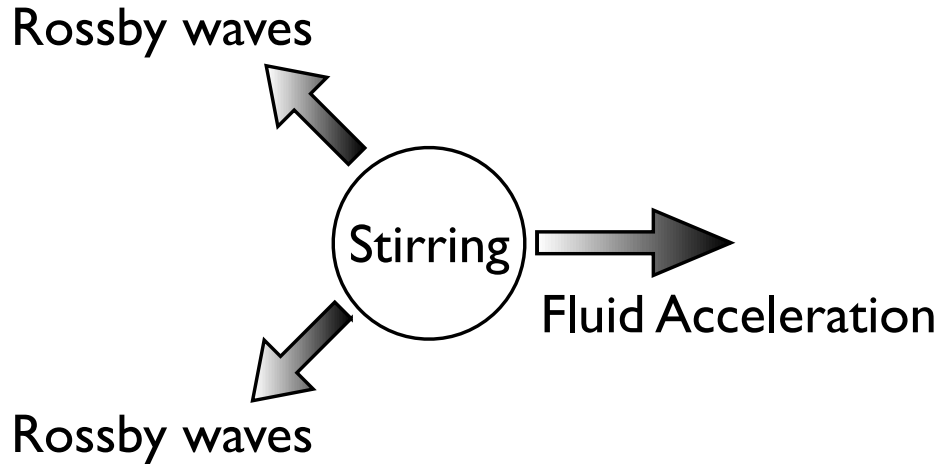


Fig. 3.6 If a region of fluid on the β -plane or on a rotating sphere is stirred, then Rossby waves will propagate westwards and away from the disturbance, and this is the direction of propagation of wave activity density. Thus, there is positive divergence of wave activity in the stirred region, and using (3.73) and (3.71) this produces a westward acceleration.

3.3.3 A numerical example

We conclude from above arguments that momentum will converge into a rapidly rotating flow that is stirred in a meridionally localized region. To illustrate this, we numerically integrate the barotropic vorticity equation on the sphere, with a meridionally localized stirring term; explicitly, the equation that is integrated is

$$\frac{\partial \zeta}{\partial t} + J(\psi, \zeta) + \beta \frac{\partial \psi}{\partial x} = -r\zeta + \kappa \nabla^4 \zeta + F_\zeta. \quad (3.75)$$

The first term on the right-hand side is a linear drag, parameterizing momentum loss in an Ekman layer. The second term removes enstrophy that has cascaded to small scales; it has a negligible impact at large scales. The forcing term F_ζ is a ‘wavemaker’ confined to a zonal strip of about 15° meridional extent, centered at about 45° N, that is statistically zonally uniform and that spatially integrates to zero. Within that region it is a random stirring with a temporal decorrelation scale of a few days and a spatial decorrelation scale corresponding to about wavenumber 8, so mimicking weather scales. Thus, it provides no net source of vorticity or momentum, but it is a source of pseudomomentum because $\overline{F\zeta} > 0$.

The results of a numerical integration of (3.75) are illustrated in Fig. 3.7 and Fig. 3.8. An eastward jet forms in the vicinity of the forcing, with westward flow on either side. The pseudomomentum stirring and dissipation that produces this flow is shown in Fig. 3.8. As expected, the dissipation has a broader distribution than the forcing, and their sum (the dot-dashed line in the figure) has the same meridional distribution as the zonal flow itself.

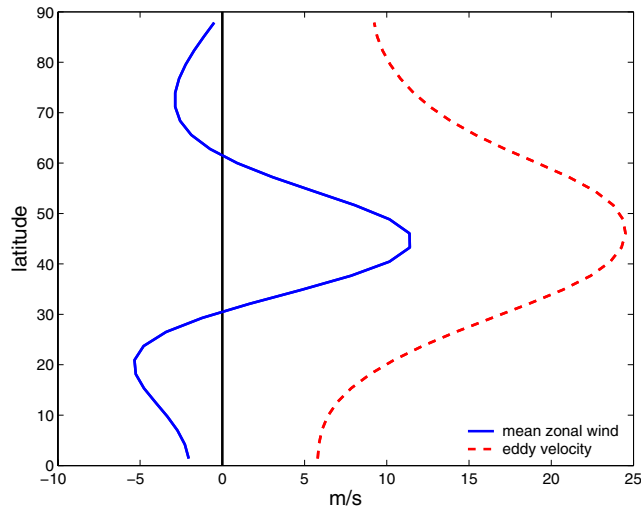


Fig. 3.7 The time- and zonally-averaged wind (solid line) obtained by an integration of the barotropic vorticity equation (3.75) on the sphere. The fluid is stirred in midlatitudes by a random wavemaker that is statistically zonally uniform, acting around zonal wavenumber 8, and that supplies no net momentum. Momentum converges in the stirring region leading to an eastward jet with a westward flow to either side, and zero area-weighted spatially integrated velocity. The dashed line shows the r.m.s. (eddy) velocity created by the stirring.

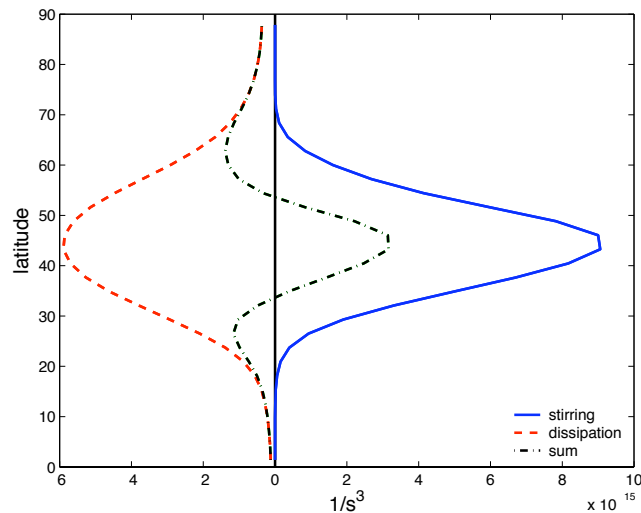


Fig. 3.8 The pseudomomentum stirring (solid line, $\overline{F'_\zeta \zeta'}$), dissipation (dashed line, $\overline{D'_\zeta \zeta'}$) and their sum (dot-dashed), for the same integration as Fig. 3.7. Because Rossby waves propagate away from the stirred region before breaking, the distribution of dissipation is broader than the forcing, resulting in an eastward jet where the stirring is centered, with westward flow on either side.

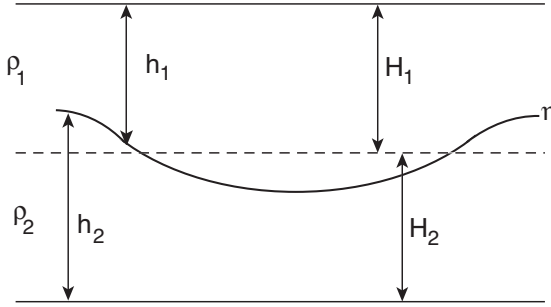


Figure 3.9 An atmosphere with two homogeneous (or isentropic) layers of mean thickness H_1 and H_2 , local thickness h_1 and h_2 , and interface η , contained between two flat, rigid surfaces.

3.4 THE FERREL CELL

We now construct a two-layer model of the mid-latitude tropospheric general circulation.

We assume that quasi-geostrophic scaling holds, and that the two fluid layers are held between two flat rigid lids.

