Mechanism of baroclinic instability based on an idealized equation in a simplest situation

Masuda, Akira
Research Institute for Applied Mechanics, Kyushu University

https://doi.org/10.15017/27123

出版情報：九州大学応用力学研究所所報. 142, pp.35-53, 2012-03. Research Institute for Applied Mechanics, Kyushu University

Mechanism of baroclinic instability based on an idealized equation in a simplest situation

Akira MASUDA*[1]

E-mail of corresponding author: *masuda@riam.kyushu-u.ac.jp*

(Received January 31, 2012)

Abstract

Baroclinic instability is investigated with emphasis on its mechanism. First a symmetric form of idealized evolution equation is derived for a two-layer flat ocean on an f-plane, which model is purified and simplified as much as possible without loss of the essence of baroclinic instability. Detailed explanation is provided to each term of the equations. The model captures well the features of baroclinic instability in a simple systematic picture, so that the equation may be called the *canonical equation* for baroclinic instability. The interpretation is based on the evolution of barotropic and baroclinic modes. In particular, the equation is reduced to the Laplace equation for two independent variables of time t and the zonal coordinate x, in the limit of long wavelengths compared with the baroclinic radius of deformation; it is related with the Cauchy-Riemann condition of complex functions. Therefore the instability mechanism becomes almost trivial and the spatial features of growing and decaying modes are interpreted by the property of complex functions (their real and imaginary parts are harmonic functions). On the other hand for short-wave disturbances, the equation is reduced to the one-dimensional wave equation. In both limits, methaphorical systems with the same equations but different physics are presented to explain why the system is unstable (growing/decaying modes) or stable (neutral waves). It also allows simple forms of analytic solutions for the growth rate and stream functions, which make it easy to understand various aspects and roles of baroclinic instability. Also simple representations are given to the meridional transport of buoyancy and the meridional circulation as for Eady's model, within the framework of a two-layer model. In addition, the role of baroclinic instability is argued in relation to Gent-McWilliams parameterization and diffusive stretching.

Key words : *baroclinic instability, canonical equation, instability mechanism, barotropic and baroclinic modes, secondary circulation, G-M parameterization and diffusive stretching*

1. Introduction

Baroclinic instability, represented by storms of the mid-latitude weather, is a well-known and familiar phenomenon not only in the atmosphere but also in the ocean. It is considered to control large-scale variability of geophysical fluids. There have been a vast amount of literature about baroclinic instability starting with Eady (1949) and Charney (1947). Now it is easy to compute the growth rate and the structure of growing modes of disturbance even for complicated and realistic situations, by numerical means. Also it has become a routine to simulate the nonlinear evolution of a system unstable to baroclinic waves.

Nevertheless, it is difficult to understand and explain adequately the mechanism of baroclinic instability and its roles in the general circulation; baroclinic instability cannot be easy, since it contains (1) the Coriolis force, (2) density stratification, and (3) three-dimensionality of space (vertical shear of zonal current, meridional gradient of density, and zonal wavenumber of disturbance), as its indispensable ingredients.

In fact, we may encounter dubious explanations even now, in spite of considerable efforts made so far. For instance, the term of *baroclinic instability* may lead one to suppose that baroclinic instability is caused directly by the baroclinic term $\nabla p \times \nabla \rho$ in the vorticity equation, where ρ denotes density, p pressure, and ∇ horizontal gradient operator. The baroclinic term $\nabla \rho \times \nabla p$, however, does not appear essential, as will be contrasted with the instability mechanism discussed in the present paper. Also the interpretation in terms of Rossby waves appear to lead to somewhat ambiguous consequences. In short, baroclinic instability is such a phenomenon that denies superficial explanation.

So it is desirable to have a variety of investigations

*1 Research Institute for Applied Mechanics, Kyushu University

and interpretation of baroclinic instability. The primary purpose here is to present an explanation of the mechanism of baroclinic instability in a simplest manner without losing the essence of baroclinic instability and mathematical rigorousness. As a byproduct, we obtain a further insight into the mathematics and physics of baroclinic instability than the usual approaches made so far give. We first propose an evolution equation for baroclinic instability, which we call the *canonical equation* of baroclinic instability. The equation is obtained in a situation simplified as much as possible. Also we argue the mechanism of instability and clarify the spatial structure of growing (decaying) or neural modes, and some roles of baroclinic instability in large-scale dynamics, on the basis of the canonical equation, which provides simple analytical formulas.

The next section describes a simplified and purified situation for baroclinic instability with notations. The third section derives the canonical equation and the meaning of each term of the equation. In the fourth section results based on the canonical equation are presented and related issues are discussed. The final section gives a summary and further discussion.

2. Formulation

There have been various models of baroclinic instability. One of the most fundamental and important among them is probably that of Eady (1949), which considers a horizontally uniform zonal flow with constant vertical shear in a continuously and uniformly stratified channel on an f-plane. In a simplified situation of the two-layer discrete model, growing modes of baroclinic instability were discussed in a paper of ours on Kuroshio frontal eddies (Masuda and Okuno 2002; see Phillips 1954 too). That paper, however, was focused on the properties of growing modes, and did not address to the mechanism (or physics) of instability, which is the primary purpose of the present article.

We make a further simplification or purification to obtain a clearer understanding (or physical interpretation) of the mechanism of baroclinic instability. In the following, the convenient notation of GFDVN is used often without definition; see Masuda (2010) for details.

Consider an ocean of a flat and zonal channel on an f-plane (Fig. 1), where the Coriolis parameter f is a positive constant (the northern hemisphere). The ocean has a meridional width L; in the zonal direction it is infinitely wide or periodic. The ocean consists of two layers with a density difference $\delta\rho$, which is quite small compared with the reference density ρ_r (which is unity for convenience), so that the Boussinesq approximation is valid. Then the reduced gravity is expressed as $g' \equiv g\delta\rho/\rho_r$, where g denotes the acceleration of grav-

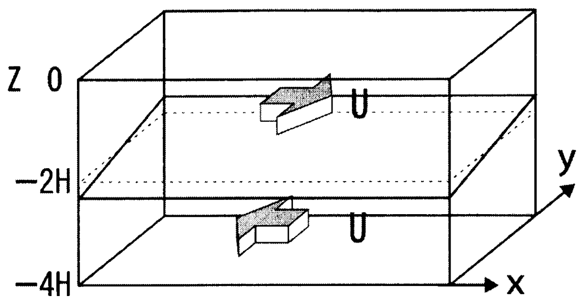

Fig. 1 A schematic view of the model ocean like a channel on an f-plane, which consists of two layers of the same average thickness $2H$. The background current is uniform and eastward in the surface layer, while westward in the bottom layer. The interface has a uniform slope due to geostrophy.

ity. Both the upper and lower layers have the same average thickness $2H$. The suffixes 1 and 2 ($i = 1, 2$) are used to refer to the upper layer and the lower layer, respectively; the upper (lower) layer is called often as the surface (bottom) layer.

We denote time by t and (eastward, northward, upward) coordinates by (x, y, z); $\boldsymbol{x} = (x, y)$. The origin of the y-axis is taken so that the southern and northern boundaries of the ocean lie at $y = \mp L/2$, respectively; the origin of the z-axis is set at the level of sea surface, so that the bottom is identical with the plane $z = -4H$. The (eastward, northward) components of horizontal current $\boldsymbol{u}$ are denoted by (u, v).

Throughout the paper we assume the quasi-geostrophic (Q-G) formulation on an f-plane (Pedlosky 1987, say), in which $\boldsymbol{u}$ is expressed as

$$u = -\frac{\partial \psi}{\partial y}, \quad v = \frac{\partial \psi}{\partial x} \quad \text{or}$$
$$\boldsymbol{u} = - \lhd \psi \qquad \text{(GFDVN)}$$

in terms of the quasi-geostrophic stream function ψ, which is related with pressure p by $f\psi = p$. As the background (abbreviated to *b.g.*) current field, we assume uniform eastward flow $U_1 = U$ in the upper layer, while uniform westward current $U_2 = -U$ in the lower layer, where U is a positive constant. Then the b.g. vertical position of the interface $z = Y$ of the two layers has a slight meridional slope of $dY/dy = 2fU/g'$ from geostrophy. Then the b.g. thickness of each layer h_i ($i = 1, 2$) is

$$\frac{h_1}{2H} = 1 - \frac{y}{2H}\frac{dY}{dy}, \quad \frac{h_2}{2H} = 1 + \frac{y}{2H}\frac{dY}{dy}.$$

We assume

$$\epsilon \equiv \frac{L}{2H}\frac{dY}{dy} \ll 1, \tag{1}$$

so that $2H$ is a good approximation to h_i $(i = 1, 2)$.

The linearized Q-G vorticity equation becomes

$$\begin{cases} \left(\frac{\partial}{\partial t} + U\frac{\partial}{\partial x}\right)\left(\nabla^2\psi_1 - \frac{F}{2}(\psi_1 - \psi_2)\right) \\ \qquad +FU\frac{\partial\psi_1}{\partial x} = 0 \\ \left(\frac{\partial}{\partial t} - U\frac{\partial}{\partial x}\right)\left(\nabla^2\psi_2 + \frac{F}{2}(\psi_1 - \psi_2)\right) \\ \qquad -FU\frac{\partial\psi_2}{\partial x} = 0 \end{cases}, \quad (2)$$

where ψ_1 and ψ_2 denote the quasi-geostrophic stream function of disturbance in the upper and lower layers, respectively; $1/F \equiv g'H/f^2$ the square of the baroclinic radius of deformation.

In the above equations, the last term originates in the meridional gradient of the b.g. thickness of each layer. It may be called the CIPT β-effect (baroclinic-Current Induced Pseudo-Topographic β effect), which has been used in a study of the southwestward movement of Meddies (Takahashi and Masuda 1998). The CIPT β may be written better as

$$\beta_t \equiv FU = \frac{f^2U}{g'H} = \frac{f}{2H}\frac{dY}{dy}. \qquad \left(\frac{\beta L}{f} = \epsilon \ll 1\right)$$

Note that CIPT β works in the opposite way between the upper and lower layers. The northward current produces negative vorticity in the surface layer, while positive in the bottom layer. This term is essential to baroclinic instability, as will be shown later.

Strictly speaking, the sea surface has a small meridonal slope to express the zonal b.g. current in the surface layer. We set it flat, however, by making the rigid lid approximation (equivalent to the assumption of an infinite barotropic radius of deformation). Also the interface between the two layers have a slope, so that the thickness of both layers h_i $(i = 1, 2)$ are not uniform. But, h_i $(i = 1, 2)$ varies only a little in comparison with the average thickness of each layer $2H$, according to the assumption (1).

The appropriate side boundary conditions are

$$v_i = \frac{\partial\psi_i}{\partial x} = 0 \quad \text{at} \quad y = \pm\frac{L}{2} \quad (i = 1, 2)$$

for the channel model. It is replaced by

$$\psi_i - \overline{\psi}_i = 0 \quad \text{at} \quad y = \pm\frac{L}{2} \quad (i = 1, 2), \qquad (3)$$

where $\overline{\bullet}$ denotes the zonal average of $\bullet$.

To remove some uncertainty peculiar to the channel model (see Appendix), we pose additional constraints of

$$\langle\psi_1 - \psi_2\rangle = 0, \qquad (4)$$

$$\langle u_1 + u_2\rangle = 0, \qquad (5)$$

where $\langle\bullet\rangle \equiv \frac{1}{L}\int_{-L/2}^{L/2} \overline{\bullet}\, dy$ denotes the horizontal average (domain average). These constraints mean the mass conservation of each layer (see Masuda 2011), and the conservation of the total zonal momentum.

Other dynamical variables are expressed in terms of ψ_i $(i = 1, 2)$. For instance the upward displacement of the interface becomes $\eta = (\psi_2 - \psi_1)f/g'$. The vertical velocity w is expressed as

$$\begin{cases} w_1 = \left(\frac{\partial}{\partial t} + U\frac{\partial}{\partial x}\right)\eta + v_1\frac{dY}{dy} \\ w_2 = \left(\frac{\partial}{\partial t} - U\frac{\partial}{\partial x}\right)\eta + v_2\frac{dY}{dy} \end{cases},$$

which yields

$$w_1 = w_2, \qquad (6)$$

since

$$U\frac{\partial\eta}{\partial x} + v_1\frac{dY}{dy} = \frac{1}{2}\frac{dY}{dy}\frac{\partial(\psi_2 - \psi_1)}{\partial x} + \frac{\partial\psi_1}{\partial x}\frac{dY}{dy}$$
$$= \frac{1}{2}\frac{dY}{dy}\frac{\partial(\psi_1 + \psi_2)}{\partial x} = -U\frac{\partial\eta}{\partial x} + v_2\frac{dY}{dy}.$$

The meridional displacement $\mathcal{Y}$ of water becomes

$$\begin{cases} \left(\frac{\partial}{\partial t} + U\frac{\partial}{\partial x}\right)\mathcal{Y}_1 = v_1 = \frac{\partial\psi_1}{\partial x} \\ \left(\frac{\partial}{\partial t} - U\frac{\partial}{\partial x}\right)\mathcal{Y}_2 = v_2 = \frac{\partial\psi_2}{\partial x} \end{cases}. \qquad (7)$$

Likewise the vertical displacement of water is

$$\begin{cases} \left(\frac{\partial}{\partial t} + U\frac{\partial}{\partial x}\right)\mathcal{Z}_1 = w_1 \\ \left(\frac{\partial}{\partial t} - U\frac{\partial}{\partial x}\right)\mathcal{Z}_2 = w_2 = w_1 \end{cases}. \qquad (8)$$

The total energy of disturbance is governed by

$$2H\frac{\partial}{\partial t}\left\langle\frac{|\nabla\psi_1|^2 + |\nabla\psi_2|^2 + (F/2)(\psi_1 - \psi_2)}{2}\right\rangle$$
$$= FUH\left\langle(\psi_1 - \psi_2)\frac{\partial(\psi_1 + \psi_2)}{\partial x}\right\rangle \qquad (9)$$
$$= 2FUH\left\langle\psi_1\frac{\partial\psi_2}{\partial x}\right\rangle = -2FUH\left\langle\psi_2\frac{\partial\psi_1}{\partial x}\right\rangle. \quad (10)$$

where we have used identities

$$\overline{r\frac{\partial r}{\partial x}} = \overline{\frac{1}{2}\frac{\partial r^2}{\partial x}} = 0, \qquad \overline{r\frac{\partial s}{\partial x}} = -\overline{s\frac{\partial r}{\partial x}} \qquad (11)$$

for zonally periodic functions r and s.

