Generation of Mean Flows and Jets on a Beta Plane and over Topography

GEOFFREY K. VALLIS AND MATTHEW E. MALTRUD

Department of Marine Sciences, University of California Santa Cruz, Santa Cruz, California

(Manuscript received 3 March 1992, in final form 27 July 1992)

ABSTRACT

This paper proposes and discusses mechanisms whereby mean flows and jets are produced by differential rotation and by topographic effects. It is shown that, in general, a mean gradient of potential vorticity not only inhibits the cascade of energy to large scales but directly produces anisotropic structures. Scalings for this are examined on the β plane using ideas from classical phenomenology. The scalings are naturally anisotropic and predict the formation of zonal flows directly through a turbulent cascade. Numerical simulations and two-point closure calculations qualitatively confirm the predictions. Also, simulations of barotropic flow on the β plane can produce zonal jet structures of exceptional persistence over many eddy turnover times.

Unsteady flow over topography generally produces a mean flow with a correlation between streamfunction and topography, with anticyclonic motion over humps. If the topography is shallow (or the flow sufficiently energetic) the mean streamfunction will be of a scale similar to that of the topography. For sufficiently steep topography, shelf (topographic Rossby) waves prevent the flow from achieving the scale of the topography, and may lead to the formation of jets parallel to but with a narrower scale than the topographic slope, superimposed on the topographically rectified mean flow. Such mechanisms can produce poleward undercurrents along oceanic boundaries and equatorward undercurrents on western boundaries. Similar mechanisms can produce alternating jets, superimposed on a mean vortex, around isolated seamounts.

1. Introduction

The production of anisotropic motion by differential rotation is a well-documented phenomenon, from theoretical, numerical, and observational standpoints (Rhines 1975; Williams 1978; others). Indeed it may be responsible for producing jets in planetary atmospheres, although the precise mechanism is not clear. A related effect, the generation of mean currents by flow over topography, is likely responsible for the less dramatic but equally interesting undercurrents in continental margins and around midocean topography; in this paper we shall see how conservation of potential vorticity leads in a simple fashion to such mean flow production. Further, particularly over steep topography such as ridges or continental slopes, the gradient of potential vorticity due to the topography can cause the mean flow produced to become jet-like. The mechanisms of these two classes of phenomena—namely, flow in differentially rotating medium and flow over topography—and their relationship to one another are the subject of this paper.

Potential vorticity gradients are, of course, ubiquitous and influential in ocean dynamics. Aside from the fact that the planetary vorticity gradient β determines the large-scale gyre structure, the “free” evolution of mesoscale eddies can hardly fail to be ultimately influenced by β because the cascade of energy to large scales will continue until the β–Rossby number is O(1). Even more pronounced will be the influence of topography since on continental slopes and over midocean ridges the potential vorticity variation produced by topographic effects can dominate the planetary contribution and lead to mean flows that can potentially transport water very large distances. The sign of the effect is such that in eastern boundary layers equatorial water is transported poleward, whereas in western boundary layers deep water can be transported equatorward. The bottom circulation can therefore be completely altered by such topographic effects and can oppose the sense of wind-driven upper-ocean circulation.

That turbulent flow on the β plane does lead to anisotropy and jet formation is well known. One candidate mechanism for the phenomena arises from a weakly nonlinear analysis of interacting Rossby waves (Rhines 1975). In this analysis, we note first the existence of a transition scale that may be taken to be \(O(U/\beta)\), where \(U\) is the rms velocity of the flow, although this particular form is not crucial to the argument. It is supposed that for scales smaller than this, turbulent effects will dominate; for larger scales, wave-like phenomena dominate. Within the wave regime, nonlinear transfer occurs through the interaction of a triad of resonant Rossby waves. The twin requirements of the production of low-frequency flow (by resonant interaction theory) and low wavenumbers (through the general tendency of geostrophic turbulence to seek the
gravest mode) lead to the generation of zonal structures, flows with a low value of the zonal wavenumber $k_z$. (The extension to the baroclinic case is straightforward: Vallis 1983.) The extreme case of beta-plane turbulence appears to be the production of intense and persistent zonal jets (Williams 1978; see also Maltrud and Vallis 1991). The scale of these jets has generally been assumed to be the transition scale, but with, as we shall see, little a priori justification. There are two immediate objections. First, Rossby waves provide no restriction on the scale of zonal flows, because the frequency of Rossby waves vanishes if $k_z = 0$. Second, zonal flows cannot in fact be produced by resonant interactions, because then the interaction coefficient vanishes. In this paper we describe another simpler mechanism for the production of zonal flow.

Topographic mean flows arise because conservation of potential vorticity leads to the production of a negative correlation (in the Northern Hemisphere) between vorticity and topography in a turbulent flow, and hence a mean flow. This simple but powerful mechanism may be the cause of various observed oceanic mean flows, in particular the poleward undercurrents ubiquitous in eastern boundary layers (Neshbya et al. 1989) and equatorward undercurrents in western boundary layers. The fact that the presence of topography implies a gradient of potential vorticity further suggests that a mechanism analogous to that producing jets in differentially rotating flow should concentrate the topographic mean flow into focused currents, and there is some observational evidence for this (Lynne and Simpson 1990).

The rest of this paper discusses the above issues in more detail, and offers a simple consistent mechanism for zonal flow production on the $\beta$ plane (or indeed in the presence of any mean potential vorticity gradient). In section 2 we provide some simple scaling arguments for the production of anisotropy. This is followed in section 3 by some numerical simulations of $\beta$-plane flow. We see that numerical simulations and a two-point statistical closure are generally consistent with the scalings, and that in the direct simulations the zonal jets produced have exceedingly long lifetimes. Indeed, they appear to be almost fixed in space, in spite of a domain and a forcing that are completely homogeneous. Section 4 contains descriptions of analogous simulations of flow over topography, and section 5 concludes.

