Physical uniqueness of higher-order Korteweg-de Vries theory for continuously stratified fluids without background shear

Shimizu, Kenji

(Citation)
Physics of Fluids, 29(10):106604-106604

(Issue Date)
2017-10

(Resource Type)
journal article

(Version)
Version of Record

Physical uniqueness of higher-order Korteweg-de Vries theory for continuously stratified fluids without background shear

Cite as: Phys. Fluids 29, 106604 (2017); https://doi.org/10.1063/1.5008767
Submitted: 23 April 2017 . Accepted: 07 October 2017 . Published Online: 26 October 2017

Kenji Shimizu

Physical uniqueness of higher-order Korteweg-de Vries theory for continuously stratified fluids without background shear

Kenji Shimizu[a)]
Kobe University, 1-1 Rokkodai-Cho, Nada-Ku, Kobe 657-0013, Japan

(Received 23 April 2017; accepted 7 October 2017; published online 26 October 2017)

The 2nd-order Korteweg-de Vries (KdV) equation and the Gardner (or extended KdV) equation are often used to investigate internal solitary waves, commonly observed in oceans and lakes. However, application of these KdV-type equations for continuously stratified fluids to geophysical problems is hindered by nonuniqueness of the higher-order coefficients and the associated correction functions to the wave fields. This study proposes to reduce arbitrariness of the higher-order KdV theory by considering its uniqueness in the following three physical senses: (i) consistency of the nonlinear higher-order coefficients and correction functions with the corresponding phase speeds, (ii) wavenumber-independence of the vertically integrated available potential energy, and (iii) its positive definiteness. The spectral (or generalized Fourier) approach based on vertical modes in the isopycnal coordinate is shown to enable an alternative derivation of the 2nd-order KdV equation, without encountering nonuniqueness. Comparison with previous theories shows that Parseval's theorem naturally yields a unique set of special conditions for (ii) and (iii). Hydrostatic fully nonlinear solutions, derived by combining the spectral approach and simple-wave analysis, reveal that both proposed and previous 2nd-order theories satisfy (i), provided that consistent definitions are used for the wave amplitude and the nonlinear correction. This condition reduces the arbitrariness when higher-order KdV-type theories are compared with observations or numerical simulations. The coefficients and correction functions that satisfy (i)-(iii) are given by explicit formulae to 2nd order and by algebraic recurrence relationships to arbitrary order for hydrostatic fully nonlinear and linear fully nonhydrostatic effects. *Published by AIP Publishing.* https://doi.org/10.1063/1.5008767

I. INTRODUCTION

Internal solitary waves (ISWs) are commonly observed on continental shelves and in marginal seas,[1–3] in estuaries,[4,5] and in lakes.[6–8] The 2nd-order Korteweg-de Vries (KdV) equation[9,10] and the Gardner (or extended KdV) equation[11,12] are often used to investigate large-amplitude ISWs, which have strong nonlinearity. However, it has been considered that the higher-order coefficients of these KdV-type equations, as well as the associated nonlinear and nonhydrostatic correction functions to the wave fields, are nonunique for continuously stratified fluids.[13] This is a great hindrance to the application of our theoretical knowledge to geophysical problems. For example, the sign of the cubic nonlinear coefficient in the Gardner equation determines the existence of amplitude-limited (the so-called "flat-top" or "thick") ISWs,[11,14] but it is apparently difficult to apply such a relationship to observed or simulated ISWs if the sign is nonunique. Although the coefficient could be "tuned" against some reference, obtaining such reference is not trivial under realistic conditions in natural water bodies. In an attempt to bridge the gap between theoretical studies based on KdV-type equations and their application to geophysical problems, this study proposes to reduce arbitrariness of the higher-order KdV theory by considering its uniqueness in physical senses.

a)kenji.shimizu.rc@gmail.com

The previous 2nd-order KdV theories for continuous stratification[13,15–18] are based on Benney's asymptotic theory.[19] Using the isopycnal-displacement amplitude $\hat{A}$ (defined later) as a prognostic variable and following the notation in Ref. 18, the 2nd-order theory yields a solution for the (vertically dependent) isopycnal-displacement field η of the form

$$\eta \sim \hat{\phi}\hat{A} + \hat{T}_n\hat{A}^2 + \hat{T}_d\hat{A}_{xx}, \tag{1}$$

where x is the horizontal coordinate (positive in the direction of wave propagation), $\hat{\phi}$ is the vertical mode of interest, and $\hat{T}_n$ and $\hat{T}_d$ are the 1st-order nonlinear and nonhydrostatic correction functions, respectively. To 2nd order, $\hat{A}$ is governed by the 2nd-order KdV equation

$$\hat{A}_t + \left(c + \hat{\alpha}\hat{A} + \hat{\alpha}_1\hat{A}^2\right)\hat{A}_x + \hat{\beta}\hat{A}_{xxx} + \hat{\beta}_1\hat{A}_{5x} + \hat{\gamma}_1\hat{A}\hat{A}_{xxx} + \hat{\gamma}_2\hat{A}_x\hat{A}_{xx} = 0, \tag{2}$$

where t is the time, c is the phase speed for long linear waves, and $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$ are coefficients. In particular, $\hat{\alpha}$ and $\hat{\alpha}_1$ are referred to as the quadratic and cubic nonlinear coefficients, respectively. Setting $\hat{\beta}_1 = \hat{\gamma}_1 = \hat{\gamma}_2 = 0$ in (2) yields the Gardner equation. Reference 13 pointed out that the asymptotic theory provides no way of uniquely determining the correction functions and the resulting higher-order coefficients. To determine them, the subsequent theories[16–18] used auxiliary conditions, which appear to be chosen based on convenience.

This study aims to extend the argument for physical uniqueness of the Gardner equation previously made in Ref. 16, and to show that the higher-order KdV theory can be made unique in three physical senses: (i) consistency of the nonlinear higher-order coefficients and correction functions with the corresponding phase speeds, which are unique for individual vertical modes, (ii) wavenumber-independence of the vertically integrated available potential energy, and (iii) its positive definiteness. The physical uniqueness (i) and (ii) can be illustrated using surface gravity waves in homogeneous fluids as an example. Using $\hat{\eta}$ as the surface-displacement amplitude and following Secs. 13.10 and 13.11 in Ref. 20, the equations governing hydrostatic fully nonlinear waves propagating into quiescent water are

$$\hat{\eta}_t + c^N(\hat{\eta})\hat{\eta}_x = 0, \tag{3a}$$

$$c^N(\hat{\eta}) = c\left(3\sqrt{1+\hat{\eta}/H} - 2\right), \tag{3b}$$

where H is the equilibrium water depth. Expanding $c^N(\hat{\eta})$ in power series around $\hat{\eta} = 0$, we get

$$\hat{\eta}_t + \left(c + \hat{\alpha}\hat{\eta} + \hat{\alpha}_1\hat{\eta}^2 + \hat{\alpha}_2\hat{\eta}^3 + \cdots\right)\hat{\eta}_x = 0. \tag{4}$$

This provides the same $\hat{\alpha}$ and $\hat{\alpha}_1$ from the KdV theory because the nonlinear ($\hat{\alpha}$) coefficients are determined independent of nonhydrostatic wave dispersion. Apparently, the series determines an arbitrary number of nonlinear coefficients. Note that the coefficients are unique because of the uniqueness of power series (e.g., Sec. 5.7 in Ref. 21). The phase speed $c^N(\hat{\eta})$ also constrains the wave fields because the conservation of volume relates the phase speed to the flow rate $Q(\hat{\eta}) = (H + \hat{\eta})\, u(\hat{\eta})$ as

$$c^N(\hat{\eta}) = \frac{dQ(\hat{\eta})}{d\hat{\eta}} \tag{5}$$

(see Sec. 2.2 in Ref. 20). We can apply a similar argument to the phase speed of linear fully nonhydrostatic waves $c^D(\kappa)$ by expanding it into power series around $\kappa = 0$ and by considering the differential equation corresponding to the dispersion relation (see Sec. 11.1 in Ref. 20), where κ is the wavenumber. However, we do not consider linear dispersion ($\hat{\beta}$) coefficients in detail because they do not depend on the auxiliary conditions (as previously pointed out for the 2nd-order theory[13,18]). To see uniqueness of the correction functions for linear fully nonhydrostatic waves, let us consider the following two expressions for the vertical displacement field η:

$$\eta(z) = \frac{\sinh\,\kappa(z+H)}{\sinh\,\kappa H}\hat{\eta}, \tag{6a}$$

$$\eta(z) = \frac{\sinh\,\kappa(z+H)}{\kappa H}\hat{\eta}, \tag{6b}$$

where z is the vertical coordinate (positive upward with origin at the equilibrium free surface). Both expressions are mathematically correct, satisfy the governing equations, and have the same expression in the long-wave limit ($\kappa \sim 0$), $\eta = H^{-1}(z+H)\hat{\eta}$. Note, however, that the potential energy $\frac{1}{2}\rho g \eta(0)^2$ (where ρ is density) is independent of wavenumber for given amplitude $\hat{\eta}$ in (6a), but not so in (6b). Naturally, (6a) is the standard form of a surface wave solution. We need to make a particular choice of the vertical structure function or the correction functions to make the potential energy independent of wavenumber. These constraints mean that the higher-order $\hat{\alpha}$ and $\hat{\beta}$ coefficients and the associated correction functions cannot be arbitrary. Although extending similar arguments to fully nonlinear and fully nonhydrostatic waves appears difficult, the uniqueness in the two limiting cases suggests that weakly nonlinear dispersion coefficients, $\hat{\gamma}_1$ and $\hat{\gamma}_2$, are more likely to be physically unique. This study investigates requirements for the above physical uniqueness for continuously stratified fluids.

Although we argue the physical uniqueness, the higher-order coefficients are in a mathematical sense nonunique because their numerical values can be changed without changing the underlying phase-speed relationships, $c^N(\hat{\eta})$ and $c^D(\kappa)$. We note four such cases here: (i) changing the normalization of vertical modes, (ii) changing the prognostic variable, (iii) changing the vertical coordinate, and (iv) near-identical transformation. The first case originates from nonuniqueness of the magnitudes of vertical modes. Apparently, different normalization of vertical modes leads to different numerical values of the corresponding "modal" variables, such as the amplitudes and the nonlinear coefficients; however, it does not affect the products $\hat{\phi}\hat{A}$, $\hat{\alpha}\hat{A}$, $\alpha_1\hat{A}^2$, etc., and hence the "physical" variables, such as the isopycnal-displacement field η and the nonlinear phase speed c^N.[16] (The caret is used to denote a modal variable whose magnitude depends on vertical-mode normalization throughout this paper.) The second case originates from a nonlinear relationship between different prognostic variables, e.g., $\hat{A}$ to the amplitude of horizontal velocity $\hat{U}$ (see, e.g., Ref. 22). The third case originates from a nonlinear relationship between the same prognostic variable in different vertical coordinates. (The explanation requires the relationships for vertical modes introduced in Sec. II, and the details are given in Appendix A.) Since the results can be put back into the original variable or coordinate using the nonlinear relationship, the second and third cases do not change the solution in a physical sense. In the fourth case, the 2nd-order coefficients are said to be nonunique because they can be modified arbitrarily using the near-identical transformation[9,10]

$$\hat{A}' = \hat{A} + \epsilon\left(\lambda_1\hat{A}^2 + \lambda_2\hat{A}_{xx} + \lambda_3\hat{A}_x\int\hat{A}dx + \lambda_4 x\hat{A}_t\right), \tag{7}$$

where $\lambda_1, \ldots, \lambda_4$ are constants determined from the original and modified coefficients, and ϵ is a small parameter representing the magnitude of nonlinearity and nonhydrostaticity. However, it does not change the solution for the original $\hat{A}$, and the original 2nd-order coefficients are required for the inverse transformation.

This paper is organized as follows. Since the analyses in this study are based on Benney's spectral (or generalized Fourier) approach,[19] we derive the fully nonlinear and fully nonhydrostatic evolutionary equations of modal amplitudes in Sec. II, by projecting the governing equations in the isopycnal coordinate onto vertical modes. Section III describes stratification used to illustrate the results. In Sec. IV, we show an alternative derivation of the 2nd-order KdV equation based on the spectral approach, which requires no auxiliary condition. The results from the previous 2nd-order KdV theories are in general different from this spectral approach but can be made consistent using a unique set of special auxiliary conditions required for the physical uniqueness (ii) and (iii). In Sec. V, we derive numerical and power-series solutions for hydrostatic fully nonlinear gravity waves by combining the spectral approach and simple-wave analysis. The solutions provide a convenient basis to investigate the physical uniqueness (i) of the proposed and previous KdV theories because the coefficients and correction functions from the proposed 2nd-order theory are the same as the corresponding terms in the power-series solution. The power-series solution also provides algebraic recurrence relationships to determine an arbitrary number of the higher-order nonlinear ($\hat{\alpha}$) coefficients and associated correction functions. This paper ends with discussion in Sec. VI and conclusions in Sec. VII. Appendices include technical details, lengthy derivations, and the derivation of a power-series solution for linear fully nonhydrostatic gravity waves for completeness.