The left-hand side of (9) is the kinetic and potential energy of disturbance, whereas the right-hand side implies the energy supply to disturbance. In order for disturbance to grow, there must be some correlation between ψ_1 and ψ_2, which will be argued later.

3. Canonical equations for baroclinic instability and the meaning of each term of the equations

In examining the linear stability, it is usual to argue the eigenvalue problem for a Fourier component of a specified zonal wavenumber k. Or, one may eliminate either stream function to obtain an equation for a single stream function, ψ_1 or ψ_2. Such methods (Drazin and Reid 1981, Pedlosky 1987), however, have been exploited so often that they would yield nothing new as regards the interpretation of baroclinic instability. We adopt a different approach here.

3.1 barotropic and baroclinic stream functions and the reduced form of barotropic vorticity equation

Let us introduce the stream functions of the barotropic mode and baroclinic mode by

$$\begin{cases} \phi \equiv \psi_1 + \psi_2 \\ \varphi \equiv \psi_1 - \psi_2 \end{cases} \leftrightarrow \begin{cases} \psi_1 = \dfrac{\phi + \varphi}{2} \\ \psi_2 = \dfrac{\phi - \varphi}{2} \end{cases},$$

respectively. The baroclinic stream function φ of disturbance represents the depression of the interface, or it is related with η by

$$\varphi = -\frac{g'}{f}\eta. \tag{12}$$

Also we define

$$\begin{cases} v_\phi \equiv \dfrac{\partial \phi}{\partial x} = v_1 + v_2 \\ v_\varphi \equiv \dfrac{\partial \varphi}{\partial x} = v_1 - v_2 \end{cases}, \quad \begin{cases} v_1 \equiv \dfrac{\partial \psi_1}{\partial x} \\ v_2 \equiv \dfrac{\partial \psi_2}{\partial x} \end{cases},$$

where v_ϕ and v_φ are the northward velocities of the disturbance barotropic and baroclinic modes, respectively. We often use suffixes ϕ and φ to refer to the borotropic and baroclinic modes, respectively.

Changing the unknowns from the pair of ψ_1 and ψ_2 to that of ϕ and φ (basis transformation), we rewrite the Q-G vorticity equation (2) in each layer as

$$\frac{\partial}{\partial(Ut)}\nabla^2\phi + \frac{\partial}{\partial x}\nabla^2\varphi = 0, \tag{13}$$

$$\frac{\partial}{\partial(Ut)}(\nabla^2 - F)\varphi + \frac{\partial}{\partial x}(\nabla^2 + F)\phi = 0, \tag{14}$$

which express the evolution of the infinitesimal disturbances of the barotropic and baroclinic modes, respectively.

Physical meaning of each term in the barotropic vorticity equation (13) is to be explained for our interpretation of baroclinic instability. The first term is trivial, but what does the second term mean? The equation says that the vertical average of vorticity, i.e., the barotropic mode of vorticity $\nabla^2\phi$, changes only through the advection of relative vorticity expressed by the second term. This is because there are no other factors producing the vertically integrated vorticity: [1] vertical stretching (horizontal convergence) of water plays no role in producing the vertically integrated vorticity on a flat ocean; [2] there is no planetary β effect; and [3] the CIPT β effect does not produce vertically integrated vorticity, since it has its origin in the vertical stretching as for [2]. In this case, therefore, the only term that can produce the barotropic vorticity $\nabla^2\phi$ is the zonal advection of the *baroclinic* vorticity $\nabla^2\varphi$ by the b.g. *baroclinic* zonal urrent. It is a most typical interaction among vertical modes: nonlinear interaction of two baroclinic modes yields a barotropic forcing (Masuda 2011, say).

As is discussed in Appendix, the operator ∇^2 is removed from this equation to give

$$\frac{\partial \phi}{\partial(Ut)} + \frac{\partial \varphi}{\partial x} = 0, \tag{15}$$

which is the *reduced* equation of the balance of barotropic vorticity. It is to be noted that the transform of basis stream functions to the barotropic and baroclinic ones has yielded such a simple representation of vorticity equation.

Looking at this equation, we may say that ϕ is produced by the advection (or zonal convergence) of φ. In terms of dynamical variables more familiar than ϕ and φ, (15) is rewritten as

$$\frac{\partial v_\phi}{\partial(Ut)} + \frac{\partial v_\varphi}{\partial x} = 0, \tag{16}$$

or equivallently as

$$\frac{\partial v_\phi}{\partial(Ut)} - \frac{g'}{f}\frac{\partial^2\eta}{\partial x^2} = 0. \tag{17}$$

Then we may say that v_φ is produced by the zonal convergence of v_φ according to (16), or by a kind of zonal diffusion of η according to (17). It should be born in mind, however, that the reduced equations (15)–(17) should be interpreted from the balance of barotropic vorticity $\nabla^2\phi$ (13). Roughly speaking, it means that the zonal advection of the *baroclinic* vorticity (expressed as $\nabla^2\varphi \sim -\varphi$) by the b.g. current U_i causes barotropic vorticity (expressed as $\nabla^2\phi \sim -\phi$).

The meaning of the baroclinic vorticity equation (14) is explained better on

$$\frac{\partial}{\partial t}\nabla^2\varphi = F\frac{\partial \varphi}{\partial t} - U\frac{\partial}{\partial x}\nabla^2\phi - UF\frac{\partial \phi}{\partial x}.$$

The left-hand side is the temporal change of baroclinic vorticity. The first term on the right-hand side is the production of baroclinic vorticity by baroclinic horizontal convergence (vertical stretching) of water. Usually this term is moved to the left-hand side and treated as a thickness term of potential vorticity. The second term on the right-hand side means the production of *baroclinic* vorticity due to the advection of *barotropic* vorticity by the b.g. *baroclinic* current. The third term means the production of baroclinic relative vorticity by the northward advection of the b.g. *baroclinic* potential vorticity (b.g. thickness) by the *barotropic* disturbance (CIPT β effect).

The form of (14) reads: the disturbance baroclinic potential vorticity changes from the baroclinic production of vorticity by [1] the advection of the barotropic vorticity by the baroclinic b.g. current and [2] the northward barotropic current under the baroclinic CIPT β effect; the latter is interpreted alternatively as the baroclinic production of potential vorticity by the northward advection of the b.g. baroclinic potential vorticity (thickness term) by the disturbance northward barotropic current v_ϕ. If we use η and v_ϕ instead of ϕ and φ, parallel to (17), (14) is expressed as

$$\frac{g'}{f}\frac{\partial}{\partial(Ut)}(\nabla^2 - F)\eta = (\nabla^2 + F)v_\phi. \quad (18)$$

Note that the reduced form is not obtained for the baroclinic vorticity equation, except for the long-wave and short-wave limits discussed later.

3.2 canonical equation for baroclinic instability when $h_1 = h_2$

Substituting the reduced barotropic vorticity equation (15) into the baroclinic vorticity equation (14), we obtain

$$\frac{\partial^2}{\partial(Ut)^2}(\nabla^2 - F)\varphi - \frac{\partial^2}{\partial x^2}(\nabla^2 + F)\varphi = 0, \quad (19)$$

which has a single unknown φ. It is equivalent to

$$\begin{aligned}
&\left(\frac{\partial}{\partial(Ut)} + \frac{\partial}{\partial x}\right)\left(\frac{\partial}{\partial(Ut)} - \frac{\partial}{\partial x}\right)\nabla^2\varphi \\
&= \left(\frac{\partial^2}{\partial(Ut)^2} - \frac{\partial^2}{\partial x^2}\right)\nabla^2\varphi \\
&= \left(\frac{\partial^2}{\partial(Ut)^2} + \frac{\partial^2}{\partial x^2}\right)F\varphi \quad (20) \\
&= \left(\frac{\partial}{\partial(Ut)} + i\frac{\partial}{\partial x}\right)\left(\frac{\partial}{\partial(Ut)} - i\frac{\partial}{\partial x}\right)F\varphi, \quad (21)
\end{aligned}$$

where $i \equiv \sqrt{-1}$ is the imaginary unit. The left-hand side is related with relative vorticity only, whereas the right-hand side is related with the stretching term alone (CIPT β included here). This equation is suggestive of further mathematical or physical aspects of baroclinic instability.

Let us call any of (19) – (21) as the *canonical equation* of baroclinic instability, because it has such s neat form that allows simple analytical consequences, as will be shown later. We see a good symmetry easily from the form of the equation. For instance, the system is invariant under the simultaneous inversion of the zonal direction and time, or the exchange of (x, t) with $(-x, -t)$; it means that, if a mode of disturbance grows with time, the disturbance with the opposite zonal character decays with time. Also we should note that (19) has an integrated form of evolution equation, which differs from the ordinary analysis, where disturbance is decomposed into individual modes specified as Fourier components with specified meridional patterns. It is to be remarked that the same form of equation holds for ϕ, ψ_1, and ψ_2.

Meanwhile from (6) we have

$$\begin{aligned}
w_1 = w_2 &= \frac{w_1 + w_2}{2} = \frac{w_\phi}{2} = \frac{\partial\eta}{\partial t} + \frac{v_\phi}{2}\frac{dY}{dy} \\
&= -\frac{fU}{g'}\left(\frac{\partial\varphi}{\partial(Ut)} - \frac{\partial\phi}{\partial x}\right). \quad (22)
\end{aligned}$$

Finally the energy equation (9) becomes

$$\begin{aligned}
&H\frac{\partial}{\partial(Ut)}\left\langle\frac{|\nabla\phi|^2 + |\nabla\varphi|^2 + F\varphi^2}{2}\right\rangle \\
&= FH\left\langle\varphi\frac{\partial\phi}{\partial x}\right\rangle = FH\left\langle\varphi v_\phi\right\rangle. \quad (23)
\end{aligned}$$

It shows that the energy of disturbance grows only when there is a positive correlation between v_ϕ and $\varphi = -g'\eta/f$. We will go back to this subject later.

It is easy to derive another conservation law

$$\frac{\partial}{\partial(Ut)}\left\langle\frac{|\nabla\phi|^2 + |\nabla\varphi|^2 + F(\varphi^2 - \phi^2)}{2}\right\rangle = 0, \quad (24)$$

since

$$\begin{aligned}
\left\langle\varphi\frac{\partial\phi}{\partial x}\right\rangle &= -\left\langle\frac{\partial\varphi}{\partial x}\phi\right\rangle = \left\langle\frac{\partial\phi}{\partial(Ut)}\phi\right\rangle \\
&= \left\langle\frac{\partial}{\partial(Ut)}\frac{\phi^2}{2}\right\rangle
\end{aligned}$$

by virtue of (11) and (15). The conservation law (24) suggests the possibility that ϕ^2 and φ^2 may grow with time infinitely.

3.3 canonical equation when $h_1 \neq h_2$

We obtain a similar equation as well when the thickness of the upper layer h_1 differs from that of the lower layer h_2 as in the ordinary two-layer model (Phillips 1954, Masuda and Okuno 2002).

We normalize the thickness of each layer by

$$H_1 \equiv \frac{\langle h_1\rangle}{\langle h_1\rangle + \langle h_2\rangle}, \quad H_2 \equiv \frac{\langle h_2\rangle}{\langle h_1\rangle + \langle h_2\rangle}.$$

A few baroclinic radii of deformation are defined as

$$\begin{cases}
\tilde{F} \equiv \dfrac{f^2}{g'(\langle h_1\rangle + \langle h_2\rangle)} \\
F_i \equiv \dfrac{f^2}{g'\langle h_i\rangle} = \dfrac{\tilde{F}}{H_i} \qquad (i = 1, 2) \\
F \equiv F_1 + F_2 = \dfrac{\tilde{F}}{H_1 H_2}
\end{cases}.$$

We define the barotropic and baroclinic stream functions by

$$\begin{cases}
\phi \equiv H_1\psi_1 + H_2\psi_2 \\
\varphi \equiv \psi_1 - \psi_2
\end{cases},$$

respectively. They are inverted to ψ_i $(i = 1, 2)$ as

$$\begin{cases} \psi_1 = \phi + H_2\varphi \\ \psi_2 = \phi - H_1\varphi \end{cases}.$$

We adopt a frame of reference moving eastward with a characteristic velocity of the b.g. current, which is the thickness-weighted average of the b.g. current. Let U be the velocity difference between the upper and lower layers. We define U_i $(i = 1, 2)$ by

$$U_1 = UH_2 > 0, \quad U_2 = -UH_1 < 0.$$

Then the vorticity equations become

$$\left(\frac{\partial}{\partial(Ut)} + H_2\frac{\partial}{\partial x}\right)\left(\nabla^2\psi_1 - F_1(\psi_1 - \psi_2)\right) + F_1\frac{\partial\psi_1}{\partial x} = 0, \tag{25}$$

$$\left(\frac{\partial}{\partial(Ut)} - H_1\frac{\partial}{\partial x}\right)\left(\nabla^2\psi_2 + F_2(\psi_1 - \psi_2)\right) - F_2\frac{\partial\psi_2}{\partial x} = 0 \tag{26}$$

for the upper and low layers, respectively.