2. Anisotropy on a $\beta$ plane: Scaling

In this section we present a number of scalings for a transition from turbulent behavior to wavelike behavior. First consider a scaling based on "classical" two-dimensional turbulence, essentially along the same lines as Kolmogorov's derivation of a dissipation scale. Let the flow be forced at a (high) wavenumber $k_f$. Energy tends to cascade to large scales and enstrophy to small scales, as illustrated in Fig. 1. Let the energy cascading rate to large scales be $\epsilon$ and the enstrophy cascading rate to small scales be $\eta$. Then in the inverse cascade regime the eddy turnover time is given by (to within a dimensionless constant)

$$\tau_e(k) = \epsilon^{-1/3}k^{-2/3}. \quad (2.1)$$

This is the only dimensionally correct scaling, assuming $\tau_e$ to be a function only of wavenumber and $\epsilon$. The inverse of $\tau_e$ defines a strain rate, or "turbulence frequency," namely,

$$\omega_T(k) = \epsilon^{1/3}k^{2/3}. \quad (2.2)$$

Now, the Rossby wave frequency is given by the well-known formula

$$\omega_R(k) = -\frac{\beta k_x}{k^2}. \quad (2.3)$$

The crossing between $\omega_T$ and $\omega_R$ is fairly sharp, as seen in Fig. 2. Thus we may be justified in supposing that for low wavenumbers Rossby waves dominate whereas for higher wavenumbers turbulence dominates, with a transition wavenumber

$$k_\beta = \left(\frac{\beta^3}{\epsilon}\right)^{1/5}, \quad (2.4)$$

obtained by equating (2.2) and (2.3), ignoring the anisotropy in (2.3). If instead of (2.2) we assume a tur-
Fig. 2. Various estimates of inverse time scales. Curve (a) sketches Rossby wave frequency \( \beta/k \) and curves (b), (c), and (d) are the estimates of inverse eddy turnover times, or "turbulent frequencies," given by \( \varepsilon^{1/3}k^{2/3} \), \( U \), and \( \xi \), respectively. The amplitude of each curve is arbitrary. The transition between waves and turbulence is, roughly, at the crossovers between (a) and (b), (c), or (d).

bulence dispersion relationship \( \omega_i = U \xi \), we recover the Rhines (1975) expression

\[
k_{\beta}^B = \sqrt{\beta/U}.
\]

(2.5)

Here, \( U \) may be interpreted as the root-mean-square velocity of the fluid. A factor of 2 sometimes finds itself, rather arbitrarily, in the denominator. If we use \( \omega_i = \xi \), where \( \xi \) is the root-mean-square vorticity, we obtain the expression of Holloway and Hendershott (1977); namely,

\[
k_{\beta}^H = \beta/\xi.
\]

(2.6)

The relationship between these expressions may be clarified by consideration of the strain rate (e.g., Kraichnan 1971)

\[
S(k) = \left[ \int_{k_0}^{k} k^2 E(k') dk' \right]^{1/2},
\]

(2.7)

where \( k_0 \) is the lowest energy-containing wavenumber. The eddy turnover time of (2.1) is the inverse of this. If the energy spectrum is given by the classical expression

\[
E(k) = C\varepsilon^{2/3}k^{-5/3},
\]

(2.8)

then substituting (2.8) in (2.7) leads to (2.1) and thence (2.4). If on the other hand most of the energy is contained at the largest scales, or suppose merely that the energy spectrum for \( k > k_0 \) falls off faster than \( k^{-3} \), then the integrand is dominated by contributions from small wavenumbers and has only a very weak wavenumber dependence. The strain rate is essentially con-

stant and equal to the square root of the total enstrophy of the flow, say \( \dot{Z} \). Equating this to the Rossby wave frequency to obtain a transition scale leads to \( k_{\beta} \sim \beta/\dot{Z}^{1/2} \), which is equivalent to (2.6). Finally, the expression (2.5) may be obtained if the local expression \( [k^3 E(k)]^{1/2} \) is used for the strain rate, and then let \( U = [k^3 E(k)]^{1/2} \) and equate this to the Rossby wave frequency. From a turbulence phenomenology point of view, this procedure has little justification, and, furthermore, the expression is not Galilean invariant. However, the fact that (2.5) is more readily evaluated than either (2.4) or (2.6) for typical oceanic flows, as well as the appeal of the simple expression \( U/k \) for the inverse turbulence time scale, makes (2.5) undoubtedly useful.

As well as a scale separation between wavelike activity and turbulent flow, the scalings in and of themselves predict the formation of anisotropy. Equating (2.2) and (2.3), that is, setting

\[
e^{1/2}k^{2/3} = -\frac{\beta k_x}{k^2}
\]

(2.9)

but retaining the anisotropy leads to the following expressions for the \( x \) and \( y \) components of the transition wavenumber:

\[
k_{xy} = \left( \frac{\beta^3}{\varepsilon} \right)^{1/5} \cos \frac{8}{5} \theta
\]

(2.10)

\[
k_{yx} = \left( \frac{\beta^3}{\varepsilon} \right)^{1/5} \sin \theta \cos \frac{3}{5} \theta.
\]

The polar coordinate is parameterized by the angle \( \theta = \tan^{-1}(k_y/k_x) \). The expressions (2.5) and (2.6) have similar anisotropic forms given by

\[
k_{\beta}^{x} = \left( \frac{\beta}{U} \right)^{1/2} \cos \frac{3}{2} \theta
\]

(2.11)

\[
k_{\beta}^{y} = \left( \frac{\beta}{U} \right)^{1/2} \cos \frac{1}{2} \theta \sin \theta,
\]

and

\[
k_{xy} = \frac{\beta}{\xi} \cos^2 \theta
\]

\[
k_{yx} = \frac{\beta}{\xi} \cos \theta \sin \theta,
\]

(2.12)

respectively.