II. EVOLUTIONARY EQUATIONS OF MODAL AMPLITUDES

The analyses in this study are based on Benney's fully nonlinear and fully nonhydrostatic spectral approach.[19] In this section, we derive the fully nonlinear and fully nonhydrostatic evolutionary equations of modal amplitudes using the standard vertical-mode (or generalized Fourier) decomposition of the wave fields[23] in the isopycnal coordinate. Similar spectral approaches have been used to study weakly nonlinear internal waves under weak topographic effects[24,25] and of linear internal waves under strong topographic effects.[26–28]

In deriving the evolutionary equations of modal amplitudes, we prefer using an isopycnal coordinate s that is a monotonic function of the background density $\rho_{\mathrm{ref}}(z)$. One of the major reasons is that the higher-order coefficients of KdV-type equations in the isopycnal coordinate can be directly compared with well-known formulae for two-layer stratified fluids. It is also more straightforward to derive the common form of KdV-type equations because isopycnal displacement appears as a prognostic variable in the isopycnal coordinate. Note that, assuming horizontally uniform background stratification, we can set $s = Z(\rho)$, which is the undisturbed height of isopycnals that corresponds to the background stratification $\rho_{\mathrm{ref}}(z)$. With this choice, the s coordinate becomes equivalent to the undisturbed height used in association with the Dubreil-Jacotin-Long (DJL) equation[29–31] and to the semi-Lagrangian vertical coordinate used in the previous higher-order KdV theories.[18,32,33]

A. Governing equations in isopycnal coordinate

We consider a stably stratified, inviscid, incompressible fluid in a nonrotating frame of reference without background currents. It is assumed that the water depth is constant and the density at rest is horizontally uniform. The governing equations in the s coordinate are derived as follows. The governing equations in the z coordinate are

$$\nabla_z \cdot \vec{u} + \frac{\partial w}{\partial z} = 0, \tag{8a}$$

$$\rho\left(\frac{D\vec{u}}{Dt}\right)_z = -\nabla_z p, \tag{8b}$$

$$\rho\left(\frac{Dw}{Dt}\right)_z = -\frac{\partial p}{\partial z} - \rho g, \tag{8c}$$

$$\left(\frac{D\rho}{Dt}\right)_z = 0, \tag{8d}$$

where $\vec{x}$ is the horizontal coordinate vector, $\vec{u}$ is the horizontal velocity vector, w is the vertical velocity, p is the pressure, ρ is the density, $\nabla = (\partial/\partial x, \partial/\partial y)$ is the horizontal differential operator, and $(D/Dt)_z = (\partial/\partial t)_z + \vec{u}\cdot\nabla_z + w(\partial/\partial z)$ is the material derivative. The subscript z on differential operators indicates that z is held fixed. Assuming free surface, the boundary conditions are given by

$$p = 0,\ \ w = \frac{Dz^t}{Dt} \ \text{ at } z = z^t(\vec{x}, t), \tag{9a}$$

$$w = 0, \quad \text{at } z = z^b. \tag{9b}$$

Hereafter, the superscripts t and b indicate values at the top and bottom of the domain, respectively. The vertical coordinate transformation is done using the standard formulae[34]

$$\frac{\partial}{\partial z} = \left(\frac{\partial s}{\partial z}\right)\frac{\partial}{\partial s}, \tag{10a}$$

$$\left(\frac{\partial}{\partial t}\right)_z = \left(\frac{\partial}{\partial t}\right)_s - \left(\frac{\partial z}{\partial t}\right)_s \left(\frac{\partial s}{\partial z}\right)\frac{\partial}{\partial s}, \tag{10b}$$

$$\nabla_z = \nabla_s - (\nabla_s z)\left(\frac{\partial s}{\partial z}\right)\frac{\partial}{\partial s}. \tag{10c}$$

Note that the above transformation and (8d) yield $(D/Dt)_z = (D/Dt)_s \equiv (\partial/\partial t)_s + \vec{u}\cdot\nabla_s$. Hereafter, temporal and horizontal derivatives are always taken with holding s fixed, and we omit the subscript s for brevity. We separate fluctuations due to wave motion from the background values and write $z(\vec{x}, s, t) = Z(s) + \eta(\vec{x}, s, t)$ and $M(\vec{x}, s, t) = M_{\mathrm{ref}}(s) + m(\vec{x}, s, t)$, where $M = p + \rho g z$ is the Montgomery potential, and $Z(s)$ and $M_{\mathrm{ref}}(s)$ are the background isopycnal height and Montgomery potential, respectively. In the s coordinate, the hydrostatic balance at rest is given by

$$0 = -\frac{\partial M_{\mathrm{ref}}}{\partial s} + g\frac{d\rho}{ds}Z. \tag{11}$$

After subtracting this hydrostatic balance, the governing equations in the s coordinate without hydrostatic assumption become

$$\frac{\partial}{\partial t}\left(\frac{\partial \eta}{\partial s}\right) = -\nabla\cdot\left(\left(\frac{dZ}{ds} + \frac{\partial \eta}{\partial s}\right)\vec{u}\right), \tag{12a}$$

$$\rho\frac{D\vec{u}}{Dt} = -\nabla m - \rho(\nabla\eta)\frac{Dw}{Dt}, \tag{12b}$$

$$\rho\left(\frac{dZ}{ds} + \frac{\partial \eta}{\partial s}\right)\frac{Dw}{Dt} = -\frac{\partial m}{\partial s} + g\frac{d\rho}{ds}\eta, \tag{12c}$$

where w is given by the kinematic relation

$$w = \frac{D\eta}{Dt}. \tag{13}$$

The terms with Dw/Dt in (12) are nonhydrostatic terms. In the s coordinate, the boundary conditions (9) become

$$m = \rho g \eta \quad \text{at} \quad s = s^t, \tag{14a}$$

$$\eta = 0 \quad \text{at} \quad s = s^b. \tag{14b}$$

The vertically integrated energy equations associated with (12) are shown in Appendix C. Assuming a steady wave, introducing a coordinate moving at the propagation speed, and deriving an equation for η, (12) yields the DJL equation in the isopycnal coordinate.[31,32]

B. Vertical modes in isopycnal coordinate

In what follows, we introduce vertical modes $\hat{\phi}$ and $\hat{\pi}$ through the relationships

$$\hat{h}\frac{d\hat{\phi}}{ds} = \frac{dZ}{ds}\hat{\pi}, \tag{15a}$$

$$\frac{c^2}{\hat{h}}\frac{d(\rho\hat{\pi})}{ds} = -\rho N^2 \frac{dZ}{ds}\hat{\phi}, \tag{15b}$$

$$c^2 \hat{h}^{-1}\hat{\pi} = g\hat{\phi} \quad \text{at} \quad s = s^t, \tag{15c}$$

$$\hat{\phi} = 0 \quad \text{at} \quad s = s^b, \tag{15d}$$

where $\hat{h}$ is a constant defined later, and $N = \sqrt{-\rho^{-1} g (dZ/ds)^{-1}(d\rho/ds)}(\geq 0)$ is the background buoyancy frequency. By combining these equations, we get an eigenvalue problem

$$c^2 \frac{d}{ds}\left(\frac{\rho}{dZ/ds}\frac{d\hat{\phi}}{ds}\right) = -\rho N^2 \frac{dZ}{ds}\hat{\phi}, \tag{16a}$$

$$c^2 \frac{1}{dZ/ds}\frac{d\hat{\phi}}{ds} = g\hat{\phi} \quad \text{at} \quad s = s^t, \tag{16b}$$

$$\hat{\phi} = 0 \quad \text{at} \quad s = s^b. \tag{16c}$$

This problem yields an infinite number of phase speeds of long linear waves c_n and the associated vertical modes $\hat{\phi}_n$, and the use of (15) yields $\hat{\pi}_n$. Except for the 2nd-order coefficients and 1st-order correction functions from previous theories (i.e., $\hat{\alpha}_1$, $\hat{\beta}_1$, $\hat{\gamma}_1$, $\hat{\gamma}_2$, $\hat{T}_n$, and $\hat{T}_d$), a lower-case subscript is hereafter used to denote a modal index, which is 0 for the barotropic mode, 1 for the 1st baroclinic mode (hereafter VM1), 2 for the 2nd baroclinic mode (VM2), etc. The vertical modes $\hat{\phi}_n$ satisfy the orthogonality relationships:

$$\int_{s^b}^{s^t} \frac{d\hat{\phi}_n}{ds}\frac{\rho}{dZ/ds}\frac{d\hat{\phi}_m}{ds} ds = \frac{\hat{\rho}}{\hat{h}_n}\delta_{nm}, \tag{17a}$$

$$\hat{\phi}_n^t \rho^t g \hat{\phi}_m^t + \int_{s^b}^{s^t} \hat{\phi}_n \rho N^2 \frac{dZ}{ds}\hat{\phi}_m ds = \frac{\hat{\rho} c_n^2}{\hat{h}_n}\delta_{nm}, \tag{17b}$$

where δ_{nm} is the Kronecker delta (=1 if $n = m$ but 0 otherwise). The orthogonalities of $\hat{\pi}_n$ can be obtained by substituting (15) into the above equations. Note that (17) for $m = n$ defines the norm of the vertical modes, and $\hat{\rho}$ and $\hat{h}$ are arbitrary normalization factors with the units of density and water depth, respectively. The caret is used to denote a modal variable whose magnitude depends on the choice of these normalization factors [i.e., the mathematical nonuniqueness (i) in the Introduction]. In this study, we calculate $\hat{h}_n$ by normalizing the maximum magnitude of $\hat{\phi}_n$ to be 1 so that the corresponding modal displacement amplitude $\hat{\eta}_n$ has the magnitude of the maximum isopycnal displacement in the vertical. In addition, we choose the sign of $\hat{\pi}_n$ to be positive at the bottom in order to avoid a sudden sign change of VM2 (and higher vertical modes) when background stratification is varied. For later convenience, we introduce the following short-hand notation for the inner product:

$$\langle a, b\rangle \equiv \frac{1}{\hat{\rho}}\left(a^t \rho^t g b^t + \int_{s^b}^{s^t} a\rho N^2 \frac{dZ}{ds} b ds\right). \tag{18}$$

C. Projection of governing equations onto vertical modes

Since the operator together with the boundary conditions in (16) is self-adjoint, the vertical modes (i.e., eigenfunctions) form a complete set of functions (e.g., Sec. 10.4 in Ref. 21). So, we can project (12) onto vertical modes by assuming generalized Fourier series of the prognostic variables,

$$\eta(\vec{x}, s, t) = \hat{\phi}_n(s)\hat{\eta}_n(\vec{x}, t), \tag{19a}$$

$$m(\vec{x}, s, t) = \frac{\rho(s)}{\hat{\rho}}\hat{\pi}_n(s)\hat{m}_n(\vec{x}, t), \tag{19b}$$

$$\vec{u}(\vec{x}, s, t) = \hat{\pi}_n(s)\hat{\vec{u}}_n(\vec{x}, t), \tag{19c}$$

$$w(\vec{x}, s, t) = \hat{\phi}_n(s)\hat{w}_n(\vec{x}, t). \tag{19d}$$

In this study, summation over all the available vertical modes applies to repeated dummy subscript indices unless otherwise stated. (If a dummy index appears more than twice within a term, it is understood that the sum of the whole term is taken on that index, instead of the sum for each pair.) Note that we define $\hat{\phi}$ and $\hat{\pi}$ as nondimensional functions so that a modal amplitude ($\hat{\eta}_n$, $\hat{m}_n$, $\hat{\vec{u}}_n$, or $\hat{w}_n$) carries the unit of the original variable. Note also that modal amplitudes (or generalized Fourier coefficients) can be calculated from known wave fields by projecting them onto the corresponding vertical modes, regardless of the strength of nonlinearity and nonhydrostaticity. For example, $\hat{\eta}_n$ is calculated from η as

$$\hat{\eta}_n = \frac{\hat{h}_n}{c_n^2}\langle\hat{\phi}_n, \eta\rangle, \tag{20}$$

which is derived from (17b) and (19a). To project (12a) and (12b) onto vertical modes, we multiply these equations by $\rho\hat{\pi}_n$ and $(dZ/ds)\hat{\pi}_n$, respectively, substitute (19), and use the orthogonality (17a) but for $\hat{\pi}_n$. The projection of (12c) is done by multiplying (12c) by $\hat{\phi}_n$, using integration by parts, and applying the surface boundary condition (14a) and the orthogonality (17b). The projection of (13) requires multiplication of the equation by $g(d\rho/ds)\hat{\phi}_n$ in the water column and by $\rho g\hat{\phi}_n$ at the surface. These procedures yield the evolutionary equations of modal amplitudes

$$\frac{\partial \hat{\eta}_n}{\partial t} = -\nabla \cdot (\hat{h}_n \hat{\vec{u}}_n) - \epsilon \nabla \cdot (\hat{N}^A_{nlm} \hat{\eta}_l \hat{\vec{u}}_m), \tag{21a}$$

$$\frac{\partial \hat{\vec{u}}_n}{\partial t} = -\frac{1}{\hat{\rho}} \nabla \hat{m}_n - \epsilon \left(\hat{N}^A_{lnm} \hat{\vec{u}}_l \cdot \nabla \hat{\vec{u}}_m + \mu \hat{N}^D_{lnm} (\nabla \hat{\eta}_l) \hat{w}_m \right), \tag{21b}$$

$$\frac{1}{\hat{\rho}} \hat{m}_n = \frac{c_n^2}{\hat{h}_n} \hat{\eta}_n + \mu \left(\hat{D}_{nm} \hat{\dot{w}}_m + \epsilon \hat{N}^D_{nlm} \hat{\eta}_l \hat{\dot{w}}_m \right), \tag{21c}$$

$$\hat{\dot{w}}_n = \frac{\partial \hat{w}_n}{\partial t} + \epsilon \hat{N}^B_{nlm} \hat{\vec{u}}_l \cdot \nabla \hat{w}_m, \tag{21d}$$

$$\hat{w}_n = \frac{\partial \hat{\eta}_n}{\partial t} + \epsilon \hat{N}^B_{nlm} \hat{\vec{u}}_l \cdot \nabla \hat{\eta}_m, \tag{21e}$$

where $\hat{\dot{w}}_n$ is the modal amplitude of the vertical acceleration Dw/Dt. The analyses later in this paper are based on these equations. Note that no sum is taken on n because it is not a dummy index. Vertical modes in this study are defined so that (21) have a form similar to the shallow-water equations. The variables

$$\hat{D}_{nm} = \frac{1}{\hat{\rho}} \int_{s^b}^{s^t} \hat{\phi}_n \rho \frac{dZ}{ds} \hat{\phi}_m ds, \tag{22a}$$

$$\hat{N}^A_{nlm} = \frac{1}{\hat{\rho}\hat{h}_l} \int_{s^b}^{s^t} \hat{\pi}_n \rho \frac{dZ}{ds} \hat{\pi}_l \hat{\pi}_m ds, \tag{22b}$$

$$\hat{N}^B_{nlm} = \frac{\hat{h}_n}{\hat{\rho} c_n^2} \left(\hat{\phi}^t_n \rho^t g \hat{\pi}^t_l \hat{\phi}^t_m + \int_{s^b}^{s^t} \hat{\phi}_n \rho N^2 \frac{dZ}{ds} \hat{\pi}_l \hat{\phi}_m ds \right), \tag{22c}$$

$$\hat{N}^D_{nlm} = \frac{1}{\hat{\rho}\hat{h}_l} \int_{s^b}^{s^t} \hat{\phi}_n \rho \frac{dZ}{ds} \hat{\pi}_l \hat{\phi}_m ds \tag{22d}$$

are modal interaction parameters due to linear (nonhydrostatic) dispersion, advection of horizontal momentum and isopycnal heaving, advection of vertical momentum, and nonlinear dispersion, respectively. Note that $\hat{N}^A_{nlm}$, $\hat{N}^B_{nlm}$, and $\hat{N}^D_{nlm}$ are nondimensional but $\hat{D}_{nm}$ has the unit of water depth. It is useful to note the relationships for index swapping, such as $\hat{D}_{mn} = \hat{D}_{nm}$, $\hat{N}^A_{mln} = \hat{N}^A_{nlm}$, $\hat{N}^A_{lnm} = \hat{h}_n^{-1} \hat{h}_l \hat{N}^A_{nlm}$, and $\hat{N}^D_{mln} = \hat{N}^D_{nlm}$. There are also identities among the interaction parameters, such as

$$\hat{N}^D_{nlm} = \hat{h}_l^{-1} \hat{D}_{nj} \hat{N}^B_{jlm}, \tag{23a}$$

$$\hat{N}^A_{nlj} \hat{N}^D_{kjm} = \hat{N}^D_{mlj} \hat{N}^B_{jnk}, \tag{23b}$$

where we note that the sum on j is taken (see Appendix B for proof). In (21), setting $\epsilon = \mu = 1$ yields the equations in the dimensional form. Small nondimensional parameters ϵ and μ indicate the order of the terms based on the following scaling. We have chosen the magnitude of $\hat{\phi}$ to be $O(1)$. We also choose $\hat{h}$ to be of the order of the water depth H so that $\hat{\pi}$ and the nondimensional parameters in (22b)–(22d) are $O(1)$. Then, we scale the other variables in (21) based on H, typical wavelength L, wave speed C, and wave-induced current speed U. These result in

$$\vec{x} = L\vec{x}^\dagger, t = (L/C)t^\dagger, \tag{24a}$$

$$c = Cc^\dagger, \hat{h} = H\hat{h}^\dagger, \hat{D}_{nm} = H\hat{D}^\dagger_{nm}, \tag{24b}$$

$$\hat{\eta} = (UH/C)\hat{\eta}^\dagger, \hat{m} = (\hat{\rho}UC)\hat{m}^\dagger, \tag{24c}$$

$$\hat{\vec{u}} = U\hat{\vec{u}}^\dagger, \hat{w} = (UH/L)\hat{w}^\dagger, \tag{24d}$$

where † denotes a nondimensional variable. Then, we get (21) in the nondimensional form with $\hat{\rho} = 1$, $\epsilon = U/C$, and $\mu = H^2/L^2$, omitting † for brevity. The vertically integrated energy equations associated with (21) are shown in Appendix C.