Multiplying (25) by H_1 and (26) by H_2, and adding together, we obtain the barotropic vorticity equation. Subtraction of (26) from (25) yields the baroclinic vorticity equation. That is, we have

$$0 = \frac{\partial\nabla^2\phi}{\partial(Ut)} + H_1H_2\frac{\partial\nabla^2\varphi}{\partial x}, \tag{27}$$

$$0 = \frac{\partial}{\partial t}\left(\nabla^2 - F\right)\varphi + \frac{\partial}{\partial x}(\nabla^2 + F)\phi + (H_2 - H_1)\frac{\partial}{\partial x}\nabla^2\varphi. \tag{28}$$

In (28) we find a term $(H_2 - H_1)\frac{\partial}{\partial x}\nabla^2\varphi$, which was absent when $H_1 \neq H_2$. It implies that *baroclinic* b.g. current advects *baroclinic* vorticity to produce not only barotropic vorticity, but also *baroclinic* vorticity; the case of $H_1 = H_2$ is an exceptional one, where the interaction between the two *baroclinic* terms yields no *baroclinic* forcing.

Again, the barotropic vorticity equation (27) is reduced to

$$0 = \frac{\partial\phi}{\partial(Ut)} + H_1H_2\frac{\partial\varphi}{\partial x}. \tag{29}$$

Substituting this reduced relation into (28) to remove ϕ, we obtain a required equation for φ

$$0 = \frac{\partial^2}{\partial(Ut)^2}\left(\nabla^2 - F\right)\varphi - \frac{\partial^2}{\partial x^2}\left[(\nabla^2 + F)H_1H_2\varphi\right] + \frac{\partial}{\partial x}\frac{\partial}{\partial(Ut)}(H_2 - H_1)\nabla^2\varphi. \tag{30}$$

Alternatively it is rewritten as

$$\left(\frac{\partial}{\partial(Ut)} + H_2\frac{\partial}{\partial x}\right)\left(\frac{\partial}{\partial(Ut)} - H_1\frac{\partial}{\partial x}\right)\nabla^2\varphi = \left(\frac{\partial^2}{\partial(Ut)^2} + H_1H_2\frac{\partial^2}{\partial x^2}\right)F\varphi. \tag{31}$$

It is obvious that (30) and (31) reduce to (19) and (21), respectively, when $H_1 = H_2 = 1/2$ with t replaced by $t/2$. The dynamical meaning of each term of (27) – (30) is thus almost the same as in the canonical equation when $H_1 = H_2$ and details are omitted.

Next let us consider a limit of $H_2 \nearrow 1$ and $H_1 \searrow 0$. We see that (27) leads to vanishing ϕ, so that $\psi_1 = \varphi$, $\psi_2 = 0$. That is, disturbance must be the surface mode from the beginning. We see then that (28) leads to

$$0 = \frac{\partial}{\partial t}(\nabla^2 - F)\varphi + U\frac{\partial}{\partial x}\nabla^2\varphi, \tag{32}$$

which governs the evolution of the surface mode. In this case no growing mode is possible, as is evident from the energy equation (23) or the conservation law (24). We observe from (32), that the potential vorticity of the surface mode is produced simply by the advection of the relative velocity by the b.g. zonal current $U_1 = U$.

It is worth mentioning that (32) can be rewritten also as

$$0 = \left(\frac{\partial}{\partial t} + U\frac{\partial}{\partial x}\right)(\nabla^2 - F)\varphi + FU\frac{\partial\varphi}{\partial x},$$

which indicates that φ behaves as baroclinic Rossby waves associated with the CIPT β $(\beta_t = FU)$ in uniform eastward zonal flow $U_1 = U$. This view of the surface mode as Rossby waves is contrasted with that based on (32).

4. Results and interpretation

Since our purpose is to understand the fundamental aspects of baroclinic instability in a simplest way, we confine our discussion to the case of $H_1 = H_2$.

4.1 growth rate of growing modes, and phase velocity of neutral modes

We first derive some fundamental properties of instability. Henceforth we often use the complex representation of real variables, where a real variable s is expressed by another complex variable s (the same notation), but the final result of real s is obtained by taking the real part of complex s; $s = \Re[s]$, where $\Re[\bullet]$ denotes the real part of $\bullet$. When we need multiplication of two real quantities r and s represented by its complex counterparts r and s, real $r \cdot s$ should be understood as

$$r \cdot s = \Re(r)\Re(s) = \Re\frac{rs^*}{2} + \Re\frac{rs}{2},$$

where $\bullet^*$ denotes the complex conjugate of $\bullet$.

Consider a mode of disturbance specified as in Appendix

$$\varphi \sim e^{ik(x - c(\boldsymbol{k})t)}\sin l\left(y + \frac{L}{2}\right), \tag{33}$$

where k is the zonal wavenumber, $l = \frac{n\pi}{L}$ the meridional wavenumber (n being a positive integer), and $\boldsymbol{k} = (k, l)$

the corresponding wavenumber vector. The complex phase velocity of the mode $c = c(\boldsymbol{k})$ is related with the complex frequency ω by $\omega(\boldsymbol{k}) \equiv k\,c(\boldsymbol{k})$. The amplitude of disturbance grows with time when $\omega_i \equiv \Im[\omega] > 0$, decays with time when $\omega_i < 0$, or stays neutral when $\omega_i = 0$, where $\Im[\bullet]$ denotes the imaginary part of $\bullet$.

[1] complex dispersion relation and critical wavenumber

Substituting (33) into the canonical equation (19) we obtain the complex dispersion relation

$$\frac{c^2(\boldsymbol{k})}{U^2} = \frac{\boldsymbol{k}^2 - F}{\boldsymbol{k}^2 + F}, \tag{34}$$

which yields

$$c = c(\boldsymbol{k}) = \begin{cases} \pm|c|i & \text{when } \boldsymbol{k}^2 \le F \\ \pm|c| & \text{when } \boldsymbol{k}^2 \ge F \end{cases}$$

and

$$0 \le \frac{|c|}{U} < 1 \quad \text{for any} \quad \boldsymbol{k}, \tag{35}$$

$$\begin{cases} \dfrac{|c|}{U} \to 1 & \text{when} \quad \boldsymbol{k}^2 \to 0 \\ \dfrac{|c|}{U} = 0 & \text{when} \quad \boldsymbol{k}^2 = F \\ \dfrac{|c|}{U} \to 1 & \text{when} \quad \boldsymbol{k}^2 \to \infty \end{cases}.$$

The behavior of the mode depends on whether c is a real number or a complex. When $\boldsymbol{k}^2 < F$, pairs of growing modes and decaying modes appear simultaneously. Likewise pairs of eastward propagating and westward propagating neutral modes appear simultaneously when $\boldsymbol{k}^2 > F$.

That is, the critical wavenumber that separates growing (decaying) modes from neutral modes is

$$|\boldsymbol{k}|_c = \sqrt{F},$$

which is just the baroclinic deformation wavenumber. Since $\boldsymbol{k} = (k, l)$ and $l = n\pi/L \ge \pi/L$, we have $|\boldsymbol{k}| \ge l$. Baroclinic instability is impossible therefore when π/L is larger than the deformation wavenumber $\sqrt{F}$. Provided that baroclinic instability has a role of adjusting a highly unstable state to a near-critical state of less active growing modes, it limits L, or the width of the frontal zone, to a value a bit larger than $\pi/\sqrt{F}$, which might have some relation with the question of how the width of the baroclinic current is determined.

[2] wavenumer at which growth rate is maximum

The growth rate $\omega_i = kc_i \equiv k\Im[c]$ is a straightforward consequence of (34). In particular the maximum growth rate in the wavenumber range of $|\boldsymbol{k}|^2 < F$ is obtained from the condition of $\partial\omega_i/\partial k = 0$. The wavenumber $k_{\max}$ for which ω_i takes the largest (positive) growth rate $\omega_{i;\max}$ is expressed explicitly as

$$\frac{k_{\max}^2}{F + l^2} = \frac{|c_{i;\max}|^2}{U^2} = \frac{\omega_{i;\max}}{U\sqrt{F + l^2}} = \sqrt{\frac{2F}{F + l^2}} - 1,$$

where $c_{i;max} \equiv \omega_{i;max}/k_{\max}$. When $l^2 \ll F$, we have

$$\frac{k_{\max}^2}{F} = \frac{|c_{i;\max}|^2}{U^2} = \frac{\omega_{i;\max}}{U\sqrt{F}} = \sqrt{2} - 1 = 0.414.$$

4.2 structure of growing (decaying) modes and neutral modes

In the two-layer model, not continuouly stratified one, the spatial structure is expressed simply by the relative amplitude and phase among ϕ, φ, and ψ_i $(i = 1, 2)$. It turns out convenient to formulate the amplitude and phase of stream functions and others relative to those of ϕ, the barotropic mode stream function. Also it is useful to take the origin of the zonal coordinate x at the location of minimum of ϕ, which corresponds to the center of the low pressure of disturbance or the storm center for growing modes of baroclinic instability.

Irrespective of whether c is real or complex, it follows from the reduced barotropic vorticity equation (15) that

$$\varphi = \frac{c}{U}\phi, \tag{36}$$

which gives ψ_i $(i = 1, 2)$ as

$$\begin{cases} \psi_1 = \dfrac{1 + c/U}{2}\phi \\ \psi_2 = \dfrac{1 - c/U}{2}\phi \end{cases}, \quad \frac{\psi_1}{\psi_2} = \frac{1 + c/U}{1 - c/U}. \tag{37}$$

The relations (36)-(37) are sufficient to determine the spatial structure of the mode of disturbance, since c is known from (34).

As for w, substituting (36) into (22), we obtain

$$w_1 = w_2 = \frac{fUk}{g'}\left(1 + \frac{c^2}{U^2}\right)i\phi.$$

Since c^2 is a real number satisfying (35), we observe that

$$\arg(w_1) = \arg(w_2) = \arg(\phi) + \frac{\pi}{2},$$

irrespective of $\boldsymbol{k}$, where $\arg(\bullet)$ denotes the argument of the complex number $\bullet$.

Except for the relative phase of w_i, the relative amplitude and phase have clear contrast and symmetry among distinct characteristics of modes (growing, decaying, and neutral), reflecting the property of c. We examine them separately.

[1] growing modes

We are concerned with growing modes most and examine them in details. The growing mode is possible when $\boldsymbol{k}^2 < F$ and the complex phase velocity becomes a pure imaginary number $c = |c|i$; growing modes do not propagate in the present frame of reference. The baroclinic stream function φ becomes

$$\varphi = +i\frac{|c|}{U}\phi, \tag{38}$$

so that

$$\begin{cases} 0 \le \dfrac{|\varphi|}{|\phi|} = \dfrac{|c|}{U} < 1 \\ \arg\left(\dfrac{\varphi}{\phi}\right) = \dfrac{\pi}{2} \Rightarrow \arg(\varphi) = \arg(\phi) + \dfrac{\pi}{2} \end{cases}$$

It is to be remembered that if $\arg(r) = \arg(s) + \pi/2$, say, the peak of r is located to the west of the location of the peak of s at a distance of $\pi/(2k)$; that relation holds for any phase.

We see that the phase difference of φ from ϕ is always $\pi/2$, while the amplitude of φ is not greater than that of ϕ. As $|\boldsymbol{k}| \nearrow \sqrt{F}$, $|\varphi|/|\phi|$ approaches 0, and the growing mode becomes almost barotropic.

As for ψ_i $(i = 1, 2)$, we have

$$\begin{cases} \psi_1 = \dfrac{1 + i|c|/U}{2}\phi \\ \psi_2 = \dfrac{1 - i|c|/U}{2}\phi \end{cases}, \quad \frac{\psi_1}{\psi_2} = \frac{1 + i|c|/U}{1 - i|c|/U}.$$

The amplitude of ψ_2 is always the same as that of ψ_1 for growing modes. It follows from (35) that

$$0 \le \theta \equiv \arg\left(\frac{\psi_1}{\phi}\right) = -\arg\left(\frac{\psi_2}{\phi}\right) < \frac{\pi}{4}.$$

The phase difference of ψ_i $(i = 1, 2)$ from ϕ varies with $\boldsymbol{k}^2$ within the range of $\pi/4$. As $|\boldsymbol{k}| \nearrow \sqrt{F}$, $\arg(\psi_i)$ approaches $\arg(\phi)$. The amplitude of ψ_i relative to $|\phi|$ changes with $\boldsymbol{k}^2$ too.

