
YUKI YASUDA AND KAORU SATO
Department of Earth and Planetary Science, University of Tokyo, Tokyo, Japan
NORIHIKO SUGIMOTO
Department of Physics, Keio University, Kanagawa, Japan

(Manuscript received 25 November 2013, in final form 3 November 2014)

ABSTRACT

By using the renormalization group (RG) method, the interaction between balanced flows and Doppler-shifted inertia–gravity waves (GWs) is formulated for the hydrostatic Boussinesq equations on the $f$ plane. The derived time-evolution equations [RG equations (RGEs)] describe the spontaneous GW radiation from the components slaved to the vortical flow through the quasi resonance, together with the GW radiation reaction on the large-scale flow. The quasi resonance occurs when the space–time scales of GWs are partially comparable to those of slaved components. This theory treats a coexistence system with slow time scales composed of GWs significantly Doppler-shifted by the vortical flow and the balanced flow that interact with each other. The theory includes five dependent variables having slow time scales: one slow variable (linear potential vorticity), two Doppler-shifted fast ones (GW components), and two diagnostic fast ones. Each fast component consists of horizontal divergence and ageostrophic vorticity. The spontaneously radiated GWs are regarded as superpositions of the GW components obtained as low-frequency eigenmodes of the fast variables in a given vortical flow. Slowly varying nonlinear terms of the fast variables are included as the diagnostic components, which are the sum of the slaved components and the GW radiation reactions. A comparison of the balanced adjustment equation (BAE) by Plougonven and Zhang with the linearized RGE shows that the RGE is formally reduced to the BAE by ignoring the GW radiation reaction, although the interpretation on the GW radiation mechanism is significantly different; GWs are radiated through the quasi resonance with a balanced flow because of the time-scale matching.

1. Introduction

Inertia–gravity waves (GWs) play an important role in the global circulation and transport by redistributing momentum and energy in the atmosphere (Alexander et al. 2010). However, because of their small scale, the actions of GWs are usually incorporated in climate models using parameterization schemes. The parameterization schemes include various assumptions because the mechanism of GW generation and the GW propagation properties are not very clear, and computational resources are limited. In particular, the dynamics of nonorographic GW generation is nonlinear and quite complicated, unlike that of orographic GW generation, and this leads to a lot of uncertainties in the parameterization schemes (e.g., Hines 1997; Alexander and Dunkerton 1999; Warner and McIntyre 2001). Thus, gaining a better understanding of the nonorographic GW dynamics is important not only for the geophysical fluid dynamics but also for the improvement of parameterization schemes.

Recent studies have clarified that GWs are spontaneously radiated from an approximately balanced flow as the flow evolves (e.g., O’Sullivan and Dunkerton 1995; Vanneste 2013; Plougonven and Zhang 2014). The amount of spontaneously radiated GWs is considered to be comparable to that of orographic GWs (e.g., Plougonven et al. 2013). Many theoretical studies on spontaneous radiation, including the present study, are
partly based on balanced-flow dynamics. Thus, we first briefly review balanced-flow dynamics.

The time evolution of a balanced flow is governed by a balanced model (BM). The BM is a system of equations that consists of a time-evolution equation for a slow variable and diagnostic formulas for fast variables (e.g., Leith 1980; Lorenz 1980; Warn et al. 1995; McIntyre and Norton 2000; Wirosoetisno et al. 2002; Saujani and Shepherd 2006; Mohebalhojeh and McIntyre 2007a,b). Here, a slow variable (e.g., potential vorticity) is a dependent variable that is zero for the GW solution but nonzero for the geostrophic solution in the linear system. In contrast, fast variables (e.g., horizontal divergence and ageostrophic vorticity) are dependent variables that are zero for the geostrophic solution but nonzero for the GW solution in the linear system. In BMs, fast variables are diagnostically obtained from the slow variable, which are called slaved components. BMs are generally obtained by imposing diagnostic relations between physical quantities such as geostrophic and hydrostatic balances on the primitive equations. The most familiar BM may be the quasigeostrophic (QG) model, in which the slow variable is QG potential vorticity (QGPV), while the slaved component is ageostrophic velocity. The QGPV is obtained from the time-evolution equation, while the ageostrophic velocity is diagnostically determined by the omega and continuity equations with a given QGPV field. GWs are completely filtered out in a BM; hence, the BM has a lower degree of freedom than the primitive equations. In other words, a BM describes the evolution on the slow manifold, which is a hypersurface in the phase space defined by the balanced relations (e.g., Vanneste 2013). In balanced-flow dynamics, the balanced flow should be distinguished from the vortical flow, because the vortical flow is a component composing the PV field, while the balanced flow (i.e., the solution of the BM) is the sum of the vortical flow and slaved components.

Theoretical studies on spontaneous GW radiation can be classified into two groups:

1) The initial state does not contain any GWs, and then GWs are radiated as the flow evolves.
2) The initial state contains GWs with quite small amplitudes, and then the GWs are amplified.

Theories classified into the second group include instabilities that occur as an interaction between balanced and imbalanced modes (e.g., Plougonven et al. 2005) and transient amplifications by a superposition of neutral modes composed of balanced and imbalanced parts (Lott et al. 2010, 2012). The classification of Lott et al. (2010, 2012) may be ambiguous, because they demonstrated that it was impossible to rigorously separate a balanced part from an imbalanced part in a constant vertical shear field. This issue will be further discussed in section 4b. In the second-group theories, GWs are not radiated if the initial state is exactly balanced.

It should be noted here that another classification is possible from a different viewpoint, as made in the review paper (Plougonven and Zhang 2014): the theories by Lott et al. (2010) and Vanneste and Yavneh (2004) are classified into the same group named, “transient generation by sheared disturbances,” while they are classified into different groups in the present paper.

The present study is classified into the first group (group 1). Roughly speaking, theories in the first group are further classified into three subgroups: Lighthill radiation, transient generation, and balance adjustment.

The Lighthill radiation theory is based on the theory of vortex sound and describes GW generation in shallow-water systems quite well (e.g., Ford et al. 2000; Sugimoto et al. 2007, 2008). The transient generation describes the generation of GWs having high ground-based frequencies in a three-dimensional system for the small Rossby number limit (Vanneste and Yavneh 2004; Ólafsdóttir et al. 2008). In contrast, the present study treats significantly Doppler-shifted GWs, whose time scales are comparable to that of the balanced flow, in a three-dimensional system for a finite but sufficiently small Rossby number. Thus, our study should be distinguished from the first and second subgroups of the first-group studies.

Balance adjustment, the third subgroup of the first group, is the mechanism for GW radiation from imbalanced sources that are produced continuously as an approximately balanced large-scale flow evolves in time (Zhang 2004). This imbalance is defined as the difference between the real flow and the balanced one. Note that the balance adjustment is not an initial-value problem, such as the classical geostrophic adjustment (or Rossby adjustment) problem (e.g., Holton 2004), Plougonven and Zhang (2007) derived a system of linear equations describing balance adjustment on the basis of a scale analysis. In their theory, residuals of the balanced-model solution are called imbalances, and the Doppler-shifted GWs, having low ground-based frequencies, are radiated from these residuals. There have been other linear theories (e.g., Snyder et al. 2009) quite similar to that of Plougonven and Zhang (2007). In the present paper, linear equations that describe the radiation of Doppler-shifted GWs from residuals of the balanced-model solution are referred to as balance adjustment equations (BAEs). The BAEs have successfully reproduced the distribution of spontaneously radiated GWs in the vortex dipole with high accuracy (Snyder et al. 2009; Wang and Zhang 2010).

However, there are a few important issues that are not elucidated quite well by the BAEs. The residuals of the balanced-model solution, which are called imbalances in
Large-scale flow field (coexistence system)

Fig. 1. Conceptual diagram of RGE theory. The vortical flow is a component composing PV, the slaved components are part of the fast variables, which are diagnostically determined by the PV distribution, and the balanced flow is the sum of the vortical flow and slaved components. GWs that are significantly Doppler shifted as a result of the vortical flow are radiated from the slaved components through the quasi resonance, while the vortical flow and slaved components (GW sources) are affected by the GW radiation reactions. In RGE theory, the large-scale-flow system is considered to be a coexistence system composed of GWs, slaved components, and vortical flow that interact with each other. These components should not be considered separately.

the BAEs, as mentioned above, depend on which BM is used to construct the BAEs. Thus, radiated GWs may significantly depend on the choice of BMs. In addition, the physical mechanism for the continuous production of the imbalance is not clear. In the BAEs, the imbalances always exist when the balanced-model solution is not an exact solution for the primitive equations. In other words, it is physically unclear why the balanced-model solution is not an exact solution. Last but not least, the large-scale flow should be continuously modified by the continuous GW radiation, as discussed by McIntyre (2009). This “adjustment” process is, however, not described in the BAEs, because they are linear equations describing only the GW generation.

In the present study, we derive a new theory that is similar to BAES, but we propose a new mechanism. The most important point in our theory is that Doppler-shifted GWs are radiated through a quasi resonance with a part of the balanced flow. More specifically, GW sources are the slaved components of the balanced flow, and GWs are radiated when space–time scales of the GWs are comparable with those of the sources. Our theory treats balance adjustment as interactions between GWs, slaved components, and vortical flow in a coexistence system, as illustrated in Fig. 1. As discussed in section 4a, our theoretical equation can be formally reduced to a BAE and will clarify that BAES actually describe the GW radiation from the slaved components. This point is discussed briefly here because it is the motivation for the present study.

The GW source of the BAE in the x direction \( F_u \) analyzed by Snyder et al. (2009) is considered as an example. The source was obtained as the residual of the QG-model solution and called residual tendency in Snyder et al. (2009). On the other hand, \( F_u \) is described using the second-order slaved components of northward velocity \( u_a^{(2)} \) and normalized pressure \( p_a^{(2)} \):

\[
F_u = \frac{\partial u_a}{\partial t} + u_a \cdot \nabla u_a + u_g \cdot \nabla u_a = f u_a^{(2)} - \frac{\partial p_a^{(2)}}{\partial x},
\]

where \( u_g = (u_g, v_g, 0) \) is the geostrophic velocity and \( u_a = (u_a, v_a, w_a) \) is the ageostrophic velocity in the QG model. The formula on the most right-hand side (RHS) can be obtained by expanding all variables with respect to Rossby number \(^1\) (Ro). As seen from this example, residuals of balanced-model solutions can be generally regarded as high-order slaved components. This fact suggests that the residuals are not imbalances but a part of the balanced flow. This point will be further discussed in section 4a.

We have developed a new theory by focusing on two important facts elucidated in previous studies. The first fact is the existence of a descent–ascent couplet on a local jet axis in a vortex dipole whose characteristic spatial scale is larger than those of GWs (see Fig. 2). Víúdez (2007) suggested that this couplet was the initial GW packet originating from the material rate of change of the ageostrophic differential vorticity. However, this suggestion does not seem consistent with the fact that this couplet exists even in the dipole in which Ro is so low that GW radiation is hardly recognized (Snyder et al. 2007). This fact means that this couplet may be attributable to high-order slaved components. We inferred from these two studies that the origin of GW sources is high-order slaved components that a local jet accompanies. The second fact is a significant Doppler shift by the vortical flow (i.e., the main part of the balanced flow). The intrinsic frequencies of GWs are usually much higher than those of the balanced flow. However, when GWs are significantly Doppler shifted by the strong vortical flow, their ground-based frequencies may be comparable to those of the balanced flow. Note that the Doppler shift is also essential for the BAES (e.g., Plougonven and Zhang 2007).

Taking the above two facts into account, we have reached a new mechanism: GWs are radiated from slaved components (i.e., part of the balanced flow)

---

\(^1\) As discussed by Warn et al. (1995), this expansion may not be mathematically valid for a long time, because both slow and fast variables may be expanded. More accurate discussion will be made in section 4a.
through a quasi resonance when the space–time scales of the GWs are partially comparable to those of the slaved components. Note that the term “quasi resonance” is used here because the frequencies of radiated GWs are not always identical to those of the slaved components. The wavelengths of the radiated GWs decrease through the wave-capture process. As described in Part II, these vertical flow couplets are a compensation flow for the strong localized jet and are produced by the vortical flow over the deformed potential temperature (θ) surfaces, which are also the slaved components. These potential temperature surfaces act as mountains, such as ones that generate orographic GWs. The RGE theory can describe two different mechanisms: the mountain-wave-like mechanism (this figure) and the velocity-variation mechanism (Fig. 3).

The RG method (Chen et al. 1994, 1996), which is a singular perturbation method, makes it possible to systematically and naturally extract time-evolution equations on slow time scales from a system containing multiple time-scale motions. Thus, this method is relevant to obtain the theoretical equations describing the interaction including quasi resonance between the vortical flow and Doppler-shifted GWs, which both have slow time scales. In the present paper, the RG method is applied to the hydrostatic Boussinesq equations on the f-plane (herein referred to as primitive equations). We use linear potential vorticity q as a slow variable and horizontal divergence δ and ageostrophic vorticity γ as fast ones. The linear theory indicates that δ and γ have high-frequency components, including (non-Doppler-shifted) GWs, but some of their nonlinear terms vary slowly in time because they consist only of q. These slowly varying nonlinear terms are mainly due to slaved...
components of balanced flow and should be distinguished from radiated GWs but should be included in the theory for the slow-time-scale motions. Thus, five dependent variables are introduced separately: Doppler-shifted fast variables ($\delta^{\text{GW}}$ and $\gamma^{\text{GW}}$), diagnostic fast variables ($\delta^{\text{diag}}$ and $\gamma^{\text{diag}}$), and a slow variable ($q$). Doppler-shifted fast variables correspond to the spontaneously radiated GWs, while the diagnostic ones correspond to the slowly varying nonlinear terms. Note again that all five variables should have slow time scales. The GW components ($\delta^{\text{GW}}$ and $\gamma^{\text{GW}}$) that include the Doppler effect by the vortical flow are obtained as the eigenmodes of the fast variables in a given vortical flow field after the renormalization is made. On the contrary, $\delta^{\text{diag}}$ and $\gamma^{\text{diag}}$ are directly and diagnostically obtained from renormalized $q$, $\delta^{\text{GW}}$, and $\gamma^{\text{GW}}$, and they correspond to the sum of the slaved components and the GW radiation reactions.