3.4.1 Equations of motion

The equations of motion are those of a two-layer Boussinesq shallow water model confined between two rigid flat surfaces, and readers who are comfortable with these dynamics may quickly skip through this section, merely glancing at the boxed equations as they pass. The momentum equations of each layer are

$$\frac{D\mathbf{u}_1}{Dt} + \mathbf{f} \times \mathbf{u}_1 = -\nabla\phi_1 \quad (3.76a)$$

$$\frac{D\mathbf{u}_2}{Dt} + \mathbf{f} \times \mathbf{u}_2 = -\nabla\phi_2 - r\mathbf{u}_2. \quad (3.76b)$$

where $\phi_1 = p_T/\rho_0$ and $\phi_2 = p_T/\rho - g'\nabla\eta$, where p_T is the pressure at the lid at the top, η is the interface displacement (see Fig. 3.9) and $g' = g(\rho_1 - \rho_2)/\rho_0$ is the reduced gravity. We have also included a simple representation of surface drag, $-r\mathbf{u}_2$, in the lowest layer, and r is a constant. We will use a constant value of the Coriolis parameter except where it is differentiated, and on zonal averaging (3.76) become

$$\frac{\partial \bar{u}_1}{\partial t} - f_0 \bar{v}_1 = \bar{v}'_1 \zeta'_1 \quad (3.77a)$$

$$\frac{\partial \bar{u}_2}{\partial t} - f_0 \bar{v}_2 = \bar{v}'_2 \zeta'_2 - r \bar{u}_2, \quad (3.77b)$$

Geostrophic balance in each layer implies

$$f_0 \mathbf{u}_{g1} = \mathbf{k} \times \nabla \phi_T, \quad f_0 \mathbf{u}_{g2} = \mathbf{k} \times \nabla \phi_T - g' \mathbf{k} \times \nabla \eta, \quad (3.78a,b)$$

where the subscript g denotes geostrophic. Subtracting one equation from the other gives

$$\boxed{f_0(\mathbf{u}_1 - \mathbf{u}_2) = g' \mathbf{k} \times \nabla \eta}, \quad (3.79)$$

now dropping the subscript g . This is thermal wind balance for this system. A temperature gradient thus corresponds to a slope of the interface height, with the interface sloping upwards toward lower temperatures, analogous to isentropes sloping up toward the pole in the real atmosphere.

The quasi-geostrophic potential vorticity for each layer is

$$q_i = \zeta_i + f - f_0 \frac{h_i}{H_i} \quad (3.80)$$

where H_i is the reference thickness of each layer, which we take to be its mean thickness. The potential vorticity flux in each layer is then

$$\overline{v'_i q'_i} = \overline{v'_i \zeta'_i} - \frac{f_0}{H_i} \overline{v'_i h'_i}. \quad (3.81)$$

Using this in (3.77) gives

$$\begin{cases} \frac{\partial \overline{u}_1}{\partial t} = \overline{v'_1 q'_1} + f_0 \overline{v}_1^* \\ \frac{\partial \overline{u}_2}{\partial t} = \overline{v'_2 q'_2} + f_0 \overline{v}_2^* - r \overline{u}_2 \end{cases}. \quad (3.82)$$

where

$$\overline{v}_i^* = \overline{v} + \frac{\overline{v'_i h'_i}}{H_i} \quad (3.83)$$

is the meridional component of the residual velocity in each layer, proportional to the *total* meridional mass flux in each layer. These are the transformed Eulerian mean (TEM) forms of the equations, first encountered in section 3.2.

In the barotropic model of section 3.3 the mean meridional velocity vanished at every latitude, a consequence of mass conservation in a single layer between two rigid flat surfaces. In a single-layer model of section ?? the mean meridional velocity was in general non zero, but the total meridional mass flux (i.e., the meridional component of the residual velocity) was zero if the domain is bounded laterally by solid walls. In the two layer model we will allow a transformation of mass from one layer to another, which is the equivalent of heating: a conversion of mass from the lower layer to the upper layer is heating, and conversely for cooling. Thus, heating at low latitudes and cooling at high latitudes leads to the interface sloping upward toward the pole. In the two-layer model the constraint that mass conservation supplies is that, assuming a statistically steady state, the total polewards mass flux summed over both layers must vanish.

The mass conservation equation for each layer is

$$\frac{\partial h_i}{\partial t} + \nabla \cdot (h_i \mathbf{u}_i) = S_i, \quad (3.84)$$

where S_i is the mass source term and we may suppose that $S_1 + S_2 = 0$ everywhere. A zonal average gives

$$\frac{\partial \overline{h}_i}{\partial t} + \frac{\partial \overline{h}_i \overline{v}_i}{\partial y} = S_i \quad (3.85)$$

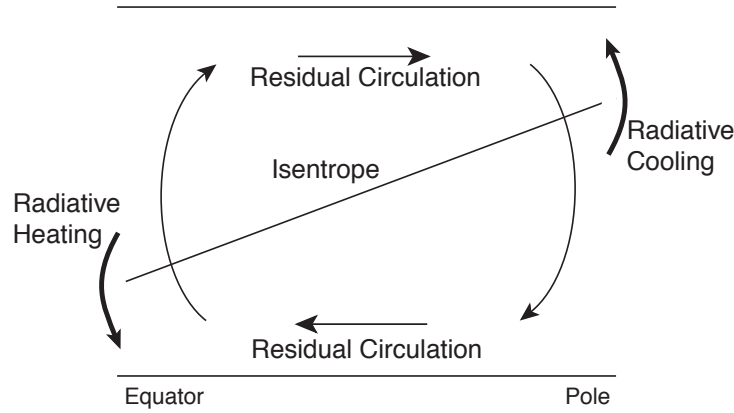


Fig. 3.10 Sketch of the zonally-averaged thermodynamics of a two-layer model. Cooling at high latitudes and heating at low leads steepens the interface upward toward the pole (thicker arrows). Associated with this there is a net mass flux — the residual flow, or the meridional overturning circulation (lighter arrows). In the tropics this circulation is accounted for by the Hadley Cell, and is nearly all in the mean flow. In midlatitudes the circulation — the residual flow — is largely due to baroclinic eddies, and the smaller Eulerian mean flow is actually in the opposite sense.

or, setting $\bar{h}_i = H_i$ and using (3.83),

$$\boxed{\frac{\partial h_i}{\partial t} + H_i \frac{\partial \bar{v}_i^*}{\partial y} = S_i} . \quad (3.86)$$

The mass source term in these equations is equivalent to heating, and let us suppose that this is such as to provide heating at low latitudes and cooling in high. This is equivalent to a conversion upper layer mass to lower layer mass at high latitudes, and conversely at low latitudes, and this can only be balanced by a poleward mass flux in the upper layer and an equatorward mass flux in the lower layer (Fig. 3.10). That is to say, an earthlike radiative forcing between equator and pole implies that *the total mass flux in the upper layer will be poleward*. This is the opposite of the mean meridional circulation of the Ferrel cell shown in Fig. ??? What's going on? Before we can answer that, let us manipulate the equations of motion and obtain a couple of useful preliminary results.

Manipulating the equations

Because the total depth of the fluid is fixed, the mass conservation equations in each layer (3.85) may each be written as an equation for the interface displacement, namely

$$\frac{\partial \eta}{\partial t} + \nabla \cdot (\eta \mathbf{u}_1) = S_1, \quad \text{or} \quad \frac{\partial \eta}{\partial t} + \nabla \cdot (\eta \mathbf{u}_2) = -S_2 \quad (3.87)$$

Because of the thermal wind equation (3.79) these equations are identical: $\mathbf{u}_1 \cdot \nabla \eta = \mathbf{u}_2 \cdot \nabla \eta$ and $S_1 = -S_2$. (If $S_1 \neq -S_2$ the flow would not remain balanced and the thermal

wind equation could not be satisfied.) The zonally-averaged interface equation may be written as

$$\frac{\partial \bar{\eta}}{\partial t} - H_1 \frac{\partial \bar{v}_1^*}{\partial y} = S, \quad \text{or} \quad \frac{\partial \bar{\eta}}{\partial t} + H_2 \frac{\partial \bar{v}_2^*}{\partial y} = S \quad (3.88)$$

where $S = -S_1 = +S_2$, consistent with the mass conservation statement

$$H_1 \bar{v}_1^* + H_2 \bar{v}_2^* = 0, \quad (3.89)$$

which states that the vertically integrated total mass flux vanishes at each latitude.