The minimum of φ is observed to the west of the center of the low pressure, because the phase difference is always $\pi/2$. On the other hand the minimum of ψ_1 (ψ_2) occurs to the west (east) of the center of the low pressure; phase difference θ is within $\pi/4$. The amplitude of ψ_1 is always the same as that of ψ_2, but $|\varphi|$ is not greater than $|\phi|$.

In the limit of long waves, where characteristics of growing modes appear typically, (36) and (37) become

$$\varphi \approx +i\phi$$

$$\begin{cases} \psi_1 \approx \dfrac{e^{+i\pi/4}}{\sqrt{2}}\phi \\ \psi_2 \approx \dfrac{e^{-i\pi/4}}{\sqrt{2}}\phi \end{cases}, \quad \psi_1 \approx i\psi_2.$$

From (7) and (8) we find

$$\mathcal{Y}_1 \approx i\frac{\phi}{2U} \approx \mathcal{Y}_2,$$

$$e^{+i\pi/4}\mathcal{Z}_1 \approx \frac{\sqrt{2}Uf}{g'}\left(1 + \frac{c^2}{U^2}\right)\mathcal{Y}_1 \approx e^{-i\pi/4}\mathcal{Z}_2.$$

Though details are omitted, the latter clarifies the Lagrangian motion $(\mathcal{Y}_i(t), \mathcal{Z}_i(t))$ to this order: in either layer, viewed from east, a water particle rotates counterclockwise on a temporally expanding ellipse such that $\overline{\mathcal{Y}_i\mathcal{Z}_i} > 0$ $(i = 1, 2)$. This motion reminds us of the observed trajectory and the indirect meridional circulation discussed later.

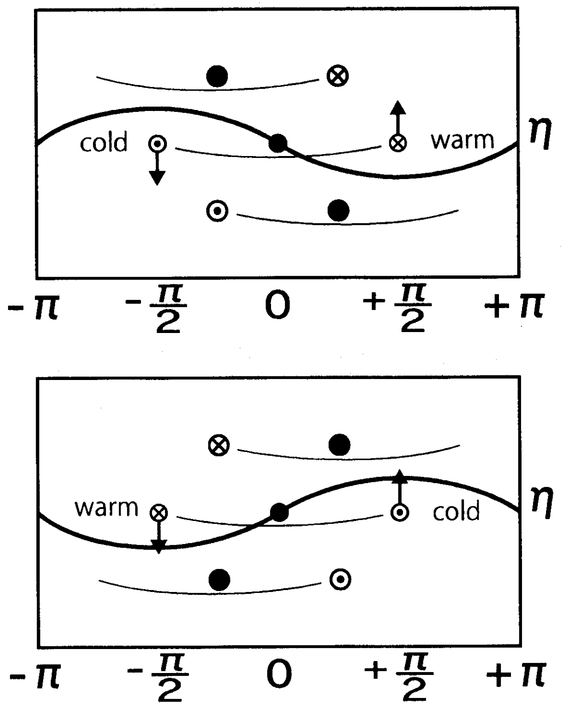

Fig. 2 A schematic view of the spatial configuration of the growing mode (top) and decaying mode (bottom). The origin of the x-axis is taken at a storm center, where the pressure around the interface is minimum. Solid circles indicate the location of pressure minimum in the surface layer, at the interface, and in the bottom layer. Correspondingly three thin curves show the zonal distribution of pressure at the surface layer, interface, and bottom layer. A thick line shows the vertical displacement of the interface. The zonally maximum northward (southward) velocity is denoted by ⊗ (⊙). The upwelling and downwelling motions are shown by upward and downward arrows, respectively. A slanted structure of pressure is evident. Note that growing modes and decayning modes have opposite relative phases. See text for details.

Figure 2 shows the spatial configuration of the growing mode (and decaying mode). In an approximate sense, ϕ and v_ϕ are considered as twice the pressure and northward velocity near the interface of the surface and bottom layers. The low pressure in the surface layer is $\theta \sim \pi/4$ westward (up-current) from the center, while θ eastward (up-current) in the bottom layer. There is a slanted configuration of ψ_i: phase of ψ is the same along a line from the upper west to the lower east. The coldest water is found where the upward displacement of the interface η is maximum; it is $\pi/2$ to the west of the center of the storm. Also we observe the maximum downwelling and the maximum southward displacement in each layer around $\pi/2$ to the west of the storm center; the minimum point of $\mathcal{Y}_1$ is a bit eastward of that of $\mathcal{Y}_2$. Though it might appear curious at first, water downwells around where the upward displacement y of the iterface is the largest. This is because, for $\boldsymbol{k}^2 < F$, positive η is induced mainly by the southward motion, which advects the northern high Y.

The phase relation above has an important meaning. As is evident from the energy equation (23), a necessary condition for growing disturbance is a positive correlation between ψ and $\partial\phi/\partial x$, or $\arg(\varphi) = \arg(\phi) + \pi/2$. This is just the exact phase relation represented by (38). This phase relation has the following physical meaning too. As is well-known and as will be shown later, growing modes transport buoyancy northward and release the potential energy stored by the meridional gradient of buoyancy:

$$\left\langle \varphi\frac{\partial\phi}{\partial x}\right\rangle = 2\left\langle \psi_1\frac{\partial\psi_2}{\partial x}\right\rangle > 0$$

which indicates that the phase of ψ_1 is shifted westward of ψ_2 and φ westward of ϕ for the growing mode. If we rewrite the above condition as

$$0 < \left\langle \varphi\frac{\partial\phi}{\partial x}\right\rangle \sim \left\langle (\psi_2-\psi_1)\frac{\partial(\psi_1+\psi_2)}{\partial x}\right\rangle \sim \left\langle \frac{\partial\psi}{\partial z}\frac{\partial\psi}{\partial x}\right\rangle$$

and considers it a discrete approximation to a continuously stratified model, it means the slanted isophase line of ψ mentioned before.

By the way, from the energy equation (23), we see that the growth rate $\omega_i = k|c|$ should satisfy

$$\omega_i\left\langle \frac{|\nabla\phi|^2+|\nabla\varphi|^2+F\varphi^2}{2}\right\rangle = FU\left\langle \varphi\frac{\partial\phi}{\partial x}\right\rangle. \quad (39)$$

It is easy to confirm (39) by using the complex dispersion relation (34) and (38).

[2] decaying modes

The decaying mode appears in the same wavenumber range $\boldsymbol{k}^2 < F$ as the growing mode. The complex phase velocity is $c = -i|c|$. In short, the decaying mode has the mirror image of the growing mode with respect to the phase as

$$\varphi = -i\frac{|c|}{U}\phi$$
$$\begin{cases} \psi_1 = \dfrac{1-i|c|/U}{2}\phi \\ \psi_2 = \dfrac{1+i|c|/U}{2}\phi \end{cases}, \quad \frac{\psi_1}{\psi_2} = \frac{1-i|c|/U}{1+i|c|/U}.$$

The relative amplitude of the decaying mode takes the opposite sign to that of the growing mode (Fig.2). These properties are anticipated from the symmetry of the canonical equation with respect to the exchange of (Ut, x) with $(-Ut, -x)$ mentioned before.

Note that the opposite sign of relative phase between the growing and decaying mode is associated with positive or negative growth rates, i.e., release or saving of the b.g. potential energy by disturbance.

[3] neutral modes

Next we examine the neutral mode ($|\boldsymbol{k}| > \sqrt{F}$). We have

$$\frac{\varphi}{\phi} = \frac{\pm|c|}{U},$$
$$\begin{cases} \dfrac{\psi_1}{\phi} = \dfrac{1\pm|c|/U}{2} \\ \dfrac{\psi_2}{\phi} = \dfrac{1\mp|c|/U}{2} \end{cases}, \quad \frac{\psi_1}{\psi_2} = \frac{1\pm|c|/U}{1\mp|c|/U}$$

for the eastward (upper sign) and westward propagating (lower sign) waves, respectively. For eastward propagating waves, the amplitude of the surface layer (bottom layer) is larger than that of the bottom layer (surface layer), as is anticipated. As for the phase, we see

$$\arg(\phi) = \arg(\varphi) = \arg(\psi_1) = \arg(\psi_2), \quad (40)$$

for which no energy is supplied to the mode.

For $\boldsymbol{k}^2 \gg F$, the coupling of the surface and lower layer is small enough, so that both layers are expected to behave independently from each other. In fact $|\psi_1|/|\psi_2|$ approaches ∞ and 0 for eastward propagting and westward propagating waves, respectively. That is, the neutural modes become the surface and bottom modes in this limit of short wavelength. Also the vorticity equations become

$$\begin{cases} \dfrac{\partial\psi_1}{\partial(Ut)} + \dfrac{\partial\psi_1}{\partial x} = 0 \\ \dfrac{\partial\psi_2}{\partial(Ut)} - \dfrac{\partial\psi_2}{\partial x} = 0 \end{cases} \Leftrightarrow \begin{cases} \dfrac{\partial\phi}{\partial(Ut)} + \dfrac{\partial\varphi}{\partial x} = 0 \\ \dfrac{\partial\varphi}{\partial(Ut)} + \dfrac{\partial\phi}{\partial x} = 0 \end{cases},$$

where the left-hand side equations show the independent evolution of surface and bottom modes. The disturbance in the surface layer moves eastward by the advection of the b.g. current $U_1 = U$ independently of the lower layer disturbance. The reversed statement is true of the bottom mode. Figure 3 shows the spatial structure of the neutral modes. One is an eastward propagating wave and the other a westward propagating wave.

Thus the canonical equation provides a quite systematic and simple picture of the spatial structure of growing (decaying) and neutral modes.

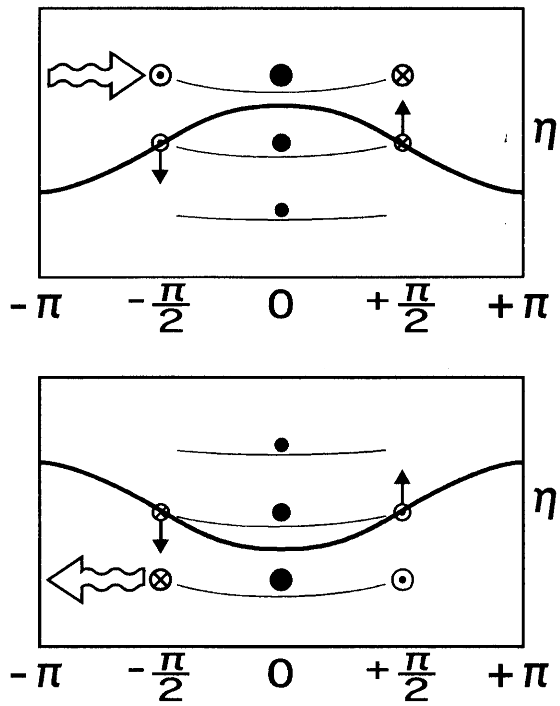

Fig. 3 The same as Fig. 2 except that it is for the neutral mode ($\boldsymbol{k}^2 \gg F$): (top) eastward propagating wave and (bottom) westward propagating wave. The amplitude is larger in the surface (bottom) layer like the surface mode (bottom mode). A broad wavy arrow indicates the propagation direction of the neutral modes. The wave propagates eastward (westward) as if it is advected by the eastward (westward) background current in the surface layer (bottom layer). The displacement of the interface takes a profile opposite to (similar to) the pressure around the interface. Note that w has the same phase relation as in the growing mode.

4.3 mechanism of instability

Our primary object is to understand the mechanism of baroclinic instability by means of the canonical equation in a most simplified situation. In particular, it is shown why the radius of deformation gives a critical scale that separates growing (decaying) modes from neutral modes.

[1] elliptic and hyperbolic equations in the limits of long- and short-waves

Since $(\nabla^2 - F)$ in the first term of (19) is negative definite, the sign of the second term of (19) is important. As was discussed before, the second term of (19) expresses the production of baroclinic vorticity by [1] the zonal advection of disturbance barotropic vorticity $\nabla^2\phi$ by baroclinic b.g. current and [2] the meridional advection of the baroclinic b.g. potential vorticity (b.g. thickness term or CIPT β) by the barotropic northward disturbance velocity v_ϕ. The two effects tend to cancel each other, because $\nabla^2 \sim -\boldsymbol{k}^2 < 0$ and $F > 0$. Thus we see the relative magnitude of ∇^2 and F becomes critical whether the disturbance grows (decays), or neutral.

For large scale disturbance $|\nabla^2| < F$, the latter dominates the former. That is, the CIPT β effect is larger than the advection of the disturbance barotropic vorticity. It will enhance the initial disturbance of φ and disturbance will grow. In fact, in the limit of $\boldsymbol{k}^2 \ll F$, the canonical equation is approximated by

$$\frac{\partial^2\varphi}{\partial(Ut)^2} + \frac{\partial^2\varphi}{\partial x^2} = 0.$$

This is an elliptic partial differential equation, or the Laplace equation for two independent variables Ut and x. That is, $\varphi(Ut, x; y)$ is a harmonic function in this notation. If $\varphi = A(Ut; y)e^{ikx}$, then $A(Ut; y)$ should be $e^{\pm kUt}$, which property is a well-known atttribute of harmonic functions. Therefore φ grows or decay with time.