Our theory clarifies that the main GW source is the slaved component that is a part of $\delta^{\text{diag}}$ and $\gamma^{\text{diag}}$. It is worth noting that the derived equations describe the tendencies of $q$ and GW sources, including the GW radiation reactions. In addition, this new theoretical system proposes a new adjustment process (i.e., resonance adjustment) that is different from the balance adjustment process (Zhang 2004), as discussed in section 4b. In the new process, GWs are radiated continuously because the GW sources are always present, and the large-scale flow becomes weak as GWs are radiated. When the flow is too weak to cause the time-scale matching between the Doppler-shifted GWs and slaved components, the GW radiation is ceased, and the weak large-scale flow without GWs is left. In this paper (Part I), we describe the formulation of the theoretical equations.

The structure of this paper is as follows. In section 2, the primitive equations are transformed before applying the RG method. In section 3, by applying RG method to the transformed primitive equations, the time-evolution equations on slow time scales are derived. These equations are referred to as renormalization group equations (RGEs). In section 4, the relation between the RGEs and BAEs is discussed. Advantages and disadvantages of the RGE system are described in comparison with the BAEs. A discussion is also made on how the wave capture is described in the RGE system. A summary and the concluding remarks are given in section 5. In appendix A, the details of the RG method are presented by using an example of an ordinary differential equation.

One possible way to confirm the validity of this new theory is to compare the predictions from this theory with the results from numerical simulations of a compressible nonhydrostatic equation model. For this purpose, quasi-steady spontaneous GW radiation in a vortex dipole is addressed in Part II. Details of the physics on the GW sources and processes of the spontaneous radiation, as in Figs. 2 and 3, are also examined in Part II.

2. Transformation of dependent variables and wavenumber representation of primitive equations

The representations of the primitive equations [i.e., hydrostatic Boussinesq equations on the $f$ plane (e.g., Vallis 2006)] are transformed before applying the RG method. After dependent variables, such as velocity, temperature, and pressure, are converted to fast and slow variables, all equations are represented in the wavenumber space. No approximations, such as geostrophic balance approximation, are needed for the manipulations in this section.

The primitive equations are nondimensionalized using the characteristic scales for large-scale flows:

\[
\frac{Du}{Ds} = v - \frac{\partial p}{\partial x},
\]

\[
\frac{Dv}{Ds} = -u + \frac{\partial p}{\partial y},
\]

\[
0 = b - \frac{\partial p}{\partial z},
\]

\[
\frac{Ro \cdot Bu\, Db}{Ds} = -w, 
\]

\[
0 = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z},
\]

where $Ro = \frac{U}{fL}$ and $Bu = \left( \frac{fL}{NH} \right)^2$.

In the Cartesian coordinates $(x, y, z)$ with $x$, $y$, and $z$, respectively, directed eastward, northward, and upward, $(u, v, w)$ represent respective components of velocity vector $\mathbf{u}$, $s$ is time, $D/Ds$ is the material derivative ($= \partial/\partial s + \mathbf{u} \cdot \nabla$), $p$ is pressure normalized by the basic-state density, and $b$ ($=g\theta/\theta_0$) is buoyancy, where $g$, $\theta_0$, and $\theta'$ are the gravitational acceleration, the constant basic potential temperature, and the perturbation of potential temperature from $\theta_0$, respectively. Nondimensional numbers Ro and Bu, respectively, represent Rossby and Burger numbers, in which $U$, $L$, and $H$ are the characteristic scales of the large-scale flow for velocity, horizontal length, and height, respectively; $f$ is a constant Coriolis parameter; and $N$ is a constant buoyancy frequency. The value of Bu is taken to be unity for simplicity, because $Bu = O(1)$ for most large-scale motions in the real atmosphere.

As in Warn et al. (1995) and Saujani and Shepherd (2006), one slow and two fast variables are introduced:
linear potential vorticity $q$ for the slow variable and horizontal divergence $\delta$ and ageostrophic vorticity $\gamma$ for the fast ones:

$$ q = \frac{\partial w}{\partial x} - \frac{\partial u}{\partial y} + \frac{\partial b}{\partial z}, \quad (8) $$

$$ \delta = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}, \quad \text{and} \quad (9) $$

$$ \gamma = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) p. \quad (10) $$

In the usual notation in which velocity and thermodynamic variables such as temperature and pressure are dependent variables, it may be hard to distinguish the components attributed to the vortical flow from those attributed to the GWs. It is better to use slow and fast variables, which are directly related to the vortical flow and GWs, respectively, when interactions between them are examined.

The primitive equations [(2)–(6)] are transformed as follows by using $q$, $\delta$, and $\gamma$:

$$ \frac{\partial q}{\partial t} = -Ro \left[ \frac{\partial}{\partial x} (u \cdot Vu) - \frac{\partial}{\partial y} (u \cdot Vu) + \frac{\partial}{\partial z} (u \cdot Vb) \right], \quad (11) $$

$$ \frac{\partial \gamma}{\partial t} + \nabla^2 \delta = -Ro \left\{ \frac{\partial^2}{\partial z^2} \left[ \frac{\partial}{\partial x} (u \cdot Vu) - \frac{\partial}{\partial y} (u \cdot Vu) \right] - \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \frac{\partial}{\partial z} (u \cdot Vb) \right\}, \quad \text{and} \quad (12) $$

$$ \frac{\partial \delta}{\partial t} - \gamma = -Ro \left[ \frac{\partial}{\partial x} (u \cdot Vu) + \frac{\partial}{\partial y} (u \cdot Vu) \right], \quad (13) $$

where $t (=\sqrt{Ro})$ is an independent time variable, and $\gamma_{zz}$ is $\gamma$ differentiated twice with respect to $z$. When $Ro$ is sufficiently small, $s$ and $t$, respectively, represent slow and fast times whose characteristic scales are significantly different. Note that the use of linear potential vorticity does not assume the linear (or geostrophic) balance. No balance assumption has been made so far. GWs are actually included in (11)–(13).

Next, a triple-cyclic boundary condition is imposed, and all dependent variables are expanded into Fourier series; for example, $q$ is expanded as follows:

$$ q(t, x) = \frac{1}{V} \sum_k q_k(t) \exp(ik \cdot x), \quad (14) $$

$$ q_k(t) = \int_V q(t, x) \exp(-ik \cdot x) \, d^3x, \quad \text{and} \quad (15) $$

$$ k = (k_1, k_2, k_3) = 2\pi(n_1/L_1, n_2/L_2, n_3/L_3), \quad (16) $$

where $L_i$, $n_i$, and $V$, respectively, represent the non-dimensional length of space, the wavenumber for the integer $i$th direction, and the non-dimensional volume of space ($=L_1L_2L_3$). Usually, velocity and thermodynamic variables are related to both slow and fast variables through elliptic partial differential equations (e.g., Saujani and Shepherd 2006). Thus, it is generally hard to explicitly represent all terms by fast and slow variables. This problem is overcome in the wavenumber space, as shown later. There is another advantage of the wavenumber representation. Generally, a slow motion with respect to both space and time is extracted from a partial differential equation system by applying the RG method (Chen et al. 1996). However, it is expected that the spontaneous radiation occurs in a wide wavenumber range because of the deformation of GWs as in the wave-capture process (Bühler and McIntyre 2005). Thus, only the slow motion with respect to time should be extracted. This can be achieved by treating partial differential equations as ordinary differential equations with respect to time in the wavenumber space.

Using the following variables,

$$ \nabla^2 \psi = q, \quad (17) $$

$$ \nabla^2 \phi = \delta, \quad \text{and} \quad (18) $$

$$ \nabla^2 \chi = \gamma_{zz}, \quad (19) $$

where $\nabla_H$ is a two-dimensional nabla operator in the $x$–$y$ plane, we obtain the simple representations for $u_k \cdot v_k$, $w_k$, and $b_k$ as follows:

$$ u_k = -ik_2 \psi_k + ik_1 \phi_k - ik_3 \chi_k, \quad (20) $$

$$ v_k = ik_1 \psi_k + ik_2 \phi_k + ik_1 \chi_k, \quad (21) $$

$$ w_k = -i \frac{k_H^2}{k_3} \phi_k, \quad \text{and} \quad (22) $$

$$ b_k = ik_3 \psi_k - i \frac{k_H^2}{k_3} \chi_k, \quad (23) $$

where $k_H = (k_1, k_2, 0)$. The representation for $p_k$ is also obtained from (4) and (23). The case of $k_3 = 0$ and that of $k_H = 0$ are ignored because they are not important for the problem addressed in this study. The terms

---

2 When linear potential vorticity is used, it is possible to construct much more accurate balanced models than the QG model. See Mohebalhojeh and Dritschel (2001) for details.
composed of \(q_k\) (i.e., \(\psi_k\)) in (20)–(23) represent the vortical flow components in the wavenumber space.

Using (17)–(23), we formally transform the primitive equations [(11)–(13)]:

\[
\frac{dq_k}{dt} = \text{Ro} \sum_m \left( I_{q_m q_k}^{\delta} q_m q_k + I_{q_m q_k}^{\gamma} q_m \gamma_k - I_{q_m \gamma_k}^{\delta} q_m \gamma_k - I_{q_m \gamma_k}^{\gamma} q_m \gamma_k \right),
\]

\[
\frac{d\gamma_k}{dt} + \omega_k^2 \delta_k = \text{Ro} \sum_m \left( I_{\gamma_m q_k}^{\delta} q_m \gamma_k + I_{\gamma_m q_k}^{\gamma} q_m \gamma_k - I_{\gamma_m \gamma_k}^{\delta} q_m \gamma_k - I_{\gamma_m \gamma_k}^{\gamma} q_m \gamma_k \right),
\]

\[
\frac{d\delta_k}{dt} - \gamma_k = \text{Ro} \sum_m \left( I_{\delta_m q_k}^{\delta} q_m \delta_k + I_{\delta_m q_k}^{\gamma} q_m \delta_k - I_{\delta_m \gamma_k}^{\delta} q_m \delta_k - I_{\delta_m \gamma_k}^{\gamma} q_m \delta_k \right),
\]

where \(\omega_k = |k| / |k|_3\). Note that \(d/dt\) represents an ordinary differentiation, because \(q_k\), \(\gamma_k\), and \(\delta_k\) depend only on time. The RHSs of (24)–(26) show triad interactions, in which \(I_{q_m q_k}^{\delta}\) and the other coefficients represent interaction coefficients (i.e., strengths of the interactions). All these coefficients depend only on wavenumbers, whose specific forms are shown in appendix B. So far, we have not given any assumptions and definitions for GWs. In our theory, GW components will be defined later on the basis of the eigendecomposition but not of the scale separation.

When Ro is so small as to enable all nonlinear terms on the RHSs of (24)–(26) to be ignored, \(q_k\) is constant, while \(\gamma_k\) and \(\delta_k\) oscillate with the frequencies of \(\pm \omega_k\). This oscillation corresponds to a plane GW in the motionless atmosphere. These GWs are not Doppler shifted, because all nonlinear terms, including the Doppler shift for GWs, are neglected. The above discussion suggests that \(q_k\) varies on the slow time scale characterized by \(s(=\text{Rot})\), while \(\gamma_k\) and \(\delta_k\) vary on the fast time scale characterized by \(t\), when Ro is small but the nonlinear terms in (24)–(26) cannot be ignored. The slow time scale is separated to a greater degree from the fast time scale for smaller Ro, indicating that high-frequency GWs are hardly excited by the vortical flow. Another important point is that the first terms on the RHSs of (25) and (26) are nonlinear terms composed only of \(q_k\). Thus, the fast variables contain not only high-frequency components but also low-frequency ones as a result of the nonlinear interactions. Roughly speaking, these slowly varying nonlinear terms are slaved components to the vortical flow and should be distinguished from GWs.

3. Application of RG method to the primitive equations

In this section, by applying the RG method to the transformed primitive equations [(24)–(26)], the time-evolution equations (RGEs) are derived. In section 3a, the assumptions used in this paper are shown and discussed. The RGEs are derived in section 3b, and their characteristics are discussed in section 3c. Separability between the diagnostic components and GW components in this RGE theory is discussed in section 3d.

The RGE system derived in this section describes the new mechanism: GWs are radiated from slaved components through a quasi resonance when the space–time scales of the GWs are partially comparable to those of the slaved components. Spontaneous GW radiation tends to occur at a later time, even when the initial state is well balanced, because GW sources (i.e., slaved components) are always present. Thus, the RGE theory indicates that the primitive equations do not have an invariant slow manifold. In the RGE theory, a flow system is considered to be a coexistence system composed of the GWs, slaved components, and vortical flow that interact with each other (Fig. 1).

a. Assumptions and their validity

The following six assumptions were made when applying the RG method:

(i) The initial state is balanced. The GWs excited are so significantly Doppler shifted by the vortical flow as to have low ground-based frequencies (i.e., slow time scales). The excitation of GWs having high ground-based frequencies that are not significantly Doppler shifted is negligible.
(ii) To take the significant Doppler shift into account, an eigenmode expansion of the fast variables in a given vortical flow is carried out after applying the RG method. Eigenmodes with eigenvalues (i.e., ground-based frequencies) that are $O(Ro)$ are referred to as GW modes. A superposition of the GW modes is considered as the spontaneously radiated GWs. When the GW modes are represented by the plane-wave bases, they are referred to as GW components ($\delta_k^G$ and $\gamma_k^G$). On the contrary, the components that are diagnostically obtained from $q_k$, $\delta_k^G$, and $\gamma_k^G$ are referred to as diagnostic components ($\delta_k^{\text{diag}}$ and $\gamma_k^{\text{diag}}$). These diagnostic components vary on the slow time scales following the $q_k$, $\delta_k^G$, and $\gamma_k^G$ variations because of the nonlinear effects.