Now, whereas (3.89) is a kinematic statement about the total mass flux, the dynamics provides a constraint on the *eddy* mass flux in each layer. Using the thermal wind relationship we have

$$\overline{(v'_1 - v'_2)\eta'} = g' \frac{\partial \overline{\eta'}}{\partial x} \eta' = 0 \quad (3.90)$$

Hence, if the upper and lower surfaces are both flat, we have that

$$\overline{v'_1 h'_1} = -\overline{v'_2 h'_2} \quad (3.91)$$

and the eddy meridional mass fluxes in each layer are equal and opposite. If the bounding surfaces are not flat, we have

$$\overline{v'_1 \eta'_T} - \overline{v'_1 h'_1} = \overline{v'_1 \eta'_B} + \overline{v'_2 h'_2} \quad (3.92)$$

instead, where η_T and η_B are the topographies at top and bottom. Eqs. (3.91) and (3.92) are *dynamical* results, and not just kinematic ones; they are equivalent to noting that the form drag on one layer due to the interface displacement is equal and opposite to that on the other, namely

$$\overline{v'_1 \eta'} = -[-\overline{v'_2 \eta'}], \quad (3.93)$$

where the minus sign inside the square brackets arises because the interface displacement is into layer one but out of layer two.

Using (3.81) (3.92) the eddy potential vorticity fluxes in the two layers are related by

$$H_1 \overline{v'_1 q'_1} + H_2 \overline{v'_2 q'_2} = H_1 \overline{v'_1 \zeta'_1} + H_2 \overline{v'_2 \zeta'_2} - f_0 \overline{v'_1 \eta'_T} + f_0 \overline{v'_1 \eta'_B}, \quad (3.94)$$

which is the layered version of the continuous result [eq. (3.30) on page 57]

$$\int_B^T \overline{v'q'} dz = \int_B^T \overline{v'\zeta'} dz + \frac{f_0}{N^2} [v'b']_B^T. \quad (3.95)$$

For flat upper and lower surfaces, and using $\overline{v_i \zeta_i} = -\partial \overline{u_i v_i} / \partial y$, (3.94) becomes

$$H_1 \overline{v'_1 q'_1} + H_2 \overline{v'_2 q'_2} = H_1 \frac{\partial}{\partial y} \overline{u'_1 v'_1} + H_2 \frac{\partial}{\partial y} \overline{u'_2 v'_2}, \quad (3.96)$$

and integrating with respect to y between quiescent latitudes gives

$$\boxed{\int [H_1 \overline{v'_1 q'_1} + H_2 \overline{v'_2 q'_2}] dy = 0}. \quad (3.97)$$

That is, the total meridional flux of potential vorticity must vanish. This is a consequence of the fact that the potential vorticity flux is the divergence of a vector field; in the continuous case

$$\overline{v'q'} = \frac{\partial \overline{u'v'}}{\partial y} - f_0 \frac{\partial}{\partial z} \frac{\overline{v'b'}}{N^2}, \quad (3.98)$$

which similarly vanishes when integrated over a volume if there are no boundary contributions.

3.4.2 Dynamics of the model

We now consider the climate, or the time-averaged statistics, of our two-layer model. The equations of motion are (3.77) or (3.82), and (3.85) or (3.86). These equations are not closed because of the presence of eddy fluxes, and in this section we make some phenomenological and rather general arguments about how these behave in order to get a sense of the general circulation. In the next section we use a specific closure to address the same problem.

Let us summarize the physical situation. The two layers of our model are confined in the vertical between two flat, rigid surfaces, and they are meridionally confined between slippery walls at high and low latitudes (the ‘pole’ and ‘equator’). The circulation is driven thermodynamically by a heating at low latitudes and cooling at high, which translates to a conversion of layer 1 fluid to layer 2 fluid at high latitudes, and the converse at low latitudes (see Fig. 3.10). This sets up an interface that slopes upwards toward the pole and, by thermal wind, a shear. This situation is baroclinically unstable, and this sets up a field of eddies, most vigorous in mid-latitudes where the temperature gradient (or interface slope) is largest. Three fields encapsulate the dynamics — the surface wind field, the meridional circulation, and the meridional temperature gradient, and our goal is to understand their qualitative structure. We note from the outset that the residual circulation is polewards in the upper layer, equatorwards in the lower layer, and that this is a thermodynamic result, a consequence of heating at low latitudes and cooling at high latitudes.

From (3.82), the steady state surface wind is given by

$$\boxed{rH_2\bar{u}_2 = H_1\overline{v'_1q'_1} + H_2\overline{v'_2q'_2} = H_1\overline{v'_1\zeta'_1} + H_2\overline{v'_2\zeta'_2}} \quad (3.99)$$

where the second equality uses (3.96). That is, *the surface wind is determined by the vertical integral of either the vorticity flux or the potential vorticity flux.*

Neglecting contributions due to the mean horizontal shear (which are small if the β -Rossby number $U/\beta L^2$ is small) the potential vorticity gradient in each layer is given by

$$\frac{\partial q_1}{\partial y} = \beta - \frac{f_0}{H_1} \frac{\partial \bar{h}_1}{\partial y} \gg 0 \quad (3.100a)$$

and

$$\frac{\partial q_2}{\partial y} = \beta - \frac{f_0}{H_2} \frac{\partial \bar{h}_2}{\partial y} \lesssim 0. \quad (3.100b)$$

In upper layer $\partial \bar{h}_1 / \partial y$ is negative so that the total potential vorticity gradient is positive and larger than β itself. In the lower layer $\partial \bar{h}_2 / \partial y$ is positive and indeed if there is to be baroclinic instability it must be as large as β in order that $\partial \bar{q} / \partial y$ change sign somewhere. Thus, although negative the potential vorticity gradient is much weaker in the lower layer. Thus, Rossby waves (meaning waves that exist because of a background gradient in potential vorticity) will propagate further in the upper layer, and this asymmetry is the key to the production of surface winds.

Now, the potential vorticity flux must be negative (and downgradient) in the upper layer, and there are various ways to see this. One is from the upper layer momentum equation (3.82a) which in a steady state gives

$$\overline{v'_1 q'_1} = -f_0 \bar{v}_1^*. \quad (3.101)$$

Because \bar{v}_1^* is polewards $f_0 \bar{v}_1^*$ is positive and the potential vorticity flux is negative in both Northern and Southern Hemispheres. Equivalently, in the upper layer the radiative forcing is increasing the potential vorticity gradient between equator and pole, so there must be an equatorward potential vorticity flux to compensate. Finally, the perturbation enstrophy or pseudomomentum equations tell us that in a steady state the potential vorticity flux is downgradient (also see section ??). This is not an independent argument, since it merely says that the enstrophy budget may be balanced through a balance between production proportional to the potential vorticity gradient and the dissipation. For similar reasons we expect the potential vorticity flux to be positive (poleward) in the lower layer.

Now, (3.97) tells us that latitudinally integrated potential vorticity flux is equal and opposite in the two layers. If the potential vorticity flux in the lower layer were everywhere equal and opposite to that in the upper layer then using (3.99) there would be no surface wind, in contrast to the observations. In fact, the potential vorticity flux is more uniformly distributed in the upper layer, and this gives rise to the surface wind observed. Let us give a couple of perspectives (on the same argument) as to why this should be so. The argument centers around the fact that the potential vorticity gradient is stronger in the upper layer, as we can see from (3.100).