These expressions are sketched in Fig. 3; they are qualitatively similar and all exhibit a characteristic dumbbell shape. Within the dumbbell, characteristic Rossby wave times are shorter than turbulent turnover times. This inhibits the transfer of energy from the turbulent regime because efficient forcing of a wavelike mode will be achieved only when the forcing frequency is comparable to its natural frequency. However, as energy cascades to larger scales the presence of Rossby
waves provides no direct inhibition to the formation of zonal flows—that is, flows with zero $x$ wavenumber. Indeed the tendency of the flow to seek the largest scale can best be achieved by the preferential excitation of zonally elongated structures, the ultimate limit of which are zonal jets extending the length of the domain. No interaction between Rossby waves is needed.

The meridional scale of such structures is not given by this simple theory. Certainly there is no a priori reason for the jet scale to be the beta scale, but nor is it realistic to expect the inverse cascade to continue to smaller and smaller $y$ wavenumber by an energy cascade along the $y$-wavenumber axis. Such a cascade will be quite inefficient since triad interactions of wavevectors in the turbulent regime will involve two almost parallel modes, resulting in a small interaction coefficient. For this reason the scale of any zonal jets can be expected to be of the order of the beta scale given by (2.4)-(2.6), but the scaling arguments do not give a precise quantitative estimate for the jet scale.

Finally, we wish to emphasize the qualitative similarity of the shapes given by the expressions given by (2.10)-(2.12). Even though there may be quantitative differences in derivation and result, the anisotropy in the beta term transcends the detailed differences in the treatments of the nonlinear term, giving a consistent overall picture of the mechanism of production of zonal flow.

3. Simulations on a $\beta$ plane

Are the scalings qualitatively or even quantitatively correct? Are the anisotropic structures predicted by the scaling actually produced in numerical simulations? Do simulations of anisotropic closure give similar results? This section addresses these questions.

a. Closure and inviscid simulations

In this subsection we perform inviscid integrations both by direct simulation and using a two-point spectral closure. The closure description and simulations may be skipped by those with little confidence in such matters, and the narrative picked up at the description of the direct simulations beginning with (3.5). For the closure model we use the eddy-damped, quasi-normal Markovian (EDQNM) scheme (Orszag 1974).

For two-dimensional inviscid turbulence, the closure evolution equation is

$$\frac{dY_k}{dt} = \sum_{k+p=q=0} \theta_{kpq} \omega_{kq} \{ Y_p Y_q - Y_q Y_p \}, \quad (3.1)$$

where

$$\omega_{kq} = \frac{2(p \times q)^2}{(k^2 p^2 q^2)} (p^2 - q^2)(k^2 - p^2) \quad (3.2)$$

and $Y_k = k^2 \langle |\psi_k|^2 \rangle$. The parameter $\theta_{kpe}$ is a phenomenological time scale, roughly the inverse of an eddy damping rate. It determines the nature of the turbulence. In pure (i.e., no waves) turbulence then a self-consistent choice is
\[ \theta_{k_p q} = \frac{1 - \exp[-t\mu_{k_p q}]}{(\mu_k + \mu_p + \mu_p)}, \]  
\[ (3.3) \]

where \( \mu_k = g(k^3 E(k))^{1/2} \) is an inverse eddy turnover time and \( \mu_{k_p q} = \mu_k + \mu_p + \mu_p \). The parameter \( g \) is an order one constant; we use \( g = 1 \).

The advantages of spectral closures of this type also perhaps more familiar real space closure are that the closure respects the inviscid invariants of energy and enstrophy and predicts the correct statistical equilibrium spectrum for unforced, inviscid flow (see also Carnevale et al. 1981). The closure does not, however, respect the conservation of vorticity on parcels and hence does not preserve arbitrary functions of vorticity.

In the presence of Rossby waves, the structure of the closure is unchanged, but the parameter \( \theta_{k_p q} \) becomes

\[ \theta_{k_p q} = \Re\frac{1 - \exp[-t(\mu_{k_p q} + i\omega_{k_p q})]}{\mu_{k_p q} + i\omega_{k_p q}}, \]  
\[ (3.4) \]

where \( \Re \) means the real part is to be taken, \( \omega_{k_p q} = \omega_k + \omega_p + \omega_q \) and \( \omega_k \) is the (unshifted) Rossby wave frequency (see Holloway and Hendershot 1977; Bartello and Holloway 1991). In a steady state this becomes

\[ \theta_{k_p q} = \Re(\mu_{k_p q} + i\omega_{k_p q})^{-1}. \]

The difference between this and (3.4) is actually rather small, except for very small times. Note that the time scales of the closure cannot be expected to quantitatively compare to those of the direct simulation because of the simplifications used to obtain (3.3) and (3.4) and the uncertainty of the parameter \( g \).

The parameter \( \theta_{k_p q} \) remains a symmetric function of all three members of a wavenumber triad, invariant to a permutation of its indices. If waves dominate over turbulence, the closure reduces to resonant interaction theory. Although no single spectrally local criterion for the dominance of waves over turbulence may be given, nonlinear interactions are inhibited when a triad contains members for which \( \omega_k \gg \mu_k \), a criterion similar to that producing (2.4). Because there will therefore be a range of wavenumbers over which the flow undergoes a transition from turbulence to waves, this implies a rather less sharp boundary between waves and turbulence than might be expected if the scalings of section 2 were interpreted too literally.

Ironically, fully anisotropic simulations of a spectral closure are more computationally intensive than are direct simulations because no efficient (e.g., transform) procedure exists for calculating the interaction coefficients. On the other hand, no ensemble of integrations is needed. The number of computations per time step increases as \( k_m^n \), where \( k_m \) is the truncation wavenumber. Still, we have performed integrations of (3.1) on a Cartesian grid with maximum wavenumber up to 32. Our goal is to see if the closure integrations and the direct simulations are able to produce transition regions similar to those of the scaling and illustrated in Fig. 3.