(iii) The amplitude of the GW component is nondimensionalized on the basis of the characteristic space–time scales of the large-scale flow, which is characterized by the balanced flow.

(iv) The characteristic length of the large-scale flow is $O(1)$, while that of the GWs is $O(Ro)$. This assumption is represented in terms of wavenumbers; $k_i = (1/Ro)k_i^G$, where $k_i^G$ is the wavenumber rescaled by the characteristic length of the GWs [i.e., $k_i^G = O(1)$].

(v) The interaction coefficients between the vortical flow and GW component and those between the diagnostic component and GW component are expanded with respect to Ro.

(vi) The interaction coefficients of $O(Ro^{-1})$ between the diagnostic component and GW component are considered only in the time-evolution equations for the GW components (or modes).

The validity of the above six assumptions is discussed in the following. Wirosoetisno et al. (2002) applied the third-order RG method to the primitive equations, but they did not include the Doppler effect on GWs. They showed that the flow field governed by the RGEs was always balanced when the initial state was balanced. Note that the primitive equations addressed in their paper were abstract equations having some properties that the hydrostatic Boussinesq equations have. The time scales of the vortical flow and non-Doppler-shifted GWs are completely separated in the regular perturbation solution. This fact means that the quasi resonance does not occur in their RGEs, and (non-Doppler-shifted) GWs are not excited in the balanced-flow field. Thus, assumption (i), which eliminates high-frequency GWs, is plausible.3

In general, the fast variables have slowly varying nonlinear terms that are distinguished from GWs such as the first terms on the RHSs of (25) and (26). Roughly speaking, these terms are regarded as slaved components. To include the slowly varying nonlinear terms separately from GWs in the theory, the RG method is applied to $\delta$ and $\gamma$ in the same way as Wirosoetisno et al. (2002). Under assumptions (i)–(vi), the distributions of these components are determined diagnostically from $q_k$, $\delta_k^G$, and $\gamma_k^G$ [see (42), (43), and (52) in later sections of the present paper]. Thus, they are referred to as diagnostic components. The slaved components are included in the diagnostic components. The diagnostic components other than the slaved components are regarded as GW radiation reactions, because they depend on $\delta_k^G$ and/or $\gamma_k^G$.

On the other hand, the GW modes are obtained by the eigenmode expansion in a given vortical flow after applying the RG method. The characteristic frequency scale of the GW modes should be $O(Ro)$ so that the derived RGEs describe the excitation of GWs through the quasi resonance with the balanced flow with a characteristic frequency scale of $O(Ro)$. Thus, the GW modes are defined as in assumption (ii). This definition has uncertainty, because the range of eigenvalues (i.e., ground-based frequencies of GW modes) is not specified. This uncertainty is discussed in Part II.

Next, we discuss the characteristic amplitude and spatial scales of the GWs to show that assumptions (iii) and (iv) are reasonable. As mentioned earlier, we consider that the GWs are radiated through the quasi resonance and the main GW sources are the slaved components composed only of $q$. This means that the characteristic magnitudes of the slaved and GW components should be of the same order. In other words, $\delta_k^G$ and $\gamma_k^G$ should be nondimensionalized on the basis of the characteristic scales of the large-scale flow, like the slaved components are.

Previous studies using numerical model simulations have indicated that the characteristic spatial scales of the GWs generated from a jet are much smaller than that of the local jet itself (O’Sullivan and Dunkerton 1995; Zhang 2004; Plougonven and Snyder 2007). This result is consistent with the scale analysis on the dispersion relation of spontaneously radiated GWs (Snyder et al. 2009). They showed that the characteristic spatial scale of the GWs should be $O(Ro)$ in the large-scale flow with the characteristic spatial scale of $O(1)$ so that their time scales were comparable by the strong Doppler effects. Note again that this time-scale matching between the large-scale flow and GWs is necessary for the quasi resonance. These are the reasons for assumptions (iii) and (iv).

3 Our theory is constructed by eliminating little-Doppler-shifted GWs at the first step. A similar theory can be constructed without using this elimination but based on the assumptions (ii)–(vi). In this case, however, it is necessary to deal with twice as many terms as for our method.
Assumptions (v) and (vi) are related to (iv). All interaction coefficients in the nondimensional primitive equations [(24)–(26)] are zeroth-order homogeneous with respect to wavenumbers (see appendix B for details). Thus, the magnitude of an interaction coefficient describing the local interaction in the wavenumber space is \(O(1)\), because all wavenumbers in the coefficient have the same order. On the other hand, for the nonlocal interactions, the magnitude of an interaction coefficient can be greater than \(O(1)\), because the magnitude of the numerator of the coefficient may be much larger than that of the denominator. In particular, the interaction between the vortical flow and GWs is considered to be nonlocal, because their characteristic wavelengths are between the vortical flow and GWs is considered to be nonlocal, because their characteristic wavelengths are largely different. Actually, it is evident from appendix B that these interaction coefficients are \(O(\text{Ro}^{-1})\), meaning that these coefficients must be expanded with respect to \(\text{Ro}\) in order to correctly obtain the regular perturbation solution (procedure A of the RG method in appendix A).

In addition, the diagnostic components include the slaved components; hence, the low-wavenumber diagnostic components are not negligible. This fact suggests that the expansion is also necessary for the interaction coefficients between the diagnostic component and GW component.

The lowest order in all expanded interaction coefficients is \(\text{Ro}^{-1}\). The coefficients of \(O(\text{Ro}^{-1})\) are important because they represent the advection of GWs. From the specific formulas in appendix B, the interaction coefficients of \(O(\text{Ro}^{-1})\) between the vortical flow and GW component are as follow:

\[
\gamma_{m_k\gamma_m}^{(1)} = \gamma_{m_k\gamma_m}^{(0)} + \frac{\text{Ro}^2}{2} \left( k_1^2 m_2 - k_2^2 m_1 \right) \frac{1}{\text{m}^2}, \quad (27)
\]

Clearly, \(I_{\gamma_{m_k\gamma_m}}^{(1)} \) and \(I_{\gamma_{m_k\gamma_m}}^{(0)} \) represent the advection of the GW component by the vortical flow. Higher-order coefficients do not include such an advection effect but represent shear effects of the vortical flow on the GW component.

The same-order interaction coefficients between the diagnostic component and GW component are obtained in the following:

\[
I_{\delta_{m_k\gamma_m}}^{(1)} + I_{\delta_{m_k\gamma_m}}^{(1)} = I_{\delta_{m_k\gamma_m}}^{(0)} = I_{\delta_{m_k\gamma_m}}^{(0)} + I_{\delta_{m_k\gamma_m}}^{(1)},
\]

\[
I_{\delta_{m_k\gamma_m}}^{(0)} = I_{\delta_{m_k\gamma_m}}^{(0)} + I_{\delta_{m_k\gamma_m}}^{(1)},
\]

\[
I_{\delta_{m_k\gamma_m}}^{(1)} = I_{\delta_{m_k\gamma_m}}^{(0)} + I_{\delta_{m_k\gamma_m}}^{(1)},
\]

\[
I_{\delta_{m_k\gamma_m}}^{(2)} = I_{\delta_{m_k\gamma_m}}^{(0)} + I_{\delta_{m_k\gamma_m}}^{(1)} = 0,
\]

\[
(29)
\]

The interaction coefficients are not symmetrical against the exchange of wavenumbers seen in subscripts. Thus, it is necessary to take into account another coefficient in which the subscript wavenumbers are exchanged with each other, as (29)–(31). In other words, it is necessary to consider a pair of products of \(\gamma_k\) and/or \(\delta_k\) with the same coefficient as in the left-hand side (LHS) of the following example:

\[
\sum_m \left[ \gamma_{m_k}^{(1)} \delta_{m_k}^{(1)} (g_{m_k}^{(1)} + \gamma_{m_k}^{(0)}) \delta_{m_k}^{(1)} \right] = \sum_m \left( I_{\gamma_{m_k}}^{(1)} + I_{\gamma_{m_k}}^{(1)} \right) \delta_{m_k}^{(1)} \gamma_{m_k}^{(0)}.
\]

Similar to (28), coefficients (30) and (31) represent the advection of the GW component by the velocity associated with \(\gamma_{m_k}^{(1)}\) and \(\gamma_{m_k}^{(0)}\). Higher-order interaction coefficients between the diagnostic component and GW component do not include such an advection effect but represent shear effects of the velocity associated with the diagnostic components.

It should be emphasized again that all expanded interaction coefficients of \(O(\text{Ro}^{-1})\) represent the advection of the GW component. The interaction coefficients of \(O(\text{Ro}^{-1})\) with “superscripts” \(q_k\) are zero in (27) and (29), because the advection of GWs is not included in the time-evolution equation for \(q_k\). As this case, it is natural that the advection of GWs is included only in the time-evolution equation for the GW component. This also means that the interaction coefficients of \(O(\text{Ro}^{-1})\) should not be included in the equations for the diagnostic component, as in assumption (vi).

b. Derivation of the renormalization group equations

In this subsection, by taking \(\text{Ro}\) as a small number, the RG method is applied to the transformed primitive equations [(24)–(26)]. First, all dependent variables (i.e., \(q_k\), \(\delta_k^{GW}\), \(\gamma_k^{GW}\), \(\gamma_k^{\text{diag}}\), and \(\gamma_k^{\text{diag}}\)) are expanded with respect to \(\text{Ro}\), and regular perturbation solutions are derived (procedure A of the RG method in appendix A). For example, \(q_k\) is expanded as \(\gamma_k = q_k^{(0)} + \text{Ro} q_k^{(1)} + \text{Ro}^2 q_k^{(2)} + \cdots\).
First, the zeroth-order regular perturbation solutions are examined. For the vortical flow, \( q^{(0)} \) is constant, by (24), and corresponds to the constant of integration, which is the object of renormalization. In other words, \( q^{(0)}_k \) will be represented by the renormalization constants and renormalized variables later.

\[
\text{Ro}\mathbf{L} = \text{Ro} \left( \begin{array}{cc} 0 & \omega_k^2/\text{Ro} \\ -1/\text{Ro} & 0 \end{array} \right) + \left( \begin{array}{c} -I_{q_{\text{m}}k_m}^{(1)} q_{\text{m}}^{(0)} \\ -I_{q_{\text{m}}k_m}^{(1)} q_{\text{m}}^{(0)} \end{array} \right),
\]

where matrix \( \mathbf{Y} \) contains both the advection (i.e., the Dopp-ler effect) and shear effects. The GW modes are obtained as eigenmodes of RoL whose eigenvalues are \( O(\text{Ro}) \). Matrix \( \mathbf{X} \) is \( O(1) \), while \( \mathbf{Y} \) is also \( O(1) \) in a high-wavenumber region.

To derive the solution of the zeroth-order GW components, the GW mode defined in assumption (ii) is formulated mathematically by using matrix \( \mathbf{L} \), which consists of the intrinsic frequency of plane GWs \( \mathbf{X} \) and the interaction terms with the vortical flow \( \mathbf{Y} \) as follows. The matrix \( \mathbf{L} \) operates on the column vector \( (\gamma_{k-m}, \delta_{k-m})^T \).

When \( \mathbf{X} \) and \( \mathbf{Y} \) are largely canceled, the eigenvalues, which correspond to the ground-based frequencies of the eigenmodes, can be reduced to \( O(\text{Ro}) \). In addition, \( \mathbf{L} \) is expanded with respect to \( \text{Ro} \) in the following:

\[
\text{Ro}\mathbf{L} = \text{Ro} \sum_{n=0} \text{Ro}^n \mathbf{L}^{(n)} = \text{Ro} \left[ \frac{1}{\text{Ro}} \left( \begin{array}{cc} -I_{q_{\text{m}}k_m}^{(1)} q_{\text{m}}^{(0)} & \omega_k^2 \\ -1 & -I_{q_{\text{m}}k_m}^{(1)} q_{\text{m}}^{(0)} \end{array} \right) \mathbf{L}^{(0)} \right] + \sum_{n=1} \text{Ro}^{n-1} \left( \begin{array}{c} -I_{q_{\text{m}}k_m}^{(a)} q_{\text{m}}^{(b)} \\ -I_{q_{\text{m}}k_m}^{(a)} q_{\text{m}}^{(b)} \end{array} \right),
\]

where the order of \( \mathbf{L}^{(n)} \) is taken to be unity. All advection effects on GWs by the vortical flow are included in \( \mathbf{L}^{(0)} \), while all shear effects of the vortical flow on GWs are included in the others \( \mathbf{L}^{(n)} \) \( (n = 1, 2, \ldots) \). The lowest-order matrix \( \mathbf{L}^{(0)} \) appears in the first-order equation, because Ro \( \mathbf{L}^{(0)} \) is \( O(\text{Ro}) \). Thus, from (25) and (26), \( \gamma_{k}^{\text{GW}(0)} \) and \( \delta_{k}^{\text{GW}(0)} \) are constant in fast time. They are constants of integration similar to \( q^{(0)}_k \) and, hence, the objects of renormalization.