I. Rossby waves and the vorticity flux

The stronger potential vorticity gradient of the upper layer is better able to support linear Rossby waves than the lower layer. Thus, the vorticity flux in the region of Rossby-wave genesis in midlatitudes will be large and positive in the upper layer, and small and negative in the lower layer. Thus, there will be more momentum convergence into the source region in the upper layer than in the lower layer, and the vertical integral of the vorticity flux largely will be roughly equal that of the upper layer. This is positive in midlatitudes and, to ensure that its latitudinal integral is zero, it is negative on either side. Using (3.99), a surface wind has the same pattern as the net vorticity flux, and so is eastward in the mid-latitude source region and westward on either side.

II. Potential vorticity flux

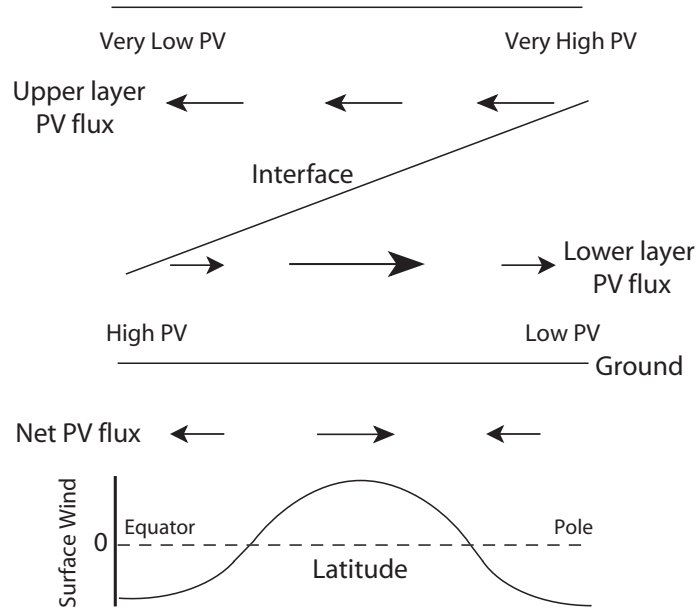


Fig. 3.11 Sketch of the potential vorticity fluxes in a two-layer model. The surface wind is proportional to their vertical integral. The PV fluxes are negative (positive) in the upper (lower) layer, but are more uniformly distributed at upper levels. The lower panel shows the net (vertically integrated) PV fluxes, and the associated surface winds.

Rossby waves are generated in the region of baroclinic instability, at approximately the same latitude in both upper and lower layers. However, because the potential vorticity gradient is larger in the upper layer than in the lower layer, Rossby waves are able to propagate more efficiently and breaking and associated dissipation will tend to be further from the source region in the upper layer than in the lower layer. Now, the pseudomomentum equation for each layer is

$$\frac{\partial P_i}{\partial t} = \frac{\partial}{\partial t} \left(\frac{\overline{q_i'^2}}{2\gamma_i} \right) = -\overline{v_i' q_i'} - \frac{\overline{D_i' q_i'}}{\gamma_i}, \quad i = 1, 2. \quad (3.102)$$

where γ_i , the potential vorticity gradient, has opposite signs in each layer. In a statistically steady state, the region of strongest dissipation is the region where the potential vorticity flux is be most negative. In the upper layer, Rossby-wave propagation allows the dissipation region to spread out from the source, whereas in the lower layer the dissipation region will be concentrated near the source. The distribution of the potential vorticity flux then becomes as illustrated in Fig. 3.12. The surface winds, being the vertical integral of the potential vorticity fluxes, are westerly in the baroclinic region and easterly to either side.

Phenomenology of a Two-layer Mid-latitude Atmosphere: a Summary

A radiative forcing that heats low latitudes and cools high latitudes will lead to a interface that slopes upward with increasing latitude, and a poleward total mass flux in the upper layer and an equatorward flux in the lower layer. The interface implies a thermal wind shear between the two layers. Neglecting relative vorticity, the potential vorticity gradients in each layer are given by

$$\frac{\partial \bar{q}_1}{\partial y} = \beta - \frac{f_0}{H_1} \frac{\partial \bar{h}_1}{\partial y} > 0 \quad \text{and} \quad \frac{\partial \bar{q}_2}{\partial y} = \beta - \frac{f_0}{H_2} \frac{\partial \bar{h}_2}{\partial y} \lesssim 0. \quad (\text{TL.1})$$

The gradient is large and positive in upper layer and small and negative in the lower layer — the gradient must change sign if there is to be baroclinic instability which we assume to be the case. This baroclinic instability generates eddy fluxes that largely determine the surface winds and the meridional overturning circulation. The zonal momentum equation in each layer is

$$\frac{\partial \bar{u}_1}{\partial t} = f_0 \bar{v}_1 + \overline{v'_1 \zeta'_1} = f_0 \bar{v}_1^* + \overline{v'_1 q'_1} \quad (\text{TL.2a})$$

$$\frac{\partial \bar{u}_2}{\partial t} = f_0 \bar{v}_2 + \overline{v'_2 \zeta'_2} - r \bar{u}_2 = f_0 \bar{v}_2^* + \overline{v'_2 q'_2} - r \bar{u}_2 \quad (\text{TL.2b})$$

In steady state the potential vorticity flux will be equatorward in the upper layer and poleward in the lower layer. Because the mass flux in each layer is equal and opposite, the surface wind is given by the vertical integral of the vorticity or potential vorticity fluxes, namely

$$r H_1 \bar{u}_1 = H_1 \overline{v'_1 q'_1} + H_2 \overline{v'_2 q'_2} = H_1 \overline{v'_1 \zeta'_1} + H_2 \overline{v'_2 \zeta'_2} \quad (\text{TL.3})$$

Because the potential vorticity gradient in the upper layer is large, this layer is more linear than the lower layer and Rossby waves are better able to transport momentum. The vorticity flux is thus stronger in the upper layer than the lower and, using (TL.3), the surface winds are positive (eastward) in the mid-latitude baroclinic zone (see Fig. 3.12). To balance the upper level midlatitude momentum flux convergence a meridional overturning circulation (a Ferrel cell) is generated. In a steady state $f_0 \bar{v}_1 = -\overline{v'_1 \zeta'_1}$ so that the zonally averaged upper level flow is equatorward. However, the total mass flux in the upper level is poleward; thus, the equatorward meridional velocity in the upper branch of the Ferrel cell is an product of an Eulerian zonal average and does not correspond to a net equatorward mass transport.