The direction simulations integrate the familiar barotropic vorticity equation

\[ \frac{\partial \zeta}{\partial t} + J(\psi, \zeta) + \beta \frac{\partial \psi}{\partial x} = 0, \]  
\[ (3.5) \]

where \( \zeta = \nabla^2 \psi \) is the vorticity and \( \psi \) is the stream-function. The code is spectral and dealiased and therefore exactly conserves energy and enstrophy to the accuracy of the time-stepping scheme, which is leapfrog. A weak Robert filter is used as necessary to control the computational mode. The initial conditions are isotropic, with energy concentrated in a ring around wavenumber 12. (Total energy = 1, total enstrophy = 290, \( \beta = 400 \).) Figures 4 and 5 illustrate the subsequent evolution of the energy spectrum. Almost instantaneously (in both closure and direct simulation) anisotropy can be detected, and the characteristic dumbbell signature of \( \beta \)-plane turbulence immediately becomes manifest and grows stronger. Energy within the wave regime remains small even as energy of the zonal flow is building up, and examination of the energy evolution of either closure or direct simulation indicates that the formation of zonal structures primarily proceeds by the direct cascade of energy into zonal flow, rather than a transfer to zonal modes from within the wave regime. Ultimately, the inviscid simulations will revert to isotropic flow because the maximum entropy state toward which the simulations evolve is determined solely by the isotropic invariants of energy and enstrophy (proof exists for the closure, see Carnevale et al. 1981). Because the statistically steady state does not therefore manifest any wave/turbulence boundary, we turn to forced-viscous simulations to see how the scales of anisotropy and transfer inhibition vary with \( \beta \).

b. Forced-viscous simulations

Forced, viscous simulations are expected to maintain an anisotropic statistical mean state even though the thermal equilibrium is isotropic. We have performed various experiments to examine both the scaling predictions and the formation of strong, persistent zonal jets. Explicitly the equation of motion is

\[ \frac{\partial \zeta}{\partial t} + J(\psi, \zeta) + \beta \frac{\partial \psi}{\partial x} = F - \kappa \zeta + D. \]  
\[ (3.6) \]

The forcing \( F \) is a random stirring at a high wavenumber, typically at about wavenumber 80 in a 256\(^2\) simulation (i.e., a simulation spectrally truncated at wavenumber 128 and therefore with approximately 256 equivalent grid points in either direction). Two frictional terms are necessary. The first, \( \kappa \zeta \), is scale independent and may physically be thought of as an Ekman drag; it is necessary to remove energy from a system. (If the effects of drag are artificially limited to a region of spectral space of lower wavenumber than \( k_p \) in an
attempt to obtain an inviscid crossover between inertial and wave effects, then typically energy simply piles up at the $\beta$ barrier and the simulation does not equilibrate, at least in simulations of hundreds of turnover times.) The second, $D$, is a hyperviscous enstrophy remover; its amplitude is chosen for each simulation to be the smallest necessary to prevent a buildup of enstrophy at the truncation wavenumber [see Maltrud and Vallis (1991) for details]. The initial condition for the following set of simulations is the end state of a steady $\beta = 0$ simulation.

The time-averaged energy spectra in the forced simulations also show clearly the characteristic figure eight or dumbbell signature (Fig. 6), implying that there is a cascade barrier at its boundary. Whether the precise barrier shape is given by (2.10), (2.11), or (2.12) is hard to determine, nor perhaps is it particularly important. The main point is that all give a qualitatively similar shape, with an enhanced zonal flow produced directly by the inverse turbulent cascade. If we define $k_g$ (as in Holloway and Hendershott 1977) as the lower boundary of a finite-width cascade barrier, then we should find that wavenumbers below $k_g$ all decrease in energy with respect to their initial state, since dissipation is still acting but turbulent transfers are slowed. For the purposes of comparison with the
scalings given by (2.4), (2.5), and (2.6), we obtain from the simulations an approximate value of $k_\beta$ where it intersects the $k_x$ axis by averaging the 2D spectrum over a small angle $\theta$ on each side of the $k_x$ axis, and following the time evolution (Fig. 7a). (The averaging is done purely to reduce the noise in the results.) The value of $k_\beta$ is then chosen to be the location where the spectrum starts becoming steeper and steeper with time. Results are plotted in Fig. 8.

It is also instructive to determine how the scale of the zonal jets ($k_{\text{jet}}$) varies with $\beta$. Given a scaling for $k_\beta$, can we predict the scale at which jets will form? Since there is no direct impedance to the cascade on the $k_y$ axis, it can be expected that $k_{\text{jet}}$ will be smaller than $k_\beta$. The jet scale is seen as the maximum in the 2D energy spectrum (Fig. 6) and also dominates the isotropic spectrum (Fig. 7b). We find that both $k_\beta$ and $k_{\text{jet}}$ clearly increase with $\beta$, with $k_{\text{jet}}$ being the smaller of the two as expected (Fig. 8). It appears that all of the analytical scalings tend to overpredict the values of $k_\beta$ and $k_{\text{jet}}$ for most of the values of $\beta$ chosen. Comparison of the functional form of the simulation curves with the scalings is not definitive, although it does seem that in these simulations $k_\beta$ increases less than linearly with $\beta$. Because of the presence of dissipation in these simulations, comparison with an inviscid criterion for

Fig. 5. Similar to Fig. 4 but now calculated using a small ensemble of direct simulations (six) of the equations of motion.
the value of $k_\theta$ may be quantitatively inaccurate, although the direct affect of dissipation does appear relatively small if one considers the time scales involved (Fig. 9).