One may think that the eigenmode representation should be introduced here because the matrices \( \mathbf{L}^{(n)} \) have been formulated. However, it should be noted that the solution for \( q^{(1)}_k \) is not described in the form of \( \sum_n \text{Ro}^n q^{(n)}_k \), but is obtained after the renormalization. Thus, the final form of \( \mathbf{L} \) is determined after renormalizing \( q_k \). The eigenmodes are obtained by eigendecomposing \( \mathbf{L} \) with the final form.

For the diagnostic components, the time derivatives in (25) and (26) are ignored because of assumption (i), and the terms whose interaction coefficients are \( O(\text{Ro}^{-1}) \) are also ignored as a result of assumption (vi). Thus, \( \delta_{k}^{\text{diag}(0)} \) and \( \gamma_{k}^{\text{diag}(0)} \) are zero.

Next, the first-order regular perturbation solutions are examined. The solution \( q^{(1)}_k \) is obtained as follows:

\[
q^{(1)}_k = \sum_m \left( q_{q_{\text{m}}k_m}^{(1)} + \sum_{m} \left( \begin{array}{c} q_{q_{\text{m}}k_m}^{(1)} q_{\text{m}}^{(0)} + q_{q_{\text{m}}k_m}^{(1)} q_{\text{m}}^{(0)} \gamma_{k-m}^{\text{GW}(0)} \delta_{k-m}^{\text{GW}(0)} q_k^{(1)} \\ q_{q_{\text{m}}k_m}^{(1)} q_{\text{m}}^{(0)} + q_{q_{\text{m}}k_m}^{(1)} q_{\text{m}}^{(0)} \gamma_{k-m}^{\text{GW}(0)} \delta_{k-m}^{\text{GW}(0)} q_k^{(1)} \end{array} \right) t. \right)
\]
The nonlinear terms consisting of $q_k^{(0)}$ only, those of $q_k^{(0)}$ and the GW component, and those of the GW components only are respectively superscribed with “balance,” “$R_{\text{GG}}$,” and “$R_{\text{GG}}$.” Here, $q_k^{R_{\text{GG}}(1)}$ and $d_k^{R_{\text{GG}}(1)}$ represent the tendencies of $q_k$ by the GW component; hence, they correspond to the GW radiation reaction on the vortical flow. The term $q_k^{R_{\text{GG}}(1)}$ includes part of the advection of $q$ by the GWs, while $d_k^{R_{\text{GG}}(1)}$ represents the formation of $q$ because of the nonlinear interaction between the GWs. Clearly, the solution $q_k^{(1)}$ is composed only of the secular terms proportional to $t$. This means that a quasi resonance with a quasi-steady forcing may occur, as zeroth-order solutions ($q_k^{(0)}, \gamma_{GW}^{(0)}$, and $\delta_{GW}^{(0)}$) are constant in fast time.

The first-order GW component satisfies the following equation:

$$
\frac{d}{dt} \left( \begin{array}{c} \gamma_k^{GW(1)} \\ \delta_k^{GW(1)} \end{array} \right) + L^{(0)} \left( \begin{array}{c} \gamma_{k-m}^{GW(0)} \\ \delta_{k-m}^{GW(0)} \end{array} \right) = F^{(1)}. 
$$

The solution is

$$
\left( \begin{array}{c} \gamma_k^{GW(1)} \\ \delta_k^{GW(1)} \end{array} \right) = -L^{(0)} \left( \begin{array}{c} \gamma_{k-m}^{GW(0)} \\ \delta_{k-m}^{GW(0)} \end{array} \right) t + F^{(1)} t, 
$$

where $F^{(1)}$ represents a forcing composed of nonlinear terms. The first row $F_1^{(1)}$ and second row $F_2^{(1)}$ of $F^{(1)}$, which are the respective forcings for $\gamma_k^{GW(1)}$ and $\delta_k^{GW(1)}$, are written as follows:

$$
F_1^{(1)} = \sum_m I_{m}^{\gamma_k} q_m^{(0)} q_{k-m}^{(0)} + \sum_m \left( I_{m}^{\gamma_k^{(0)} \delta_m^{GW(0)}} + I_{m}^{\gamma_k^{(1)} \delta_m^{GW(0)}} + I_{m}^{\gamma_k^{(0)} \gamma_m^{GW(0)}} \right) 
$$

$$
+ I_{m}^{\gamma_k^{(1)} \delta_m^{GW(0)}} + I_{m}^{\gamma_k^{(0)} \gamma_m^{GW(0)}} \right) 
$$

$$
F_2^{(1)} = \sum_m I_{m}^{\delta_k} q_m^{(0)} q_{k-m}^{(0)} + \sum_m \left( I_{m}^{\delta_k^{(0)} \delta_m^{GW(0)}} + I_{m}^{\delta_k^{(1)} \delta_m^{GW(0)}} + I_{m}^{\delta_k^{(0)} \gamma_m^{GW(0)}} \right) 
$$

$$
+ I_{m}^{\delta_k^{(1)} \delta_m^{GW(0)}} + I_{m}^{\delta_k^{(0)} \gamma_m^{GW(0)}} \right) 
$$

Here, the following summation symbol is introduced:

$$
\sum_{m} I_{m}^{\gamma_k^{(1)} \delta_m^{GW(0)}} + \sum_{m} I_{m}^{\gamma_k^{(1)} \gamma_m^{GW(0)}} 
$$

where $a^{\gamma_k^{(1)}}, \delta_m^{GW(0)},$ and the others are arbitrary complex numbers. This summation symbol means that the term whose superscripts are exchanged with each other is added when the superscripts are different.

Similar to $q_k^{(1)}$, the first-order GW components consist of the secular terms only; hence, the quasi resonance with the quasi-steady forcing $F^{(1)}$ may occur. It is worth noting here that, in general, not all secular terms contribute to the RGEs. After renormalization, it becomes clear which secular term modifies the amplitude of the GW component.

The diagnostic components $\gamma_k^{\text{diag}(1)}$ and $\delta_k^{\text{diag}(1)}$ are obtained as follows, by ignoring the time-derivative terms and the terms with interaction coefficients of $O(\text{Ro}^{-1})$ in (25) and (26), as for the zeroth-order case:
The renormalization constant $\gamma_k^{\text{diag}(1)}$ is given by

$$
\gamma_k^{\text{diag}(1)} = -\sum_m \delta_m^{\delta_k q_m} q_m^{(0)} \delta_k q_m^{(0)} - \sum_m (\delta_m^{\delta_k q_m} q_m^{(0)} \delta_k q_m^{(0)}) + \delta_k^{R_{\text{GW}(0)} q_m^{(0)} \delta_k q_m^{(0)} + \delta_k^{R_{\text{GW}(0)} q_m^{(0)} \delta_k q_m^{(0)}} + \delta_k^{R_{\text{GW}(0)} q_m^{(0)} \delta_k q_m^{(0)}})
$$

where a superscript “slave” denotes the term corresponding to the slaved component in the balanced model.

Next, the secular terms are renormalized using the RG method (Chen et al. 1994, 1996). All $\tau$ in the secular terms are transformed into $t = (t - \tau) + \tau$ (procedure B of the RG method). To cancel the divergence due to $\tau$ at each order of $\text{Ro}$, renormalization constants and renormalized variables are introduced (procedure C of the RG method). For example, by introducing the renormalization constant $q_k^{(1)}$ and renormalized variable $q_k^{(0)}$, the secular term $\delta_k q_k^{(0)}$ is represented by $q_k^{(0)} = q_k^{(0)} + \text{Ro} q_k^{(1)}$. Similar manipulations are performed for the GW components. Note that as the diagnostic components do not have secular terms, such manipulations are not necessary.\(^4\)

First, the linear potential vorticity is renormalized. Determining $q_k^{(1)}$ so that the divergence due to $\tau$ is removed, we obtain the renormalized perturbation solution as follows (procedure D of the RG method):

$$
q_k(t, \tau) = q_k^{(0)} + \text{Ro} \cdot [q_k^{\text{balance}(1)}(\tau) + q_k^{R_{\text{GW}(1)}}(\tau)] + \text{Ro} q_k^{(1)}(\tau) + \text{Ro}^2 q_k^{(2)}(\tau) + \ldots
$$

where each variable with a tilde is composed of the renormalized variables instead of the zeroth-order ones (e.g., $q_k^{\text{balance}(1)} = \sum_m q_m^{\delta_k q_m} q_m^{(0)} q_m^{(0)}$). The RG and $q_k^{(1)}$ are obtained by $\partial q_k^{(1)}/\partial \tau = 0$ and $t = \tau$ (procedure E of the RG method):

$$
\frac{d q_k^{(1)}}{d t} = \text{Ro} \cdot [q_k^{\text{balance}(1)} + q_k^{R_{\text{GW}(1)}} + q_k^{R_{\text{GW}(1)}}] + \text{Ro}^2 q_k^{(2)} + \ldots
$$

The RGE for the GW components is derived by determining the renormalization constants (procedures D and E of the RG method):

$$
\frac{d}{d t} \left( \begin{array}{c}
\gamma_k^{\text{GW}} \\
\delta_k^{\text{GW}}
\end{array} \right) + \text{Ro} \frac{d}{d t} \left( \begin{array}{c}
\gamma_k^{\text{GW}} \\
\delta_k^{\text{GW}}
\end{array} \right) = \text{Ro} \tilde{F}_p^{(1)} + \text{Ro}^2
$$

where $\tilde{F}_p^{(1)}$ represents the nonlinear forcing projected onto the eigenmode. The third term on the RHS of (50) represents the variation of $q_k^{(1)}$ due to the change in the basis functions (i.e., eigenmodes). The order of this third term is Ro because $\frac{d q_k^{(1)}}{d t} = O(\text{Ro})$ by (46); hence, all terms on the RHS of (50) are $O(\text{Ro})$. This result means that the amplitudes of the GW modes vary in slow time $s = \text{Ro} t$, similar to $\tilde{q}$. By inversely

\(^4\)The second-order solutions for the diagnostic components have secular terms. These terms are automatically renormalized when renormalizing $q_k^{\text{GW}}$, $\gamma_k^{\text{GW}}$, and $\delta_k^{\text{GW}}$. See appendix D for details.
transforming $\mathbf{g}$ obtained by (50), the solutions for the GW components represented by the plane-wave bases ($\gamma_k^{GW}$ and $\delta_k^{GW}$) can be obtained. It should be emphasized again that the GWs are simply obtained from this eigenmode expansion and that the spectral decomposition or the spatial-scale separation are not used to identify GWs in the present study.\footnote{In addition, note that the assumption (iv) with respect to the GW characteristic scale has not been used to obtain the GW modes. As shown in Part II, the GW modes, which are obtained as eigenmodes, contain motions with intermediate spatial scales as well as short ones [$\sim O(\text{Ro})$]. Here, the intermediate scale means the spatial scale between the characteristic scale of the balanced flow and the characteristic scale of GWs.}

Last, the diagnostic components are obtained from (42) and (43) using $\tilde{q}_k$, $\tilde{\gamma}_k^{GW}$, and $\tilde{\delta}_k$, instead of $\tilde{q}_k^{(0)}$, $\tilde{\gamma}_k^{GW(0)}$, and $\tilde{\delta}_k^{GW(0)}$:

\[
\tilde{\delta}_k^\text{diag} = \text{Ro} \tilde{\gamma}_k^\text{diag} + O(\text{Ro}^2) \quad \text{and} \quad 
\tilde{\gamma}_k^\text{diag} = \text{Ro} \tilde{\gamma}_k^\text{diag} + O(\text{Ro}^2).
\] (52)

In this way, we have finally obtained the system of first-order RGEs that is composed of the time-evolution equations for the linear potential vorticity [(46)] and the GW mode [(50)], and the formulas giving the diagnostic components [(52)]. Note that all variables vary in slow time $s (=\text{Ro})$. See appendixes C and D for the derivations of the second-order regular perturbation solutions and RGEs [(D1), (D3), and (D5)], respectively. All slaved components in the second-order balanced model are shown in appendix E.

c. Characteristics of the RGE system

In this subsection, we discuss the GW sources and reactions of the GW radiation in the RGE system. An important point is that the GW source $\mathbf{P}_p^{(1)}$ on the RHS of (50) includes the term proportional to the slaved components:

\[
\mathbf{P}_p^{\text{slave}(1)} = \mathbf{P}^{-1} \left( \omega_k^\text{slave(1)} \tilde{q}_k^\text{slave(1)} \right).
\] (53)

The other terms of the GW source ($\mathbf{P}_p^{(1)} - \mathbf{P}_p^{\text{slave}(1)}$) are classified into two groups; one is the advection of the GWs by the velocity associated with the diagnostic components, which corresponds to the terms of the second summation in the first lines of (38) and (39), and the other is the deformation of the sources by the GW radiation, which corresponds to the terms of the second lines in (38) and (39). The slaved components are considered to be the largest of the GW sources, because they are composed only of $\tilde{q}_k$. Note that $\tilde{q}_k$ is larger than $\delta_k^\text{diag}$, $\gamma_k^\text{diag}$, $\gamma_k^{GW}$, and $\delta_k^{GW}$ for most large-scale motions. The slaved components are generally not zero, because they are diagnostically obtained from the vortical flow. In other words, the slaved components are part of the balanced flow. This fact means that the GW sources, including the slaved components, are usually not zero.

