 Remarks on Charney’s Note on Geostrophic Turbulence  

KA KIT TUNG  
Department of Applied Mathematics, University of Washington, Seattle, Washington

WENDELL T. WELCH*  
Department of Geology and Geophysics, Yale University, New Haven, Connecticut

11 February 2000 and 18 January 2001

ABSTRACT
Charney in 1971 generalized results for two-dimensional (2D) turbulence to quasigeostrophic (QG) turbulence and obtained two results that have important implications for the atmosphere. The first is an attempt to prove that, similar to 2D turbulence, energy in QG turbulence goes only upscale in the net. The second is a demonstration that 3D QG motion in terms of a 3D wavenumber in a stretched coordinate is isomorphic to 2D turbulence. Charney’s proofs are shown here to be problematic.

1. Introduction
Charney’s (1971) note, “Geostrophic Turbulence,” is generally credited as laying the foundation for the subject. Although three-dimensional (3D) in nature, largescale motion in the atmosphere and oceans satisfying geostrophic scaling was shown to have more in common with the 2D turbulence of Kraichnan (1967) than with the 3D turbulence of Kolmogorov (1941a,b). As Charney (1971) demonstrated, the existence of a scalar invariant, the “pseudo-potential vorticity,” in addition to the energy invariant, provides a powerful constraint on energy transfers in quasigeostrophic (QG) turbulence, which is absent in 3D turbulence.

Charney’s work was probably motivated by the observation available at the time (e.g., Wiin-Nielsen 1967), which showed an apparent $k^{-3}$ power-law behavior in the energy spectrum for horizontal wavenumbers $k$ in the synoptic scales (zonal wavenumbers 7–18), and its similarity to the $k^{-3}$ spectrum predicted by Kraichnan (1967) for 2D turbulence for wavenumbers higher than the excitation wavenumber. Charney’s note contains two main results:

- It attempts to prove that energy cascades upscale in the net in QG turbulence, similar to 2D turbulence.
- There is a demonstration of isomorphism between QG and 2D turbulence, and consequently the observed $k^{-3}$ spectrum over the synoptic scales was explained using Kraichnan’s (1967) theory on isotropic and homogeneous 2D turbulence.

Both of these results contain major flaws, mathematical in nature in result 1 and quantitative in result 2. Although parts of the problem we will discuss may be known to some—for example, Merilees and Warn (1975) pointed out that the result of Fjortoft (1953) on the direction of energy cascade in 2D turbulence, on which Charney relied, was in error—the implications for large-scale atmospheric turbulence probably have not been fully appreciated. Unproven “folklore” in 2D turbulence concerning the direction of energy cascades [see comments by Eyink (1996)] are often carried over to QG turbulence in the atmosphere without further proof. In this note, we point out some of the problem areas.

2. Mathematical aspects of Charney’s proof
Large-scale atmospheric flows satisfying QG scaling conserve what Charney (1971) called “pseudo-potential vorticity” (which we will call “potential vorticity”),

$$q = \nabla \hat{\psi} + \frac{f_{0}}{\rho} \left( \frac{\rho}{N^2} \psi \right) + \beta y,$$


© 2001 American Meteorological Society
where $\psi$ is the streamfunction, $\bar{\rho}(z)$ is the mean density, $f$ and $\beta$ are Coriolis parameters, and $N$ is the Brunt–Väisälä frequency. Its conservation equation takes the form

$$\frac{\partial}{\partial t}q + J[\psi, q] = 0. \quad (1)$$

It is seen that (1) is analogous to the 2D vorticity equation

$$\frac{\partial}{\partial t} \nabla_h \psi + J[\psi, \nabla_h \psi] = 0, \quad (2)$$

and so one may expect that QG flows have behaviors analogous to 2D flows.