Examination of the 2D energy spectra typically reveals a peak at $k_x = 0$, implying strong zonal motions. Figure 10a shows a plot of time versus latitude of the zonally averaged zonal velocity from a simulation.

**Fig. 7.** (a) The truncated angular average of the 2D spectrum from the same simulation seen in Fig. 6, with the average taken over an angle of $\theta = \pi/6$ on either side of the $k_x$ axis. The solid curve is the initial condition taken from an identical simulation except that $\beta = 0$. The dashed curves are taken from two equally spaced subsequent times in the evolution of the flow. The cascade barrier is seen to be at approximately $k = 13$. (b) Angular average of the 2D spectrum seen in Fig. 6 showing the location of $k_{jet}$ to be approximately at $k = 10$. 
forced at quite high wavenumber, typically close to 80. The production of zonal jets was generally enhanced by a scale separation between the forcing and the \( \beta \) scale. Forcing at lower wavenumbers of the same order as \( k_B \) appears to produce somewhat less persistent structures due to direct phase scrambling from the forcing, but the structures are still very long lived. The latitude of the jets is remarkably persistent. In all such simulations we have performed, the time-averaged flow appears to be stable, using the criterion that \( \beta - \frac{\partial^2 U}{\partial y^2} \) remain one-signed (Kuo 1949), where \( U \) is the zonally averaged zonal velocity (Fig. 10b).

Also of interest in Fig. 10b is the clear difference in the shape of the jets depending on their direction, with the eastward jets becoming much sharper and narrower (i.e., they have a larger magnitude of \( \partial^2 U / \partial y^2 \)) than the westward jets. This difference is consistent with the Kuo stability criterion: westward jets must remain sufficiently broad because \( \partial^2 U / \partial y^2 \) is bounded by \( \beta \). No such bound exists for eastward currents, however, since they will remain stable no how large the magnitude of \( \partial^2 U / \partial y^2 \) becomes. This difference may also be explained in terms of a hydraulic theory of zonal jets on the \( \beta \) plane (Armi 1989). The tendency for a zonal current to either broaden or narrow depends on the value of the Rossby–Froude number,

\[
R_{OF} = \frac{u_m}{\alpha' \beta a^2},
\]

where \( u_m \) is the maximum speed in the jet, \( a \) is the half-width of the jet, and \( \alpha' \) is a constant that depends surprisingly weakly on the jet profile. According to this theory, westward jets will always tend to be broad, independent of the value of \( R_{OF} \). Eastward jets will also be broad if \( R_{OF} > 1 \), but will narrow if \( R_{OF} < 1 \). In our simulations, \( R_{OF} \) is always less than unity. If we assume that the scale of the jets is directly related to \( k_B \), that is, \( k_{jet} = \gamma k_B \), where \( \gamma \) is an order one (or smaller) constant, then the width of a typical jet is simply half a wavelength, \( \pi / \gamma k_B \). Thus we get

\[
R_{OF} = \frac{4\gamma^2 k_B^2 u_m}{\alpha' \beta \pi^2}.
\]

Now, if we assume that \( k_B \) is given by the scaling (2.5) and that \( U = \sqrt{\pi u_m} \) and \( \alpha' = 0.24 \) (the latter two come from assuming a sinusoidal velocity profile), we find that \( R_{OF} \) does not depend directly on \( \beta \) or the energy of the flow, but is simply a constant that is less than unity for \( \gamma < 1 \),

\[
R_{OF} = \frac{8.3\gamma^2}{\pi^{3/2}},
\]

thus implying that eastward jets will tend to be rela-
tively narrow. The assumption that $k_t$ scales as $\sqrt{\beta}$ (which seems to agree fairly well with simulations) is necessary to derive this particular result; in general we may expect (3.9) to have a weak beta dependence. The assumption about the velocity profile simply changes the value of the constant by a small amount.

4. Mean flow generation by topographic interaction

In this section we discuss the interaction of eddy motion with topography, typically of a larger scale than the energy-containing eddies, while restricting ourselves to flow on the $f$ plane. There are two rather distinct effects we wish to describe. The first simply pertains to the initial generation and control of mean flows by topography through vortex stretching, or the conservation of potential vorticity. This is the subject of the next subsection. The second effect (subsection 4b) pertains to the formation of jets over topography, arising because of the presence of a gradient of potential vorticity.

\textbf{a. Elementary theoretical considerations and flow over shallow topography}

The quasigeostrophic equation

$$\frac{Dq}{Dt} = 0,$$

(4.1)

where $q = \nabla^2 \psi + fh'/H$, or the more primitive shallow water version

$$\frac{D}{Dt} \frac{\xi + f}{h} = 0$$

(4.2)

both express the conservation of a form of potential vorticity. Here, $h$ is the total height of the water column, $H$ its mean height, and $h'$ the height of the topography. If a column of water is passively moved over topography, then all other things being equal, the relative vorticity must compensate for the change in height of the water column, leading to anticyclonic motion over humps or cyclonic flow over valleys. (In the rest of the paper our notation will parochially apply only to
Northern Hemisphere parity, in which anticyclonic motion corresponds to negative relative vorticity). The mechanism whereby this occurs is (relative) vortex stretching by the planetary vorticity. Thus, we may expect a negative correlation between relative vorticity and topography. The production of this correlation depends on the “passive” nature of the flow over topography. That is, we imagine a fluid column being moved up or down a topographic slope more-or-less randomly by the action of the velocity field produced by other eddies. On the other hand, if a single, isolated vortex interacts with a topographic feature, then a positive vortex will in fact “climb” a hill in a spiraling motion through the propensity of a positive vortex to self-ad- vector in a pseudonorthwesterly direction (Carnevale et al. 1988). Even though the vortex weakens as it boot- straps its way up the hill, this effect leads to a positive correlation between vorticity and topography. This mechanism will tend to dominate if the vortex is iso- lated from other eddies. In a vigorous eddy field we should expect the former mechanism to dominate and a negative correlation to arise between vorticity and topography.