In this theory, it is considered that the quasi resonance could occur and generate the GWs from the slaved components, because the GW modes have the same-order ground-based frequencies as the balanced flow. The exact frequency matching between the GW modes and their sources is not necessary, which is different from a pure resonance. This is because $\mathbf{P}_p^{(1)}$ is diagnostically determined by the projection of $\mathbf{F}_p^{(1)}$ onto the GW eigenmode, and the frequency of $\mathbf{F}_p^{(1)}$ is generally not related to the frequency (i.e., eigenvalue) of the GW mode. More importantly, this fact means that the spatial scales as well as time scales of GWs need to be partially comparable to those of the GW sources (i.e., slaved components).

It is also important that the nonlinear terms containing the GW components can be regarded as the GW radiation reactions in (35), (38), (39), (42), (43), (46), and (52). The existence of these terms means that the variations in the vortical flow, the diagnostic components, and the GW sources are affected by the spontaneously radiated GWs (i.e., the reaction of the GW radiation). The GW sources also include an implicit reaction through the slaved components, which are varied by the reaction on $q$. In particular, the reactions on $q$ (i.e., $q_R^{(0)}(1)$ and $q_R^{(0)}(1)$ in (46)) are possibly the most important reactions of GW radiation in the RGE theory, because they describe the variation in the large-scale (vortical) flow due to the GWs.

It is necessary to confirm the validity of the RGE theory, because we made several assumptions to construct the theory. Here, it is verified on the basis of the forms of $q_R^{(0)}(1)$ and $q_R^{(0)}(1)$ that the RGEs likely describe the reaction of GW radiation on the large-scale flow. The difference in the reaction on the vortical flow between the RGE (46) and the original equation for $q$ [(24)] is attributed to the form of interaction coefficients. The interaction coefficients in $q_R^{(0)}(1)$ (i.e., $I_{q_R^{(0)},q_R^{(0)}}^a$, $I_{q_R^{(0)},q_R^{(0)}}^b$, and $I_{q_R^{(0)},q_R^{(0)}}^c$) are the same as those in the original equation [(24)]. On the contrary, the interaction coefficients in $q_R^{(0)}(1)$ (i.e., $I_{q_R^{(0)},q_R^{(0)}}^a$ and $I_{q_R^{(0)},q_R^{(0)}}^b$) are the leading terms of those (i.e., $I_{q_R^{(0)},q_R^{(0)}}^a$ and $I_{q_R^{(0)},q_R^{(0)}}^b$) in the original equation [(24)]. Note that the superscripts (0) emphasize this fact. The spatial distribution and magnitude of $q_R^{(0)}(1)$ in the vortex dipole hardly change if the leading coefficients are replaced with the original ones (not shown). Thus, it is considered that the RGE theory describes the GW radiation reaction on the large-scale vortical flow reasonably well. Note that
the balanced flow in the RGE system (i.e., the sum of the vortical flow and slaved components) asymptotically approaches one in the balanced model, as the GW components approach zero (see appendix E). The difference between the two balanced flows is due to the GW radiation reaction.

In the linear theory, GWs do not have potential vorticity (PV), which indicates that a PV field is generally hardly affected by GWs in the real atmosphere. However, a finescale spatial structure of PV, which is likely due to the GWs, has been recognized slightly above the tropopause at midlatitudes in the general circulation model in which GWs are explicitly resolved (Miyazaki et al. 2010). The RGE theory likely describes such fine PV structures as the reactions of GW radiation. This also suggests that it may sometimes be hard to extract and/or remove GWs from observation and/or numerical model data.

Finally, two remarks need to be made regarding the application of the RGEs to numerical calculations. First, it is necessary to change the basis functions from the eigenmodes to the plane waves at each step when integrating the RGEs. This is because most processes, such as the GW radiation reactions on $\hat{q}_k$ in (46), the GW sources in (38) and (39), and the diagnostic components in (42) and (43), are represented by the plane-wave bases, while the time-evolution equation for the GWs [(50)] is described by the eigenmode bases. Second, the elements of the GW-mode set may vary in time because the eigenvalues also vary in time. However, this is not a severe problem, because it is expected that $\delta_{k,GW}$ and $\gamma_{GW}^*$ composed of the GW modes change almost continuously in slow time.

d. Separability between diagnostic and GW components on RGE theory

A major difference in our treatment from previous studies is that the Doppler shift has been taken into account for the GW components (or modes). The GW modes having low ground-based frequencies are mainly composed of high wavenumbers so as to be affected by a significant Doppler shift. In contrast, the diagnostic components are mainly composed of the slaved components having low wavenumbers. However, the spatial scales of GWs need to be partially comparable to those of diagnostic (or slaved) components in order that a quasi resonance could occur, as discussed in the previous subsection. In other words, projections of the GW sources on some GW eigenmodes need to be present for the quasi resonance [i.e., $F^{(1)}_p \neq 0$ in (50) or $F^{(2)}_p \neq 0$ in (D5)]. As will be examined in Part II, low-wavenumber GW components are essential for the spontaneous radiation, whose spatial scales are comparable to those of the slaved components. This fact suggests that some of the GW components and diagnostic components may hardly be distinguished, as discussed by McIntyre (2009), although both components have been assumed to be separable in our theory. In this subsection, we discuss again the validity of the separability and also discuss an effect of the “inseparability” between GW components and diagnostic (or slaved) components on the RGE theory.

The diagnostic components are mainly composed of the slaved components that are proportional to the products of $\hat{q}_k$. This fact suggests that the diagnostic components consist mainly of low wavenumbers; hence, the Doppler effect on them is quite weak. Thus, the GW components that are largely Doppler shifted are expected to include few diagnostic components. This means that the diagnostic components and GW components are usually separable with sufficient accuracy.

Next, we discuss changes in the RGE theory due to the inseparability. The separability may not be certain when “characteristic” spatial scales of diagnostic components (determined by the wavenumbers at the peaks of energy density spectra) are comparable to those of GW components. It is considered that the first-order diagnostic components are separable from the GW components more clearly than the second-order diagnostic ones, because the first-order ones consist of lower-order nonlinear interactions (or terms).

In the first-order RGE system, only the first-order diagnostic components are present. This fact indicates that the separability between diagnostic and GW components is generally valid, compared with the separability in the second-order RGE system, in which the second-order diagnostic components as well as first-order ones are present. When the first-order diagnostic components are not separable from the GW ones in the first-order RGE system, the first-order diagnostic ones and slaved ones cannot be defined. This means that all terms including the first-order diagnostic components are taken to be zero. In addition, the main GW sources in (38) and (39) cannot be replaced with the first-order slaved components; for example, the first term in (38) remains $\sum_m F^{(0)}_k \delta_{\omega k,m} q_{m}$, and it cannot be replaced with $\delta_{\omega k}^2 \delta_{k,\text{slav}(1)}$. Thus, the quasi resonance between GWs and first-order slaved components considered in the separable case needs to be interpreted as that between GWs and nonlinear sources composed only of $q$. In other words, GWs are radiated by the nonlinear self-interactions of vortical flow.

In the second-order RGE system, the first- and second-order diagnostic components are present. The separability between the second-order diagnostic and GW components may not be certain, compared with that between the first-order diagnostic and GW components. Thus, the former is
discussed further. When the second-order diagnostic components are not separable from the GW ones in the second-order RGE system, the second-order diagnostic ones, as well as the second-order slaved ones, cannot be defined. This means that all terms including the second-order diagnostic components are taken to be zero. In addition, some main GW sources in (D5) cannot be replaced with the second-order slaved components: for example, the first term in the first row of (D8), which is the main part of the GW source $F_p$ in (D5), remains $\hat{\gamma}_{\text{GW}}^{(2)}$, and it cannot be replaced with $\omega_k \delta_k^{\text{slave(2)}}$. Thus, the quasi-resonance between GWs and second-order slaved components considered in the separable case needs to be interpreted as that between GWs and nonlinear sources that are composed of $q$ and the first-order slaved component.

In this way, only the interpretation on the GW radiation may be changed depending on the separability, while it is considered that the first- and second-order RGE systems can be derived regardless of the separability. In the present paper and in Part II, we consider that the first- and second-order diagnostic components may be changed depending on the separability, and it cannot be replaced with $\omega_k \delta_k^{\text{slave(2)}}$. Thus, the quasi-resonance between GWs and second-order slaved components considered in the separable case needs to be interpreted as that between GWs and nonlinear sources that are composed of $q$ and the first-order slaved component.

The higher-order diagnostic components (higher than second order) may include significant high-wavenumber components because of multiple nonlinear interactions (or high-order nonlinear terms). This fact suggests a possible inseparability between diagnostic and GW components; hence, higher-order RGES (higher than second order) may not be able to be derived in the same way as used for the first- and second-order RGES.

4. Discussion

In this section, the relations between our new theory and the theories proposed in previous studies are discussed. In section 4a, it is shown that the RGE can be formally reduced to a BAE (e.g., Plougonven and Zhang 2007). The advantages and disadvantages of the RGE system are compared with those of BAЕs. In section 4b, the adjustment process described by the RGE system is compared with the balance adjustment process (Zhang 2004). Section 4c features a discussion on how the wave capture (Bühler and McIntyre 2005) is described in the RGE system.

a. Relation to balance adjustment equations

BAЕs are linear equations that describe the GW radiation from the residuals of the balanced-model solution. In contrast, RGЕs are nonlinear equations that describe time evolutions of both the large-scale flow ($q$, $\gamma_{\text{const(1)}}$, and $\delta_{\text{const(1)}}$) and the GWs ($\gamma_{\text{GW}}$ and $\delta_{\text{GW}}$). The nonlinear terms in the RGE for the GWs are regarded as the GW radiation reactions on the GW sources, as discussed in section 3c. The RGE is formally reduced to a BAE if these nonlinear terms in the GW sources are neglected. To demonstrate this, the second-order RGE is compared with the BAE derived from the first-order BM. The first-order BM is obtained by eliminating all $O(\text{Ro}^2)$ terms in the second-order BM (see appendix E). The first-order BM derived using the RG method corresponds to the QG model (Wisconsino et al. 2002).

All terms that are proportional to the GW component in the GW sources are neglected in the second-order RGE represented by the plane-wave bases. This manipulation is equivalent to neglect the GW radiation reactions and the advection of the GWs by the velocity associated with the slaved components. A similar discussion can be made when these advection terms remain. Equation (D5) is transformed into the following linear equation:

$$
\frac{d}{dt} \begin{pmatrix} \gamma_{\text{GW}}^k & \gamma_{\text{GW}}^{k-m} \\ \delta_{\text{GW}}^k & \delta_{\text{GW}}^{k-m} \end{pmatrix} + \text{Ro} \begin{pmatrix} \omega_k \delta_k^{\text{slave(1)}} & \gamma_{\text{GW}}^{k-m} \\ -\gamma_k^{\text{slave(1)}} & -\text{Ro}^2 \left( \omega_k \delta_k^{\text{slave(2)}} + \delta_k^{\text{slave(2)}} \right) \end{pmatrix} + O(\text{Ro}^3),
$$

where $\text{Ro}$ is equal to $\text{Ro}^{\text{const(1)}}$. $\gamma_{\text{GW}}^{(0)}$ represents the advection by the vortical flow and the intrinsic frequency of plane GWs, while $\omega_k^{\text{const(1)}}$ represents the shear effects of the vortical flow on GWs:

$$
\hat{\mathbf{L}} = \begin{pmatrix}
-\left( \frac{I_m^{\gamma_k^{(0)}} + I_m^{\gamma_k^{(0)}}}{\text{Ro}} \right) - \left( \frac{\gamma_k^{(0)}}{\text{Ro}} \right) & - \left( \frac{\gamma_k^{(0)}}{\text{Ro}} \right) - \left( \frac{\gamma_k^{(0)}}{\text{Ro}} \right) \\
- \left( \frac{I_m^{\gamma_k^{(0)}}}{\text{Ro}} \right) - \left( \frac{\gamma_k^{(0)}}{\text{Ro}} \right) & \left( \frac{\gamma_k^{(0)}}{\text{Ro}} \right) + \left( \frac{\gamma_k^{(0)}}{\text{Ro}} \right)
\end{pmatrix} + \begin{pmatrix} 0 & \omega_k^2 \\ -\frac{1}{\text{Ro}} & 0 \end{pmatrix}.
$$
From now on, the original RGE system for the large-scale flow [(D1) and (D3)] and the GWs [(D5)] is distinguished from the linearized RGE for the GWs [(54)] by using the terms “original RGE system” and “linearized RGE,” respectively.

Next, we carry out similar manipulations in the derivation of the BAES (Plougonven and Zhang 2007; Snyder et al. 2009; Wang and Zhang 2010) and obtain the BAE represented in the wavenumber space. The terms δ_k and γ_k are assumed to be the sums of the first-order slaved components [(E2) and (E3)] and the eddy components correspond to the secondary flow in Plougonven and Zhang (2007):

\[
\delta_k = Ro\delta_k^{\text{slave(1)}} + \delta_k^{\text{eq(2)}} \quad \text{and} \quad \gamma_k = Ro\gamma_k^{\text{slave(1)}} + \gamma_k^{\text{eq(2)}}. \tag{57}
\]

The eddy components generally consist of both GW and high-order slaved components. By substituting (57) into the primitive equations [(25) and (26)] while also neglecting the products of eddy components, we obtain the BAE:

\[
\begin{aligned}
\frac{d}{dt}(\gamma_k^{\text{eq(2)}}) + Ro \left[ \begin{array}{c}
-\frac{\partial}{\partial \gamma_k^{\text{eq(1)}}} - \frac{\partial}{\partial \gamma_k^{\text{eq(1)}}}
\end{array} \right] & + \left( \begin{array}{c}
\frac{\omega_k^2}{Ro}
\end{array} \right) \left( \begin{array}{c}
\gamma_k^{\text{eq(2)}}
\end{array} \right) + O(Ro^3) \\
&= -Ro \frac{d}{dt}(\gamma_k^{\text{eq(2)}}) + Ro^2 \left( \begin{array}{c}
\sum_m (\lambda_m \gamma_k^{\text{eq(1)}}) + \frac{1}{Ro} \left( \begin{array}{c}
-\frac{\partial}{\partial \gamma_k^{\text{eq(1)}}}
\end{array} \right) \right) + O(Ro^3). \tag{58}
\end{aligned}
\]

The time derivative in the second line of (58) is transformed into the slaved components (i.e., the first term in the third line) using (E1)–(E3).