Charney (1971) obtained an energy equation by multiplying Eq. (1) by $-\bar{\rho} \psi$ and integrating over $x, y$, and $z$ (The range of $z$ is semi-infinite, from 0 to $\infty$; the $x$ domain is periodic; the solution vanishes at two points in $y$):

$$\frac{d}{dt} E = \frac{d}{dt} \int \int \int \left[\nabla_h \psi \cdot \nabla_h \psi + \frac{f^2}{N^2} \psi^2 \right] \bar{\rho} \, dx \, dy \, dz$$

$$= \int \int \frac{f^2}{N^2} \nabla \psi \cdot \nabla \psi \bar{\rho} \, dx \, dy = 0. \quad (3)$$

In arriving at energy conservation, Eq. (3), Charney used the vertical boundary conditions

$$\bar{\rho} \psi \vert_{-\infty} \to 0 \text{ as } z \to \infty \quad \text{and} \quad \psi \vert_{z=0} = 0 \text{ at } z = 0. \quad (4)$$

The upper-boundary condition in (4) equivalent to the assumption that the energy density vanishes at infinity. Without it there could be energy leakage to infinity and so there would not be energy conservation.$^1$

Charney defined a 3D elliptic operator $L$ by

$$L(\psi) = \nabla_h \psi + \frac{f^2}{\bar{\rho} N^2} [\psi_z],$$

and obtained the equation for the conservation of potential entrophy by multiplying Eq. (1) by $\bar{\rho} L(\psi)$ and integrating over all $x, y,$ and $z$ [although the $z$ integration is not necessary (see Salmon 1998)]:

$$\frac{d}{dt} F = \frac{d}{dt} \int \int \left[ L(\psi)^2 \bar{\rho} \right] dx \, dy \, dz = 0. \quad (5)$$

Using the two derived conservation laws, (3) and (5), Charney then proceeded to derive the result that energy flows upscale. His (excerpted) argument follows:

"'Now $L$ is a self-adjoint elliptic operator with a complete orthonormal set of eigenfunctions $\psi_m$ and eigenvalues $\lambda_m (m = 1, 2, \ldots)$ . . . . ."

"By virtue of the completeness property, we may set

$$\psi = \sum_{m} a_m \psi_m,$$

where

$$L(\psi_m) = -\lambda_m \psi_m. \quad (6)$$

"Substituting . . . we obtain

$$2E = \sum_{m} \lambda_m a_m^2 = \sum_{m} b_m = \text{constant}$$

$$2F = \sum_{m} \lambda_m a_m^2 = \sum_{m} \lambda_m b_m = \text{constant}$$

"It then follows that

$$\sum_{m} b_m < \frac{1}{\lambda_m} \sum_{m} \lambda_m b_m < \frac{2F}{\lambda_m},$$

i.e., that $\sum_{m} b_m$ approaches zero with increasing $M$, and an energy cascade is impossible. All the other theorems pertaining to energy exchange among spectral components in two-dimensional flow may now be shown to apply to three dimensional quasigeostrophic flow as well . . . . ."

There are several problems with Charney’s proof. We will discuss the first in this section. It turns out that there is only one eigenfunction satisfying (6) and the boundary conditions (4). That eigenfunction is $\psi_m = 0$; that is, it is a strictly 2D flow. Consequently, Charney did not prove anything more than what had already been shown for 2D flows by Fjortoft (1953). We can demonstrate this by explicitly solving Eq. (6) via separation of variables. Write

$$\psi_m(x, y, z) = \phi_m(z) e^{i k_x x + i k_y y}$$

and substitute into Eq. (6) to yield (with $k^2 = k_x^2 + k_y^2$)

$$f_0^2 \frac{d}{dz} \left( \bar{\rho}(z) \frac{d}{dz} \phi_m \right) + \bar{\rho}(z)(\lambda_m - k^2)\phi_m = 0. \quad (9)$$

For $N^2$ constant (assumed later by Charney) and $\bar{\rho}(z) = \bar{\rho}(0)e^{-zh}$, Eq. (9) can be solved explicitly as

$$\phi_m(z) = [A_m e^{-\alpha_m z} + B_m e^{\alpha_m z}] e^{i2Hz},$$

where

$$\alpha_m^2 = \frac{N^2}{f_0^2} (k^2 - \lambda_m) + \frac{1}{4H^2}.$$
claimed. It is of the Sturm–Liouville form, but it is not a regular Sturm–Liouville system because the domain is semi-infinite, and one of the coefficients vanishes at one of the end points (viz. \( z = \infty \)) (see, e.g., Birkhoff and Rota 1969). The upper boundary, \( z = \infty \), is a singular point of Eq. (9). The reason that Charney’s system has only one eigenvalue is because his boundary condition (4) at the singular point overspecifies the problem. There should have also been a continuous spectrum