III. STRATIFICATION

For illustration purposes, we primarily use the hyperbolic tangent stratification in the undisturbed height coordinate ($s = Z$),

$$\rho(Z) = \hat{\rho} + \frac{\Delta\rho}{2} \left(1 - \tanh \frac{Z + h}{d} \right), \tag{25}$$

where $\Delta\rho(>0)$, h, and $d(>0)$ are free parameters. The variable h roughly represents the pycnocline depth but is allowed to be <0 or $>H$. Note that this stratification approaches two-layer stratification when $d \ll H$. Also, under the Boussinesq approximation $\Delta\rho/\hat{\rho} \ll 1$, the buoyancy frequency N approaches constant when $H \ll d$ and exponential when $h \ll d$ or $H - d \ll h$.

To explore the parameter space, it is convenient to note the following scaling property.[35] If $(\hat{h}_n, \hat{D}_{nm}, \hat{\eta}_n)$ and $(c_n, \hat{\vec{u}}_n)$ are solutions to the problems in Secs. IV and V under the parameters (Z, h, d, H), $\Delta\rho/\hat{\rho}$, and κ, then $a^{-2}(\hat{h}_n, \hat{D}_{nm}, \hat{\eta}_n)$ and $(ab)^{-1}(c_n, \hat{\vec{u}}_n)$ are the corresponding solutions for the baroclinic modes under the scaling $a^{-2}(Z, h, d, H)$, $b^{-2}\Delta\rho/\hat{\rho}$, and $a^2\kappa$, assuming the Boussinesq approximation. The same scaling but $a^{-1}(c_n, \hat{\vec{u}}_n)$ applies to the barotropic mode. Note that the scaling does not modify $\hat{N}^A_{nlm}$, $\hat{N}^B_{nlm}$, and $\hat{N}^D_{nlm}$, and that the scaling with a also works under non-Boussinesq conditions. So, using small $\Delta\rho/\hat{\rho}$, it is sufficient to vary h/H and d/H for fixed H and $\Delta\rho/\hat{\rho}$ to explore the parameter space. We use $H = 1$ and $\Delta\rho/\hat{\rho} = 10^{-3}$ in our calculation. Figure 1 shows some example stratification used later.

For analytical calculation, we use two-layer stratification. The subscripts I and II denote the upper and lower layers, and h_I and h_{II} are the upper and lower layer thicknesses, respectively. Note that nonhydrostaticity induces vertical dependence within each layer, and we need vertical modes to express it.

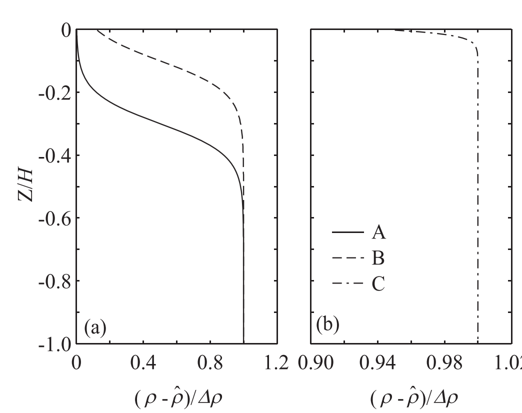

FIG. 1. Examples of hyperbolic tangent density profiles. (a) Stratification A ($h/H = 0.3$, $d/H = 0.1$) and stratification B ($h/H = 0.1$, $d/H = 0.1$), and (b) stratification C ($h/H = -0.04$, $d/H = 0.0288$).

So we introduce negligibly weak linear stratification within each layer and write density ρ as

$$\rho(Z) = (\rho_{II} - \Delta\rho\mathscr{H}(Z + h_I))\left(1 - \frac{N^2}{g}Z\right), \quad (26)$$

where $\mathscr{H}(Z)$ is the Heaviside function, $\Delta\rho = (\rho_2 - \rho_1)$ is the density difference across the interface, and N is assumed to be negligibly small.

IV. ALTERNATIVE 2ND-ORDER KdV THEORY

In this section, we derive the 2nd-order KdV equation by applying the standard asymptotic expansion used in the KdV theory to the evolutionary equations of modal amplitudes (21). This spectral approach has an advantage that there is no need to solve differential equations for the correction functions, which have nonunique solutions. The theory is applied to two-layer stratification for checking the results against well-known formulae,[22,36,37] and to continuous stratification for comparison with the previous theories.[13,16,18] We then investigate the relationship between the proposed and previous theories.

A. Derivation of 2nd-order KdV equation using spectral approach

We use (21) in the nondimensional form ($\hat{\rho} = 1$), assume the KdV scaling ($\epsilon = \mu$), and introduce coordinates moving at the long linear wave speed $(\xi,\tau) = (x - c_p t, \epsilon t)$. We focus on a particular mode (say, pth mode) and assume asymptotic expansion of the form

$$\hat{\eta}_p = \hat{\eta}_p^{(0)} + \epsilon\hat{\eta}_p^{(1)} + \epsilon^2\hat{\eta}_p^{(2)} + \cdots, \quad (27a)$$

$$\hat{u}_p = \hat{u}_p^{(0)} + \epsilon\hat{u}_p^{(1)} + \epsilon^2\hat{u}_p^{(2)} + \cdots, \quad (27b)$$

$$\hat{\eta}_q = \epsilon\hat{\eta}_q^{(1)} + \cdots, \quad (27c)$$

$$\hat{u}_q = \epsilon\hat{u}_q^{(1)} + \cdots. \quad (27d)$$

Hereafter, the subscript $q(\neq p)$ is used to indicate modes other than the pth mode. Deleting $\hat{m}_n$, $\hat{\omega}_n$, and $\hat{w}_n$ from (21), using the identity (23a), substituting the above expansion into the resulting equation, and collecting like terms, we get

$$\frac{\partial}{\partial\xi}\left(-c_p\hat{\eta}_n^{(k)} + \hat{h}_n\hat{u}_n^{(k)} + \hat{F}_n^{(k)}\right) + \frac{\partial\hat{\eta}_n^{(k-1)}}{\partial\tau} = 0, \quad (28a)$$

$$\frac{\partial}{\partial\xi}\left(c_n\hat{\eta}_n^{(k)} - \frac{c_p}{c_n}\hat{h}_n\hat{u}_n^{(k)} + \hat{G}_n^{(k)}\right) + \frac{\hat{h}_n}{c_n}\frac{\partial\hat{u}_n^{(k-1)}}{\partial\tau} = 0, \quad (28b)$$

where $n = p,q$ and it is assumed that $\hat{\eta}_n^{(-1)} = \hat{u}_n^{(-1)} = \hat{\eta}_q^{(0)} = \hat{u}_q^{(0)} = 0$. $\hat{F}_n^{(k)}$ and $\hat{G}_n^{(k)}$ are functions determined by the solutions up to $O(\epsilon^{k-1})$ and are given in Appendix D. Assuming $\hat{\eta}_n, \hat{u}_n \to 0$ as $\xi \to \pm\infty$, this problem is solved as follows. For $n = p$ and $k = 0$, (28) yields the relationship for long linear waves $\hat{u}_p^{(0)} = (c_p/\hat{h}_p)\hat{\eta}_p^{(0)}$. For $n = p$ and $k = 1$, (28) yields the 1st-order KdV equation and the 1st-order velocity for the pth mode

$$\hat{u}_p^{(1)} = \frac{c_p}{\hat{h}_p}\hat{\eta}_p^{(1)} - \frac{\hat{\alpha}_p^{(1)}}{6\hat{h}_p}\hat{\eta}_p^{(0)}\hat{\eta}_p^{(0)} + \frac{\hat{\beta}_p^{(1)}}{\hat{h}_p}\frac{\partial^2\hat{\eta}_p^{(0)}}{d\xi^2}, \quad (29)$$

where the 1st-order coefficients are given by

$$\hat{\alpha}_p^{(1)} = \frac{3c_p}{2\hat{h}_p}\hat{N}_{ppp}^A, \quad (30a)$$

$$\hat{\beta}_p^{(1)} = \frac{1}{2}c_p\hat{h}_p\hat{D}_{pp}. \quad (30b)$$

To obtain the 2nd-order solution, we first determine $\hat{\eta}_q^{(1)}$ and $\hat{u}_q^{(1)}$ from (28) for $n = q(\neq p)$ and $k = 1$ as

$$\hat{\eta}_q^{(1)} = \frac{c_p^2}{c_p^2 - c_q^2}\left(\frac{3\hat{N}_{qpp}^A}{2\hat{h}_p}(\hat{\eta}_p^{(0)})^2 + \hat{h}_q\hat{D}_{qp}\frac{\partial^2\hat{\eta}_p^{(0)}}{d\xi^2}\right), \quad (31a)$$

$$\hat{u}_q^{(1)} = \frac{c_p}{c_p^2 - c_q^2}\left(\frac{c_p^2 + 2c_q^2}{2}\frac{\hat{N}_{qpp}^A}{\hat{h}_p\hat{h}_q}(\hat{\eta}_p^{(0)})^2 + \hat{D}_{qp}c_p^2\frac{\partial^2\hat{\eta}_p^{(0)}}{d\xi^2}\right), \quad (31b)$$

where no sum is taken on p because it is the single mode of interest. Substituting the 1st-order KdV equation obtained at the previous order, (29) and (31) into (28) for $n = p$ and $k = 2$, and deleting $\hat{\eta}_p^{(2)}$ and $\hat{u}_p^{(2)}$, we get the expression for $\partial\hat{\eta}_p^{(1)}/\partial\tau$. Combining it with the 1st-order KdV equation, putting $\hat{\eta} = \hat{\eta}^{(0)} + \epsilon\hat{\eta}^{(1)} + \epsilon^2\hat{\eta}^{(2)}$ and using the original coordinates, we get the 2nd-order KdV equation

$$\frac{\partial\hat{\eta}_p}{\partial t} + \left(c_p + \hat{\alpha}_p^{(1)}\hat{\eta}_p + \hat{\alpha}_p^{(2)}\hat{\eta}_p^2\right)\frac{\partial\hat{\eta}_p}{\partial x} + \hat{\beta}_p^{(1)}\frac{\partial^3\hat{\eta}_p}{dx^3} + \hat{\beta}_p^{(2)}\frac{\partial^5\hat{\eta}_p}{dx^5} + \hat{\gamma}_p^{(2a)}\hat{\eta}_p\frac{\partial^3\hat{\eta}_p}{dx^3} + \hat{\gamma}_p^{(2b)}\frac{\partial\hat{\eta}_p}{\partial x}\frac{\partial^2\hat{\eta}_p}{dx^2} \sim 0, \quad (32)$$

where the 2nd-order coefficients are given by

$$\hat{\alpha}_p^{(2)} = -\frac{(\hat{\alpha}_p^{(1)})^2}{6c_p} + \frac{3}{4}\sum_{q\neq p}\frac{c_p(\hat{N}_{ppq}^A)^2}{\hat{h}_p\hat{h}_q}\frac{5c_p^2 + 4c_q^2}{c_p^2 - c_q^2}, \quad (33a)$$

$$\hat{\beta}_p^{(2)} = \frac{3(\hat{\beta}_p^{(1)})^2}{2c_p} + \frac{1}{2}\sum_{q\neq p}c_p\hat{h}_p\hat{D}_{pq}^2\hat{h}_q\frac{c_p^2}{c_p^2 - c_q^2}, \quad (33b)$$

$$\hat{\gamma}_p^{(2a)} = \frac{7}{3}\frac{\hat{\alpha}_p^{(1)}\hat{\beta}_p^{(1)}}{c_p} + \sum_{q\neq p}\frac{3c_p^3\hat{D}_{pq}\hat{N}_{qpp}^A}{c_p^2 - c_q^2} - \frac{1}{2}\hat{N}_{ppp}^D c_p\hat{h}_p, \quad (33c)$$

$$\hat{\gamma}_p^{(2b)} = \frac{31}{6}\frac{\hat{\alpha}_p^{(1)}\hat{\beta}_p^{(1)}}{c_p} + \sum_{q\neq p}\frac{6c_p^3\hat{D}_{pq}\hat{N}_{qpp}^A}{c_p^2 - c_q^2} - \hat{N}_{ppp}^D c_p\hat{h}_p. \quad (33d)$$

Adding the 0th- and 1st-order solutions for η and u yields

$$\eta \sim \hat{\phi}_p\hat{\eta}_p + \hat{\phi}_p^{N(1)}\hat{\eta}_p^2 + \hat{\phi}_p^{D(1)}\frac{\partial^2\hat{\eta}_p}{dx^2}, \quad (34a)$$

$$u \sim \frac{c_p}{\hat{h}_p}\left(\hat{\pi}_p\hat{\eta}_p + \hat{\pi}_p^{N(1)}\hat{\eta}_p^2 + \hat{\pi}_p^{D(1)}\frac{\partial^2\hat{\eta}_p}{dx^2}\right), \quad (34b)$$

where the 1st-order correction functions are given by

$$\hat{\phi}_p^{N(1)} = \frac{3}{2}\sum_{q\neq p}\frac{\hat{N}^A_{ppq}}{\hat{h}_p}\frac{c_p^2}{c_p^2 - c_q^2}\hat{\phi}_q, \tag{35a}$$

$$\phi_p^{D(1)} = \sum_{q\neq p}\hat{D}_{pq}\hat{h}_q\frac{c_p^2}{c_p^2 - c_q^2}\hat{\phi}_q, \tag{35b}$$

$$\hat{\pi}_p^{N(1)} = -\frac{\hat{\alpha}_p^{(1)}}{6c_p}\hat{\pi}_p + \frac{1}{2}\sum_{q\neq p}\frac{\hat{N}^A_{ppq}}{\hat{h}_q}\frac{c_p^2 + 2c_q^2}{c_p^2 - c_q^2}\hat{\pi}_q, \tag{35c}$$

$$\hat{\pi}_p^{D(1)} = \frac{\hat{\beta}_p^{(1)}}{c_p}\hat{\pi}_p + \sum_{q\neq p}\hat{h}_p\hat{D}_{pq}\frac{c_p^2}{c_p^2 - c_q^2}\hat{\pi}_q. \tag{35d}$$

Once the coefficients and correction functions are determined, the subscript p can be dropped as in (1) and (2) because only the pth mode appears in the problem. Note that no auxiliary condition is required in the above derivation.