It should be emphasized here that all second-order slaved components in (58) are seemingly identical to those in the second-order BM, but they are not exactly identical. The symbol = formally is used in (58) to emphasize this fact. However, the differences in the second-order slaved components between (58) and the second-order BM are considered to be quite small, because the difference in δ_k between the first- and second-order BMs is almost negligible for most large-scale motions. It is worth noting that only these small differences in the second-order slaved components are the “true” imbalances. Here a true imbalance means an imbalance other than any slaved components. In the BAES, all the sources in (58) including the slaved components of the balanced flow are called imbalances. This may not be appropriate, because the imbalance is originally defined as the difference between the real flow and balanced one. As shown in Part II, GWs spontaneously radiated in a vortex dipole can be reproduced by the linearized RGE with high accuracy. This result implies that the main GW sources are not the true imbalances but the slaved components.

In summary, the linearized RGE (54) and the BAE (58) are linear equations that describe the GW radiation from the sources mainly composed of the slaved components. In particular, as shown in Part II, two main GW sources in the linearized RGE (i.e., γ_k^{\text{slave(1)}} and δ_k^{\text{slave(1)}}) are also contained in the BAE (58). Thus, it has been shown that the RGE theory can be regarded as an extension of the BAE theory (Plougonven and Zhang 2007; Snyder et al. 2009; Wang and Zhang 2010).

However, this fact does not mean that the only difference between BAES and our theory is in the interpretation of GW sources. The most essential difference is in the
proposed mechanism of GW radiation. In the BAE theory, it is considered that GWs are radiated from the residuals of the balanced-model solution that are called “imbalances.” In contrast, in our theory, GWs are considered to be radiated through the quasi resonance with the slaved components of the balanced flow. Note again that the slaved components are a part of the balanced flow. In other words, the RGE theory physically clarifies that imbalances (i.e., GWs) can be produced in a balanced-flow field because of the time-scale matching (or, equivalently, the quasi resonance). To treat a quasi resonance correctly, the RGE theory requires the eigendecomposition of \( \mathbf{L} \) even when an equation is linear, although the eigendecomposition is not performed for the linearized RGE (54) to compare with the BAE (58). In addition, it should be also emphasized here that in the linearized (and original) RGE, all slaved components can be GW sources if a quasi resonance occurs. This point is significantly different from the BAE theory that shows that only the second-order slaved components are regarded as the sources. As shown in Part II, the most significant source in vortex dipoles is one of the first-order slaved components \( \gamma_{k}^{(1)} \) that is not included in the source in the BAE (58). The RGE theory elucidates two important facts in physical nature that are inherent in the BAEs: GW sources are the slaved components, and GW radiation is due to the quasi resonance.

In addition, the original RGE system can describe the GW radiation reaction on the large-scale flow, unlike the BAEs. However, there are a few disadvantages in the original and linearized RGEs. The most critical one is that it may be difficult to apply the RGEs to high-resolution observation and/or numerical model data, unlike with the BAEs. There are two reasons for this. First, the Fourier expansion is used to construct the RGE theory, which suggests that it may be impossible for the RGEs to describe GW radiation near the ground, such as in a surface dipole (Snyder et al. 2007). Second, the eigendecomposition of \( \mathbf{L} \) is necessary when integrating the RGEs. This eigendecomposition generally requires a lot of computational resources. The applicability to observation and/or numerical model data is an advantage of the BAEs.

Another disadvantage is that the RGEs cannot describe the GW radiation from the true imbalances, because imbalances are not originally included in the RGE system. In addition, the RGEs cannot describe the radiation of GWs having high ground-based frequencies. Some studies have shown that significant GWs including high-frequency ones are radiated during the initial adjustment (e.g., Snyder et al. 2007, 2009). This initial adjustment (i.e., three-dimensional Rossby adjustment) cannot be addressed using the RGE theory. The BAEs may have the potential to describe the GW radiation from the true imbalances and to address such initial adjustments, although a detailed investigation is needed.

\[ \text{b. Adjustment process} \]

Our RGE theory proposes a new adjustment process associated with spontaneous GW radiation, which is very different from the balance adjustment theory (Zhang 2004). We call this new adjustment process “resonance adjustment.” In the balance adjustment process, it is considered that imbalances are continuously produced by the time evolution of an approximately balanced large-scale flow. GWs are radiated from these imbalances. This cycle is continuously repeated until the large-scale flow becomes so weak that the GW radiation is hardly recognized.

To explain the new adjustment process using our theory (i.e., resonance adjustment), let us consider that an approximately balanced large-scale flow is sufficiently strong at first. GW sources are always present in the large-scale flow field, because the GW sources are slaved components to the vortical flow; namely, they are part of the balanced flow. Note that the slaved components are diagnostically obtained by the PV distribution. The strong large-scale flow makes GWs significantly Doppler shifted, leading to the time-scale matching between slaved components and Doppler-shifted GWs that is needed for the quasi resonance. This quasi resonance is the excitation mechanism of GWs. The large-scale flow becomes weak by the GW radiation, because the GWs take the energy away from the balanced flow. Subsequently, the Doppler shift becomes insignificant, and the quasi resonance rarely occurs. On the whole, the large-scale flow weakens and the GW radiation is ceased. This scenario can answer the question as to why the spontaneous GW radiation is sometimes continuous in time, as in a vortex dipole (Snyder et al. 2007; Vúdez 2007), and it also explains why the balanced-model solution is usually not an exact solution for the primitive equations.

The time-scale matching was also addressed by Lott et al. (2010, 2012). They proposed that spontaneous radiation can be considered as transient amplifications by a superposition of neutral modes composed of balanced and imbalanced parts by analyzing the linear primitive equations with a constant vertical shear flow in an infinite domain. In their papers, the Doppler effect is infinitely stronger as the altitude gets higher; hence, the balanced wave solution is considered to be smoothly connected with the imbalanced wave (i.e., GW) solution at higher altitudes. However, in the real atmosphere, a region with a strong basic flow is confined rather than extended infinitely in space. Thus, time-scale matching should be realized locally, not globally, in space. In the present paper, this local time-scale matching has been
achieved using the eigenmodes of the fast variables in
a given vortical flow, which are composed of some plane-
wave components, such as wave packets, but are not
identical to the one plane-wave solution.

c. Expression of wave capture

Spontaneously radiated GWs may be captured in
a large-scale flow through the process of “wave capture”
(Bühler and McIntyre 2005; Plougonven and Snyder
2005; Wang et al. 2010). Wave capture is important for
wave–mean flow interactions and wave propagation
into the middle atmosphere. GWs in a large-scale flow
having significant deformation are contracted by the
advection, meaning that the GW group velocities as-
ymptotically approach the large-scale flow velocity, and
finally the GWs are captured in the large-scale flow. The
BAEs can describe the wave-capture process (Snyder
et al. 2009; Wang and Zhang 2010; Wang et al. 2010). As
described earlier, the RGE can be formally reduced to
the BAE, indicating that the RGE can also describe the
wave-capture process, although the process is not ex-
licitly described in the RGE system.

In the RGE system, a GW packet is regarded as a su-
perposition of eigenmodes determined by the large-
scale vortical flow. This means that the advection effect
on the GW packet is contained in the eigenmodes and
eigenvalues. The initial amplitudes of the eigenmodes
composing the GW packet are determined by the GW
sources, and each eigenmode evolves in time according
to its eigenvalue (i.e., ground-based frequency). The
above description makes it clear that the structure and
evolution of GW packets depend not only on large-scale
flow but also on the initial GW sources, although the
wave capture is implicitly described in the RGE system.
Bühler and McIntyre (2005) showed that GWs generally
break near the edge of a capture region as a result of
eddy (or turbulent) diffusion. In Part II, by integrating
the RGE with diffusion, it will be shown that the RGE
can describe the wave capture and that the structure of
GW packets in the dipole is affected by the GW sources
as well as the background flow.

5. Summary and concluding remarks

We have developed a new theory on spontaneous GW
radiation that deals with the coexistence system com-
posed of GWs, slaved components, and vortical flow that
interact with each other (Fig. 1). The essence of this
theory is that GW radiation is due to a quasi resonance
between the largely Doppler-shifted GWs and the
slaved components (i.e., part of the balanced flow) that
occurs when the space–time scales of the GWs are par-
tially comparable to those of the slaved components.

The theoretical equations were derived from the
hydrostatic Boussinesq equations on the f plane using
the renormalization group (RG) method (Chen et al.
1994, 1996), which is a singular perturbation method.
In general, the RG method makes it possible to ex-
tract the slowly varying components systematically
and naturally from a system containing multiple time-
scale motions. The derived equations, called re-
normalization group equations (RGEs), describe not
only the spontaneous GW radiation from the slaved
components through the quasi resonance but also the
GW radiation reactions on the balanced flow and on the
GW sources.

It was shown that the RGE for the time evolution of
GWs can be formally reduced to a balance adjustment
equation (BAE) (e.g., Plougonven and Zhang 2007) if
the GW radiation reactions on the GW sources are
neglected. This fact suggests that the RGE can de-
scribe the wave-capture process like the BAE can,
although the process is not explicitly described. The
RGE theory clarified two important facts on the
physical nature inherent in the BAEs. First, the GW
sources, which were called imbalances in the BAEs,
are composed mainly of the slaved components (i.e.,
a part of the balanced flow). Second, the GW radia-
tion mechanism is the quasi resonance with the slaved
components. Thus, it is not generally appropriate to
call all GW sources in the BAEs (i.e., residuals of the
balanced-model solution) imbalances, because im-
balances are generally defined as the difference be-
tween the real flow and the balanced one. On the
other hand, there are a few disadvantages in the RGE
theory. The most critical one is that the RGEs are
hard to use for analyses of observation and/or nu-
merical model data.

The RGE theory provides a new scenario of the ad-
justment process (i.e., resonance adjustment). GWs are
radiated continuously through the quasi resonance, be-
cause the GW sources (i.e., slaved components) are al-
ways present, and then the large-scale flow weakens
because of the GW radiation. When the flow becomes
too weak to make GWs significantly Doppler shifted,
the time scales of GWs cannot be matched with those of
slaved components. Eventually, quasi resonance rarely
occurs, and the GW radiation is ceased.

In Part II, the validity of the RGE theory derived
in this paper will be confirmed using numerical sim-
ulations. Moreover, the physical interpretations on
the GW sources, as in Figs. 2 and 3, are discussed in
detail.

There is an interesting issue regarding the GW ra-
diation reaction. As discussed in section 3c, possibly
the most important reactions are represented by $\dot{q}^{RGe(1)}$
and $\tilde{q}^{R_{00}(1)}$, because their low-wavenumber components represent the variation in the large-scale (vortical) flow caused by the GWs. However, as noted in section 2, the RGEs cover the motions with various spatial scales. To examine only such large-scale variations due to the GW radiation, it may be effective to further renormalize the derived RGEs with respect to space so as to extract large-scale structures in a flow field.

Our theory is in contrast to the theories proposed by Vanneste and Yavneh (2004) and Ólafsdóttir et al. (2008), in which the exponential asymptotics has been used for the small $Ro$ (Rossby number) limit. According to their theories, the amplitude of high-ground-based-frequency GWs excited from a balanced flow is proportional to $\exp(-\alpha/\text{Ro})$, where $\alpha$ is a constant real number. On the contrary, we showed that low-ground-based-frequency GWs could be excited whose amplitudes are proportional to the power of $\text{Ro}$. The difference in the GW frequencies is attributed to the degree of Doppler shift. Note that although Vanneste and Yavneh (2004) and Ólafsdóttir et al. (2008) considered the Doppler shift, the Doppler shift did not affect the GW radiation, because both GWs and balanced flow were Doppler shifted in the same way. These results suggest that the GW amplitude with respect to $\text{Ro}$ depends on the degree of Doppler shift. In Part II, by using the numerical model data, it will be shown that the Doppler-shifted GW amplitude in the vortex dipole is proportional to the power of $\text{Ro}$ in the range of $0.15 \leq \text{Ro} \leq 0.4$.

**Acknowledgments.** This study was supported partly by Grant-in-Aid for Research Fellow (258466), by Grant-in-Aid for Scientific Research (A) (25247075), and by Grant-in-Aid for the Young Scientists (B) (25800265) of the Ministry of Education, Culture, Sports and Technology (MEXT), Japan, and by the Leading Graduate Course for Frontiers of Mathematical Sciences and Physics (FMSP) program at the University of Tokyo.

**APPENDIX A**

**Example Application of the RG Method**

In this section, an application of the RG method (Chen et al. 1994, 1996) is explained using an ordinary differential equation. The RG method is a singular perturbation method that has several features, as follow: (i) The derived differential equations referred to as renormalization group equations (RGEs) describe the slow variation of the system with respect to time and/or space. (ii) Many singular perturbation methods can be regarded as RG methods [i.e., the boundary layer method, WKB method, multiple-scale method, and reductive perturbation method; see Bender and Orszag (1999) for a detailed discussion on these methods]. (iii) No particular preparations, such as the proper introduction of multiple time scales in the multiple-scale method, are required. (iv) The solution is sometimes more accurate than that obtained by using another method, such as the boundary layer method.