On the other hand, disturbance of short wavelengths follows a quite different evolution. When $|\nabla^2|$ dominates F, baroclinic zonal advection of baroclinic vorticity is dominant, which reduces ϕ and works as a restoring effect. Then disturbance oscillates in time, or propagates eastward or westward. In fact, the limiting form of the canonical equation becomes

$$\frac{\partial^2\varphi}{\partial(Ut)^2} - \frac{\partial^2\varphi}{\partial x^2} = 0,$$

i.e., the one-dimensional wave equation.

The present explanation translates each mathematical property to its physical counterpart, so that there can be no ambiguity. The mechanism is universal in mathematical sense, irrespective of the details of the dynamics in concern. It is obvious that the disturbance may grow when the evolution equation becomes the Laplace equation for independent variables of Ut, x.

[2] interpretation in more physical terms

A more physical interpretation is obtained if we use η and v_ϕ instead of ϕ and φ. This interpretation is based on the equations

$$\begin{cases} \dfrac{\partial v_\phi}{\partial(Ut)} = \dfrac{g'}{f}\dfrac{\partial^2\eta}{\partial x^2} \\ \dfrac{g'}{f}\dfrac{\partial}{\partial(Ut)}(\nabla^2 - F)\eta = (\nabla^2 + F)v_\phi \end{cases}.$$

Suppose a depression of interface $[-\eta] > 0$ occurs near $kx = \pi/2$, or consider $\eta \sim \cos(kx + \pi/2) = -1$; see the top panel of Figure 2. Since $\partial^2\eta/\partial x^2 \sim -k^2\eta > 0$ there, the first equation shows that the northward barotropic velocity v_ϕ increases there. The feedback to $[-\eta]$ by the increased v_ϕ is described by the second equation. Evidently $(\nabla^2 - F)\eta$ in the left-hand side has the same sigh as $[-\eta]$ irrespective of the scale of disturbance. In contrast, the sign of $(\nabla^2 + F)v_\phi$ in the right hand depends on the scale of disturbance.

Two cases are to be examined separately therefore. We first consider a large scale of depression of the interface ($[-\eta] > 0$) such that $|\nabla^2| < F$, where $(\nabla^2 + F)v_\phi$ has the same sign as v_ϕ. Then the increased v_ϕ enhances the baroclinic potential vorticity (thickness term), which is approximated by $\frac{g'F}{f}[-\eta]$. Thi is, the initial depression of interface $[-\eta]$ increases further by this positive feedback. That explains why the initial small depression of the interface grows with time.

When the scale of initial depresion of the interface is smaller than the baroclinic radius of deformation, however, $|\nabla^2 v_\phi|$ is larger than $|Fv_\phi|$, and the net effect is reversed; the initial depression is restored to the original level. This is the scheme of oscillation. Dsturbance therefore does not grow, but oscillates, or propagates zonally.

[3] interpretation based on metaphorical systems for long- and short-scale limits

The reason why disturbance grows (decays) or stays neutral according as its horizontal scale is given by the following metaphors as well. The limits of long waves and short waves respectively become

$$\begin{cases} \dfrac{\partial\phi}{\partial(Ut)} + \dfrac{\partial\varphi}{\partial x} = 0 \\ \dfrac{\partial\varphi}{\partial(Ut)} \pm \dfrac{\partial\phi}{\partial x} = 0 \end{cases} \leftrightarrow \begin{cases} \dfrac{\partial p}{\partial t} + \dfrac{\partial u}{\partial x} = 0 \\ \dfrac{\partial u}{\partial t} \pm \dfrac{\partial p}{\partial x} = 0 \end{cases}$$

where, + represents the short-wave limit, and − the long-wave limit. The right-hand side expresses the governing equations of the metaphorical system. It is obtained by replacing ϕ, φ on the left-hand side by p and u. In the metaphorical systems, we may consider p as the pressure, or the surface elevation, and u the zonal velocity of shallow water waves, where the depth and gravitational acceleration are assumed to be unity.

For the short-wave limit, the metaphorical system holds usual physics of water waves. The result is the ordinary wave equation for nondispersive neutral waves. On the other hand, for the long-wave limit, the metaphorical system has a curious property: according to the lower equation, high p does not pushes out, but pulls in water. Such a system must be unstable, of course.

[4] limit of long-waves and analytic functions

In the long-wave limit, also the baroclinic vorticity equation is reduced to

$$\frac{\partial\varphi}{\partial(Ut)} = \frac{\partial\phi}{\partial x},$$

which implies that barotropic northward current under the CIPT β pushes down the interface to cause baroclinic vertical stretching. We see from this, in particular, that $w_1 = w_2 = 0$ by virtue of (22); it merely reflects the northward advection of the b.g. interface Y by barotropic northward current v_ϕ.

Putting the reduced vorticity equations of barotropic and baroclinic modes together, we have

$$\frac{\partial\varphi}{\partial(Ut)} = \frac{\partial\phi}{\partial x}, \quad \frac{\partial\phi}{\partial(Ut)} = -\frac{\partial\varphi}{\partial x}. \tag{41}$$

Addition and subtraction of the two equations of (41) yield

$$\frac{\partial\psi_1}{\partial(Ut)} = \frac{\partial\psi_2}{\partial x}, \quad \frac{\partial\psi_2}{\partial(Ut)} = -\frac{\partial\psi_1}{\partial x}. \tag{42}$$

Both (41) and (42) are identified with the Cauchy-Riemann condition of analytical functions. We observe therefore both $\varphi + i\phi$ and $\psi_1 + i\psi_2$ are analytical functions of $(Ut) + ix$. It follows directly from this that every Fourier component of e^{ikx} should be

$$\begin{cases} \varphi + i\phi \sim e^{\pm k([Ut]+ix)} \\ \psi_2 + i\psi_1 \sim e^{\pm k([Ut]+ix)} \end{cases}.$$

The positive and negative signs indicate the growing and decaying disturbance, respectively. Any mode therefore decays or grows with time, so long as it has a zonal dependence like $e^{\pm ikx}$. Only from this analytical property of $\psi + i\phi$ and $\psi_1 + i\psi_2$, it is easy to obtain the phase relation

$$\arg(\varphi) - \arg(\phi) = \arg(\psi_1) - \arg(\psi_2) = \pm\frac{\pi}{2}$$

for growing modes (upper sign) and decaying modes (lower sign) of e^{ikx}, respectively. That is, for growing modes, the minimum of ψ_1 (ψ_2) is to the west (east) of the minimum of ϕ; the minimum of φ is to the west of the minimum of ϕ, which we saw in the previous subsection from the direct calculation.

[4] further investigation on the long-wave and short-wave limits

Long-wave and short-wave limits have been argued in a few ways in this subsection. It is worth mentioning that they have been suggested by the alternative forms of canonicl equations (20) and (21). If $|\nabla^2| \gg F$, for instance, the canonical equation immediately reduces to the wave-equation. If $\nabla^2 \ll F$, on the other hand, (21) suggests an analytical relation of stream functions.

Finally energetics and conservation law of (23) and (24) are discussed in these limits. In the long-wave limit, we see

$$\frac{\partial}{\partial t}\frac{\langle \phi^2 - \psi^2 \rangle}{2} = 0,$$

which shows that the orbit on the $(\langle\phi^2\rangle, \langle\varphi^2\rangle)$-plane is a hyperbolic curve. That is, the energy $\langle\phi^2\rangle + \langle\varphi^2\rangle$ will go away to ∞ for $t \to \pm\infty$. On the other hand, the short-wave limit has a contrasted feature. The energy equation becomes

$$\frac{\partial}{\partial t}\frac{\langle \phi^2 + \psi^2 \rangle}{2} = 0,$$

so that the orbit is a circle on the $(\langle\phi^2\rangle, \langle\varphi^2\rangle)$-plane.

4.4 artificial surface and bottom modes and their resonance as Rossby waves

When we rewrite (2) as

$$\begin{cases} \left(\frac{\partial}{\partial t} + U\frac{\partial}{\partial x}\right)\left(\nabla^2 - \frac{F}{2}\right)\psi_1 + FU\frac{\partial \psi_1}{\partial x} \\ \quad = -\left(\frac{\partial}{\partial t} + U\frac{\partial}{\partial x}\right)\frac{F}{2}\psi_2 \\ \left(\frac{\partial}{\partial t} - U\frac{\partial}{\partial x}\right)\left(\nabla^2 - \frac{F}{2}\right)\psi_2 - FU\frac{\partial \psi_2}{\partial x} \\ \quad = -\left(\frac{\partial}{\partial t} - U\frac{\partial}{\partial x}\right)\frac{F}{2}\psi_1 \end{cases},$$

it reads that the surface or bottom layer is forced by the right-hand side due to the disturbance of the other layer. The right-hand sides are interpreted as the interaction or coupling with the other layer.

If the stream function of bottom (surface) layer is forced to be zero artificially, the evolution equation becomes that of the surface (bottom) layer alone. We call such motions as the *artificial* surface (bottom) mode; the artificial surface mode is described as Rossby waves due to the CIPT β advected by the b.g. eastward current U_1. The dispersion relations for artificial surface and bottom modes agree with those for the corresponding Rossby waves:

$$\begin{cases} c_1 - U = -\dfrac{FU}{k^2 + F/2} = -\dfrac{\beta_t}{k^2 + F/2} < 0 \\ c_2 + U = +\dfrac{FU}{k^2 + F/2} = +\dfrac{\beta_t}{k^2 + F/2} > 0 \end{cases}. \quad (43)$$

The artificial surface (bottom) mode propagates westward (eastward) in the frame of reference moving with U_1 (U_2),

However, if we rewrite (43) as

$$-1 \le \frac{c_2}{U} = -\frac{c_1}{U} = \frac{F/2 - k^2}{F/2 + k^2} \le 1,$$

we observe that $|c_1|$ and $|c_2|$ are within the range of U in the reference system moving with the average velocity ($U_\phi = 0$) of the surface and bottom layers.

Baroclinic instability is interpreted sometimes in terms of the resonance between those Rossby waves described above as the artificial surface and bottom modes. Let us examine how adequate the interpretaion of resonance is.

We see that resonance occurs only when $c_1 = c_2 = 0$, which condition is satisfied at wavenumber $k_{res}/\sqrt{F} = 1/\sqrt{2} = 0.707$, roughly equal to $k_{\max}/\sqrt{F} \approx \sqrt{\sqrt{2} - 1} = 0.644$ in the previous subsection. Such a close correspondence should be considered accidental, however. The argument based on the resonance between those artificial surface and bottom mode is not correct, although it may be suggestive. Note that the artificial surface and bottom modes neither satisfy the full equations for the two-layer system so far discussed nor represent the correct spatial structure of growing modes. Artificial modes and their resonance may be used as a heuristic tool.

4.5 secondary transport by growing modes and energetics

In this subsection we are concerned with the growing mode of $l = \pi/L$ only.

[1] meridional transport of water and buoyancy

The deviation of the thickness of each layer δh_i due to disturbance is expressed by η as

$$-\delta h_1 = \delta h_2 = \eta = -\frac{f}{g'}\varphi = -\frac{f}{g'}\frac{|c|}{U}i\phi,$$

where

$$\phi = \Re\left[a_\phi e^{ikx + \omega_i t}\cos ly\right], \quad (44)$$

a_ϕ being the initial amplitude of ϕ. Then the zonal mean of the northward transport of water in the surface layer due to the growing mode is

$$\begin{aligned} \overline{v_1\delta h_1} &= -\overline{\frac{\partial \psi_1}{\partial x}\eta} = \frac{f}{2g'}\overline{\frac{\partial(\phi + \varphi)}{\partial x}\varphi} \\ &= \frac{f\omega_i}{4g'U}|a_\phi|^2 e^{2\omega_i t}\cos^2 ly = A\cos^2 ly > 0, \end{aligned}$$

where

$$A \equiv \frac{f\omega_i}{4g'U}|a_\phi|^2 e^{2\omega_i t} > 0$$

is the time-dependent magnitude of $\overline{v_1\delta h_1}$. Obviously we have

$$\overline{v_2\delta h_2} = \overline{\frac{\partial \psi_2}{\partial x}\eta} = -\overline{v_1\delta h_1} < 0.$$

Thus the growing mode transports water northward in the upper layer and southward in the lower layer.

The quantity defined by

$$v_{\parallel;i} \equiv \frac{\overline{v_i\delta h_i}}{h_i} \quad (i = 1, 2)$$

is the eddy-induced transport velocity of Gent et al. (1995). We observe

$$v_{\parallel;i} \cdot \frac{dh_i}{dy} < 0 \quad (i = 1, 2), \tag{45}$$

which shows that the eddy-induced transport velocity by baroclinic instability is directed from the thicker toward the thinner regions of the layer in concern. We will discuss it and Gent-McWilliams parameterization (Gent and McWilliams 1991) later. For this case we have

$$-v_{\parallel;2} = v_{\parallel;1} = \frac{A}{h_1} \cos^2 ly. \tag{46}$$

The water transport above is related with buoyancy transport as

$$g'\overline{v_1 \delta h_1} - g'\overline{v_2 \delta h_2} = 2g' h_1 \overline{v}_{\parallel;1} = 2g' A \cos^2 ly.$$

In either layer, buoyancy or heat is transported northward by growing modes of disturbance.