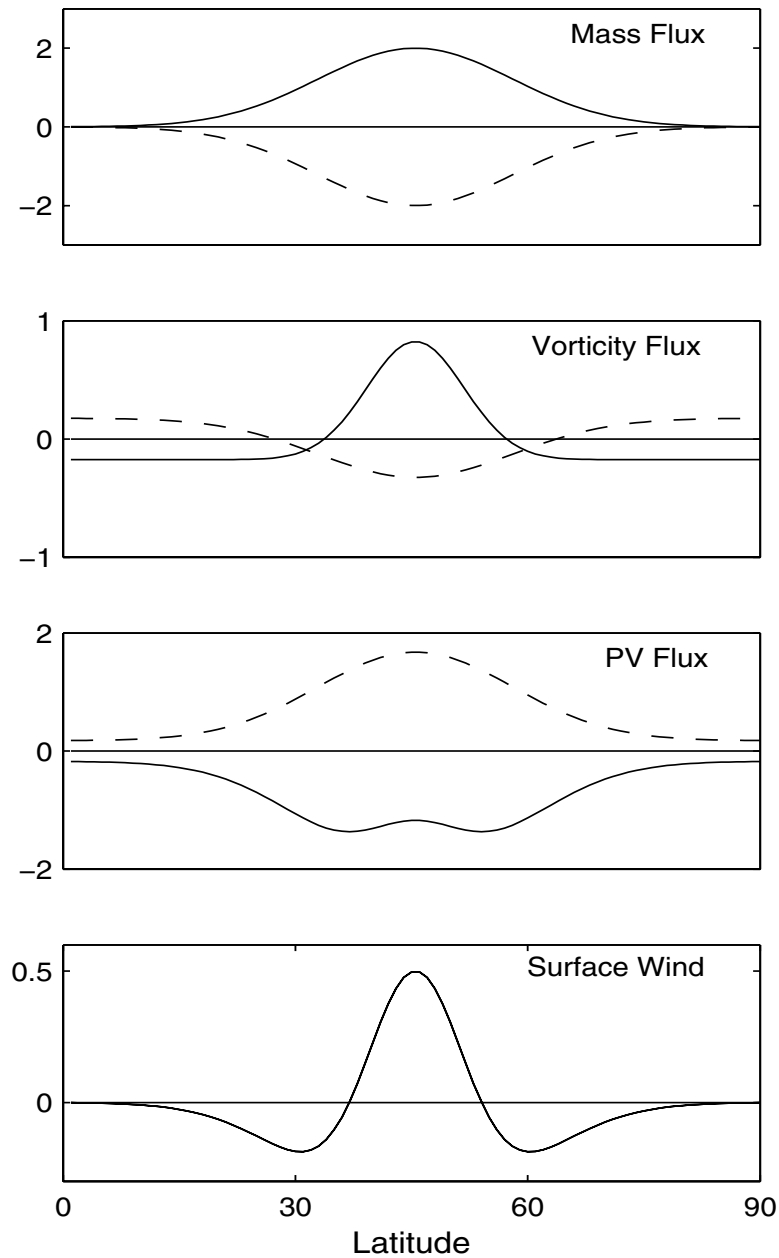


Fig. 3.12 Schema of the eddy fluxes in a two-layer model of an atmosphere with a single mid-latitude baroclinic zone. The upper layer fluxes are solid lines, the lower layer fluxes are dashed. The lowest panel shows the sum of the lower and upper layer vorticity fluxes (or, the same thing, the sum of the potential vorticity fluxes), which is proportional (when the surface friction is a linear drag) to the surface wind. The fluxes satisfy the various relationships and integral constraints of section ?? but are otherwise idealized.

Momentum balance and the overturning circulation

From thermodynamic arguments we deduced that the residual circulation is ‘direct’, meaning that warm fluid rises in low latitudes, moves poleward aloft, and returns near the surface. In low latitudes where eddy effects are small the zonally averaged Eulerian circulation circulates in the same way, giving us the Hadley Cell. In midlatitudes, we may determine the Eulerian circulation from the momentum equation, (3.77). In the upper layer the balance is between the vorticity flux and the Coriolis term, namely

$$f_0 \bar{v}_1 = -\overline{v'_1 \zeta'_1} < 0. \quad (3.103)$$

That is, *the mean Eulerian flow is equatorward*, and this is the upper branch of the Ferrel cell. Note that the Eulerian circulation is in the opposite sense to the residual circulation.

In lower layer the vorticity fluxes are weak and the balance is largely between the Coriolis force on the meridional wind and the frictional force on the zonal wind. If the upper layer flow is polewards the lower layer flow must be equatorwards by mass conservation, and so the zonal wind is positive (eastwards); that is

$$r \bar{u}_2 \approx f_0 \bar{v}_2 = -\frac{H_1}{H_2} f_0 \bar{v}_1 > 0. \quad (3.104a,b)$$

where the second inequality follows by mass conservation of the Eulerian flow.

In terms of the TEM form of the equations, (3.82), the corresponding balances in the center of the domain are:

$$f_0 \bar{v}_1^* = -\overline{v'_1 q'_1} > 0, \quad (3.105a)$$

and

$$r \bar{u}_2 = f_0 \bar{v}_2^* + \overline{v'_2 q'_2} = -f_0 \frac{H_1}{H_2} \bar{v}_1^* + \overline{v'_2 q'_2} = \frac{H_1}{H_2} \overline{v'_1 q'_1} + \overline{v'_2 q'_2} > 0, \quad (3.105b)$$

using mass conservation. An example of the dynamical balances of the two-layer model is given in Fig. 3.12)

3.5 THE ANTACTIC CIRCUMPOLAR CHANNEL

We now take a closer look at the Antarctic Circumpolar Current (ACC) itself, with less of a focus on how the ACC connects the rest of the worlds ocean and more of a focus on its own internal dynamics. This current system, sketched in Fig. 3.13, differs from other oceanic regimes primarily in that the flow is, like that of the atmosphere, predominantly zonal and re-entrant. The two obvious influences on the circulation are the strong, eastward winds (the ‘roaring forties’, the ‘furious fifties’, and the ‘screaming sixties’) and the buoyancy forcing associated with the meridional gradient of atmospheric temperature and radiative effects which cause ocean cooling at high latitudes and warming at low. Providing a detailed description of the resulting flow is properly the province of numerical models, and here our goals are much more modest, namely to describe some of the basic dynamical mechanisms that determine the structure and transport of the system.⁵

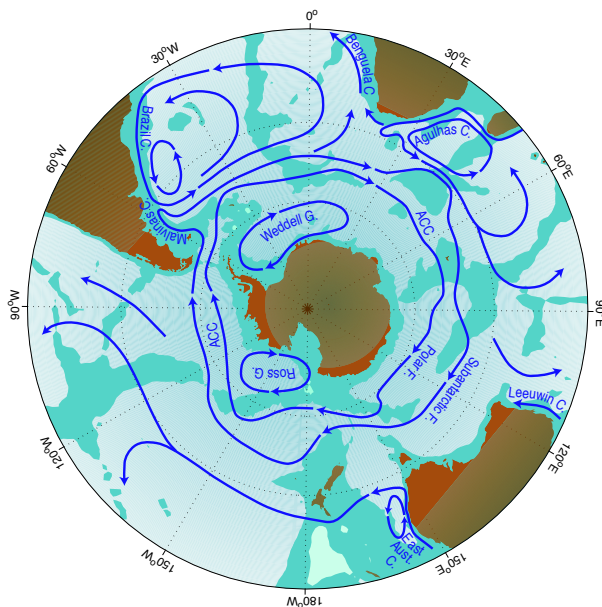


Figure 3.13 Schema of the major currents in the Southern Ocean. Shown are the South Atlantic subtropical gyre, and the two main cores of the ACC, associated with the Polar front and the sub-Antarctic front.⁴

3.5.1 Steady and eddying flow

Consider again the simplified geometry of the Southern Ocean as sketched in Fig. ???. The ocean floor is flat, except for a ridge (or ‘sill’) at the same longitude as the the gyre walls; this is a crude representation of the topography across the Drake Passage, that part of the ACC between the tip of South America and the Antarctic Peninsula. In the planetary geostrophic approximation, the steady response is that of nearly vertical isopycnals in the area above the sill, as illustrated in Fig. ??. Below the sill a meridional flow can be supported and the isotherms spread polewards, as illustrated in numerical solutions using the primitive equations (Fig. 3.14).⁶

The stratification of the non-eddy simulation is similar to that predicted by the idealized model illustrated in Fig. ??. However, the steep isotherms within the channel contain a huge amount of available potential energy (APE), and the flow is highly baroclinically unstable, and if baroclinic eddies are allowed to form, the solution is dramatically different: the isotherms slump, releasing that APE and generating mesoscale eddies. An important conclusion is that *baroclinic eddies are of leading-order importance in the dynamics of the ACC*. A dynamical description of the ACC without eddies would be *qualitatively* in error, in much the same way as would a similar description of the mid-latitude troposphere (i.e., the Ferrel Cell). (However, it is less clear whether these eddies are important in the interaction of the ACC with the rest of the worlds oceans.) These eddies transfer both heat and momentum, and much of the rest of our description will focus on their effects.