An equation similar to the Van der Pol equation is considered as an example:

$$\frac{d^2x}{dt^2} + x = \epsilon \left(1 - \left[x^2 + \left(\frac{dx}{dt}\right)^2\right]^{3/2}\right) \frac{dx}{dt}. \quad (A1)$$

The Van der Pol equation is a well-known equation describing the time evolution of a nonlinear oscillator. Its solution asymptotically approaches a closed periodic orbit in the phase space (i.e., a limit cycle). The LHS of (A1) corresponds to a harmonic oscillator with a period of 1, while the RHS corresponds to a forcing term that is negative when the amplitude of $x$ is larger than 1 and positive when the amplitude of $x$ is smaller than 1. Here, $\epsilon$ is a small parameter.

The RG method is applied following procedures A–E. In procedure A, the regular perturbation solution is derived. When $x$ is expanded to $x = x^{(0)} + \epsilon x^{(1)} + \cdots$, the first-order perturbation solution is obtained:

$$x^{(0)} = R_0 \sin(t + \Theta_0) \quad \text{and} \quad (A2)$$

$$x^{(1)} = \frac{(1 - R_0^2)R_0}{2} t \sin(t + \Theta_0), \quad (A3)$$

where $R_0$ and $\Theta_0$ are constants of integration. The first-order homogeneous solution is omitted for clarity (Kunihiro 1995, 1997). The first-order solution [(A3)] consists only of a secular term proportional to $t$, which diverges as $t \to \infty$. When the first-order solution becomes comparable to the zeroth-order solution [$t = O(1/\epsilon)$], the assumption on the perturbation theory (i.e., $x_0 \gg \epsilon x_1$) breaks down, and the perturbation solution becomes invalid. This is because the resonance with the forcing composed of the zeroth-order solution occurs in the first-order perturbation equation. Clearly, the sign of the forcing term in the original equation [(A1)] varies depending on the amplitude of $x$; hence, the divergence with respect to $t$ is inappropriate. The above example indicates that in the regular
perturbation method it is impossible to describe the effects of feedback from high-order to low-order perturbations. This is the main reason for the existence of secular terms.

In procedure B, to solve the problem attributed to the secular term \( t \) (the cause for the divergence) is transformed into \( t' = (t - \tau) + \tau \), where \( \tau \) is an arbitrary real number. In procedure C, the renormalized variables \((\tilde{R}, \tilde{\Theta})\) are introduced as follows: \( R_0 = \tilde{R}(\tau) + \sum_n e^n a^{(n)}(\tau) \) and \( \Theta_0 = \tilde{\Theta}(\tau) + \sum_n e^n b^{(n)}(\tau) \). In procedure D, the renormalization constants \( a^{(n)} \) and \( b^{(n)} \) are determined so that the divergence due to \( \tau \) is removed at each order: \( a^{(1)} = - (1 - \tilde{R}^2) \tilde{R} \tau / 2 \) and \( b^{(1)} = 0 \) from (A2) and (A3). Thus, the renormalized perturbation solution is obtained as follows:

\[
x(t, \tau) = \tilde{R} \sin(t + \tilde{\Theta}) + \epsilon \left(1 - \frac{\tilde{R}^2}{2}\right) \frac{(t - \tau) \sin(t + \tilde{\Theta})}{(t - \tau)^2} + O(\epsilon^2). \quad (A4)
\]

In procedure E, since \( \tau \) does not appear in the original equation [(A1)], \( \partial x / \partial \tau \) should be zero for any \( t \). The RGEs describing the time evolution are obtained from \( \partial x / \partial \tau = 0 \).

\[
\frac{d\tilde{R}}{d\tau} = \epsilon \left(1 - \frac{\tilde{R}^2}{2}\right) + O(\epsilon^2) \quad \text{and} \quad (A5)
\]

\[
\frac{d\tilde{\Theta}}{d\tau} = 0 + O(\epsilon^2). \quad (A6)
\]

Finally, \( t \) is substituted into \( \tau \) because \( \tau \) is arbitrary. The solution \( x = \tilde{R} \sin(t + \tilde{\Theta}) + O(\epsilon^2) \) is obtained. Here, \( \tilde{R} \) and \( \tilde{\Theta} \) are obtained by solving (A5) and (A6) as

\[
\tilde{R} = \frac{R(0)}{\sqrt{e^{-\epsilon t} + R(0)^2(1 - e^{-\epsilon t})}} + O(\epsilon^2 t) \quad \text{and} \quad (A7)
\]

\[
\tilde{\Theta} = \Theta(0) + O(\epsilon^2 t), \quad (A8)
\]

where \( R(0) \) and \( \Theta(0) \) are determined by the initial conditions. The amplitude \( \tilde{R} \) approaches 1 with respect to the slow time \( \epsilon t \), which is different from the fast time \( t \) of the harmonic oscillator. This fact suggests that the original equation [(A1)] has a limit cycle, which was also the case in the Van der Pol equation. In fact, when the Poincaré–Bendixson theorem is applied, it is shown that this system has a limit cycle. In this way, it is possible to extract the slow motion systematically and naturally by applying the RG method. It is also shown that the RGEs can be regarded as the envelope equations from a geometric point of view (Kunihiro 1995, 1997).

The procedures for applying the RG method are summarized as follows:

A) The regular perturbation solution for \( x(t) \) is derived.

B) The independent variables \( t \), which cause the divergence of the secular terms, are transformed into \( t' = (t' - \tau') + \tau' \), where \( n \) is a natural number and \( \tau' \) is an arbitrary real number.

C) Renormalized variables and renormalization constants are introduced. The constants of integration in the zeroth-order solution are represented by these quantities.

D) The renormalization constants are determined so that the divergence due to \( \tau \) is removed at each order, and the renormalized perturbation solution \( x(t, \tau) \) is obtained.

E) The RGEs are derived by \( \partial x(t, \tau) / \partial \tau = 0 \) for any \( t \).

The approximate solution for the original equation is obtained by using the solutions for the RGEs and substituting \( t \) into \( \tau \).

APPENDIX B

Specific Forms of Interaction Coefficients

In this section, all 18 interaction coefficients in (24)–(26) are specified. The potentials \( \psi_k, \phi_k, \) and \( \chi_k \), respectively related to \( q_k, \delta_k \), and \( \gamma_k \), are used to represent formulas simply [see also (17)–(23)]:

\[
\psi_k = \frac{q_k}{|k|^2}, \quad (B1)
\]

\[
\phi_k = -\frac{\delta_k}{|k_H|^2}, \quad \text{and} \quad (B2)
\]

\[
\chi_k = -\frac{k_3^2 \gamma_k}{|k|^4 |k_H|^2}. \quad (B3)
\]

It is worthwhile to discuss the relations of the above potentials with other potentials used in the previous studies on the dynamics of balanced flows. Muraki et al. (1999) introduced the potentials \( \Phi, F, G; \) \( \Phi \) corresponds to \( \psi \) (slow variable), while \( F \) and \( G \) correspond to linear superpositions of \( \phi \) and \( \chi \) (fast variables) in the present paper. On the other hand, Dritschel and Viúdez (2003) introduced the vector potential \( \phi \); its horizontal components correspond to \( (G, -F) \) in Muraki et al. (1999) and, hence, the linear superposition of \( \phi \) and \( \chi \) (fast variables) in the present paper, while its vertical component...
corresponds to $\Phi$ in Muraki et al. (1999) and, hence, $\psi$ (slow variable) in the present paper. These examples show that there are degrees of freedom in the choice of slow and fast variables. In the present paper, we have used $(q, \delta, \gamma)$ with the respective potentials $(\psi, \phi, \chi)$, because they are, respectively, related to the physical quantities of linear potential vorticity, horizontal divergence, and ageostrophic vorticity.

The following interaction coefficients are represented using the above potentials [(B1)–(B3)]. All coefficients are zeroth-order homogeneous with respect to wavenumbers. For example, both the numerator and denominator of $I_{q,\omega,q}$ in (B4) are the sums of the fourth power of wavenumbers, although its denominator is hidden in the product of the potentials $\psi_m \psi_l$, where $l$ denotes $k - m$:

\[
I_{q,\omega,q}^{\delta} q_m q_l = \frac{1}{V} \left[ k^2_k l^2_k (l^2_m + l^2_l) + m^2_m l^2_l + m^2_k l^2_k - l^2_k m^2_k - l^2_m m^2_m \right] \psi_m \psi_l, \quad (B7)
\]

\[
I_{q,\omega,q}^{\gamma} q_m q_l = \frac{1}{V} \left[ l^2_k m^2_m + m^2_k m^2_m - l^2_k l^2_m - l^2_k l^2_m \right] \psi_m \psi_l, \quad (B8)
\]

\[
I_{q,\omega,q}^{q} q_m q_l = \frac{1}{V} \left[ k^2_k l^2_k (l^2_m + l^2_l) - m^2_m l^2_l + m^2_k l^2_k - l^2_k m^2_k - l^2_m m^2_m \right] \psi_m \psi_l, \quad (B9)
\]

\[
I_{q,\omega,q}^{\delta} q_m q_l = \frac{1}{V} \left[ k^2_k l^2_k (l^2_m + l^2_l) - m^2_m l^2_l - m^2_k l^2_k - l^2_k m^2_k - l^2_m m^2_m \right] \psi_m \psi_l, \quad (B10)
\]

\[
I_{q,\omega,q}^{\gamma} q_m q_l = \frac{1}{V} \left[ l^2_k m^2_m + l^2_k m^2_m - l^2_k l^2_m - l^2_k l^2_m \right] \psi_m \psi_l, \quad (B11)
\]

\[
I_{q,\omega,q}^{\gamma} q_m q_l = \frac{1}{V} \left[ k^2_k l^2_k (l^2_m + l^2_l) - m^2_m l^2_l - m^2_k l^2_k - l^2_k m^2_k - l^2_m m^2_m \right] \psi_m \psi_l, \quad (B12)
\]
\[
I_{\delta_{m_1}}^{\delta_{m_1}} \delta_{m_1} \gamma_1 = \frac{1}{V} \left[ k_2 m_1 l_1^2 - k_1 m_2 l_2^2 + m_1 l_2 (k_2 m_2 - k_1 l_1) - \frac{m_1^2}{m_3} (k_2 m_1 - k_1 m_2) \right] \phi_m \psi_1,
\]
(B13)

\[
I_{\delta_{m_1}}^{\delta_{m_1}} \delta_{m_1} \gamma_1 = \frac{1}{V} \left[ \frac{k_2^2 l_1}{k_3 l_3} \left( m_1 l_1 + m_2 l_2 - \frac{m_2^2}{m_3} l_3 \right) + k_1 k_2 (m_1 l_2 + m_2 l_1) - m_1 l_2 (k_2 m_2 - k_1 l_1) \right. \\
+ m_2 l_2 (k_2 l_2 - k_1 m_1) - \frac{m_1^2}{m_3} l_3 (k_1 l_2 - k_2 l_1) \right] \phi_m \psi_1,
\]
(B14)

\[
I_{\delta_{m_1}}^{\delta_{m_1}} \delta_{m_1} \gamma_1 = \frac{1}{V} \left[ -l_1 l_2 (k_1 m_2 + k_2 m_1) \\
+ m_2 l_2 (k_1 m_2 + k_2 l_1) \right] \phi_m \psi_1.
\]
(B15)

\[
I_{\delta_{m_1}}^{\delta_{m_1}} \delta_{m_1} \gamma_1 = \frac{1}{V} \left[ -l_1 l_2 (k_1 m_2 + k_2 m_1) \\
+ m_1 l_2 (k_1 m_2 + k_2 l_1) \right] \phi_m \psi_1.
\]
(B16)

\[
I_{\delta_{m_1}}^{\delta_{m_1}} \delta_{m_1} \gamma_1 = \frac{1}{V} \left[ -l_1 l_2 (k_1 m_2 + k_2 m_1) \\
+ m_1 l_2 (k_1 m_2 + k_2 l_1) \right] \phi_m \psi_1.
\]
(B17)

\[
I_{\delta_{m_1}}^{\delta_{m_1}} \delta_{m_1} \gamma_1 = \frac{1}{V} \left[ l_1 (k_2 m_2 - k_2 m_1) \\
+ l_2 (k_1 m_2 - k_2 m_1) + l_1 l_2 (k_2 m_2 - k_1 m_1) \\
+ \frac{m_1}{m_3} l_3 (k_1 m_2 - k_1 m_1) \right] \psi \phi_1.
\]
(B18)

\[
I_{\delta_{m_1}}^{\delta_{m_1}} \delta_{m_1} \gamma_1 = \frac{1}{V} \left[ \frac{l_1}{l_2} m_1 k_1 + \frac{l_2}{l_3} m_2 k_2 + m_1 l_2 (k_1 m_2 + k_2 l_1) \\
+ \frac{m_1}{l_3} m_3 k_H \cdot m_H \right] \phi_m \psi_1.
\]
(B19)

\[
I_{\delta_{m_1}}^{\delta_{m_1}} \delta_{m_1} \gamma_1 = \frac{1}{V} \left[ -l_1 l_2 [k_2 m_1 + m_2 (k_1 l_2 + k_2 m_1)] \\
+ l_1 [k_2 m_2 + m_1 (k_1 m_2 + k_2 l_1)] \\
+ \frac{m_1 l_3}{m_3} l_2 (k_1 l_2 - k_2 l_1) \right] \phi_m \psi_1, \text{ and}
\]
(B20)

\[
I_{\delta_{m_1}}^{\delta_{m_1}} \delta_{m_1} \gamma_1 = \frac{1}{V} \left[ l_1 l_2 (k_1 m_2 + k_2 m_1) \\
- m_1 l_2 (k_1 m_2 + k_2 l_1) \right] \phi_m \psi_1.
\]
(B21)