Thirdly there is no mean meridional transport of zonal momentum. That is, the Reynolds stress vanishes. If we use (11), it is easy to confirm that

$$\overline{u_i v_i} = -\overline{\frac{\partial \psi_i}{\partial y} \frac{\partial \psi_i}{\partial x}} = 0 \quad (i = 1, 2), \tag{47}$$

because u_i and v_i is out of phase for disturbance of the form of (44); there is no horizontal shear in the b.g. current.

[2] form drag at the interface and energetics

The force exerted by the surface layer to the bottom layer is expressed as

$$\begin{aligned} \overline{p_1 \frac{\partial \eta}{\partial x}} &= \overline{f\psi_1 \frac{\partial \eta}{\partial x}} = -f\overline{\frac{\partial \psi_1}{\partial x}\eta} = f\overline{v_1 \delta h_1} \\ &= f h_1 v_{\parallel;1} > 0. \end{aligned} \tag{48}$$

The force exerted by the bottom layer to the surface layer is expressed as

$$-\overline{p_2 \frac{\partial \eta}{\partial x}} = -\overline{f\psi_2 \frac{\partial \eta}{\partial x}} = f\overline{\frac{\partial \psi_1}{\partial x}\eta} = -f\overline{v_1 \delta h_1} < 0,$$

which is opposite to (48) of course. Thus the surface layer is pushed westward by the bottom layer via form drag and vice versa.

In association with the form drag, let us consider the work done by the form drag to the mean flow. The interfacial form drag exerts a force opposite to the b.g. current ($U_1 = U$, $U_2 = -U$), on each layer. Then the mean work done by the form drag to the mean flow may be estimated as

$$-2fU\overline{v_1 \delta h_1} = -2fUA \cos^2 ly < 0,$$

which would reduce the kinetic energy of the b.g. current. This loss of kinetic energy of the mean current is connected to the release of the b.g. potential energy associated with the meridional transport of buoyancy as

$$2fU \left\langle p_1 \frac{d\eta}{dx} \right\rangle = FUH \left\langle \frac{\partial \phi}{\partial x} \varphi \right\rangle, \tag{49}$$

the right hand of which is exactly the same as the energy supply to the growing mode. It is to be born in mind that the right-hand is the total energy supply to disturbance. The transfer of energy to the growing mode of disturbance is expressed in either way: the mechanical energy supply via interfacial form drag or release of potential energy by the eddy-tansport of buoyancy.

Baroclinic instability releases the potential energy of the b.g. field (meridional gradient of buoyancy), which is expressed as the northward transport of buoyancy. That is associated with the transport of zonal momentum from the surface to lower layers. From the energetic viewpoint, this assertion seems valid as we have seen above in (49).

[3] note on the momentum balance

The balance of zonal momentum is more complicated than the argument of interfacial form drag suggests. There is a subtle issue peculiar to the rotating system, where the Coriolis force complicates the dynamics and the conservation of momentum is not guaranteed *a priori*.

From (48), one may suppose that the growing mode transfers the zonal momentum of the surface layer to the lower layer by means of form drag. According to (48) this transfer is positive at every y. This does not imply, however, that the lower layer increases its zonal mean zonal momentum as much as (48).

The form drag to the surface layer is associated with the northward water transport $\overline{v_1 \delta h_1}$. The Coriolis force acting on this northward transport is eastward and cancels out the form drag, which is westaward. Therefore the form drag itself does not corresponds to the reduction of the zonal mean zonal current in the surface layer. We must be more careful.

Although we have been concerned with the Q-G dynamics primarily, it is necessary here to take into consideration the *ageostrophic* component of velocity in order to discuss the zonal mean zonal momentum. We define

$$\begin{cases} \boldsymbol{u}_{ge;i} \equiv -\dfrac{1}{f} \lhd p_i \\ \boldsymbol{u}_{ag;i} \equiv \boldsymbol{u} - \boldsymbol{u}_{ge;i} \end{cases} \quad (i = 1, 2),$$

where the suffixes ge and ag denote the geostrophic and ageostrophic components, respectively. Note that $\overline{v}_i = \overline{v}_{ag;i}$, because the zonal mean of the geostrophic zonal current vanishes identically:

$$\overline{v}_{ge;i} = -\frac{1}{f}\overline{\frac{\partial p_i}{\partial x}} = 0.$$

The deceleration of the zonal mean eastward current of the surface layer would be due to negative $f\overline{v}_{ag;1}$, not directly by the interfacial form drag in this viewpoint.

In the next subsection we examine how $\overline{v}_{ag;i}$ is determined by perturbation analysis.

4.6 secondary meridional circulation

Now let us investigate the secondary meridional circulation induced by a growing mode as was investigated in the last subsection. To take into account the ageostrophic component of v, we expand the current field in a series of terms of the increasing orders with respect to small amplitude. For example, the horizontal current field is expressed as

$$\begin{aligned} \boldsymbol{u} &= \boldsymbol{u}^{(0)} + \boldsymbol{u}^{(1)} + \boldsymbol{u}^{(2)} + \cdots \\ &= \boldsymbol{U} + \boldsymbol{u}^{(1)} + \boldsymbol{u}^{(2)} + \cdots, \end{aligned}$$

where the shoulder suffix (j) indicates that the variable is the j-th order one. The zeroth-order represents the b.g. field and the first-order is assumed to be the geostrocphic current of the growing mode:

$$\begin{aligned} \phi^{(1)} &= a_\phi e^{ik(x-c(\boldsymbol{k})t)} \cos ly, \quad \varphi^{(1)} = i(|c|/U)\phi^{(1)}, \\ \psi_1^{(1)} &= (1 + i|c|/U)\frac{a_\phi}{2} e^{ik(x-c(\boldsymbol{k})t)} \cos ly, \\ \psi_2^{(1)} &= (1 - i|c|/U)\frac{a_\phi}{2} e^{ik(x-c(\boldsymbol{k})t)} \cos ly, \\ \boldsymbol{u}_i^{(1)} &= - \lhd \psi_i^{(1)} \qquad (i = 1, 2). \end{aligned}$$

Also we have

$$\begin{aligned} \psi_i &= \psi_i^{(1)} + \psi_i^{(2)} + \cdots, \\ \eta &= Y + \eta^{(1)} + \eta^{(2)} + \cdots, \\ \eta^{(1)} &= -\frac{f}{g'}\varphi^{(1)} = -\frac{f|c|}{g'U} i\phi^{(1)}. \end{aligned}$$

Then the discussion in the previous subsection yields

$$\begin{cases} -h_2 v_{\|;2}^{(2)} = h_1 v_{\|;1} \equiv -\overline{\eta^{(1)} v_1^{(1)}} = A\cos^2 ly \\ v_{\|;2}^{(2)} \approx -v_{\|;1}^{(2)} \\ v_{\|;\varphi}^{(2)} \equiv v_{\|;1}^{(1)} - v_{\|;2}^{(1)} \approx 2v_{\|;1}^{(1)} \\ v_{\|;\phi}^{(2)} \equiv v_{\|;1}^{(1)} + v_{\|;2}^{(1)} \approx 0, \end{cases}.$$

[1] second-order zonal mean current $\overline{u}_{ge}^{(2)}$ and $\overline{v}_{ag}^{(2)}$

There must be *ageostrophic* velocity component for the generation of the zonal mean zonal current $\overline{u}$ in the second-order. We have the zonal mean of the zonal momentum equation as

$$\frac{\partial \overline{u}_{ge;i}^{(2)}}{\partial t} - f\overline{v}_{ag;i}^{(2)} \approx -\frac{\partial}{\partial y}\left(\overline{u_{ge;i}^{(1)} v_{ge;i}^{(1)}}\right) = 0 \quad (i = 1, 2)$$

by virtue of (47). It follows immediately that

$$\overline{v}_{ag;i}^{(2)} = \frac{1}{f}\frac{\partial \overline{u}_{ge;i}^{(2)}}{\partial t};$$

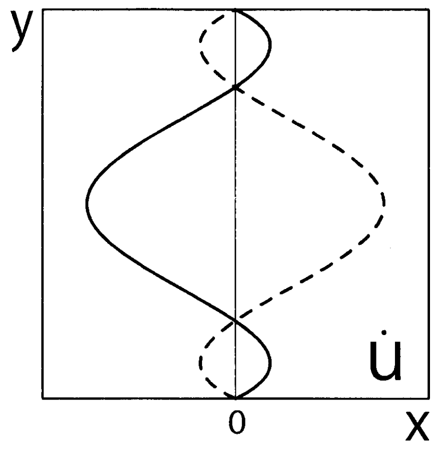

Fig. 4 Secondary acceleration of eastward zonal mean zonal current in the surface layer (solid line) and bottom layer (dashed line) induced by growing modes for the case of $\sqrt{F}L/2 = 3$. It corresponds to $f\overline{v}_{ag}^{(2)}$, so that it vanishes along the southern and northern boundaries. See text for details.

the acceleration of $\overline{u}_{ge;i}$ is due to $f\overline{v}_{ge;i}$ rather than the form drag itself. The side boundary condition of $\overline{u}_{ge;i}^{(2)}$ becomes

$$\frac{\partial \overline{u}_{ge;i}^{(2)}}{\partial t} = 0 \quad \text{on} \quad y = 0,\ L. \tag{50}$$

The equation of mass balance gives

$$\begin{cases} +\dfrac{\partial \overline{\eta}^{(2)}}{\partial t} = \dfrac{\partial}{\partial y}\left(h_1 v_{\|;1}^{(2)} + h_1 \overline{v}_{ag;1}^{(2)}\right) \\ -\dfrac{\partial \overline{\eta}^{(2)}}{\partial t} = \dfrac{\partial}{\partial y}\left(h_2 v_{\|;2}^{(2)} + h_2 \overline{v}_{ag;2}^{(2)}\right) \end{cases}.$$

Since $h_1 v_{\|;1}^{(2)} = -h_1 v_{\|;2}^{(2)}$, addition of the above two equations gives the mass conservation condition of the secondary circulation.

$$0 = h_1 \overline{v}_{ag;1}^{(2)} + h_2 \overline{v}_{ag;2}^{(2)}.$$

On the other hand subtraction yields

$$\begin{aligned} 2\frac{\partial \overline{\eta}^{(2)}}{\partial t} &\approx h_1 \frac{\partial v_{\|;\varphi}^{(2)}}{\partial y} + h_1 \frac{\partial \overline{v}_{ag;\varphi}^{(2)}}{\partial y} \\ &\approx h_1 \frac{\partial v_{\|;\varphi}^{(2)}}{\partial y} + \frac{h_1}{f}\frac{\partial}{\partial t}\frac{\partial \overline{u}_{ge;\varphi}^{(2)}}{\partial y} \\ &= h_1 \frac{\partial v_{\|;\varphi}^{(2)}}{\partial y} - \frac{h_1}{f}\frac{\partial^2}{\partial y^2}\frac{\partial \overline{\varphi}^{(2)}}{\partial t}, \end{aligned} \tag{51}$$

where $\overline{v}_{ge;\varphi}^{(2)} \equiv \overline{v}_{ge;1}^{(2)} - \overline{v}_{ge;1}^{(2)}$. Because $g'\overline{\eta}^{(2)}/f = -\varphi^{(2)} = \psi_2^{(2)} - \psi_1^{(2)}$ and $F = f^2/g'H = 2f^2/g'h_1 = 2f^2/g'h_2$, (51) is written as

$$\left(\frac{\partial^2}{\partial y^2} - F\right)\frac{\partial \overline{\varphi}^{(2)}}{\partial t} = f\frac{\partial v_{\|;\varphi}^{(2)}}{\partial y}, \tag{52}$$

which represents the balance of the zonal mean second-order Q-G baroclinic vorticity; it is the curl of the baroclinic forcing of interfacial drag that drives the zonal mean boroclinic potential vorticity in the second-rder.

Differentiation with respect to y gives

$$\left(\frac{\partial^2}{\partial y^2}-F\right)\frac{\partial\overline{u}_{\varphi}^{(2)}}{\partial t}=-f\frac{\partial^2\overline{v}_{\|;\varphi}^{(2)}}{\partial y^2}=\frac{4fl^2A}{h_1}\cos 2ly,$$

the solution of which becomes

$$f\overline{v}_{ag;\varphi}^{(2)}=\frac{\partial}{\partial t}\overline{u}_{\varphi}^{(2)}$$
$$=-\frac{f}{2H}\frac{4l^2}{F+4l^2}A\left[\cos 2ly+\frac{\cosh\sqrt{F}y}{\cosh\sqrt{F}L/2}\right].\ (53)$$

The $\cosh\sqrt{F}y$ term of (53) expresses the correction to satisfy the side boundary condition (50).

Figure 4 shows $\partial\overline{u}_i^{(2)}/\partial t$, the secondary acceleration of the b.g. zonal mean zonal current induced by growing modes for the case of $\sqrt{F}L/2=3$. It is the same as $f\overline{v}_{ag;i}^{(2)}$. The left-hand side of (53) is negative in central latitudes of the channel, and a little positive near the side boundaries.

This secondary circulation decelerates the zonal mean zonal current of the surface layer in the principal part, whereas somewhat accelerates close to the side boundaries. Nevertheless (53) may be associated with the assertion that boroclinic instability transfers eastward momentum of the surface layer to the bottom layer. This is partly because the zonal mean current surely decreases in the principal area of the surface layer, and partly because the total amount of surface zonal momentum decreases with time as follows.