There exist two theories that quantify the correlation in special circumstances—“maximum entropy” and “minimum enstrophy.” Neither of these build in the material conservation of $q$ on parcels, only being aware of the inviscid conservation of quadratic invariants. Thus, both theories can be expected to fail if higher-order invariants are indeed important. However, the conservation of the quadratic invariants alone is at least able to lead to vorticity–topography correlations. The maximum entropy theory (Salmon et al. 1976) strictly applies only to an unforced inviscid fluid. If we pop- ulate a topographic domain with a field of eddies, then the time-averaged flow can be predicted by statistical mechanical arguments to be

$$\langle \psi_k \rangle = \frac{-h_k}{a + bk^2}, \quad (4.3)$$

where $a$ and $b$ are parameters determined by the initial values of energy and enstrophy.

Minimum enstrophy, on the other hand, assumes that in real geostrophic turbulence enstrophy decays much faster than energy owing to the cascade of enstrophy to small scales (where it may be dissipated) and the trapping of energy at large scales. The time-averaged state should then be approximately given by a minimum enstrophy state for a given energy. A variational calculation then yields a time-averaged flow of form similar to (4.3). One might seek to justify minimum enstrophy by the following argument: The non- linear interactions always act to try to maximize en- strophy, and so the sense of nonlinear transfer of enstrophy will be the same in viscous flows as in inviscid flows. Hence, there will be a cascade of enstrophy to small scales, where it may be dissipated, producing a tendency toward a minimum enstrophy state.

Both of these mechanisms seek to quantify the time mean state produced by the interaction of an eddy field with topography. (The underlying mechanism produc- ing the flow–topography correlation is of course the conservation of potential vorticity as a parcel is passively advected over topography.) In forced dissipative situations, neither theory can be expected to quantitatively predict the correct time mean flow. However, this superficial failing should not of itself concern us, and in any case a closure theory of flow over topography, which contains these tendencies and is valid in more general circumstances, can be con- structed (e.g., Herring 1977). But in the next subsection we show how another mechanism arises for flow over steep topography, which qualitatively affects the scale of the mean flow and its correlation with topography.

For “shallow” topography, in a sense to be subsequent quantified, the sense and scale of the mean flow in both forced-dissipative and freely decaying flow can be predicted following these ideas. We have performed a number of numerical integrations of both decaying and forced-dissipative flow over fairly idealized topography, considering first the simple case of barotropic eddy activity over a single ridge. The decay experiments involve the numerical integration of the barotropic quasigeostrophic equations from random initial conditions, with a viscous term that acts pri- marily on the higher wavenumbers. In the forced-dis- sipative cases, forcing is typically white noise, concen- trated between wavenumbers 10 and 14, and a linear, scale-independent Ekman drag is introduced to remove energy at the largest scales. The scale separation be- tween forcing and topography mimics the scale sepa- ration between deformation radius and large topo- graphic features in the ocean. Results for both forced- dissipative and decay experiments demonstrate the same qualitative mechanism, and we will only present the results from the forced simulations.

For small to moderate values of the topographic slopes the mean streamfunction is closely correlated with the topography. With a meridional ridge, a predominantly northward flow is produced on the western slope (Fig. 11a) and southward flow on the eastern slope. (The orientation of a ridge is irrelevant on an $f$ plane, so more generally the flow on either slope is pseudowestward; that is, facing downstream, higher values of mean potential vorticity are on the right.) There is no mean cross-slope flow, so that in this sit- uation the imposition of a wall with free-slip boundary conditions along the ridge would make little essential difference, except possibly in how the global momentum invariants are maintained. It is well documented that in eastern boundary currents of ocean gyres pole- ward flowing undercurrents (Neshyba et al. 1989) fre- quently occur. These regions are also generally quite
Fig. 11. Maps of the time-averaged N-S velocity along a simple meridional ridge on the $f$-plane for various heights of topography in a barotropic forced-dissipative simulation, forced near wavenumber 12. The topography is proportional to $-\cos x$, where $x$ is the E–W coordinate, so it is peaked at the center line, with no $y$ variation. Dark shades indicate positive velocities with black being highest positive values and white the highest negative values. Each map has an accompanying plot of the along-topography ($y$) average velocity as a function of cross-topography coordinate $x$. (a) The amplitude of the topography is $h = 20$. Note the flow has approximately the same scale as the topography. (b) $h = 200$. Now, jets appear at $k = 3$ superimposed on the broad background flow seen in (a). The jet peak at $x/(2\pi) = 0.65$ is strong enough to reverse the background flow. (c) $h = 1000$. Jets appear at $k = 12$ superimposed on the background flow.
baroclinically unstable, suggesting that a possible mechanism for the poleward flow is the production of a mean flow by the above eddy-topographic interaction. Note that no smaller-scale or "rough" topography is actually necessary to produce a mean flow in the first instance, although its importance cannot be ruled out (Holloway 1987; Brink 1986; Haidvogel and Brink 1986).

Is such an effect large enough to produce noticeable currents in the ocean? Consider the California Current, with a slope of a few hundred kilometers extent. Then moving a zero vorticity parcel halfway up the slope produces a relative vorticity of approximately $\xi \sim f h / H \sim 0.5 f$. The magnitude of the poleward velocity produced by this is approximately $V \sim L 0.5 f \approx 5$ m s$^{-1}$, for $L \sim 100$ km where $L$ is the cross-slope scale of the motion. This conservative and very rough estimate is obviously far greater than is either observed or can be produced on energetic considerations, and the consequent restoring force is responsible for topographic Rossby waves. Nevertheless, it indicates the potential of the mechanism and that the magnitude of the mean flow is likely to be a stronger function of the eddy intensity than the topography itself.

b. Flow over steeper topography

If the amplitude of the topography is too large for a parcel to traverse it with the energy available to it, we may expect the mean flow to still be oriented parallel to topographic contours, but to become organized into bands. Suppose for simplicity that the topography is a simple sinusoid. Then both the maximum entropy and minimum enstrophy theories predict a resulting flow of similar scale. However, any flow in which the streamlines are parallel to the topography contours is a (inviscid) solution. Indeed as the topography gets larger, the mean flow does not maintain the same scale as the topography, but rather multiple jetlike structures form, superimposed on the broad background flow (Fig. 11b).