**APPENDIX C**

**Derivation of the Second-Order Regular Perturbation Solutions**

In this section, we derive the second-order regular perturbation solutions. For a simple description, a symbol expressing the coefficient of the power of \( t \) is introduced as follows:

\[ a = b \text{ when } a = bt^n + c \text{ and } \]
\[ a = b = 0 \text{ when } a = b, \]
(C1) (C2)

where \( n \) is an arbitrary integer, and \( a, b, \) and \( c \) are arbitrary constant complex numbers. The symbol \( \sum \text{swap} \) is redefined as the sum of all terms whose superscripts and underlines are replaced simultaneously with each other as shown in the example below:

\[
\sum \text{swap} \left( c^{x_1(1)} d^{y_2(2)} \right) = \sum \left( c^{x_1(1)} d^{y_2(2)} + c^{y_2(2)} d^{x_1(1)} \right).
\]
(C3)

The second-order regular perturbation solution for \( q_k^{(2)} \) is as follows:
These secular terms are automatically absorbed into the above solutions have secular terms proportional to $t$. In contrast to the first-order solutions [(42) and (43)], the second-order solutions for the diagnostic components are as follows:

$$ q^{(2)}_k = \sum_m \left( I_{\alpha m} q^{(0)}_{\alpha m} q^{(1)}_{\alpha m} + I_{\alpha m} q^{(0)}_{\alpha m} \delta_{\alpha m} q^{(1)}_{\alpha m} + I_{\alpha m} q^{(0)}_{\alpha m} q^{(1)}_{\alpha m} \right) + \frac{1}{\omega_k} \sum_m \left( I_{\alpha m} q^{(0)}_{\alpha m} q^{(1)}_{\alpha m} + I_{\alpha m} q^{(0)}_{\alpha m} \delta_{\alpha m} q^{(1)}_{\alpha m} + I_{\alpha m} q^{(0)}_{\alpha m} q^{(1)}_{\alpha m} \right) \frac{t}{2} + \sum_m \left( I_{\alpha m} q^{(0)}_{\alpha m} q^{(1)}_{\alpha m} + I_{\alpha m} q^{(0)}_{\alpha m} \delta_{\alpha m} q^{(1)}_{\alpha m} + I_{\alpha m} q^{(0)}_{\alpha m} q^{(1)}_{\alpha m} \right) \frac{t}{2}$$

The second-order solutions for the diagnostic components are as follows:

$$ \gamma^{(2)}_{\text{diag}} = -\sum_m \left( I_{\alpha m} q^{(0)}_{\alpha m} q^{(1)}_{\alpha m} + I_{\alpha m} q^{(0)}_{\alpha m} \delta_{\alpha m} q^{(1)}_{\alpha m} + I_{\alpha m} q^{(0)}_{\alpha m} q^{(1)}_{\alpha m} \right) + \frac{1}{\omega_k} \sum_m \left( I_{\alpha m} q^{(0)}_{\alpha m} q^{(1)}_{\alpha m} + I_{\alpha m} q^{(0)}_{\alpha m} \delta_{\alpha m} q^{(1)}_{\alpha m} + I_{\alpha m} q^{(0)}_{\alpha m} q^{(1)}_{\alpha m} \right) \frac{t}{2} + \sum_m \left( I_{\alpha m} q^{(0)}_{\alpha m} q^{(1)}_{\alpha m} + I_{\alpha m} q^{(0)}_{\alpha m} \delta_{\alpha m} q^{(1)}_{\alpha m} + I_{\alpha m} q^{(0)}_{\alpha m} q^{(1)}_{\alpha m} \right) \frac{t}{2}$$

and

$$ \delta^{(2)}_k = \frac{1}{\omega_k} \sum_m \left( I_{\alpha m} q^{(0)}_{\alpha m} q^{(1)}_{\alpha m} + I_{\alpha m} q^{(0)}_{\alpha m} \delta_{\alpha m} q^{(1)}_{\alpha m} + I_{\alpha m} q^{(0)}_{\alpha m} q^{(1)}_{\alpha m} \right) + \frac{1}{\omega_k} \sum_m \left( I_{\alpha m} q^{(0)}_{\alpha m} q^{(1)}_{\alpha m} + I_{\alpha m} q^{(0)}_{\alpha m} \delta_{\alpha m} q^{(1)}_{\alpha m} + I_{\alpha m} q^{(0)}_{\alpha m} q^{(1)}_{\alpha m} \right) \frac{t}{2} + \sum_m \left( I_{\alpha m} q^{(0)}_{\alpha m} q^{(1)}_{\alpha m} + I_{\alpha m} q^{(0)}_{\alpha m} \delta_{\alpha m} q^{(1)}_{\alpha m} + I_{\alpha m} q^{(0)}_{\alpha m} q^{(1)}_{\alpha m} \right) \frac{t}{2}$$

In contrast to the first-order solutions [(42) and (43)], the above solutions have secular terms proportional to $t$. These secular terms are automatically absorbed into the renormalization constants of $q_k$, $\gamma^{GW}_k$, and $\delta^{GW}_k$ by applying the RG method.

The GW components satisfy the following equation:
The formulas giving the diagnostic components are as follows:

$$
\gamma_k^{\text{diag}} = \text{Ro} \gamma_k^{\text{diag}(1)} + \text{Ro}^2 \gamma_k^{\text{diag}(2)} + O(\text{Ro}^3) \quad \text{and} \quad \delta_k^{\text{diag}} = \text{Ro} \delta_k^{\text{diag}(1)} + \text{Ro}^2 \delta_k^{\text{diag}(2)} + O(\text{Ro}^3).
$$

The first-row components of nonlinear forcings ($G^{(2)}$ and $F^{(2)}$) are as follows:

$$
\frac{d}{dt} \begin{pmatrix} g_k^{(2)} \\ \delta_k^{(2)} \end{pmatrix} = G^{(2)} + F^{(2)}. \tag{C7}
$$

The RGE for the GW components is expressed as

$$
\frac{d}{dt} \begin{pmatrix} g_k^{\text{GW}(2)} \\ \delta_k^{\text{GW}(2)} \end{pmatrix} + L^{(0)} \begin{pmatrix} \gamma_k^{\text{GW}(1)} \\ \delta_k^{\text{GW}(1)} \end{pmatrix} + L^{(1)} \begin{pmatrix} \gamma_k^{\text{GW}(0)} \\ \delta_k^{\text{GW}(0)} \end{pmatrix} = G^{(2)} + F^{(2)}. \tag{C8}
$$

The second-row components are obtained similarly. The solutions are in the following:

$$
\begin{align*}
G^{(2)}_1 &= \sum_m \left[ I_m \sqrt{C_2} \right] \gamma_m^{(2)} \delta_k^{(2)} + I_m \sqrt{C_2} \gamma_m^{(1)} \delta_k^{(1)} + I_m \sqrt{C_2} \gamma_m^{(0)} \delta_k^{(0)} + I_m \sqrt{C_2} \gamma_m^{(1)} \delta_k^{(1)} + I_m \sqrt{C_2} \gamma_m^{(0)} \delta_k^{(0)} \\
F^{(2)}_1 &= \sum_m \left[ I_m \sqrt{C_2} \right] \gamma_m^{(2)} \delta_k^{(2)} + I_m \sqrt{C_2} \gamma_m^{(1)} \delta_k^{(1)} + I_m \sqrt{C_2} \gamma_m^{(0)} \delta_k^{(0)} + I_m \sqrt{C_2} \gamma_m^{(1)} \delta_k^{(1)} + I_m \sqrt{C_2} \gamma_m^{(0)} \delta_k^{(0)} 
\end{align*}
$$

The second-row components are obtained similarly. The solutions are in the following:

$$
\begin{align*}
G^{(2)}_1 &= \sum_m \left[ I_m \sqrt{C_2} \right] \gamma_m^{(2)} \delta_k^{(2)} + I_m \sqrt{C_2} \gamma_m^{(1)} \delta_k^{(1)} + I_m \sqrt{C_2} \gamma_m^{(0)} \delta_k^{(0)} + I_m \sqrt{C_2} \gamma_m^{(1)} \delta_k^{(1)} + I_m \sqrt{C_2} \gamma_m^{(0)} \delta_k^{(0)} \\
F^{(2)}_1 &= \sum_m \left[ I_m \sqrt{C_2} \right] \gamma_m^{(2)} \delta_k^{(2)} + I_m \sqrt{C_2} \gamma_m^{(1)} \delta_k^{(1)} + I_m \sqrt{C_2} \gamma_m^{(0)} \delta_k^{(0)} + I_m \sqrt{C_2} \gamma_m^{(1)} \delta_k^{(1)} + I_m \sqrt{C_2} \gamma_m^{(0)} \delta_k^{(0)} 
\end{align*}
$$

The RGE for the GW components is expressed as

$$
\frac{d}{dt} \begin{pmatrix} g_k^{\text{GW}(2)} \\ \delta_k^{\text{GW}(2)} \end{pmatrix} = \text{Ro} \left[ L^{(0)} + \text{Ro} L^{(\text{const} (1))} \right] \begin{pmatrix} \gamma_k^{\text{GW}(1)} \\ \delta_k^{\text{GW}(1)} \end{pmatrix} + \text{Ro} \tilde{F}^{(1)} + \text{Ro}^2 \tilde{F}^{(2)} + O(\text{Ro}^3). \tag{D4}
$$

In the second-order RGE, $L$ is equal to $L^{(0)} + \text{Ro} L^{(\text{const} (1))}$, which is different from the first-order RGE, where $L = L^{(0)}$. Equation (D4) is represented in terms of the eigenmodes of $L$, as follows:

$$
\frac{d}{dt} \begin{pmatrix} g_k \\ \delta_k \end{pmatrix} = -\text{Ro} \tilde{g} + \text{Ro} \tilde{F}^{(1)} + \text{Ro}^2 \tilde{F}^{(2)} \tag{D5}
$$

Similar to the first-order RGE, $\tilde{g}$, $\tilde{F}^{(1)}$, and $\tilde{F}^{(2)}$ are obtained by transforming the basis functions.
\[ \Omega = \text{diag}(\tilde{\Omega}_1, \tilde{\Omega}_2, \ldots) = \mathbf{P}^{-1} \mathbf{L} \mathbf{P}, \]
\[ \tilde{\mathbf{g}} = \mathbf{P}^{-1} \left( \mathbf{\gamma}_k^{\text{GW}} \right) \], and
\[ \mathbf{F}_p^{(1)\text{or}(2)} = \mathbf{P}^{-1} \mathbf{F}_p^{(1)\text{or}(2)}. \]

The solutions for the GW components (\( \gamma_k^{\text{GW}} \) and \( \delta_k^{\text{GW}} \)) are obtained by inversely transforming \( \tilde{\mathbf{g}} \), calculated by (D5) as follows:

\[ \gamma_k^{\text{GW}} = \gamma_k + O(\text{Ro}^3) \quad \text{and} \quad \delta_k^{\text{GW}} = \delta_k + O(\text{Ro}^3). \]

The forcing \( \mathbf{F}_p^{(2)} \) contains the terms composed only of the slaved components in the following. Each slaved component is shown in appendix E.

**APPENDIX E**

**Balanced Model Derived by the Second-Order RG Method**

The second-order balanced model is derived by using the RG method. This model is obtained by substituting \( \delta^{\text{GW}}, \gamma^{\text{GW}} = 0 \) into the second-order RGEs in appendix D [(D1), (D3), and (D5)]. This manipulation corresponds to the fact that the Doppler shift is assumed to be so weak that there are no GW modes [i.e., the orders of all eigenvalues are larger than \( O(\text{Ro}) \)]. The derived model is the same as a model obtained by applying the method of Wirosotoisno et al. (2002) to the hydrostatic Boussinesq equations on the \( f \) plane. In the balanced model, \( \tilde{q} \) is given by the time-evolution equation, while \( \tilde{\delta} \) and \( \tilde{\gamma} \) are determined diagnostically. In other words, \( \tilde{\delta} \) and \( \tilde{\gamma} \) are slaved to \( \tilde{q} \).

\[ \frac{dq_k}{dt} = \text{Ro} \sum_m \left( \int_{q_k^{\text{slave}(1)}} \hat{q}_m \hat{q}_k \right) + \text{Ro}^2 \sum_m \left( \int_{q_k^{\text{slave}(1)}} \hat{q}_m \hat{q}_k \right) + \text{Ro}^2 \sum_m \frac{\text{swap}}{\omega_k} \left( \int_{q_k^{\text{slave}(1)}} \hat{q}_m \hat{q}_k \right) \]

where \( q_k = \tilde{q}_k + O(\text{Ro}^3) \).
\[
\delta_k = \text{Ro} \left( \frac{1}{\omega_k^2} \sum_m \tilde{q}^\text{slave}_m \delta_{m,k-m} \right) + \text{Ro}^2 \frac{1}{\omega_k^2} \sum_m \tilde{q}^\text{balance}_m \delta_{m,k-m} + \text{Ro}^2 \sum_m \left( \frac{1}{\omega_k^2} \tilde{q}^\text{slave(1)}_m \delta_{m,k-m} + \frac{1}{\omega_k^2} \tilde{q}^\text{slave(2)}_m \delta_{m,k-m} \right) + O(\text{Ro}^3).
\]

**REFERENCES**