B. Comparison with previous theories

To check the results in Sec. IV A, we first applied the theory to two-layer stratification and compared the resulting 2nd-order coefficients with the well-known formulae[22,36,37] for surface and internal waves. The details are given in Appendix E. The coefficients agree with the previous formulae under the Boussinesq approximation.

Using the hyperbolic tangent stratification (25), we also compared the 2nd-order coefficients from the proposed theory with those from one of the previous theories that (effectively) uses the isopycnal coordinate.[18] To see the sensitivity to auxiliary conditions, however, we applied different conditions by adding some multiple of the vertical mode of interest $\hat{\phi}_p$ to the 1st-order nonlinear and nonhydrostatic correction functions.[13] We considered three auxiliary conditions: (i) zero isopycnal-displacement correction at the depth where the magnitude of $\hat{\phi}_p$ is maximum (hereafter condition ZD), (ii) zero volume-transport correction at the depth where $\hat{\phi}_p$ is maximum (ZT), and (iii) zero correction of the vertical gradient of isopycnal displacement at the bottom (ZG). The conditions (i), (ii), and (iii) were chosen following Refs. 18, 16, and 13, respectively. Note that Ref. 16 also assumed zero volume-transport correction at the surface, but it is impossible to set both this condition and (ii) in the theory in Ref. 18. Note also that Ref. 13 used the condition (iii) assuming that it is modified later by comparison with numerical simulations. We nonetheless use these three conditions to examine the sensitivity.

Numerical methods used for the computation are explained in Appendix F. We chose the number of vertical cells $n_L = 200$ to ensure reasonable resolution with a thin stratified region when $h/H < d/H$ or $1 - d/H < h/H$ [e.g., Fig. 1(b)]. Some care was necessary because part of the water column can be almost homogeneous with the hyperbolic tangent stratification (25). As seen in the two-layer case in Appendix E, a homogeneous region supports degenerate modes, all of which have zero phase speed. To calculate such modes, we included negligibly weak linear stratification (density difference of $10^{-9}\hat{\rho}$ between the surface and bottom) as in the two-layer case. Since degenerate modes make the modal expansion inefficient, we simply included all the available modes ($n_M = n_L = 200$, where n_M is the number of modes) to ensure the convergence.

For comparisons with the previous theory, we focused on the cubic nonlinear coefficient, which is most sensitive to the choice of the auxiliary condition (Fig. 2). We use the notation $\hat{\alpha}_1$ for the nonlinear coefficient under an arbitrary auxiliary condition from the previous theory, which includes $\hat{\alpha}_p^{(2)}$ from the proposed theory as a special case as shown in Subsection IV C. For VM1, $\hat{\alpha}_1$ under the conditions ZT and ZG show skewed patterns with positive values over a wider range of conditions than in the proposed theory [Figs. 2(a), 2(c), and, 2(d)]. In contrast, $\hat{\alpha}_1$ under the condition ZD overall has a pattern similar to the proposed theory [Figs. 2(a) and 2(b)]. Furthermore, the two agree perfectly in the two-layer limit $d/H \to 0$ (although the results are still different at $d/H = 0.01$ in the figures) and symmetric stratification

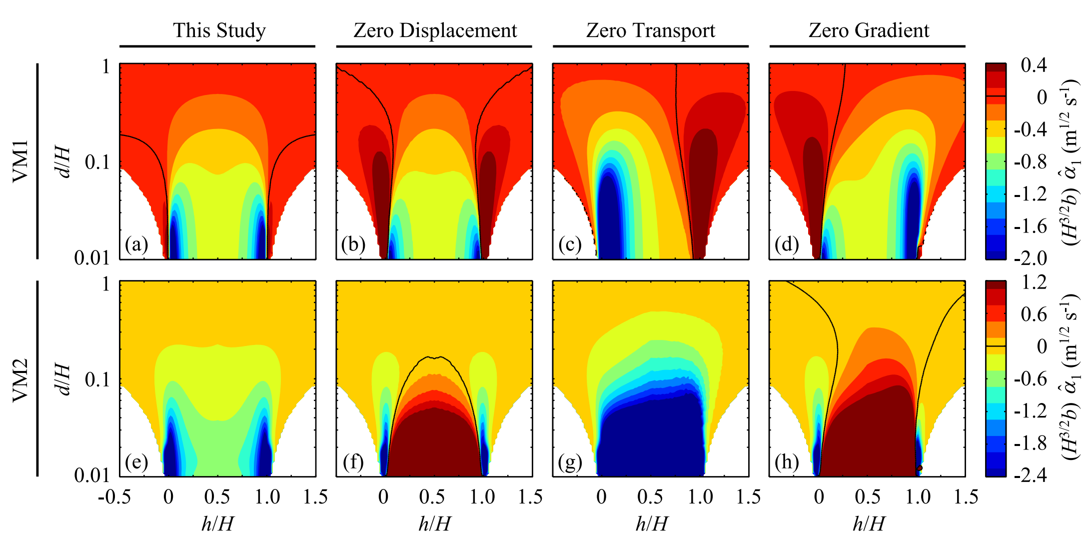

FIG. 2. Comparison of normalized cubic nonlinear coefficient $\hat{\alpha}_1$ calculated under different auxiliary conditions ($n_L = n_M = 200$). Stratification is hyperbolic tangent (25). Upper row: 1st baroclinic mode, lower row: 2nd baroclinic mode, 1st column: proposed 2nd-order theory [which implicitly assumes (38)], 2nd column: zero isopycnal-displacement correction, 3rd column: zero volume-transport correction, and 4th column: zero correction of vertical gradient of isopycnal displacement at the bottom. Solid lines show isopleths for $\hat{\alpha}_1 = 0$. Normalization uses $b = \sqrt{10^{-3}\hat{\rho}/\Delta\rho}$.

$h/H = 0.5$. However, when the buoyancy frequency is close to being monotonic ($h/H \lesssim 0.0$ or $h/H \gtrsim 1.0$), this auxiliary condition yields larger $\hat{\alpha}_1$ than the proposed theory, which can result in sign difference. For VM2, the conditions ZD and ZG yield positive $\hat{\alpha}_1$, which does not occur in the proposed theory [Figs. 2(e), 2(f), and 2(h)]. $\hat{\alpha}_1$ under the condition ZT is always negative [Fig. 2(g)], but the magnitude is much larger than the proposed theory for $d/H \lesssim 0.1$. These substantial differences in $\hat{\alpha}_1$, particularly the sign differences, are the motivation of this study because they pose difficulties in the application of the 2nd-order KdV equation and the Gardner equation to realistic continuous stratification.

C. Relationship between proposed and previous 2nd-order KdV theories

The higher-order coefficients and correction functions from the previous theory[18] are in general different from those in Sec. IV A but can be made the same using special auxiliary conditions. These conditions can be found by considering the vertically integrated available potential energy P associated with (12),

$$P = \frac{\hat{\rho}}{2} \langle \eta, \eta \rangle . \tag{36}$$

Note that P is positive definite irrespective of nonlinearity and nonhydrostaticity (or wavenumber) from this definition. P is also closely related to the norm of isopycnal displacement field $\sqrt{\langle \eta, \eta \rangle}$. Substituting (1) and (34a) into the above equation, we get, in the nondimensional form

$$P = \frac{1}{2} \left\langle \hat{\phi}_p, \hat{\phi}_p \right\rangle \hat{A}^2 + \epsilon \left\langle \hat{\phi}_p, \hat{T}_n \right\rangle \hat{A}^3 + \epsilon \left\langle \hat{\phi}_p, \hat{T}_d \right\rangle \hat{A} \frac{\partial^2 \hat{A}}{dx^2} + O(\epsilon^2), \tag{37a}$$

$$P = \frac{1}{2} \left\langle \hat{\phi}_p, \hat{\phi}_p \right\rangle \hat{\eta}_p^2 + O(\epsilon^2). \tag{37b}$$

The latter relationship is a special case of Parseval's theorem [its general form is (C2a) in Appendix C]. Note that the definitions of $\hat{A}$ and $\hat{\eta}_p$ are in general different. For example, $\hat{A}$ is often chosen as vertical displacement of the isopycnal whose equilibrium depth corresponds to zero-crossing of the correction functions,[16,18] whereas $\hat{\eta}_p$ is defined as a generalized Fourier coefficient (20). Note also that (37b) guarantees P to be positive definite and does not involve horizontal derivatives of the wave field (or wavenumber), but it is not the case with (37a). Consistency between these expressions requires

$$\left\langle \hat{\phi}_p, \hat{T}_n \right\rangle = 0, \tag{38a}$$

$$\left\langle \hat{\phi}_p, \hat{T}_d \right\rangle = 0. \tag{38b}$$

These conditions are satisfied in (35a) and (35b) because they do not contain a component proportional to $\hat{\phi}_p$. By applying (38a) to the theory in Ref. 18 using the orthogonality of the form (17), it is confirmed that the resulting cubic nonlinear coefficients are the same as those in Figs. 2(a) and 2(e).

For linear waves with slowly varying amplitude, (38b) is the physical requirement for P to be independent from wavenumber. This result can be extended to general wavenumber as follows. Using the superscripts $\circ$ and D to denote, respectively, the values at a reference wavenumber κ° and arbitrary wavenumber κ, substituting $\eta = \hat{\eta}_p^\circ \hat{\phi}_p^\circ$ and $\eta = \hat{\eta}_p^D \hat{\phi}_p^D$ (no sum on p) separately into (36), and requiring $\hat{\eta}_p^D = \hat{\eta}_p^\circ$ from linearity, we get $\left\langle \hat{\phi}_p^D, \hat{\phi}_p^D \right\rangle = \left\langle \hat{\phi}_p^\circ, \hat{\phi}_p^\circ \right\rangle$. This suggests that we can conveniently use the long-wave limit ($\kappa^\circ = 0$) as a reference, where the amplitude is defined as a generalized Fourier coefficient (20). Using (17b), we get

$$\left\langle \hat{\phi}_p^D, \hat{\phi}_p^D \right\rangle = \frac{c_p^2}{\hat{h}_p} \quad \text{(no sum on } p). \tag{39}$$

Alternatively, by writing $\hat{\phi}_p^D = \hat{\phi}_p + \Delta\hat{\phi}_p$, we get

$$2 \left\langle \hat{\phi}_p, \Delta\hat{\phi}_p \right\rangle + \left\langle \Delta\hat{\phi}_p, \Delta\hat{\phi}_p \right\rangle = 0 \quad \text{(no sum on } p). \tag{40}$$

This generalizes (38b) to an arbitrary wavenumber. If we apply this to surface waves in homogeneous fluids, it reduces to $\Delta\hat{\phi}_p(Z = 0) = 0$ because $\hat{\phi}_0(Z = 0) = 1$ for the choice of vertical-mode normalization in this study. This condition is satisfied by (6a) in the Introduction but not by (6b). The power-series solution for linear fully nonhydrostatic gravity waves with the above auxiliary condition, derived in Appendix G, yields the same $\hat{\beta}_p^{(1)}$, $\hat{\beta}_p^{(2)}$, $\hat{\phi}_p^{D(1)}$, and $\hat{\pi}_p^{D(1)}$ from the 2nd-order theory proposed in Sec. IV, (30b), (33b), (35b), and (35d).

It is desirable that P is positive definite in the KdV theory, as in the original definition (36). When the amplitude is chosen as a generalized Fourier coefficient $\hat{\eta}_p$, Parseval's theorem guarantees the quadratic form in general (see Appendix C), and (37a) is the corresponding requirement in the 2nd-order theory. The previous theories based on Benney's asymptotic theory do not require the amplitude under an arbitrary auxiliary condition $\hat{A}$ to be $\hat{\eta}_p$. This leads to an undesirable result that P is not positive definite, but this needs to be weighted against other deficiencies in the asymptotic theory (see Sec. VI).

The results in this subsection show that the wavenumber-independence and positive definiteness of P [i.e., the physical uniqueness (ii) and (iii) in the Introduction] offer a set of special auxiliary conditions that uniquely determine the higher-order coefficients and associated correction functions. The special auxiliary conditions (38) can be used in the previous 2nd-order KdV theory to yield the same results as the spectral approach in Sec. IV A. In Sec. V, we compare the proposed and previous theories against hydrostatic fully nonlinear solutions to investigate requirements for the physical uniqueness (i).

V. FULLY NONLINEAR SIMPLE WAVES

In this section, we briefly revisit the method of simple waves (e.g., Secs. 6.8 and 13.10 in Ref. 20, or Sec. 2.9 in Ref. 38) using surface waves in homogeneous fluids as an example and extend it to continuously stratified fluids. Then, we illustrate the numerical and power-series solutions for continuously stratified fluids and compare the results with the proposed and previous 2nd-order theories.