\[
k^2 + \frac{f_0^2}{4H^2N^2} < \lambda_m < \infty
\]

corresponding to the eigenfunctions

\[
\phi_m(z) = C_m e^{\frac{2\pi m}{D}z}
\]

where

\[
k^2 = \frac{N^2}{f_0^2} (\lambda_m - k^2) - \frac{1}{4H^2}.
\]

These eigenfunctions satisfy the less restrictive upper condition

\[
\overline{p} \phi_m(z) \text{ bounded as } z \to \infty.
\]

This set was eliminated by Charney’s boundary condition (4).

The mathematical situation is analogous to the Legendre functions, governed by the following singular Sturm–Liouville system:

\[
\frac{d}{dx} \left( 1 - x^2 \frac{d}{dx} \right) \phi(x) + \lambda \phi(x) = 0, \quad -1 < x < 1,
\]

subject to the proper boundary condition

\[
\phi(x) \text{ bounded as } x \to \pm 1.
\]

The eigenvalues are \( \lambda = m(m + 1) \), \( m = 0, 1, 2, \ldots \) and the complete set of orthonormal eigenfunctions are \( P_m(x) \), where

\[
P_0(x) = 1, \quad P_1(x) = x, \quad P_2(x) = \frac{3}{2} x^2 - \frac{1}{2},
\]

\[
P_3(x) = \frac{5}{2} x^3 - \frac{3}{2} x, \quad \text{etc.}
\]

If, instead, the more restrictive boundary condition

\[
\frac{d}{dx} \phi(x) = 0 \quad \text{at } x = \pm 1
\]

is used, all but one of the eigenfunctions are eliminated. The only eigenfunction that remains is the “barotropic” solution

\[
P_0(x) = 1.
\]

Given this problem, is there any way to rectify Charney’s solution? A complete set of eigenfunctions could in theory be constructed by adding the continuous eigenvalues and eigenfunctions to the barotropic mode. However, these radiating solutions do not satisfy energy conservation. Without energy conservation [Eq. (3)] Charney’s proof also fails. There does not appear to be a way to remedy the problem for the semi-infinite domain.

In oceanic applications, the vertical domain is bounded and a rigid-lid upper boundary condition can be imposed. For the atmosphere, such a lid can possibly be justified for motion trapped in the troposphere, for which case the height of the lid, \( D \), is to be the height of the tropopause. (It is not physically justified in the atmospheric case to take the limit \( D \to \infty \), because the upper stratosphere and mesosphere are very dissipative due to the presence of breaking planetary and gravity waves.) With this rigid upper boundary, the vertical eigenfunctions are (Flierl 1978; Hua and Haidvogel 1986)

\[
\phi_m(z) = \cos \left( \frac{m\pi}{D} z \right), \quad m = 0, 1, 2, 3, \ldots
\]

These eigenfunctions form a complete set. Equation (8) then is correct, but we note that it is still not a proof of upscale energy cascade. Equation (8) is nothing more than a statement about an a priori condition for convergence of the infinite series representation of \( \overline{p} \) and hence of \( E \) and \( F \). It must hold for any time, including the initial time. It does not imply that energy will tend to flow from high to low wavenumbers. The same criticism applies to Fjørtoft’s proof for 2D turbulence, which Charney’s proof mirrored.

3. Discussion

There is a common misconception that the relationship between enstrophy, \( G \), and energy, \( E \), that is, \( G(k) = k^2 E(k) \) for 2D and a similar relationship for QG turbulence, decides the direction of energy cascade and that the direction is upscale in the net. This line of argument originated with Fjørtoft on 2D turbulence with his “triad interaction” proof (separate from his “convergence” proof discussed in section 2). The triad interaction proof from 2D turbulence can plausibly be carried over to QG turbulence if one replaces enstrophy by potential enstrophy, as Charney seems to have done. However, Charney stated: “Fjørtoft found that a transfer of energy from one wavenumber to a higher one must be accompanied by still more energy toward a lower wavenumber,” but it is precisely this finding by Fjørtoft that was in error. The corrected statement was given by Merilees and Warn (1975) and it reads: “energy and enstrophy in a 2D non-divergent flow cascade both to lower and higher wavenumbers,” but “the majority of interactions are such that more energy flows to and from
smaller wavenumbers while more enstrophy flows to and from larger wavenumbers” (emphases added).