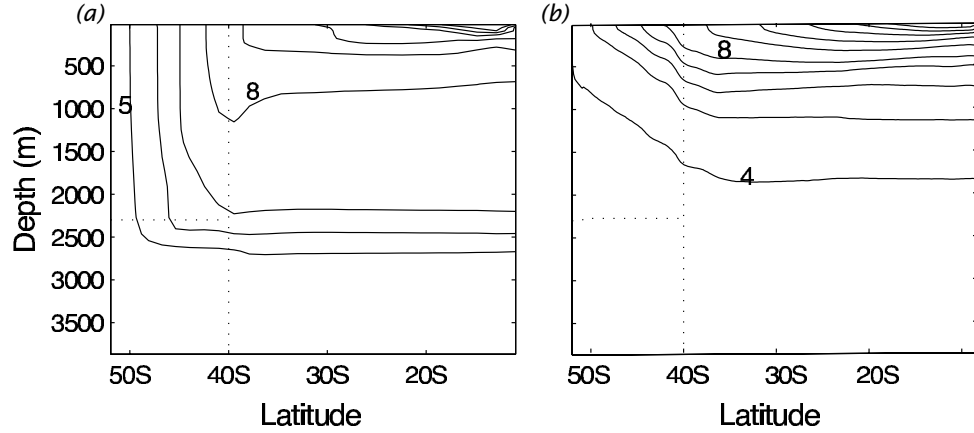


Fig. 3.14 The zonally-averaged temperature field in numerical solutions of the primitive equations in a domain similar to that of Fig. ?? (except that here the channel and sill are nestled against the polewards boundary). Panel (a) shows the steady solution of a diffusive model with no baroclinic eddies, and (b) shows the time-averaged solution in a higher resolution model that allows baroclinic eddies to develop. The dotted lines show the channel boundaries and the sill.⁷

3.5.2 Vertically integrated momentum balance

The momentum supplied by the strong eastward winds must somehow be removed. Presuming that lateral transfers of momentum are small the momentum must be removed by fluid contact with the solid earth at the bottom of the channel. Thus, let us first consider the vertically integrated momentum balance in a channel, without regard to how the momentum might be vertically transferred. We begin with the frictional-geostrophic balance

$$\mathbf{f} \times \mathbf{u} = -\nabla \phi + \frac{\partial \tilde{\boldsymbol{\tau}}}{\partial z}, \quad (3.106)$$

where $\tilde{\boldsymbol{\tau}}$ is the kinematic stress (and henceforth we drop the tilde). Integrating over the depth of the ocean gives [c.f. (??)]

$$\mathbf{f} \times \hat{\mathbf{u}} = -\nabla \hat{\phi} - \phi_b \nabla \eta_b + \boldsymbol{\tau}_w - \boldsymbol{\tau}_f, \quad (3.107)$$

where $\boldsymbol{\tau}_w$ is the stress at the surface (due mainly to the wind), and $\boldsymbol{\tau}_f$ is the frictional stress at the bottom. A hat denotes a vertical integral and ϕ_b is the pressure at $z = \eta_b$, where η_b is the z -coordinate of the bottom topography.

The x -component of this equation is just

$$f \hat{v} = -\frac{\partial \hat{\phi}}{\partial x} - \phi_b \frac{\partial \eta_b}{\partial x} + \tau_w^x - \tau_f^x, \quad (3.108)$$

and on integrating around a line of latitude the term on the left-hand side vanishes by mass conservation and we are left with

$$\oint [\phi_b \frac{\partial \eta_b}{\partial x} + \tau_w^x - \tau_f^x] dx = 0. \quad (3.109)$$

The first term is the form drag, encountered in sections ?? and ??, and observations and numerical simulations indicate that it is this, rather than the frictional term τ_f^x , that predominantly balances the wind stress in the zonally and vertically integrated momentum balance.⁸ We return to the question of why this should be so later.

The vorticity balance is also dominated by a balance between bottom pressure torque and wind stress curl. Taking the curl of (3.107) gives

$$\beta \hat{v} = \mathbf{k} \cdot \nabla \phi_b \times \nabla \eta_b + \text{curl}_z \boldsymbol{\tau}_w - \text{curl}_z \boldsymbol{\tau}_f. \quad (3.110)$$

Now, on integrating over an area bounded by a latitude circle and applying Stokes's theorem the β term vanishes by mass conservation and we regain (3.109). This means that Sverdrup balance, in the usual sense of $\beta v \approx \text{curl}_z \boldsymbol{\tau}_w$, cannot hold in the zonal average: the left-hand side vanishes but the right-hand side does not. The same could be said for the zonal integral of (3.110) across a gyre, but the two cases do differ: In a gyre Sverdrup balance can (in principle) hold over most of the interior, with mass balance being satisfied by the presence of an intense western boundary current. In contrast, in a channel where the dynamics are zonally homogeneous then v must be, on average, zero at *all* longitudes and form drag and/or frictional terms must balance the wind-stress curl in a given water column. Sverdrup balance is thus a less useful foundation for channel dynamics (at least zonally homogeneous ones) than it is for gyres. Of course, the real ACC is *not* zonally homogeneous, and may contain regions of polewards Sverdrup flow balanced by equatorwards flow in boundary currents along the eastern edges of sills and continents, and the extent to which Sverdrup flow is a leading-order descriptor of its dynamics is partly a matter of geography.

We cannot in general completely neglect nonconservative frictional terms, on two counts. First, they are the means whereby kinetic energy is dissipated. Second, if there is a contour of constant orographic height encircling the domain (i.e., encircling Antarctica) then the form drag will vanish when integrated along it. However, the same integral of the wind stress will not vanish, and therefore must be balanced by something else. To see this explicitly, write the vertically integrated vorticity equation, (3.110), in the form

$$\beta \hat{v} + J(\phi_b, \eta_b) = \text{curl}_z \boldsymbol{\tau}_w - \text{curl}_z \boldsymbol{\tau}_f. \quad (3.111)$$

If we integrate over an area bounded by a contour of constant orographic height (i.e., constant η_b) then both terms on the left-hand side vanish, and the wind-stress along that line must be balanced by friction. In the real ocean there may be no such contour that is confined to the ACC — rather, any such contour would meander through the rest of the ocean; indeed, no such confined contour exists in the idealized geometry of Fig. ??.