A. Revisiting simple waves in homogeneous fluids

The simple-wave problem for homogeneous fluids is based on the shallow-water equations,

$$\frac{\partial}{\partial t}\begin{bmatrix}\eta\\ u\end{bmatrix} = -\begin{bmatrix}u & H+\eta\\ c^2H^{-1} & u\end{bmatrix}\frac{\partial}{\partial x}\begin{bmatrix}\eta\\ u\end{bmatrix}, \tag{41}$$

where $c = \sqrt{gH}$, and g is the acceleration due to gravity. In this subsection, η represents the surface displacement for brevity. The fully nonlinear phase speed (eigenvalue) c^N and the corresponding left eigenvector $\tilde{\boldsymbol{l}}$ associated with (41) are

$$c^N_\pm = u \pm c\sqrt{1+\eta/H}, \tag{42a}$$

$$\tilde{\boldsymbol{l}}_\pm = \begin{bmatrix}\pm\dfrac{c}{H\sqrt{1+\eta/H}}\\ 1\end{bmatrix}. \tag{42b}$$

Signs are vertically ordered throughout this study. The tilde is used to denote a variable whose magnitude depends on the normalization of left eigenvectors. Multiplying (41) by $\tilde{\boldsymbol{l}}^T_\pm$ from the left and putting the equation in the characteristic form, we get

$$\frac{d\tilde{R}_\pm}{dt} = \frac{du}{dt} \pm \frac{c}{H\sqrt{1+\eta/H}}\frac{d\eta}{dt} = 0 \quad \text{on characteristic} \quad \frac{dx}{dt} = c^N_\pm, \tag{43}$$

where $\tilde{R}_\pm$ is the Riemann invariant. Hereafter, the superscript T denotes the matrix transpose.

In simple-wave problems, only one of the Riemann invariants varies across the characteristics. For example, we may consider a straight water flume with a uniform width in $x > 0$, initially at rest but forced at $x = 0$ for $t > 0$. In this case, all the characteristics associated with negative sign emanate from the region of rest (Fig. 3), so $\tilde{R}_-$ has the same value everywhere. Integrating (43) for $\tilde{R}_-$ provides the expression for u in terms of η,

$$u(\eta) = \int_0^\eta \frac{c}{H\sqrt{1+\eta'/H}}d\eta' = 2c\left(\sqrt{1+\frac{\eta}{H}} - 1\right). \tag{44}$$

Substituting this into the upper row of (41) (i.e., the conservation of volume) and using (5) yields (3). Alternatively,

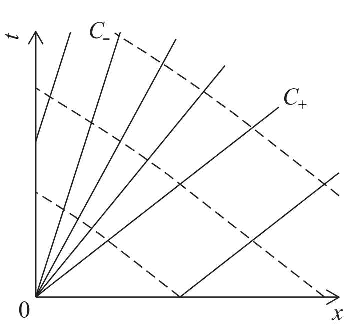

FIG. 3. Schematic of characteristics in a simple-wave problem. Initial condition is $\eta = u = 0$ everywhere. Boundary condition at $x = 0$ is $\eta = 0$, $u = 0$ for $t < 0$ but constant negative u for $t > 0$. Solid and dashed lines show right- and left-propagating characteristics, respectively.

the phase speed can be obtained by substituting the above $u(\eta)$ into (42a).

Simple waves considered above may appear too special compared to the KdV theory. However, the simple-wave assumption is roughly equivalent to introducing a coordinate moving at a constant speed (see Sec. 10.3 in Ref. 20), which is assumed either explicitly or implicitly by the leading-order balance $\partial\eta/\partial t \sim -c(\partial\eta/\partial x)$ in the KdV theory.

The example above implies two important aspects for the application of simple-wave analysis to gravity waves in continuously stratified fluids. First, the simple-wave assumption provides a way of expressing fully nonlinear solution in terms of one prognostic variable. Second, since only one degree of freedom is allowed in the simple-wave analysis, the form of the governing equations determines the type of waves that can be investigated by the method. For our purpose, the evolutionary equations of modal amplitudes (21) are suitable because their linear free-wave solutions are horizontally propagating waves with modal structure in the vertical, whereas the standard governing equations in the height coordinate (8) or in the isopycnal coordinate (12) are unsuitable because their linear free-wave solutions are obliquely ascending and descending internal waves.

B. Simple waves in continuously stratified fluids

We now extend the fully nonlinear simple-wave analysis in Sec. V A to gravity waves in continuously stratified fluids. To simplify notation in the following analysis, we write the hydrostatic fully nonlinear version of (21) ($\epsilon = 1$ and $\mu = 0$) in a matrix form following previous studies.[39,40] To do so, we consider the first n_M vertical modes and introduce the following $n_M \times 1$ column vectors and $n_M \times n_M$ matrices,

$$(\hat{\boldsymbol{\eta}})_n = \hat{\eta}_n, \tag{45a}$$
$$(\hat{\boldsymbol{u}})_n = \hat{u}_n, \tag{45b}$$
$$(\mathbf{C})_{nm} = c_n\delta_{nm}, \tag{45c}$$
$$(\hat{\mathbf{H}})_{nm} = \hat{h}_n\delta_{nm}, \tag{45d}$$
$$(\hat{\mathbf{U}}(\hat{\boldsymbol{u}}))_{nm} = \hat{N}^A_{nml}\hat{u}_l, \tag{45e}$$
$$(\hat{\mathbf{Y}}(\hat{\boldsymbol{\eta}}))_{nm} = \hat{N}^A_{nlm}\hat{\eta}_l. \tag{45f}$$

Hereafter, a parenthesis with subscript indices is used to denote a matrix component. Then, we write (21) for $\epsilon = 1$ and $\mu = 0$ as

$$\frac{\partial\hat{\boldsymbol{v}}}{\partial t} = -\hat{\mathbf{K}}(\boldsymbol{v})\frac{\partial\hat{\boldsymbol{v}}}{\partial x}, \tag{46}$$

where $\hat{\boldsymbol{v}} = [\hat{\boldsymbol{\eta}}^T\ \hat{\boldsymbol{u}}^T]^T$ is a prognostic variable column vector, and

$$\hat{\mathbf{K}}(\hat{\boldsymbol{v}}) = \begin{bmatrix}\hat{\mathbf{U}} & \hat{\mathbf{H}}+\hat{\mathbf{Y}}\\ \mathbf{C}^2\hat{\mathbf{H}}^{-1} & \hat{\mathbf{U}}^T\end{bmatrix}. \tag{47}$$

Note that (46) has a form similar to (41) and reduces to it for homogeneous fluids (using $\hat{N}^A_{000} = \hat{N}^B_{000} = \hat{N}^D_{000} = 1$). The wave fields are given by

$$\eta = \hat{\boldsymbol{\phi}}^T\hat{\boldsymbol{\eta}}, \tag{48a}$$

$$u = \hat{\boldsymbol{\pi}}^T\hat{\boldsymbol{u}}, \tag{48b}$$

where

$$(\hat{\boldsymbol{\phi}})_n = \hat{\phi}_n(s), \tag{49a}$$

$$(\hat{\boldsymbol{\pi}})_n = \hat{\pi}_n(s). \tag{49b}$$

To conduct simple-wave analysis, we need left eigenvectors of (47). They are calculated from the eigenvalue problem

$$c^N \tilde{\boldsymbol{l}}^H = \tilde{\boldsymbol{l}}^H \hat{\mathbf{K}}. \tag{50}$$

Hereafter, the superscript H denotes the conjugate transpose. This problem yields $2n_M$ sets of fully nonlinear phase speed c_r^N and the associated left eigenvector $\tilde{\boldsymbol{l}}_r$. Note that Eq. (50) is written to allow complex variables because c_r^N and $\tilde{\boldsymbol{l}}_r$ may appear as conjugate pairs. Under weak nonlinearity, the solutions consist of right- and left-propagating waves for each vertical mode n. Under moderately strong nonlinearity, this remains the case for lower vertical modes, but it becomes difficult to relate the solutions to the modal index n for higher modes. Therefore, to label the solutions, we simply use the indices r and s that vary from 1 to $2n_M$. Multiplying (46) by $\tilde{\boldsymbol{l}}_r^H$ from the left, we get an equation for the Riemann invariant $\tilde{R}_r$, which remains constant along each characteristic $dx/dt = c_r^N$. Writing $\tilde{\boldsymbol{R}} = [\tilde{R}_1, \ldots, \tilde{R}_{2n_M}]^T$ and

$$\tilde{\mathbf{L}} = [\tilde{\boldsymbol{l}}_1, \ldots, \tilde{\boldsymbol{l}}_{2n_M}], \tag{51}$$

the Riemann invariants are collectively defined through the differential relationships (see Sec. 5.3 in Ref. 20),

$$d\tilde{\boldsymbol{R}} = \tilde{\mathbf{L}}(\hat{\boldsymbol{v}})^H d\hat{\boldsymbol{v}}. \tag{52}$$

To integrate the above differential relationship, we assume that only $\tilde{R}_{+p}$ varies across the characteristics, where the subscript $+p$ indicates a particular (pth) lower vertical mode propagating towards positive x. The rest of the Riemann invariants are assumed to be uniform everywhere. So, all the equations in (52) except for the one for $\tilde{R}_{+p}$ can be used to determine $d\hat{\boldsymbol{\eta}}$ and $d\hat{\boldsymbol{u}}$ as functions of $d\hat{\eta}_p$. However, for simplifying notation, understanding the properties of the problem, and making the numerical integration scheme more robust, it is more convenient to write this problem using $\tilde{R}_{+p}$ as the independent variable

$$\tilde{\mathbf{L}}(\hat{\boldsymbol{v}})^H \frac{d\hat{\boldsymbol{v}}}{d\tilde{R}_{+p}} = \boldsymbol{I}_{+p}, \tag{53}$$

where $(\boldsymbol{I}_{+p})_r = \delta_{+p,r}$. The analysis in this section is based on this differential equation, integrated together with (50) and (51) from the initial condition $\hat{\boldsymbol{\eta}} = \boldsymbol{O}$ and $\hat{\boldsymbol{u}} = \boldsymbol{O}$. Note that the integration is possible as long as $(c_{+p}^N, \tilde{\boldsymbol{l}}_{+p})$ are real and $\tilde{\mathbf{L}}^H$ is invertible. In particular, note that the solution for the $+p$th mode remains real when some of the other modes appear as conjugate pairs. Except for simple cases such as (44), this integration has to be done numerically. Some details regarding the (numerical) integration are given in Appendix F. The resulting relationship between $\tilde{R}_{+p}$ and $\hat{\eta}_p$ can be used to express the solution as a function of $\hat{\eta}_p$. Then, we can write the pth row in (46) in the conservative form and apply the kinematic relationship between the phase speed and flow rate (5) to $\hat{Q}_p(\hat{\eta}_p) = \left(\hat{h}_p\delta_{pm} + \hat{N}^A_{plm}\hat{\eta}_l(\hat{\eta}_p)\right)\hat{u}_m(\hat{\eta}_p)$, yielding

$$\frac{\partial \hat{\eta}_p}{\partial t} + c_{+p}^N(\hat{\eta}_p)\frac{\partial \hat{\eta}_p}{\partial x} = 0. \tag{54}$$

This is the continuous stratification version of (3). The phase speed is of course the same as that from (50), but the use of (5) shows the constraint between the phase speed and the wave fields. We see that the fully nonlinear phase speed is unique unless the $+p$th eigen solution to (50) is degenerate, and that the phase speed is a function of the amplitude of a single vertical mode. This allows applying our phase-speed argument for constraining the nonlinear ($\hat{\alpha}$) coefficients of KdV-type equations, made in the Introduction, to gravity waves in continuously stratified fluids.

Although solutions to (50) and (53) cannot be written in a closed form in general, we can derive algebraic recurrence relationships to calculate the power-series solution, which yields an arbitrary number of the nonlinear coefficients of KdV-type equations, $\hat{\alpha}^{(k)}$, and the associated correction functions, $\hat{\phi}^{N(k)}$ and $\hat{\pi}^{N(k)}$. The details are given in Appendix H. The derivation shows that no auxiliary condition is required when the power-series solution is written in $\hat{\eta}_p$, although an auxiliary condition is required when $\tilde{R}_{+p}$ appears explicitly. Analytically solving the problem around $\hat{\eta}_p = 0$ up to 2nd order yields the same $\hat{\alpha}_p^{(1)}$, $\hat{\alpha}_p^{(2)}$, $\hat{\phi}_p^{N(1)}$, and $\hat{\pi}_p^{N(1)}$ from the 2nd-order theory in Sec. IV, (30a), (33a), (35a), and (35c).

C. Application to hyperbolic tangent stratification

For the hyperbolic tangent stratification (25), the direct numerical solutions were calculated from (50), (51), and (53) and the power-series solutions from (H2), (H3b), and (H7). Examples using the three density profiles in Fig. 1 are shown in Figs. 4 and 5. For VM1 and VM2, the use of the lowest 10 modes in the numerical integration typically yielded reasonable results. For example, the phase speeds typically converged within ≈3% for $|\hat{\eta}_p|/\hat{h}_p \lesssim 2$. In general, the convergence was slower for larger $|\hat{\eta}_p|$, for smoother stratification (larger d/H), near transition to linear instability, and when $d\hat{\boldsymbol{v}}/d\tilde{R}_{+p} \approx \boldsymbol{O}$ in (53) (i.e., near a sink or node using the terminology in the mathematics of nonlinear differential equations; see, e.g., Sec. 18.4 in Ref. 21). The convergence of $\hat{\alpha}_p^{(2)}$ from the power-series solution was typically faster: within ≈3% using the first 3 and 5 modes for VM1 and VM2, respectively. Using the straightforward numerical implementation of (H2), (H3b), and (H7), the number of reliable terms in the power-series solution was limited to ≈15 by the computational accuracy.