More formally, the mechanism is just that which causes jets on the $\beta$ plane: the addition of topography to the barotropic quasigeostrophic equation is identical to the addition of a beta effect (possibly spatially dependent) and if the horizontal variation of the topography is fairly uniform the effects are identical. We therefore expect alternating jets to form along the topography, with a scale approximately given by the wavenumber

$$k_h = \frac{\hat{h} x}{L},$$

(4.4)

where $\hat{h} = fh/H$, or any of the analogous forms given in section 2, with the topographic variation playing the role of the planetary gradient of potential vorticity.

Figure 11 indicates how the banded structure becomes enhanced as either the topography gets bigger or the eddy field weaker. If the scale given by the inverse of (4.4) is smaller than the scale of the topography, the qualitative form of neither the maximum entropy nor the minimum enstrophy state can be achieved. (Since the flow is viscous, this does not demonstrate "nonergodicity" in the sense of Shepherd (1987), although it may be loosely analogous to it.)

As expected with the analogy to the $\beta$-plane case, the scale of the jets becomes smaller as the height of the topography is increased (Fig. 11c). Due to the fairly low resolution of this simulation ($64^2$), the strength of the jets in Fig. 11c is being limited by friction. As we increase the resolution (Fig. 12), the jets form at a scale that is largely unaffected by the direct action of viscous diffusion. The jet persistence is also enhanced somewhat (as in the case of the $\beta$ plane) if the scale of the forcing is much smaller than the scale of the jets. One aspect of the topographic flows seen to a greater extent here than in the $\beta$-plane flows is some degree of latitudinal migration of the jets. This is especially the case in the vicinity of the peaks and valleys where the topographic height varies the slowest, that is, where the topographic $\beta$ effect is small, and where the jets may lose their coherence and disappear. On the $\beta$ plane, of course, the flow is degenerate in the sense that there can be no correlation between latitude (i.e., value of ambient potential vorticity) and the local value of relative vorticity, by homogeneity, unless the initial conditions are remembered. This symmetry is lost in the topographic case, which accounts for the preferential broad (i.e., cross-topographic scale) pseudo westward
flow on either side of the ridge. However, once formed, the jets on the $\beta$ plane are, in fact, rather persistent.

c. Flow around seamounts and other irregular topography

A similar mechanism can produce alternating flow around more complicated topography, such as seamounts or ridges that are not straight. Again both forced and decaying experiments have been performed for these cases. For smaller values of the topography the mean flow is able to achieve the grave scale of the topography. For larger values the flow again shows a banded structure. In the case of flow around a circular hump, a map of the magnitude of the time-averaged velocity (Fig. 13) shows that circular bands of alternating direction are superimposed on the large-scale flow, which may result in reversals of the flow direction. In a similar manner, flow along an S-shaped ridge will either broadly follow the topography or form bands parallel to the topographic contours (Fig. 14a). Figure 14b shows the velocity vector field and clearly indicates that the superimposed jets may become strong enough to cause flow reversals.

d. Internally generated instability and deep countercurrents in baroclinic flows

Here we show that in a slightly more realistic situation the interaction of an internally generated insta-

---

**Fig. 13.** Map of the magnitude of the time-averaged velocity around a circular seamount that is Gaussian in cross section. Black indicates highest speed. Variations in shade reflect the presence of alternating circular jets superimposed on the background flow, either slowing it down (lighter shade) or speeding it up (darker shade).

**Fig. 14.** (a) Map of the time-averaged meridional velocity along an S-shaped ridge that is sinusoidal in cross section. Black indicates highest positive speed; white indicates highest negative speed. The jets are superimposed on the large-scale background flow and follow the topographic contours. (b) Velocity vectors of the flow seen in (a). One of the jets is strong enough to create a clear reversal in the background flow.
from a northward flowing countercurrent along the slope. The resulting mean shear, which is primarily due to the southward upper-level flow over more or less quiescent deeper layers, is baroclinically unstable, resulting in a sea of eddies that interact with each other, the topography, and the mean flow. The collection of mechanisms maintaining the upper-level flow, and hence the mean shear, is not immediately relevant to the interaction of the eddies with the topography, and so a simple model of such a situation is to consider the interaction of a uniform shear over a meridional ridge. In the absence of a full ocean model (and the host of other phenomena obscuring the central issue which would entail) there is no fully self-consistent way to determine the upper-level flow, so this is best imposed. Similar arguments apply to flow over the Mid-Atlantic Ridge or western boundary currents. The shear is chosen to be baroclinically unstable, thereby generating eddies that may interact with the topography. Our goal is to see if this may produce a countercurrent in the lower layer.