Integrated over the width of the ocean, the rate of temporal change of baroclinic zonal momentum becomes

$$2Hf\int_{-L/2}^{L/2}\overline{v}_{ag;\varphi}^{(2)}\,dy=2H\frac{\partial}{\partial t}\int_{-L/2}^{L/2}\overline{u}_{\varphi}^{(2)}\,dy$$
$$=-\frac{2f}{\sqrt{F}}\frac{4l^2}{F+4l^2}A\tanh\frac{\sqrt{F}L}{2}<0,$$

which is due to the correction term. Without the correction term or when $l^2\ll F$, it is zero or quite small in comparison with the momentum transfer associated with the form drag $fAL/2$.

This is the *net* change of the zonal mean zonal current due to baroclinic instability within the framework of the secondary perturbation argument. The picture much differs from that envisaged straightforward from the interfacial form drag; in a rotating system, acceleration of current follows the exerted force neither in magnitude nor in direction. Recall the North Equatorial Counter Current flowing eastward under the trade wind blowing westward.

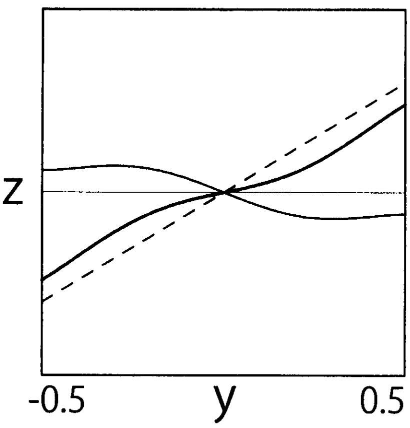

Fig. 5 A schematic distribution of secondary zonal mean vertical displacement of the interface as a function of y for the case of $\sqrt{F}L/2=3$. A thin line shows the secondary one $\overline{\eta}^{(2)}$, a dashed line the b.g. Y, and a thick solid line the sum $Y+\eta^{(2)}$, though they are largely exaggerated.

The secondary change of the vertical displacement of the interface is expressed as

$$\frac{\partial\overline{\eta}^{(2)}}{\partial t}=\int_0^y\frac{f}{g'}\frac{\partial\overline{u}_{ge;\varphi}^{(2)}}{\partial t}\,dy$$
$$=-\frac{4l^2}{F+4l^2}\frac{\sqrt{F}A}{2}\left(\frac{\sqrt{F}}{2l}\sin 2ly+\frac{\sinh\sqrt{F}y}{\cosh\sqrt{F}L/2}\right).$$

It tends to flatten the basic meridional slope of the interface Y in central latitudes, whereas slightly sharpens near the side boundaries; $\partial\eta^{(2)}/\partial t\neq 0$ at $y=\pm L/2$ due to the correction term. Figure 5 gives an exmaple of the secondary change of the vertical displacement of the interface $\eta^{(2)}$ and the resulting meridional slope of the interface $Y+\eta^{(2)}$ for the case of $\sqrt{F}L/2=3$.

[2] stream function of the meridional circulation

Using the above distribution of $\overline{v}_{ag;\varphi}^{(2)}$ we can compose the stream function $\chi=\chi^{(2)}$ for the meridional circulation; $\chi^{(2)}$ is defined so that

$$\overline{w}_{ag}^{(2)}=+\frac{\partial\chi^{(2)}}{\partial y},\quad \overline{v}_{ag}^{(2)}=-\frac{\partial\chi^{(2)}}{\partial z}.$$

If we know $\overline{v}_{ag;\varphi}^{(2)}$, it is easy to deduce that

$$\chi^{(2)}=\frac{h_1-|z+h_1|}{2}\overline{v}_{ag;\varphi}^{(2)}$$

by using the conditions

$$\begin{cases}\dfrac{\partial\overline{v}_{ag}^{(2)}}{\partial z}=0\quad\text{for}\quad -4H<z<0\\ \qquad\text{except at the interface } z=-h_1\\ \overline{w}_{ag}^{(2)}=0\quad\text{at}\quad z=0,\ -4H\end{cases}$$

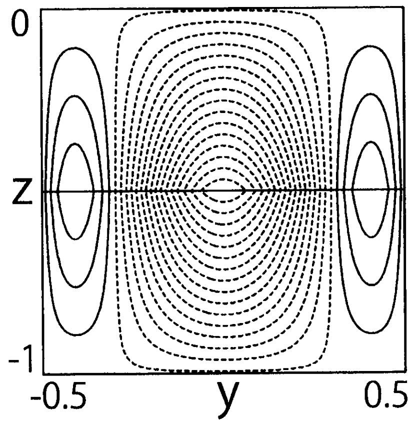

Fig. 6 Secondary meridional circulation induced by growing modes for the case of $\sqrt{F}L/2 = 3$: contours of the stream function $\chi^{(2)}$ on the y,z-plane with the abscissa y and ordinate z. A solid line shows a positive contour and a dotted line a negative one. The contour interval is 1/10 of the absolutely maximum value.

The vertical velocity is linear with respect to z in each layer

$$\overline{w}^{(2)}_{ag} = \frac{h_1 - |z + h_1|}{2}\frac{\partial \overline{v}^{(2)}_{ag;\varphi}}{\partial y},$$

where (53) is to be substituted into $\overline{v}^{(2)}_{ag;\varphi}$.

Figure 6 shows the secondary meridional circulation by contours of the stream function $\chi^{(2)}$ for the case of $\sqrt{F}L/2 = 3$. There are three cells, among which the central counterclockwise cell is called the indirect circulation. It must be born in mind, however, that the three-cell circulation above may not explain the real three-cell structure of atmosphere. Nonlinear processes of growing modes are believed to play a more critical role than in the pertubation analysis.

4.7 baroclinic instability, G-M parameterization and diffusive stretching

The most important role of baroclinic instability has been said to moderate the large-scale meridional gradient of buoyancy. Apropos of this effect of baroclinic instability, a brief note is added here: intimate relations are pointed out among G-M parameterization (Gent and McWilliams 1991), diffusive stretching (Masuda and Uehara 1992), and vertical viscosity.

Let us consider a slow and locally averaged dynamics, in which baroclinic instability is treated as eddies of smaller scales in time and space. Our concern is the average over several times of such events. In this subsection $\overline{\bullet}$ means the local average of $\bullet$, not the zonal mean. Let $\psi = \psi(\boldsymbol{x}, z, t)$ be the three-dimensional Q-G stream function, $b \equiv -g(\rho - \rho_r)/\rho_r$ buoyancy, and $N^2 = N^2(z) \equiv d\langle b\rangle/dz$ the b.g. buoyancy frequency.

In the quasi-steady (statistically steady) state of this situation, we have the hydrostatic equation

$$\overline{b} = \frac{\partial \overline{p}}{\partial z} = f\frac{\partial \overline{\psi}}{\partial z} \tag{54}$$

and the thermal-wind relation (relation of baroclinic geostrophy)

$$f\frac{\partial \overline{\boldsymbol{u}}}{\partial z} = -\triangleleft\, \overline{b}. \tag{55}$$

In association with the baroclinic instability, we consider eddy-induced horizontal diffusion of density, where the diffusion coefficient is denoted by $\kappa_h = \kappa_h(z)$; it represents G-M parameterization and diffusive stretching. Likewise let $\nu_v = \nu_v(z)$ be the vertical eddy viscosity, representing the interfacial form drag. Their mutual relations are argued below.

[1] baroclinic instability and G-M parameterization

From (45) we see that, within a layer of the same density, eddies arising from baroclinic instability transport water from thicker to thinner regions. This is the content of G-M parameterization (Gent and McWilliams 1995). It is expressed in a vector form as

$$h_i \boldsymbol{u}_{\|;i} \approx -\kappa_G \nabla h_i \qquad \text{(thickness diffusion)},$$

where κ_G mimics the diffusion coefficient. In a continuously stratified case it is rewritten as

$$\frac{\partial \overline{h}}{\partial \overline{b}}\boldsymbol{u}_{\|} \approx -\kappa_G \nabla_{\overline{b}}\frac{\partial \overline{h}}{\partial \overline{b}}, \tag{56}$$

where $\nabla_{\overline{b}}$ is the horizontal gradient operator with $\overline{b}$ fixed and $\overline{h}$ is identified with z. Since

$$\nabla_{\overline{b}}\overline{h} = -\frac{\partial \overline{h}}{\partial \overline{b}}\nabla \overline{b},$$

(56) becomes

$$\begin{aligned}\frac{\partial \overline{h}}{\partial \overline{b}}\boldsymbol{u}_{\|} &\approx -\kappa_G \frac{\partial}{\partial \overline{b}}\nabla_{\overline{b}}\overline{h} = \kappa_G \frac{\partial}{\partial \overline{b}}\left(\frac{\partial \overline{h}}{\partial \overline{b}}\nabla \overline{b}\right) \\ &= \kappa_G(z)\frac{\partial \overline{h}}{\partial \overline{b}}\frac{\partial}{\partial z}\left(\frac{\partial \overline{h}}{\partial \overline{b}}\nabla \overline{b}\right) \\ &= \frac{\partial \overline{h}}{\partial \overline{b}}\frac{\partial}{\partial z}\left(\frac{\kappa_G}{\frac{\partial \overline{b}}{\partial \overline{h}}}\nabla \overline{b}\right) = \frac{\partial \overline{h}}{\partial \overline{b}}\frac{\partial}{\partial z}\left(\frac{\kappa_G}{N^2}\nabla \overline{b}\right).\end{aligned}$$

We find

$$\boldsymbol{u}_{\|} \approx \frac{\partial}{\partial z}\left(\frac{\kappa_G}{N^2}\nabla \overline{b}\right). \tag{57}$$

This is not equal to (46); G-M parameterization simply assumes (56) or (57) as an overall quasi-steady approximation to (46), which is for a single event of baroclinic instability. Thus we may interpret $\boldsymbol{u}_{\|}$ as the velocity to express the horizontal transport or diffusion of buoyancy. In short G-M parameterization is a device to implement this role of baroclinic instability in a coarse resolution numerical model of ocean general circulation.

[2] equivalence of diffusive stretching and vertical viscosity (interfacial friction)

The horizontal diffusion of buoyancy related with G-M parameterization and baroclinic instability is shown to be equivalent with the vertical viscosity as follows. Using (54) we have

$$\frac{\partial \overline{b}}{\partial t} + N^2 \overline{w} = \kappa_h \nabla^2 \overline{b} + \text{ other terms.}$$
$$= \kappa_h \nabla^2 \left(f \frac{\partial \overline{\psi}}{\partial z} \right) + \text{ other terms.}$$

Then the contribution to the production of vorticity by stretching associated with horizontal diffusion becomes

$$\frac{\partial \nabla^2 \overline{\psi}}{\partial t} \sim f \frac{\partial \overline{w}}{\partial z} + \text{other terms.}$$
$$\sim \frac{\partial}{\partial z} \left(\frac{f^2 \kappa_h}{N^2} \frac{\partial \nabla^2 \overline{\psi}}{\partial z} \right) + \text{other terms.} \qquad (58)$$

This effect of horizontal diffusion on vorticity balance is called the *diffusive stretching*, which was discussed first in clarifying the structure of thermohaline circulation (Masuda and Uehara 1992, Mizuta and Masuda 1998).

On the other hand

$$\frac{\partial \nabla^2 \overline{\psi}}{\partial t} \sim \frac{\partial}{\partial z} \left(\nu_v \frac{\partial \nabla^2 \overline{\psi}}{\partial z} \right) + \text{other terms} \qquad (59)$$

indicates the usual effect of vertical viscosity on vorticity. Both (58) and (59) play the same role in the Q-G dynamics. The effect of horizontal diffusion of density (buoyancy) can be replaced by vertical viscosity, if

$$\frac{\nu_v}{\kappa_h} = \frac{f^2}{N^2}. \qquad (60)$$

That is, vertical viscosity or interfacial drag is equivalent to diffusive stretching due to horizontal diffusion.

In fact the production of vorticity by vertical viscosity becomes

$$\frac{\partial}{\partial z} \left[\nu_v \frac{\partial \nabla^2 \overline{\psi}}{\partial z} \right] = \frac{\partial}{\partial z} \left(\nu_v \frac{\partial \nabla \times \overline{\boldsymbol{u}}}{\partial z} \right)$$
$$= \nabla \times \frac{\partial}{\partial z} \left(\nu_v \frac{\partial \overline{\boldsymbol{u}}}{\partial z} \right) = -\nabla \times \frac{\partial}{\partial z} \left(\frac{\nu_v}{f} \lhd \overline{b} \right)$$
$$= \frac{\partial}{\partial z} \left(\frac{\nu_v}{f} \nabla^2 \overline{b} \right) = \frac{\partial}{\partial z} \left(\frac{f \kappa_h}{N^2} \nabla^2 \overline{b} \right),$$

which agrees with (58) if (60) is satisfied.