For stratification A, the fully nonlinear phase speed for VM1, c_{+1}^N, shows initial increase and subsequent decrease, as in two-layer stratification[41] [Fig. 4(a)]. For VM2, c_{+2}^N shows a similar behaviour, although the peak is located at $-\hat{\eta}_2/\hat{h}_2 = 2.07$ and the decrease does not appear in Fig. 4(a). In contrast, c_{+0}^N (for the barotropic mode) increases monotonically. The power-series solutions around $\hat{\eta}_p = 0$ approach the corresponding results from numerical integration with increasing number of the terms n_S for small $|\hat{\eta}_p|$ but diverge from them for $|\hat{\eta}_p|$ larger than the radius of convergence of the power series. The d'Alembert ratio test (see Sec. 5.7 in Ref. 21) suggests that the radius of convergence in $\hat{\eta}_p/\hat{h}_p$ is roughly 0.9 and 1.0 for the barotropic mode and VM1, respectively. The radius could not be determined for VM2 because of the slower convergence and the limited number of reliable terms

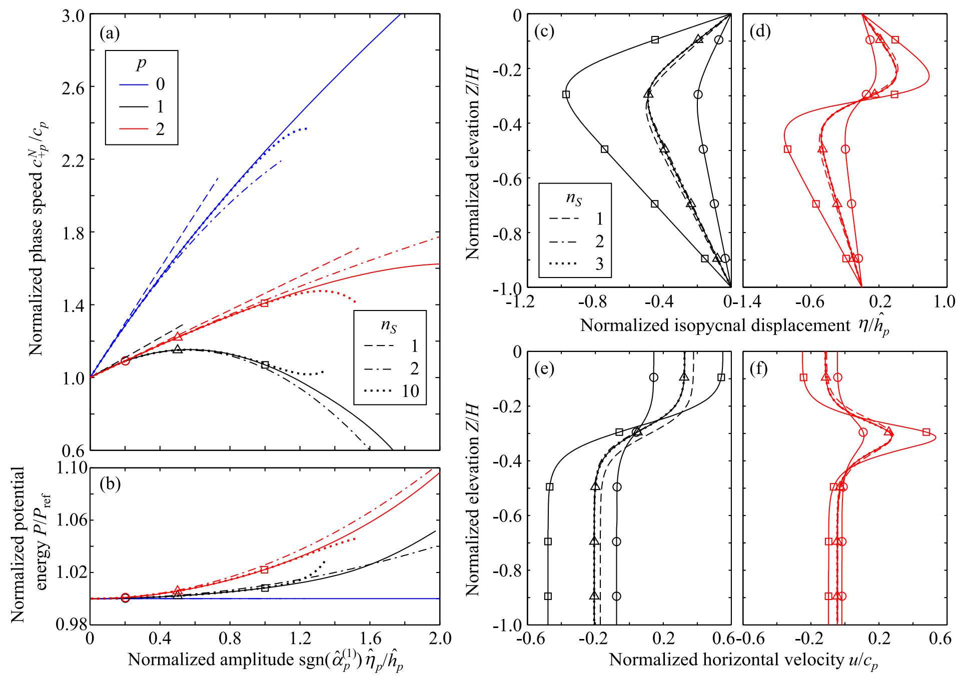

FIG. 4. Fully nonlinear simple-wave solutions for barotropic ($p = 0$), 1st baroclinic ($p = 1$), and 2nd baroclinic ($p = 2$) modes under stratification A [see Fig. 1(a)]. (a) Normalized phase speeds as a function of normalized modal displacement amplitude, (b) normalized vertically integrated available potential energy, (c) normalized isopycnal displacements for $p = 1$, (d) those for $p = 2$, (e) normalized horizontal velocities for $p = 1$, and (f) those for $p = 2$. Solid lines show results of direct numerical integration, whereas broken lines show power-series solutions around the linear case ($\hat{\eta}_p = 0$) with up to n_Sth power term [see (H1); note that vertical structure for $n_S = 1$ is linear, and that for $n_S = 2$ with 1st-order correction]. Colored symbols show correspondence between phase speed and vertical structure. Note that $\hat{\alpha}_p^{(1)} > 0$ for $p = 0$ but <0 for $p = 1,2$. Computation is done with $n_L = 200$ and $n_M = 10$. Normalization uses $P_{\rm ref} = \frac{1}{2}\hat{\rho}c_p^2\hat{h}_p^{-1}\hat{\eta}_p^2$, and $H^{-1/2}bc_p = (3.13, 0.0402, 0.0127)$ and $H^{-1}\hat{h}_p = (0.999, 0.204, 0.0415)$ for $p = (0,1,2)$, where $b = \sqrt{10^{-3}\hat{\rho}/\Delta\rho}$.

in the power-series solution. (Note that the maximum amplitude of solitary waves for surface waves[42] or internal waves[35] does not restrict the amplitude range of the hydrostatic solutions considered here; e.g., dam break problem in Sec. 13.10 in Ref. 20.)

Unlike stratification A, $\hat{\alpha}_1^{(2)}$ is positive for VM1 under stratification C, and c^N_{+1} increases monotonically at least until $-\hat{\eta}_1/\hat{h}_1 = 4$ [Fig. 5(a)]. Such monotonic increase typically happens for density profiles with monotonically increasing or decreasing buoyancy frequency. c^N_{+2} changes a little until the solution for VM2 reaches a sink [red cross in Fig. 5(a)]. The power-series solutions show similar convergence characteristics to stratification A, except that c^N_{+2} diverges from the directly computed solution at smaller $|\hat{\eta}_p|$.

Unlike the weakly nonlinear asymptotic solution, the power-series solution can be calculated around any $\hat{\eta}_p$. Using the result of numerical integration at $-\hat{\eta}_p/\hat{h}_p = 1$, the power-series solution around that $\hat{\eta}_p$ was calculated for stratification B [Fig. 5(b)]. The solutions correctly approach the corresponding results of numerical integration around $-\hat{\eta}_p/\hat{h}_p = 1$.

The vertically integrated available potential energy P is essentially given by the linear relationship $P_{\rm ref} = \frac{1}{2}\hat{\rho}c_p^2\hat{h}_p^{-1}\hat{\eta}_p^2$ for the barotropic mode, and deviated relatively little from $P_{\rm ref}$ for VM1 and VM2 even under strong nonlinearity [Fig. 4(b)]. The deviation is always positive because of Parseval's theorem (C2a) and the excitation of modes other than the mode of interest by nonlinear interaction.

For the hyperbolic tangent stratification, the linear modes of interest plus the 1st-order corrections provide a good approximation to the vertical structure of the wave fields even under moderately strong nonlinearity [e.g., Figs. 4(c)–4(f)].

The above examples illustrate that the higher-order coefficients $\hat{\alpha}^{(k)}$ and correction functions $\hat{\phi}^{N(k)}$ and $\hat{\pi}^{N(k)}$ from the power-series solution provide the correct asymptotic behaviour of the hydrostatic fully nonlinear solution, and that they are physically unique unless the mode of interest is a degenerate eigen solution to (50). This in turn confirms that $\hat{\alpha}_p^{(1)}$, $\hat{\alpha}_p^{(2)}$, $\hat{\phi}_p^{N(1)}$, and $\hat{\pi}_p^{N(1)}$ from the 2nd-order theory proposed in Sec. IV A are physically unique.

D. Physical uniqueness of nonlinear correction in previous theories

The hydrostatic fully nonlinear solution derived in this section shows not only the physical uniqueness (i) (i.e., consistency of the nonlinear coefficients and correction functions with the corresponding fully nonlinear phase speeds) of the 2nd-order theory proposed in Sec. IV but also that of the previous theories. This can be shown as follows. As in Sec. IV, we use the previous theory[18] because it (effectively) uses the isopycnal coordinate. Using the notation in the Introduction for the previous theory ($\hat{A}$, c, $\hat{\alpha}$, $\hat{\alpha}_1$, $\hat{T}_n$) and that in Sec. IV A for the proposed theory ($\hat{\eta}_p$, c_p, $\hat{\alpha}_p^{(1)}$, $\hat{\alpha}_p^{(2)}$, $\hat{\phi}_p^{N(1)}$), the relationships between relevant variables are

$$\hat{T}_n \sim \hat{\phi}_p^{N(1)} + a\hat{\phi}_p, \quad (55a)$$

$$\hat{\alpha}_1 \sim \hat{\alpha}_p^{(2)} + a\hat{\alpha}_p^{(1)}, \quad (55b)$$

$$\hat{A} + a\hat{A}^2 \sim \hat{\eta}_p, \quad (55c)$$

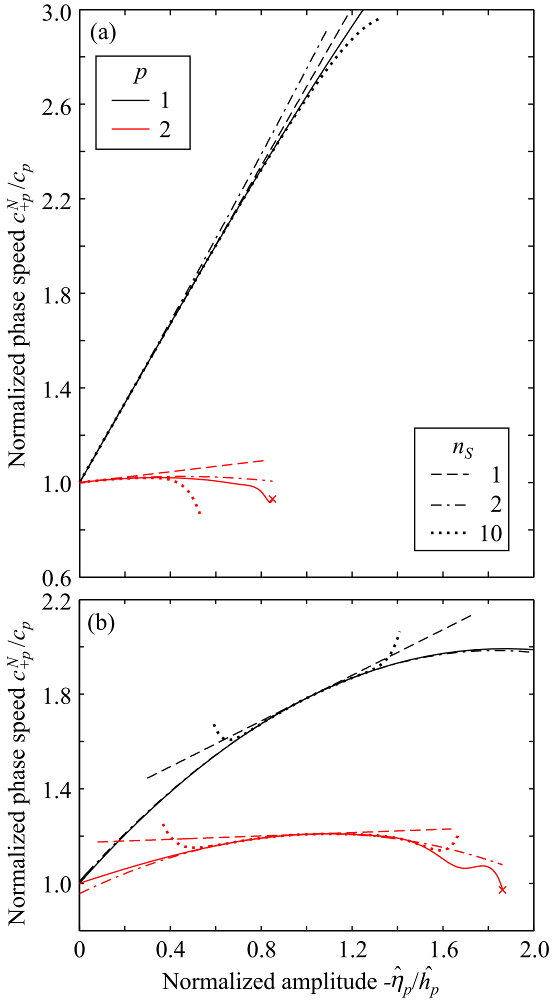

FIG. 5. Normalized phase speeds from fully nonlinear simple-wave solutions for 1st baroclinic ($p = 1$) and 2nd baroclinic ($p = 2$) modes. (a) Stratification C and (b) stratification B (see Fig. 1). Solid lines show results of direct numerical integration, whereas broken lines show power-series solutions around the linear case ($\hat{\eta}_p = 0$) in (a) and $-\hat{\eta}_p/\hat{h}_p = 1$ in (b) with up to n_Sth power term. The numerical solution is unreliable near sinks (red crosses). Computation is done with $n_L = 200$ and $n_M = 10$. For $p = (1, 2)$, normalization uses $H^{-1/2}bc_p = (0.002\,39, 0.001\,06)$ and $H^{-1}\hat{h}_p = (0.0324, 0.0140)$ for stratification C but $H^{-1/2}bc_p = (0.0255, 0.009\,79)$ and $H^{-1}\hat{h}_p = (0.131, 0.0411)$ for stratification B, where $b = \sqrt{10^{-3}\hat{\rho}/\Delta\rho}$.

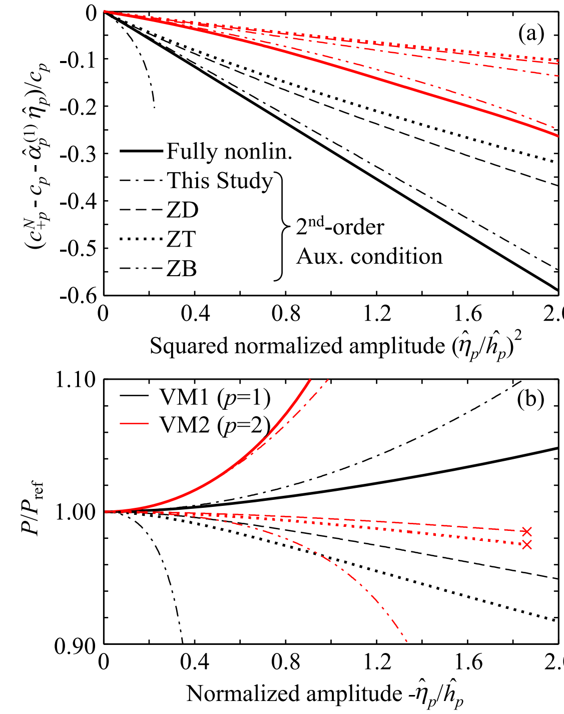

FIG. 6. Comparisons of different auxiliary conditions. (a) Normalized deviation of phase speed c^N_{+p} from 1st-order theory, and (b) normalized vertically integrated available potential energy P. Stratification is stratification B [see Fig. 1(a)]. Red crosses indicate sinks. Note that the slope at $\hat{\eta}_p = 0$ in (a) is the normalized cubic nonlinear coefficient when amplitude is defined as $\hat{\eta}_p$ since the horizontal axis is the squared normalized amplitude. Normalization uses linear phase speed c_p and P in linear limit, $P_{\mathrm{ref}} = \frac{1}{2}\hat{\rho}c_p^2\hat{h}_p^{-1}\hat{\eta}_p^2$. Abbreviations are as follows: Fully nonlin.: numerical solution for fully nonlinear simple waves, this study: proposed 2nd-order theory, ZD: previous 2nd-order theory with zero isopycnal-displacement correction condition, ZT: zero volume-transport correction condition, and ZB: zero correction of vertical gradient of isopycnal displacement at the bottom. See the caption of Fig. 5 for other computational conditions, and Sec. IV B for details of different auxiliary conditions. For reference, cubic nonlinear coefficients under (this study, ZD, ZT, ZB) are (−0.40, −0.10, −1.2, 0.048) for VM1 and (−0.39, −0.24, −0.65, −0.17) for VM2.

where a is a coefficient determined by the auxiliary condition for $\hat{T}_n$.[13] The last relationship is obtained by projecting (1) onto the mode of interest, and a special case of (7). Substituting the last two relationships into the power-series expansion of $c^N(\hat{\eta})$ in the proposed theory [i.e., (H8a) with $\hat{\eta}_p^\circ = 0$] and (37b), we get

$$c^N(\hat{A}) \sim c + \hat{\alpha}\hat{A} + \hat{\alpha}_1\hat{A}^2, \tag{56a}$$

$$P(\hat{A}) \sim \frac{c^2}{2\hat{h}}\hat{A}^2\left(1 + 2a\hat{A}\right). \tag{56b}$$

Although not proved analytically, numerical evaluation of $c^N(\hat{A})$ from the previous theory and $c^N(\hat{\eta}_p)$ from the proposed theory shows that they are asymptotically equivalent, even when the sign of $\hat{\alpha}_1$ is opposite from $\hat{\alpha}^{(2)}_p$ [see the slopes at $\hat{\eta}_p = 0$ in Fig. 6(a)]. This confirms that the previous theories not only satisfy the physically uniqueness (i) irrespective of the auxiliary condition as shown in Ref. 16 but also provide the correct asymptotic expansion of the hydrostatic fully nonlinear solution. We can expect this to be true to higher-order nonlinearity because we can substitute arbitrary $\hat{\eta}_p(\hat{A})$ into $c^N(\hat{\eta}_p)$ from the power-series solution and because the coefficients of an asymptotic series are unique for a given form of the series (see Sec. 3.4 in Ref. 43). In other words, nonuniqueness of the 2nd-order nonlinear coefficient originates from the change of prognostic variable used in the asymptotic expansion [i.e., the mathematical nonuniqueness (ii) in the Introduction]. Similarly, P under arbitrary auxiliary conditions are asymptotically equivalent. For example, the derivative of P/P_{ref} ($P_{\mathrm{ref}} = \frac{1}{2}\hat{\rho}c_p^2\hat{h}_p^{-1}\hat{\eta}_p^2$ is P in the linear limit) with respect to $\hat{\eta}_p$ is zero at $\hat{\eta}_p = 0$ irrespective of the auxiliary condition [Fig. 6(b)]. Note, however, that P is positive definite only when $\hat{A} = \hat{\eta}_p$ [see (37)]. Note also that asymptotically equivalent results under different auxiliary conditions have different accuracy and convergence characteristics (Sec. 3.5 in Ref. 43) as seen in Fig. 6; they are separate issues from physical uniqueness. An important result here is that it is essential to use consistent definitions (or the same auxiliary condition) for the amplitude $\hat{A}$ and the cubic nonlinear coefficient $\hat{\alpha}_1$ to make $c^N(\hat{A})$ the correct asymptotic expansion of the fully nonlinear phase speed.