3.5.3 Form drag and baroclinic eddies

How does the momentum put in at the surface by the wind stress make its way to the bottom of the ocean where it may be removed by form drag? We saw in section ?? that one mechanism is by way of a mean meridional overturning circulation, with an upper branch in the Ferrel cell and a lower branch at the level of the sill, with no meridional

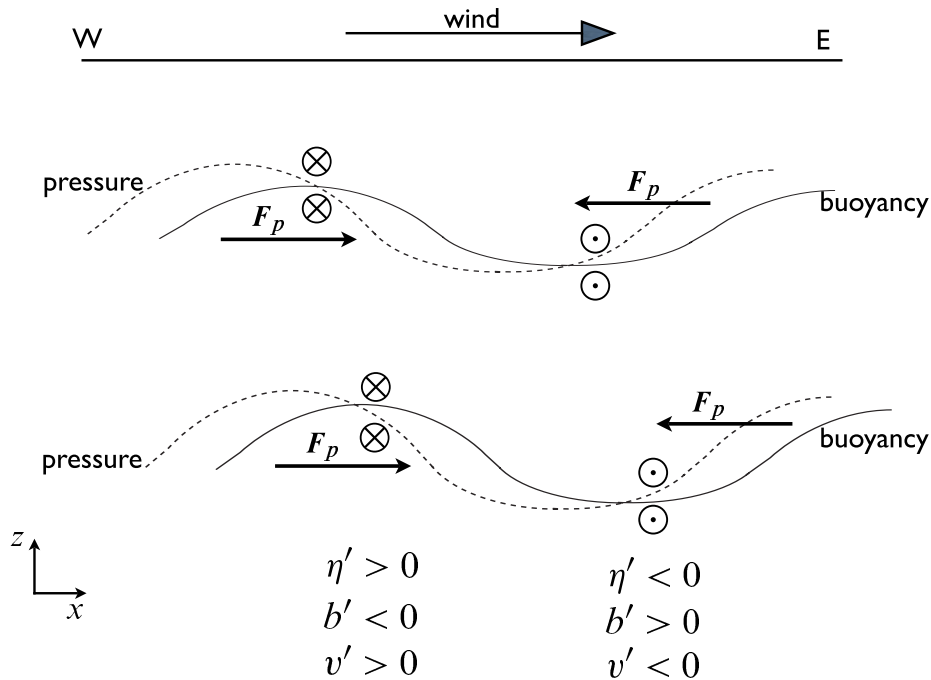


Fig. 3.15 Eddy fluxes and form drag in a Southern Hemisphere channel, viewed from the south. In this example, cold (less buoyant) water flows equatorwards and warm water polewards, so that $\overline{v'b'} < 0$. The pressure field associated with this flow (dashed lines) provides a form-drag on the successive layers, F_p , shown. At the ocean bottom the westward form drag on the fluid arising through its interaction with the orography of the sea-floor is equal and opposite to that of the eastward wind-stress at the top. The mass fluxes in each layer are given by $\overline{v'h'} \approx -\partial_z(\overline{v'b'}/N^2)$. If the magnitude of buoyancy displacement increases with depth then $\overline{v'h'} < 0$, providing a polewards mass flux that could balance the equatorwards mass flux in the Ekman layer.

flow between. However, the presence of baroclinic eddies changes things in two related ways:

- (i) Eddy form drag can pass momentum vertically within the fluid
- (ii) Eddies can allow a net meridional mass flux.

Momentum dynamics of layers

Let us first model the channel as a finite number of fluid layers, each of constant density and lying one on top of the other — a ‘stacked shallow water’ model, equivalent to a model expressed in isopycnal coordinates. The wind provides a stress on the upper layer which sets it into motion, and this in turn, via the mechanism of form drag, provides a stress to the layer below, and so on until the bottom is reached. The lowest layer then equilibrates via form drag with the bottom topography or via Ekman friction, and the general mechanism is illustrated in Fig. 3.15.

Recalling the results of section ??, the form drag at a layer interface is given by

$$\tau_i = -\eta_i \overline{\frac{\partial p_i}{\partial x}} = -\rho_0 f \overline{\eta_i v_i} \quad (3.112)$$

where p_i is the pressure and η_i the displacement at the i 'th interface (i.e., between the i 'th and $i + 1$ layer as in Fig. 3.16), and the overbar denotes a zonal average. If we define the averaged meridional transport in each layer by

$$V_i = \int_{\eta_i}^{\eta_{i-1}} \rho_0 v \, dz \quad (3.113)$$

then, neglecting the meridional momentum flux divergence (as explained in the next subsection), the time- and zonally-averaged zonal momentum balance for Boussinesq layers of fluid are:

$$-f \overline{V}_1 = \tau_w - \tau_1 = \eta_1 \overline{\frac{\partial p_1}{\partial x}} + \tau_w, \quad (3.114a)$$

$$-f \overline{V}_i = \tau_{i-1}^x - \tau_i = -\eta_{i-1} \overline{\frac{\partial p_{i-1}}{\partial x}} + \eta_i \overline{\frac{\partial p_i}{\partial x}}, \quad (3.114b)$$

$$-f \overline{V}_N = \tau_{N-1} - \tau_N = -\eta_{N-1} \overline{\frac{\partial p_{N-1}}{\partial x}} + \eta_b \overline{\frac{\partial p_b}{\partial x}} - \tau_f, \quad (3.114c)$$

where the subscripts 1, i and N refer to the top layer, an interior layer, and the bottom layer, respectively. Also, η_b is the height of bottom topography and τ_w is the zonal stress imparted by the wind which, we assume, is confined to the uppermost layer. The term τ_f represents drag at the bottom due to Ekman friction, but we have neglected any other viscous terms or friction between the layers.

The vertically integrated meridional mass transport must vanish, and thus summing over all the layers (3.114) becomes

$$0 = \tau_w - \tau_f - \tau_N \quad (3.115)$$

or, noting that $\tau_N = -\overline{\eta_b \partial p_b / \partial x}$,

$$\boxed{\tau_w = \tau_f - \eta_b \overline{\frac{\partial p_b}{\partial x}}}. \quad (3.116)$$

Thus, the stress imparted by the wind (τ_w) may be communicated vertically through the fluid by form drag, and ultimately balanced by the sum of the bottom form stress (τ_N) and bottom friction (τ_f). Note that it is also a type of form drag that leads to the momentum balance in the steady model of section ??; in that case, the southwards return flow is nestled against the sill, and the associated Coriolis force is balance by a pressure force against the sill.

Momentum dynamics in height coordinates

We now look at these same dynamics in height coordinates, using the quasi-geostrophic TEM formalism (it may be helpful to review section 3.2 before proceeding). We write the zonally-averaged momentum equation in the form

$$-f_0 \bar{v}^* = \nabla_m \cdot \mathcal{F} + \frac{\partial \tau}{\partial z} \quad (3.117)$$

where \bar{v}^* is the residual meridional velocity, τ is the zonal component of the kinematic stress (wind induced and frictional) and \mathcal{F} is the Eliassen-Palm flux, which satisfies

$$\nabla_m \cdot \mathcal{F} = -\frac{\partial}{\partial y} \overline{u'v'} + \frac{\partial}{\partial z} \left(\frac{f_0}{N^2} \overline{v'b'} \right) = \overline{v'q'}. \quad (3.118)$$

The stress is typically important only in an Ekman layer at the surface and in a frictional layer at the bottom. Now, if the horizontal velocity and buoyancy perturbations are related by $v' \sim b'/N$ (see chapter 2) then the two terms comprising the potential vorticity flux scale as

$$\frac{\partial}{\partial y} \overline{u'v'} \sim \frac{v'^2}{L_e}, \quad \frac{\partial}{\partial z} \left(f_0 \frac{\overline{v'b'}}{\bar{b}_z} \right) \sim \frac{v'^2}{L_d} \quad (3.119)$$

where L_e is the scale of the eddies and L_d is the deformation radius. If the former is much larger than the latter, as we might expect in a field of developed geostrophic turbulence (and as is indeed observed in the ACC), then the potential vorticity flux is dominated by the buoyancy flux [so also justifying the neglect of the lateral momentum fluxes in (3.114)] and (3.117) becomes

$$-f_0 \bar{v}^* \approx \frac{\partial \tau}{\partial z} + \frac{\partial}{\partial z} \left(f_0 \frac{\overline{v'b'}}{\bar{b}_z} \right). \quad (3.120)$$

If we integrate this equation over the depth of the channel the term on the left-hand side vanishes and we have

$$\tau_w = \tau_f - \left[f_0 \frac{\overline{v'b'}}{\bar{b}_z} \right]_{-H}^0, \quad (3.121)$$

where τ_w is the wind stress and τ_f is the frictional stress at the bottom. As we noted in section ??, the vertical component of the EP flux is equivalent to a form stress acting on a fluid layer, and (3.121) expresses essentially the same momentum balance as (3.116) (where it was assumed that the stress at the top arises from the wind). Thus, the EP flux expresses the passage of momentum vertically through the water column, and it is removed at the bottom through frictional stresses and/or form drag with the orography.