To use Fjørtoft’s (own) (mistyped) example, if we have three different scales \( l_1 > l_2 > l_3 \) in the ratios \( l_1/l_2 = 2; l_1/l_3 = 2 \), then we can find the ratio of energy change in the longest and shortest scales to be \( \Delta E_s / \Delta E_l = 4 \). However, both \( \Delta E_s \) could be positive or both negative; in the latter case, the longest scale loses more energy downscale to the intermediate scale than the shortest scale loses upscale to the intermediate scale, yielding a net downslope energy flow. Thus such triad-based reasoning for either 2D or QG turbulence can not determine the sign (direction) of net energy transfer. The fact that Charney’s (and Fjørtoft’s) proofs cannot indicate the downscale energy flow. Thus such triad-based reasoning for either 2D or QG turbulence can therefore talk about “dissipative” and “inertial” scales. However, both D = Charney’s (1975), Batchelor (1953), Salmon (1998) for 2D turbulence can carry over to explain the same in QG. However, both Charney’s (1971) was interested in, that is, the equilibrium energy spectrum of the atmosphere, is a forced dissipative system. There the direction of energy transfer can depend on the spectral location of forcing and dissipation.

Charney also noted a similarity between Eqs. (1) and (2). Equation (2), for 2D motion, conserves vorticity, \( \nabla \cdot \psi \), whose Fourier spectral component is

\[-(k_x^2 + k_y^2)\hat{\psi}(k_x, k_y) = -k_z^2\hat{\psi}(k).\]  

(10)

Equation (1) for QG motion, conserves, along horizontal trajectories, the potential vorticity, \( \nabla \cdot \psi + (f_x \psi - \partial_z \theta)/(N^2)\), whose spectral component in 3D is

\[-(k_x^2 + k_y^2 + k_z^2 f_x^2 / N^2)\hat{\psi}(k_x, k_y, k_z, f_0 / N) = -k_z^2\hat{\psi}(k).\]  

(11)

Charney used this isomorphism to say that Kraichnan’s argument for the \( k^{-3} \) spectrum in 2D turbulence could be carried over to explain the same in QG. However, crucial to this is the assumption of isotropy; in the QG case it means that \( k_x^2, k_y^2, \) and \( k_z^2 f_x^2 / N^2 \) must be comparable in magnitude. Charney had shown previously (Charney 1947; Charney and Drazin 1961) that synoptic waves forced by baroclinic instability are mostly trapped in the troposphere. Letting the density scale height \( H \) be the maximum vertical scale for these waves (Held 1978), and letting \( k_x = 2\pi/L_x \) and \( k_y = 2\pi/L_y \), we need, with \( N f_0 \sim 100 \),

\[L_x, L_y < \frac{H}{f_0} = L_R \sim 700 \text{ km}\]

for isotropy. Therefore, distortion is realizable in a 3D atmosphere only for horizontal scales of motion much less than 700 km. However, the \( k^{-1} \) part of the observed spectrum occurs for horizontal scales longer than 1000 km (see Charney’s Fig. 1; Nastrom and Gage 1985; Wiin-Nielsen 1967).

Incidentally, the scales of motion with zonal wavenumbers 7–13 are the “energy injection” scales and therefore do not satisfy the conditions for an “inertial subrange” of Kraichnan. See Welch and Tung (1998) for an alternative explanation of their slope.

Acknowledgments. The research is supported in part by National Science Foundation, Climate Dynamics Program, under Grant ATM 9813770. We thank Drs. Rick Salmon, Michael Montgomery, Theodore Shepherd, and an anonymous reviewer for helpful comments on earlier versions of the manuscript.

REFERENCES