VI. DISCUSSION

This study extended the argument for physical uniqueness of the higher-order KdV theory for continuously stratified fluids[16] and considered the uniqueness in the following three physical senses: (i) consistency of the nonlinear higher-order coefficients and correction functions with the corresponding phase speeds, (ii) wavenumber-independence of the vertically integrated available potential energy P, and (iii) its positive definiteness. This study confirmed not only that the cubic nonlinear coefficient from the previous theories[13,15–18] is physically unique in the sense (i) under an arbitrary auxiliary condition[16] but also that the previous theories provide correct asymptotic expansion of the hydrostatic fully nonlinear phase speed. This study also showed that there is a unique set of special auxiliary conditions (38) for P to be wavenumber-independent and positive definite. Combining the spectral (or generalized Fourier) approach based on vertical modes and the simple-wave analysis in the isopycnal coordinate, we also derived hydrostatic fully nonlinear gravity wave solutions, as well as algebraic recurrence relationships to determine an arbitrary number of the nonlinear ($\hat{\alpha}$) and linear nonhydrostatic ($\hat{\beta}$) coefficients and associated correction functions. To my knowledge, these are the original contributions of this study.

It is worth considering implications of the results in this study to large amplitudes beyond the formal applicability of the asymptotic theory (i.e., $\epsilon \gtrsim 1$) because the Gardner equation is often used to analyze ISWs with such large amplitudes.[5,16] On the one hand, nonuniqueness of the cubic nonlinear coefficient offers flexibility of "tuning" the Gardner equation for a wide amplitude range beyond the formal applicability.[13] Although obtaining a reference for such tuning is not trivial under realistic conditions in oceans and lakes, this flexibility may be preferable to imposing positive definiteness of P in some application, assuming that the error in P is not too large [e.g., Fig. 6(b)]. On the other hand, defining the amplitude as generalized Fourier coefficients $\hat{\eta}_p$ is a special case that offers several advantages. First, it is expected to have a wider applicable range because $\hat{\eta}_p$ is defined for arbitrary strength of nonlinearity [see (20)] and is a good prognostic variable to characterize the fully nonlinear solution (Sec. V), whereas the amplitude under an arbitrary auxiliary condition $\hat{A}$ is defined through an asymptotic relationship (1) and may become unrealistically large or cease to exist under strong nonlinearity [see (55c)]. In addition, Parseval's theorem guarantees wavenumber-independence and positive definiteness of P in general. Second, $\hat{\eta}_p$ leads to more intuitive results because it makes the asymptotic theory consistent with the generalized Fourier theory (Sec. IV C). For example, the 2nd adiabatic (or integral) invariant of the Gardner equation, $\int_{-\infty}^{+\infty} \hat{A}^2 dx$, is equivalent to the conservation of available potential energy only if $\hat{A} = \hat{\eta}_p$ [see (37)]. These suggest that the special auxiliary condition (38a) might be useful as a default condition when the Gardner equation is not "tuned" against some reference. In contrast to the above nonlinear case, there appears to be no benefit in replacing (38b) with an arbitrary auxiliary condition (and in introducing wavenumber dependence of P) for linear nonhydrostatic waves. For the 2nd-order KdV equation, however, it appears necessary to use the same auxiliary condition for $\hat{T}_n$ and $\hat{T}_d$ [(38) or otherwise] because a common definition of $\hat{A}$ needs to be used in the equation. In particular, wavenumber independence of P appears to require not only (38b) but also (38a).

In the Introduction and Sec. IV B, it is mentioned that nonuniqueness of the cubic nonlinear coefficient poses difficulty in applying the Gardner equation to the observed or simulated ISWs. If the positive definiteness of P is not imposed, there is no unique choice for the cubic nonlinear coefficient $\hat{\alpha}_1$; however, this study showed that the consistent definitions (or the same auxiliary condition) have to be used for the amplitude $\hat{A}$ and the coefficient $\hat{\alpha}_1$, in order to make the phase speed of the Gardner equation the correct asymptotic expansion of the fully nonlinear phase speed [i.e., the physical uniqueness (i)]. This has already been shown in Ref. 16, but it is a point worth emphasizing. Once $\hat{A}$ is calculated from observed or modeled wave fields for comparison with solutions to the Gardner equation, that definition of amplitude uniquely determines the coefficient $\hat{\alpha}_1$. It appears common to calculate an ISW amplitude from the observed or simulated vertical displacement of a particular isopycnal but to calculate $\hat{\alpha}_1$ by setting the nonlinear correction function $\hat{T}_n$ to be zero at the depth of maximum isopycnal displacement or volume transport. When applied to variable depths, for example, continental shelf/slope topography, the two choices are inconsistent from a view point of the physical uniqueness (i) because they use different auxiliary conditions.

As far as I am aware, this is the first study that derived the phase speed and associated vertical structure of hydrostatic fully nonlinear gravity waves for individual vertical modes in continuously stratified fluids. It extends the fully nonlinear simple-wave analysis for homogeneous (e.g., Sec. 13.10 in Ref. 20) and two-layer stratified[41] fluids to continuously stratified fluids. The fully nonlinear solution would enable the construction of simplified fully nonlinear models in the future. For example, by allowing two Riemann invariants to vary, it would be possible to obtain the equations for bi-directional fully nonlinear waves using $\hat{\eta}_p$ and $\hat{u}_p$ as the prognostic variables. Furthermore, adding the linear nonhydrostatic term in (21) to the equation for $\hat{u}_p$ would yield the Boussinesq equation for continuous stratification. It would also be worth comparing (21) and the Miyata-Choi-Camassa (MCC) model[22,44] or the phenomenological uni-directional equation[41] for two-layer stratification, and extending these weakly nonhydrostatic, fully nonlinear models to continuous stratification. Note that, neglecting all the baroclinic (and degenerate) modes and noting $\hat{N}^B_{000} = 1$, (21) is already the so-called Green-Naghdi equation,[45,46] the homogeneous counterpart of the MCC model. These extensions are in progress.

VII. CONCLUSIONS

Using the spectral (or generalized Fourier) approach based on vertical modes and the simple-wave analysis in the isopycnal coordinate, this study proposed to reduce arbitrariness of the higher-order KdV theory by considering its uniqueness in the following three physical senses: (i) consistency of the nonlinear higher-order coefficients and correction functions with the corresponding phase speeds, (ii) wavenumber-independence of the vertically integrated available potential energy, and (iii) its positive definiteness. Both the 2nd-order theory proposed in Sec. IV and the previous theories satisfy (i), provided that consistent definitions are used for the wave amplitude and the nonlinear correction. This condition reduces the arbitrariness when higher-order KdV-type theories are compared with observations or numerical simulations. Although flexibility of 'tuning' the higher-order coefficients may be preferred in some applications, the physical uniqueness (ii) and (iii) offers special auxiliary conditions that make the previous asymptotic KdV theories consistent with the spectral approach and uniquely determine the higher-order coefficients and associated correction functions.

ACKNOWLEDGMENTS

I thank two anonymous referees for constructive comments and criticisms and Steve Buchan for proof-reading.

APPENDIX A: DEPENDENCE OF MODAL AMPLITUDES ON VERTICAL COORDINATE

The choice of the vertical coordinate affects modal amplitudes (i.e., $\hat{\eta}, \hat{m}, \hat{\vec{u}}$, and $\hat{w}$), their evolutionary equations, and the higher-order coefficients of KdV-type equations. As an example, let us consider the formula to calculate $\hat{w}_n$ from a known vertical velocity field w,

$$\hat{w}_n = \frac{\hat{h}_n}{c_n^2}\left\langle \hat{\phi}_n, w \right\rangle, \tag{A1}$$

where the inner product is defined in (18). Vertical modes in the z coordinate $\hat{\phi}_n(z)$ have the same functional form as those in the undisturbed height coordinate Z because the governing equation has the same form [i.e., $(dZ/ds)^{-1}(d/ds) \to d/dZ$ and $(\rho, Z) \to (\rho_{\mathrm{ref}}, z)$ in (16a)], and $Z(\rho)$ and $\rho_{\mathrm{ref}}(z)$ refer to the same stratification. However, since $z = Z + \eta$, we have $w(z) = w(Z) + \epsilon(\partial w/\partial Z)\eta(Z) + \cdots$ in the nondimensional form based on the scaling (24). So, the above formula in the z coordinate becomes

$$(\hat{w}_n)_{z\text{-coord}} \sim \hat{w}_n + \epsilon \frac{\hat{N}^B_{nml}}{\hat{h}_m}\hat{\eta}_l \hat{w}_m \tag{A2}$$

to 1st order. Such a nonlinear relationship between modal amplitudes in different vertical coordinates leads to the dependence of the higher-order coefficients of KdV-type equations on the choice of the vertical coordinate. Although the dependence on the vertical coordinate does not appear explicitly in the evolutionary equations of modal amplitudes (21), the form of the equations depends on the vertical coordinate because the original governing equations have different forms [cf. (8) and (12)].

Note that we need to use a particular vertical coordinate, such as the isopycnal coordinate or the semi-Lagrangian coordinate used in previous higher-order KdV theories[18,32,33] to obtain results consistent with multilayer formulation.

APPENDIX B: PROOF OF THE IDENTITY RELATIONSHIPS (23)

The identity relationships among the interaction coefficients (23) are derived as follows. First, let us consider (23a). The projection of $\hat{\pi}_l\hat{\phi}_m$ onto $\hat{\phi}$ yields $\hat{\pi}_l\hat{\phi}_m = \hat{\phi}_j\hat{N}^B_{jlm}$, and substituting this into the definition of $\hat{N}^D_{nlm}$ (22d) yields (23a). To prove (23b), let us consider

$$\int_{s^b}^{s^t} \hat{\pi}_n \rho \frac{dZ}{ds}\hat{\pi}_l\hat{\phi}_k\hat{\phi}_m ds. \tag{B1}$$

The projection of $\hat{\pi}_l\hat{\pi}_n$ onto $\hat{\pi}$ yields $\hat{\pi}_l\hat{\pi}_n = \hat{\pi}_j\hat{h}_l\hat{h}_j^{-1}\hat{N}^A_{nlj}$ and substituting it into (B1) yields the right-hand-side of (23b). Similarly, the projection of $\hat{\pi}_n\hat{\phi}_k$ onto $\hat{\phi}$ yields $\hat{\pi}_n\hat{\phi}_k = \hat{\phi}_j\hat{N}^B_{jnk}$, and substituting it into (B1) yields the left-hand-side of (23b).

APPENDIX C: ENERGETICS IN ISOPYCNAL AND MODAL COORDINATES

The vertically integrated available potential and kinetic energy equations associated with (12) can be obtained as follows. To derive the potential energy equation, we multiply (12a) by m, apply integration by parts to the both sides, and substitute (12c), (13), and (14). To derive the kinetic energy equation, we take the dot product of $((\partial Z/\partial s) + (\partial\eta/\partial s))\,\vec{u}$ and (12b), multiply (12c) by w, use (12a), and add the resulting equations. These yield

$$\frac{\partial P}{\partial t} + \nabla\cdot\vec{J}_P = -C_{PK}, \tag{C1a}$$

$$\frac{\partial K}{\partial t} + \nabla\cdot\vec{J}_K = C_{PK}, \tag{C1b}$$

$$P = \frac{1}{2}\left(\rho^t g(\eta^t)^2 + \int_{s^b}^{s^t} \rho N^2 \frac{dZ}{ds}\eta^2 ds\right), \tag{C1c}$$

$$K = \int_{s^b}^{s^t} \frac{\rho}{2}\left(\frac{dZ}{ds} + \frac{\partial\eta}{\partial s}\right)\left(|\vec{u}|^2 + w^2\right) ds, \tag{C1d}$$

$$\vec{J}_P = \int_{s^b}^{s^t} m\left(\frac{dZ}{ds} + \frac{\partial\eta}{\partial s}\right)\vec{u} ds, \tag{C1e}$$

$$\vec{J}_K = \int_{s^b}^{s^t} \frac{\rho}{2}\left(\frac{dZ}{ds} + \frac{\partial\eta}{\partial s}\right)\left(|\vec{u}|^2 + w^2\right)\vec{u} ds, \tag{C1f}$$

$$C_{PK} = \int_{s^b}^{s^t} \rho\left(\frac{dZ}{ds} + \frac{\partial\eta}{\partial s}\right)\frac{\partial\eta}{\partial t}\frac{Dw}{Dt} ds - \int_{s^b}^{s^t}\left(\frac{dZ}{ds} + \frac{\partial\eta}{\partial s}\right)\vec{u}\cdot\nabla m ds. \tag{C1g}$$

Adding (C1a) and (C1b) yields the conservation of total energy $E = P + K$. The energetics can also be expressed in terms of modal amplitudes. The most straightforward way of deriving the expressions is by applying the modal decomposition to P, K, $\vec{J}_P$, and $\vec{J}_K$. The same results can be obtained by manipulating the modal amplitude equations (21) and using the identity among the interaction coefficients (23). These yield

$$P = \frac{\hat{\rho}}{2}\left(\hat{\eta}_n \frac{c_n^2}{\hat{h}_n}\hat{\eta}_n\right), \tag{C2a}$$

$$K = \frac{\hat{\rho}}{2}\left[\hat{\vec{u}}_n \cdot \left(\hat{h}_n\delta_{nm} + \hat{N}^A_{nlm}\hat{\eta}_l\right)\hat{\vec{u}}_m + \hat{w}_n\left(\hat{D}_{nm} + \hat{N}^D_{nlm}\hat{\eta}_l\right)\hat{w}_m\right], \tag{C2b}$$

$$\vec{J}_P = \hat{m}_n\left(\hat{h}_n\delta_{nm} + \hat{N}^A_{nlm}\hat{\eta}_l\right)\hat{\vec{u}}_m, \tag{C2c}$$

$$\vec{J}_K = \frac{\hat{\rho}}{2}\left[\hat{N}^A_{nkm}\left(\hat{\vec{u}}_n\cdot\left(\hat{h}_k\delta_{jk} + \hat{N}^A_{klj}\hat{\eta}_l\right)\hat{\vec{u}}_m\right)\hat{\vec{u}}_j + \hat{N}^D_{nkm}\left(\hat{w}_n\left(\hat{h}_k\delta_{jk} + \hat{N}^A_{klj}\hat{\eta}_l\right)\hat{w}_m\right)\hat{\vec{u}}_j\right]. \tag{C2d}$$

Note that decomposing the energetics into individual modal components is not straightforward due to nonlinear effects (although it could be done using the approach in Refs. 47 and 40). However, the relationships for P remain linear in the isopycnal coordinate, and Parseval's theorem holds for the isopycnal displacement field η under full nonlinearity and nonhydrostaticity [compare (C1c) and (C2a)].