[3] interfacial drag and horizontal diffusion of density

We first note that (46) is written in a vector form as

$$\frac{\partial (\overline{p \nabla \eta})}{\partial z} = -f \neg \boldsymbol{u}_{\parallel} \qquad (61)$$

in a continuously stratified case. If we admit G-M parameterization (57), the right-hand side of (61) is related to the meridional gradient of $\overline{b}$ as

$$-f \neg \boldsymbol{u}_{\parallel} = -\frac{\partial}{\partial z} \left(\frac{f \kappa_G}{N^2} \lhd \overline{b} \right). \qquad (62)$$

On the other hand the upward flux of momentum is due to the interfacial form drag. If it is expressed via vertical viscosity, the net force by the interfacial form drag is expressed as

$$-f \neg \boldsymbol{u}_{\parallel} = \frac{\partial (\overline{p \nabla \eta})}{\partial z} = \frac{\partial}{\partial z} \left(\nu_v \frac{\partial \overline{\boldsymbol{u}}}{\partial z} \right)$$
$$= -\frac{\partial}{\partial z} \left(\frac{\nu_v}{f} \lhd \overline{b} \right) \qquad (63)$$

from the thermal-wind relation (55). In order for (63) to agree with (62), ν_v must be equal to $\kappa_G f^2/N^2$. If we further assume (60), κ_G may be identified with κ_h. Thus the interfacial form drag is translated into the vertical viscosity, horizontal diffusion, and G-M parameterization, through the physics of baroclinic instability.

Another suggestive relation is obatined as follows. Using (57), we see the curl of the net force associated with the interfacial form drag is expressed as

$$\nabla \times (-f \neg \boldsymbol{u}_{\parallel}) = f \nabla \cdot \boldsymbol{u}_{\parallel}$$
$$= \frac{\partial}{\partial z} \nabla \cdot \left(\frac{f \kappa_G}{N^2} \nabla \overline{b} \right) = \frac{\partial}{\partial z} \left(\frac{f \kappa_G}{N^2} \nabla^2 \overline{b} \right)$$
$$= \frac{\partial}{\partial z} \left(\frac{f^2 \kappa_G}{N^2} \frac{\partial \nabla^2 \overline{\psi}}{\partial z} \right),$$

which is the diffusive stretching (58) due to the ordinary horizontal diffusion of buoyancy if $\kappa_h = \kappa_G$.

Finally equation (52), which expresses the generation of the zonal mean circulation, is interpreted as a discretized version of the baroclinic component of

$$\frac{\partial}{\partial t} \left(\nabla^2 \overline{\psi} + \frac{\partial}{\partial z} \left[\frac{f^2}{N^2} \frac{\partial \overline{\psi}}{\partial z} \right] \right)' = f \nabla \cdot \boldsymbol{u}_{\parallel}. \qquad (64)$$

The right-hand side of (64) represents the diffusive stretching mentioned above, due to horizontal transport of buoyancy or interfacial form drag associated with baroclinic instability. This equation is valid regardless of the assumption of G-M parameterization.

We thus have observed an intimate relationship or consistency among baroclinic instability, diffusive stretching, P-M parameterization, interfacial form drag, and the Coriolis force acting on the ageostrophic component of velocity. If horizontal diffusion of buoyancy κ_h is ascribed to baroclinic instability or mass transport like G-M parameterization, all are related with one another as above. Plausible though it may be, the argument in this subsection is considered tentative, unless we have more reliable evidence to support it. There are a lot of subtle and misleading issues in this subject of eddy-transport of mass, heat (buoyancy), momentum, and others. Those will be addressed to in future.

5. Summary and discussion

A simplest situation of baroclinic instability was studied to understand its mechanism. A neat form of evolution equation was obtained, which is called the canonical equation for baroclinic instability. It describes the coupling of barotropic and baroclinic modes, rather than that of surface and bottom layers. Each term of the equation has a clear physical meaning therefore in the sense of vertical modes. Also the mechanism of instability is interpreted in terms of the feedback between two vertical modes. We find, in particular, the equation is reduced to the harmonic equation (Laplace equation) for two independent variables of time t and zonal coordinate x, in the limit of long-wave disturbance; it is related with the Cauchy-Riemann condition of complex functions. Therefore the instability mechanism becomes almost trivial and the spatial configurations of growing and decaying modes are interpreted by the property of complex functions. On the other hand for short-wave disturbances, the equation is reduced to the wave equation in a one-dimensional space.

The canonical equation gives explicit simple solutions for the complex phase velocity and associated modes of stream functions. Such solutions make it easy to understand various aspects and roles of baroclinic instability. The resulting picture is quite systematic and simple compared with ordinary models.

However, growth rates and spatial structure of the growing mode obtained are not necessarily new. Rather, it is worth mentioning that the canonical equation reproduces almost all results of Eady's model of continuously stratified fluids, in a simplest two-layer model. In other words, we can understand basic aspects and roles of baroclinic instability in a purified form.

The indispensable ingredients of baroclinic instability and the features of disturbance are summarized as follows.

[1] Baroclinic instability occurs in a stably stratified fluid in a rotating system. There must be background current with vertical shear, which is associated with the horizontal gradient of buoyancy; it produces the necessary CIPT β effect.

[2] Baroclinic instability needs no planetary β effect. Stratification is necessary, but Boussinesq fluid and only two layers are sufficient. Geographical or topographic variabilities are not necessary.

[3] Disturbance of scales larger than the baroclinic radius of deformation becomes a growing or decaying mode. For growing (decaying) modes, the minimum pressure of each layer is found upflow (downflow) side of the storm center at the interface. Positive feedback between barotropic and baroclinic modes plays a key role in the present interpretation.

[4] Growing modes level off the horizontal difference of buoyancy to release the potential energy of the background field. It has not been easy to explain that the growing mode has a horizontal scale larger than the baroclinic radius of deformation. It becomes easy, however, through the canonical equation, in which baroclinic radius comes into play in a most natural way.

A few variations of the canonical equation provided us with some unexpected mathematics of baroclinic instability. Also intimate relations were pointed out among baroclinic instability, G-M parameterization, diffusive stretching, and interfacial form drag. Nonlinear aspects of baroclinic instability was not explored, however. They are quite important in the real role of baroclinic instability in general circulation. For instance, the momentum balance in the southern ocean has often been controversial as Hidaka's dilemma. There remain a lot of such subtle problems to be inquired further.

References

1) Eady, E. T.(1949): Long waves and cyclone waves, *Tellus*, **1**, 35-52.

2) Charney, J. G.(1947): The dynamics of long Long waves in a baroclinic westerly current, *J. Meteor.*, **4**, 135-162.

3) Courant, R. and Hilbert, D. (1953): Methods of Mathematical Physics, volume 1, Interscience Publishers, New York, 560 pp.

4) Drazin. P. G. and Reid, W. H. (1981): Hydrodynamic Stability, Cambride University Press, Cambridge, 527 pp.

5) Gent, P. R., McWilliams, J. C. (1990): Isopycnal mixing in ocean circulation models, *J. Physical Oceanogr.*, **20**, 150-155.

6) Gent, P. R., Willebrand, J., McDougall, T. J., McWilliams, J. C. (1995): Parameterizing eddy-induced tracer transports in ocean circulation models, *J. Physical Oceanogr.*, **25**, 463-474.

7) Masuda, A. (2010): Vector notations suitable for geophysical fluid dynamics with examples and applications, *Rep. Res. Inst. Appl. Mech., Kyushu Univ.*, **138**, 1-12.

8) Masuda, A. (2011): Vertical modes of quasi-geostrophic flows in an ocean with bottom topography — evolution equation and energetics —, *Rep. Res. Inst. Appl. Mech., Kyushu Univ.*, **141**, 28-46.

9) Masuda, A., Okuno, A. (2002): Kuroshio frontal eddies as baroclinic instability, *Rep. Res. Inst. Appl. Mech., Kyushu Univ.* (in Japanese), **122**, 25-36.

10) Masuda, A., Uehara, K. (1992): A reduced-gravity model of the abyssal circulation with Newtonian cooling and horizontal diffusion, *Dee-Sea Res.*, **39**, 1453-1479.

11) Mizuta, G., Masuda, A.. (1998): Three-dimensional structure of thermohaline circulation steered by bottom topography, *J. Phys. Ocenogr.*, **28**, 1979-1998.

12) Pedlosky, J. (1987): Geophysical Fluid Dynamics, 2-nd edn., Springer-Verlag, 710 pp.

13) Phillips, N. A. (1954): Energy transformations and meridional circulations associated with simple baroclinic waves in a two-level, quasi-geostrophic model, *Tellus*, **4**, 273-286.

14) Takahashi, J. and A. Masuda (1998): Mechanisms of the southward translation of Meddies, *J. Oceanogr.*, **54**, 669-680.

Appendix

A1. meridional modes

Let $\hat{\psi}_k = \hat{k}(y,t)$ be a Fourier-component of zonal waveumber k for a well-behaved function $\psi = \psi(\boldsymbol{x}, t)$:

$$\hat{\psi}_k \equiv \int \mathrm{e}^{-ikx}\psi\, dx.$$

We may expand $\hat{\psi}_k = \hat{\psi}_k(y,t)$ in terms of certain meridional modes, which are defined through the following eigenvalue problem

$$\begin{cases} \text{o.d.e.}: & \dfrac{d^2}{dy^2}s = -l^2 s \quad -L/2 < y < L/2 \\ \text{b.c.}: & s = 0 \quad \text{at} \quad y = \pm\dfrac{L}{2} \end{cases},$$

where $l^2 > 0$ is the required eigenvalue and s the eigenfunction associated with l^2.

The eigenvalues are $l^2 = n^2\pi^2$, n being positive integers, and the associated eigenfunctions $\sin n\pi y/L$ (say, Courant and Hilbert 1953, Masuda 2011). Thus we may expand $\hat{\psi}_k$ as

$$\hat{\psi}_k = \sum_{n=1}^{\infty} \check{\psi}_{k,n}(t) \sin\frac{n\pi}{L}\left(y + \frac{L}{2}\right),$$

irrespective of k. Finally the normal mode to consider has a form of

$$\mathrm{e}^{ikx}\sin l\left(y + \frac{\pi}{2}\right), \quad l = \frac{n\pi}{L}.$$

In particular for the gravest mode of $n = 1$, we may use

$$\mathrm{e}^{ikx}\sin l\left(y + \frac{\pi}{2}\right) = \mathrm{e}^{ikx}\cos ly.$$

A2. Reduced form of barotropic vorticity equation

In Section 3, the barotropic vorticity equation (13) is reduced to (15). This is almost obvious and guaranteed as follows. We first decompose the disturbance fields into their zonal mean and the residues

$$\phi = \overline{\phi} + (\phi - \overline{\phi}), \quad \varphi = \overline{\varphi} + (\varphi - \overline{\varphi}),$$
$$\text{s.t.} \quad \overline{\phi - \overline{\phi}} = 0, \quad \overline{\varphi - \overline{\varphi}} = 0.$$

Note that this procedure is supposed for a zonally periodic channel; if the domain is infinite in the x-direction, both $\overline{\phi}$ and $\overline{\varphi}$ would vanish.

Then zonal average of (13) and (14) yields

$$\begin{cases} \dfrac{\partial^2}{\partial y^2}\dfrac{\partial\overline{\phi}}{\partial t} = 0 \\ \left(\dfrac{\partial^2}{\partial y^2} - F\right)\dfrac{\partial\overline{\varphi}}{\partial t} = 0 \end{cases},$$

the solutions of which become

$$\begin{cases} \overline{\phi} = {}^{\exists}a(t)y + {}^{\exists}c(t) \\ \overline{\varphi} = {}^{\exists}\alpha(t)\sinh\sqrt{F}y + {}^{\exists}\gamma(t) \end{cases}.$$

Since a spatially constant term is meaningless for ϕ, we put $c(t) = 0$. Also we have $\gamma(t) = 0$, in order to conserve the mass of each layer (4).

The framework of Q-G dynamics does not determine the remaining coefficients of $a(t)$ and $\alpha(t)$. From the total conservation of zonal momentum (5), however, we see

$$0 = \frac{\partial}{\partial t}\int_{-L/2}^{L/2}(H_1\overline{u}_1 + H_2\overline{u}_2)\, dy = \frac{\partial}{\partial t}\int_{-L/2}^{L/2}\frac{\partial\overline{\phi}}{\partial y}\, dy$$

which gives $a(t) = {}^{\exists}a_0$. Then it follows that

$$0 = \frac{\partial\overline{\phi}}{\partial(Ut)} = \frac{\partial\overline{\phi}}{\partial(Ut)} + \frac{\partial\overline{\varphi}}{\partial x}. \tag{A1}$$

The remainders are expanded in Fourier integrals as

$$\phi - \overline{\phi} = \int \hat{\phi}_k(y,t)\mathrm{e}^{ikx}\, dx,$$
$$\varphi - \overline{\varphi} = \int \hat{\varphi}_k(y,t)\mathrm{e}^{ikx}\, dx.$$

For each $k \neq 0$, we have from (13)

$$\nabla^2\left[\left(\frac{\partial}{\partial(Ut)}\hat{\phi}_k(y,t) + ik\hat{\varphi}_k(y,t)\right)\mathrm{e}^{ikx}\right] = 0.$$

To satisfy the boundary conditions (3), the quantity in the bracket should vanish. Thus we obtain

$$0 = \frac{\partial}{\partial(Ut)}(\phi - \overline{\phi}) + \frac{\partial}{\partial x}(\varphi - \overline{\varphi}). \tag{A2}$$

Addition of (A1) and (A2) yields

$$\frac{\partial}{\partial(Ut)}\phi + \frac{\partial}{\partial x}\varphi = 0,$$

reduced form of the barotropic vorticity equation.