We use a two-layer model. Let the mean upper-layer flow be $U$ and the lower-layer flow be zero. Then the quasigeostrophic potential vorticity equations for two layers (of equal thickness) may be written

$$\frac{\partial q_1}{\partial t} + J(\psi_1, q_1) + U\frac{\partial}{\partial x} \nabla^2 \psi_1 + \frac{1}{2} \lambda^2 U \frac{\partial \psi_1}{\partial x} = D_1,$$

$$\frac{\partial q_2}{\partial t} + J(\psi_2, q_2) - \lambda^2 U \frac{\partial \psi_2}{\partial x} = D_2,$$  

(4.5a)  

(4.5b)

where $q_1 = \nabla^2 \psi_1 + \lambda^2 / 2 (\psi_2 - \psi_1)$, $q_2 = \nabla^2 \psi_2 + \lambda^2 / 2 (\psi_1 - \psi_2) + h$ are the potential vorticities in each layer of the relative flow. (On the $f$ plane the $x$ and $y$ coordinates are interchangeable.) Here, $\lambda$ is an inverse deformation radius and $h$ the bottom topography. The terms $D_1$ and $D_2$ represent frictional effects; we use a biharmonic and an Ekman friction in the lower layer to absorb enstrophy and energy, respectively. Explicitly $D_1 = -\nu \nabla^4 \xi$ and $D_2 = -\nu \nabla^4 \xi - k \xi$. Typical midlatitude values for the first deformation radius are typically of order 30–50 km, and the cross-slope scale of the topography is typically an order of magnitude larger. We choose the value of the inverse deformation radius to be wavenumber 30 to reflect such a scale separation, and the topography to be a simple ridge oriented parallel to the mean flow. With this orientation of slope, the mean flow does not directly interact with the topography.

In the absence of beta, topography, or friction, any imposed shear is linearly unstable, although the topography acts, like beta, to stabilize the shear. We impose a linearly unstable shear, and integrate the model sufficiently long to obtain a statistically steady state with stable means. Figure 15 shows typical instantaneous and time-averaged streamfunctions, and in Fig. 16 we plot the mean upper- and lower-level alongslope velocity profiles for a selection of topographies and imposed shears. Along both sides of the ridge rectified mean current is produced. The sense of this flow is pseudowestward, just as in the barotropic experiments. This translates to poleward flow along western continental slopes or the western sides of meridional ridges.

Fig. 15. Typical instantaneous and time-averaged streamfunctions in an unstable two-layer simulation of flow along a ridge. The topography is a single sinusoid in the zonal ($x$) direction, with no meridional variation.
and equatorward flow along eastern continental slopes. The rectified current is largely barotropic, because it is of a much larger scale than the baroclinic instability scale and energy transfer to large scales tends to be barotropic.

For large topographic slopes the countercurrents become organized into a banded structure, just as in β-plane turbulence and may dominate the sense of the mean flow. Rough estimates may be obtained for the width of such jets using the scalings of section 2. If \( U \sim 10 \text{ cm s}^{-1} \) and the topographic slope is \( \sim 10 \text{ m km}^{-1} \) then the expected width of the countercurrent will be of order 1–10 km. This is a rather rough estimate, but it indicates that the topographic slope is sufficiently large that we do not expect a countercurrent generated by eddy–topographic interactions to be as large as the scale of the slope itself. The presence of multiple alongslope currents is the extreme limit of this, for which there is some observational evidence (see various articles in Neshyba et al. 1989). More observations, either in continental slope or midocean ridge regions, would obviously be useful.

5. Discussion

Scaling relationships and certain theoretical arguments indicate that the inverse energy cascade of geostrophic turbulence will be inhibited by a mean gradient of potential vorticity. Irrespective of the details of any particular scaling theory, the transition wavenumber between “waves” and “turbulence” is quite anisotropic: a characteristic dumbbell shape (in spectral space) is predicted (Fig. 3) into which energy transfer is inhibited, and this feature is robustly produced both by direct numerical simulation and two-point closure. There is no inhibition on the formation of zonal scales, essentially because Rossby waves have zero frequency for such structures. Indeed, since energy can penetrate only slowly into the wave regime, zonal and near-zonal flow will be preferentially excited as energy cascades to the gravest scales. The production of zonal jets can therefore result as a direct consequence of an anisotropic turbulent inverse energy cascade.

If the forcing scale is spectrally removed from (i.e., at a smaller scale than) the β scale, rather intense and very persistent zonal jets can be produced. These last many eddy turnover times, and once formed appear to be almost fixed in space, barely drifting in the meridional direction. The scale and strength of these jets appears to be such as to keep the sign of \( q_y = \beta - \partial^2 u / \partial y^2 \) positive, and hence the jets are, by this criterion, stable. Eastward jets are noticeably sharper than westward jets, consistent with the stability criterion and with a hydraulic theory of β-plane jets.

The quasi-passive advection of vorticity over topography will lead to a negative correlation between vorticity and topography. Over ridges or continental slopes this results in a pseudo westward mean flow; that is, poleward (equatorward) mean flows on the western (eastern) sides of meridional ridges. One may conjecture that this is the cause of the almost ubiquitous poleward undercurrents in eastern boundary currents. If the topography is sufficiently steep, then a second effect becomes noticeable, namely, the concentration of the mean flow into narrow currents. This occurs through the topographic β effect—just as the more familiar β effect due to differential rotation produces zonal jets. Possible locations for such phenomena are on continental slopes and midocean ridges, and it would be interesting to look for the presence of mean, and possibly multiple, currents flowing more or less parallel to the topography in regions of eddy activity.

Regarding the direct application of these ideas to oceanic flows, let us make the following points. First, the use of quasigeostrophy is formally valid only for shallow topography, clearly not satisfied on continental
slopes. A tractable theoretical framework including order-one topography may be difficult to construct, but numerical simulations with a primitive equation or intermediate model are possible. Second, oceanic boundaries are somewhat different from ridges because of the strict condition of no normal flow at the boundary. Finally, attempts should be made to determine how this type of theory of mean flow over slopes can be observationally distinguished from the qualitatively different type of model described by McCready (1981).

Acknowledgments. This work was supported by the ONR (N00014-90-J-1618) and the NSF (ATM 8914004). Comments from B. Cushman-Roisin, G. Holloway, T. Shepherd, and a reviewer were most helpful in revising the manuscript.

REFERENCES