Mass fluxes and thermodynamics

Associated with the form drag is a meridional mass flux, which in the layered model appears as V_i in each layer. The satisfaction of the momentum balance at a particular latitude goes hand-in-hand with the satisfaction of the mass balance. Above any topography the Eulerian mean momentum equation is (with quasi-geostrophic scaling and neglecting eddy momentum fluxes),

$$f_0 \bar{v}_a = \tau \quad (3.122)$$

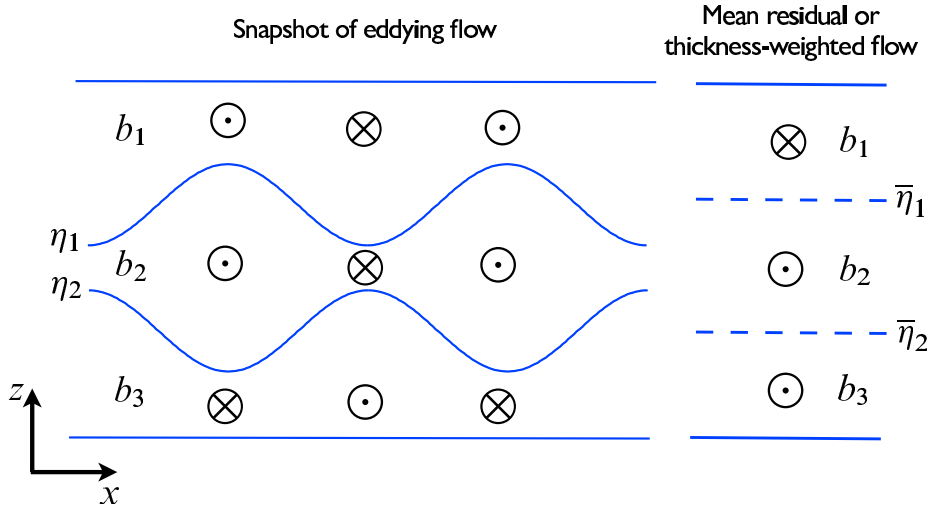


Fig. 3.16 A schema of the meridional flow in an eddying channel. The eddying flow may be organized (for example by baroclinic instability) such that, even though at any given level the Eulerian meridional flow may be small, there is a net flow in a given isopycnal layer. The residual (\bar{v}^*) and Eulerian (\bar{v}) flows are related by $\bar{v}^* = \bar{v} + \overline{v'h'}/\bar{h}$; thus, the thickness-weighted average of the eddying flow on the left gives rise to the residual flow on the right, where $\bar{\eta}_i$ denotes the mean elevation of the isopycnal η_i .

where \bar{v}_a is the ageostrophic meridional velocity and τ the zonal stress. That is, all the zonally-averaged meridional flow is ageostrophic and, in this approximation, it is non-zero only near the surface (i.e., equatorwards Ekman flow) and below the level of the topography, where it can be supported by friction and/or form drag. Even in an eddying flow, the Eulerian circulation is primarily confined to the upper Ekman layer and a frictional or topographically interrupted layer at the bottom, as illustrated in Fig. 3.17. This is a perfectly acceptable description of the flow, and is not an artifact in any way. However, and analogously to the atmospheric Ferrel Cell (section ?? and ??), if the flow is unsteady this circulation does not necessarily represent the flow of water parcels, nor does it imply that water parcels cross isopycnals, as might be suggested by the leftmost panel of Fig. 3.17. That flow is better represented by the residual flow, or the thickness weighted flow, and as sketched in Fig. 3.16 there can be a net meridional flow in a given layer (i.e., of a given water mass type) *even when the net meridional Eulerian flow at the level of mean height of the layer is zero*.

The vertically integrated mass flux must of course vanish, and even though one component of this — the equatorwards Ekman flow — is determined mechanically, the overall sense of the residual circulation cannot be determined by the momentum balance alone: thermodynamic effects play a role. The thermodynamic equation may be written in TEM form as

$$\frac{\partial \bar{b}}{\partial t} + J(\psi^*, \bar{b}) = Q[b] \quad (3.123)$$

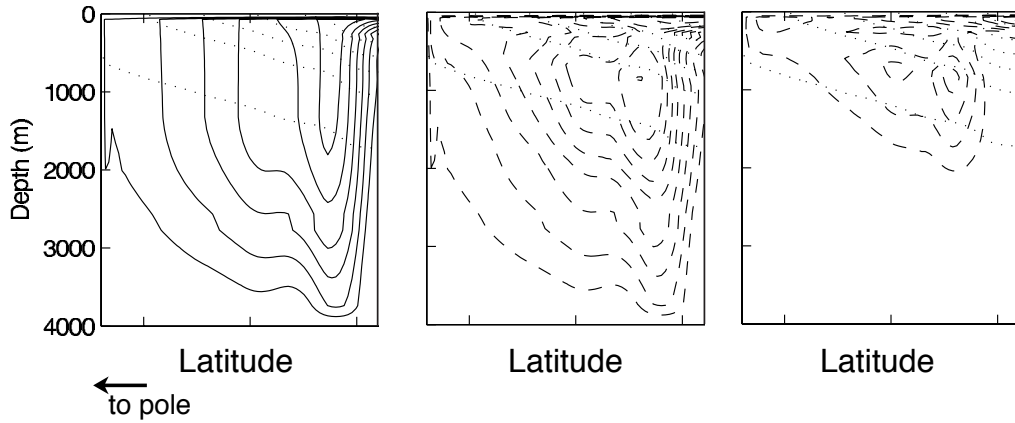


Fig. 3.17 The meridional circulation in the re-entrant channel of an idealized, eddying numerical model of the ACC (as in Fig. 3.14, but showing only the region south of 40°S). Left panel, the zonally averaged Eulerian circulation. Middle panel, the eddy induced circulation. Right panel, the residual circulation. Solid lines represent a clockwise circulation and dashed lines, anticlockwise. The faint dotted lines are the mean isopycnals. Over much of the channel the model ocean is losing buoyancy (heat) to the atmosphere and so the net, or residual, flow at the surface is polewards.

where $J(\psi^*, \bar{b}) = (\partial_y \psi^*)(\partial_z \bar{b}) - (\partial_z \psi^*)(\partial_y \bar{b}) = \bar{v}^* \partial_y \bar{b} + \bar{w}^* \partial_z \bar{b}$, ψ^* is the streamfunction of the residual flow and $Q[b]$ represents heating and cooling, which occurs mainly at the surface. In the ocean interior and in statistically steady state we therefore have

$$J(\psi^*, \bar{b}) = 0, \quad (3.124)$$

the general solution of which is $\psi^* = G(\bar{b})$ where G is an arbitrary function. That is, the residual flow is along isopycnals, and this is approximately satisfied by the numerical solution shown in Fig. 3.17.

At the surface, however, the flow is generally not adiabatic, because of heat exchange with the atmosphere. In the simulations shown there is a net heat flux, and consequent buoyancy loss, from the ocean to the atmosphere at the polewards edge of the domain. Heat balance is then achieved by a polewards residual flow of warmer fluid at the surface and sinking at the the highest latitudes. If there were no surface fluxes, and the flow were everywhere adiabatic, then we can expect the residual circulation to vanish. Note that the sense of the subsurface circulation determines how the form drag varies with depth; if the residual flow were zero for example then, either from (3.114) or from (3.120), we see that the form drag must be constant with depth.