APPENDIX D: EXPRESSIONS FOR $\hat{F}_n^{(k)}$ AND $\hat{G}_n^{(k)}$

The variables $\hat{F}_n^{(k)}$ and $\hat{G}_n^{(k)}$ in (28) are given by

$$\hat{F}_p^{(0)} = \hat{G}_p^{(0)} = 0, \tag{D1a}$$

$$\hat{F}_p^{(1)} = \hat{N}^A_{ppp}\hat{\eta}_p^{(0)}\hat{u}_p^{(0)}, \tag{D1b}$$

$$\hat{G}_p^{(1)} = \frac{1}{2}\frac{\hat{h}_p}{c_p}\hat{N}^A_{ppp}(\hat{u}_p^{(0)})^2 + c_p\hat{h}_p\hat{D}_{pp}\frac{\partial^2\hat{\eta}_p^{(0)}}{d\xi^2}, \tag{D1c}$$

$$\hat{F}_q^{(1)} = \hat{N}^A_{qpp}\hat{\eta}_p^{(0)}\hat{u}_p^{(0)}, \tag{D1d}$$

$$\hat{G}_q^{(1)} = \frac{1}{2c_q}\hat{N}^A_{qpp}\hat{h}_p(\hat{u}_p^{(0)})^2 + \frac{\hat{h}_q}{c_q}\hat{D}_{qp}c_p^2\frac{\partial^2\hat{\eta}_p^{(0)}}{d\xi^2}, \tag{D1e}$$

$$\hat{F}_p^{(2)} = \hat{N}^A_{ppp}\left(\hat{\eta}_p^{(0)}\hat{u}_p^{(1)} + \hat{\eta}_p^{(1)}\hat{u}_p^{(0)}\right) + \hat{\eta}_p^{(0)}\hat{N}^A_{ppq}\hat{u}_q^{(1)} + \hat{h}_p\hat{u}_p^{(0)}\hat{N}^A_{ppq}\frac{1}{\hat{h}_q}\hat{\eta}_q^{(1)}, \tag{D1f}$$

$$\begin{aligned}\hat{G}_p^{(2)} = {} & \frac{\hat{h}_p}{c_p}\hat{N}^A_{ppp}\hat{u}_p^{(0)}\hat{u}_p^{(1)} + \frac{\hat{h}_p}{c_p}\hat{u}_p^{(0)}\hat{N}^A_{ppq}\hat{u}_q^{(1)} \\ & + c_p\hat{h}_p\hat{D}_{pp}\left(\frac{\partial^2\hat{\eta}_p^{(1)}}{d\xi^2} - \frac{2}{c_p}\frac{\partial}{\partial\xi}\frac{\partial\hat{\eta}_p^{(0)}}{\partial\tau}\right) \\ & + c_p\hat{h}_p\hat{D}_{pq}\frac{\partial^2\hat{\eta}_q^{(1)}}{d\xi^2} \\ & - \hat{h}_p^2\hat{N}^D_{ppp}\left(2\hat{u}_p^{(0)}\frac{\partial^2\hat{\eta}_p^{(0)}}{d\xi^2} + \frac{\partial\hat{u}_p^{(0)}}{\partial\xi}\frac{\partial\hat{\eta}_p^{(0)}}{\partial\xi}\right) \\ & + c_p\hat{h}_p\hat{N}^D_{ppp}\left(\hat{\eta}_p^{(0)}\frac{\partial^2\hat{\eta}_p^{(0)}}{d\xi^2} + \frac{1}{2}\frac{\partial\hat{\eta}_p^{(0)}}{\partial\xi}\frac{\partial\hat{\eta}_p^{(0)}}{\partial\xi}\right),\end{aligned} \tag{D1g}$$

where no sum is taken on $p(\neq q)$ because it is the single mode of interest, but the sum on q is taken in (D1f) and (D1g).

APPENDIX E: APPLICATION OF PROPOSED 2ND-ORDER KdV THEORY TO TWO-LAYER STRATIFICATION

To check the results in Sec. IV A, we applied the proposed 2nd-order KdV theory to two-layer stratification and compared the resulting 2nd-order coefficients with the well-known formulae[22,36,37] for surface and internal waves. Substituting the two-layer density profile (26) into the eigenvalue problem for vertical modes (16a), noting that N is assumed to be negligibly small, and solving the problem for each layer, we get

$$\hat{\phi}_n = \begin{cases} \dfrac{c\hat{A}}{N}\left(\sin\left(\dfrac{N}{c}Z\right) + \dfrac{cN}{g}\cos\left(\dfrac{N}{c}Z\right)\right) \\ \dfrac{c\hat{B}}{N}\sin\left(\dfrac{N}{c}(Z+H)\right)\end{cases}, \tag{E1}$$

where $\hat{A}$ and $\hat{B}$ are integration constants. In this appendix, upper and lower expressions after an opening curly bracket are conditional for the upper layer ($-h_I < Z < 0$) and the lower layer ($-H < Z < -h_I$), respectively. The continuity of $\hat{\phi}_n$ and pressure at the interface yield the relationship between $\hat{A}$ and $\hat{B}$,

$$\hat{A}\left(-S_I + \frac{cN}{g}C_I\right) = \hat{B}S_{II}, \tag{E2}$$

and the dispersion relationship

$$\left(C_IC_{II} - \frac{\rho_I}{\rho_{II}}S_IS_{II}\right)\left(\frac{cN}{g}\right)^2 - (S_IC_{II} + C_IS_{II})\left(\frac{cN}{g}\right) + \frac{\Delta\rho}{\rho_{II}}S_IS_{II} = 0, \tag{E3}$$

where we use the short-hand notation

$$C_J = \cos\left(\frac{Nh_J}{c}\right), \quad S_J = \sin\left(\frac{Nh_J}{c}\right), \tag{E4}$$

and $J = I, II$ for the upper and lower layers, respectively. For $N \to 0$ and finite c, (E3) reduces to the standard dispersion relationship. For $N \to 0$ and $c \to 0$ but finite N/c, (E3) reduces to $S_IS_{II} \sim 0$. In this limit, we see from (E2) that $S_I = 0$ and $S_{II} = 0$ correspond to localized solutions within the upper and lower layers, respectively. We denote these solutions with modal indices q_I and q_{II}. These solutions are degenerate ($c_{q_I} \sim c_{q_{II}} \sim 0$), but the associated vertical modes are finite and required for the completeness of the modal expansion (19). To leading order under the Boussinesq approximation, the solutions are the celerities

$$c_0^2 \sim gH, \quad c_1^2 \sim \frac{\Delta\rho g}{\hat{\rho}}\frac{h_Ih_{II}}{H}, \tag{E5a}$$

$$c_{q_I}^2 \sim c_{q_{II}}^2 \sim 0, \tag{E5b}$$

and the associated modal structure,

$$\hat{\phi}_0 \sim \frac{Z+H}{H}, \qquad \hat{\phi}_1 \sim \begin{cases} -\dfrac{Z}{h_I} \\ \dfrac{Z+H}{h_{II}}\end{cases}, \tag{E6a}$$

$$\hat{\phi}_{q_I} \sim \begin{cases} -\sin\dfrac{q_I\pi}{h_I}Z \\ 0 \end{cases}, \quad \hat{\phi}_{q_{II}} \sim \begin{cases} 0 \\ \sin\dfrac{q_{II}\pi}{h_{II}}(Z+H). \end{cases} \tag{E6b}$$

Other necessary parameters are calculated from (15), (17a), and (22),

$$\hat{h}_0 \sim H, \quad \hat{h}_1 \sim \frac{h_I h_{II}}{H}, \quad \hat{h}_{q_J} \sim \frac{2h_J}{(q_J\pi)^2}, \tag{E7a}$$

$$\hat{D}_{00} \sim \hat{D}_{11} \sim \frac{H}{3}, \quad \hat{D}_{01} \sim \frac{H + h_{II}}{6}, \tag{E7b}$$

$$\hat{D}_{0q_I} \sim \frac{h_I}{q_I\pi}\left(1 - \frac{h_{II}}{H}(-1)^{q_I}\right), \tag{E7c}$$

$$\hat{D}_{0q_{II}} \sim -\frac{1}{q_{II}\pi}\frac{h_{II}^2}{H}(-1)^{q_{II}}, \tag{E7d}$$

$$\hat{D}_{1q_J} \sim -\frac{h_J}{q_J\pi}(-1)^{q_J}, \tag{E7e}$$

$$\hat{N}^A_{000} \sim \hat{N}^A_{011} \sim 1, \tag{E7f}$$

$$\hat{N}^A_{111} \sim \frac{h_I - h_{II}}{H}, \tag{E7g}$$

$$\hat{N}^A_{001} \sim \hat{N}^A_{00q_J} \sim \hat{N}^A_{11q_J} \sim 0, \tag{E7h}$$

$$\hat{N}^D_{000} \sim \frac{1}{3}, \quad \hat{N}^D_{111} \sim 0. \tag{E7i}$$

We substitute these into (30) and (33) and use the following relationships related to the Riemann zeta function (e.g., Sec. 5.9 in Ref. 21 or Sec. 23.2 in Ref. 48),

$$\sum_{q=1}^{\infty}\frac{1}{q^4} = \frac{\pi^4}{90}, \quad \sum_{q=1}^{\infty}\frac{(-1)^{q-1}}{q^4} = \frac{7\pi^4}{720}. \tag{E8}$$

The results are the coefficients for surface waves,

$$\hat{\alpha}_0^{(1)} \sim \frac{3c}{2H}, \tag{E9a}$$

$$\hat{\beta}_0^{(1)} \sim \frac{1}{6}cH^2, \tag{E9b}$$

$$\hat{\alpha}_0^{(2)} \sim -\frac{3c}{8H^2}, \tag{E9c}$$

$$\hat{\beta}_0^{(2)} \sim \frac{19}{360}cH^4, \tag{E9d}$$

$$\hat{\gamma}_0^{(2a)} \sim \frac{5}{12}cH, \tag{E9e}$$

$$\hat{\gamma}_0^{(2b)} \sim \frac{23}{24}cH \tag{E9f}$$

and those for internal waves,

$$\hat{\alpha}_1^{(1)} \sim \frac{3}{2}c_1\frac{h_I - h_{II}}{h_I h_{II}}, \tag{E10a}$$

$$\hat{\beta}_1^{(1)} \sim \frac{1}{6}c_1 h_I h_{II}, \tag{E10b}$$

$$\hat{\alpha}_1^{(2)} \sim -\frac{3c_1}{h_I^2 h_{II}^2}\left(\frac{(h_I - h_{II})^2}{8} + h_I h_{II}\right), \tag{E10c}$$

$$\hat{\beta}_1^{(2)} \sim c_1 h_I^2 h_{II}^2\left(\frac{1}{24} + \frac{1}{90}\frac{h_I^3 + h_{II}^3}{h_I h_{II} H}\right), \tag{E10d}$$

$$\hat{\gamma}_1^{(2a)} \sim \frac{7}{12}c_1(h_I - h_{II}), \tag{E10e}$$

$$\hat{\gamma}_1^{(2b)} \sim \frac{31}{24}c_1(h_I - h_{II}). \tag{E10f}$$

These coefficients agree with those from previous studies[22,36,37] under the Boussinesq approximation. Note that under non-Boussinesq conditions, the results of the proposed theory with free surface deviate from those in the previous studies assuming a rigid lid because layer densities appear in different ways in the dispersion relationship (E3).

APPENDIX F: NUMERICAL METHODS

For numerical computation of vertical modes, we integrate (15)–(17) and (22) with respect to s over each vertical cell, yielding the equations effectively for multilayer stratification. Then, we follow Ref. 27 to calculate vertical modes and the interaction parameters. This discretization keeps the identity relationships (23) in the discretized space. We use n_L number of vertical cells with a uniform height and the first n_M vertical modes. $n_L = 200$ is used throughout this paper to ensure reasonable resolution in a thin stratified region when $h/H < d/H$ or $1 - d/H < h/H$ [e.g., Fig. 1(b)].

Numerical integration of (53) together with (50) and (51) requires attention. Five important points are noted here. First, we need to be able to specify the $+p$th solution out of the solutions to (50). In general, this is not trivial because two eigen solutions to (50) can become very close during the integration. Fortunately, we did not encounter such difficulty for the lowest three vertical modes we considered. Second, the numerical integration does not provide the value of $\tilde{R}_{+p}$ corresponding to $\hat{\eta}_p = 0$ (i.e., the integration constant), but it is sufficient to define $\tilde{R}_{+p}$ as the deviation from this constant value for our purpose. Third, since $\tilde{\mathbf{L}}$ is a matrix consisting of left eigenvectors, $\tilde{\mathbf{L}}^H$ is invertible except in rather exceptional cases in which some of the left eigenvectors are not linearly independent. Fourth, we terminate the numerical integration when c^N_{+p} becomes complex-valued (i.e., linear instability) or near a sink. Results of the numerical integration are sensitive to the numerical scheme near these terminal points in general; however, using $\tilde{R}_{+p}$ as the independent variable gave more robust results near a sink. Fifth, the numerical integration is also sensitive to the numerical scheme near the beginning of the integration. The power-series solution derived in Appendix H can be used to set the initial condition at small $\tilde{R}_{+p}$ for accuracy and efficiency.

APPENDIX G: POWER-SERIES SOLUTION FOR LINEAR FULLY NONHYDROSTATIC WAVES IN CONTINUOUSLY STRATIFIED FLUIDS

For fully nonhydrostatic sinusoidal waves, it is well-known that the phase speed c^D and the associated vertical structure of isopycnal displacement $\hat{\phi}^D$ can be determined as solutions to an eigenvalue problem [e.g., Eq. (6.10.2) in Ref. 23]. The corresponding problem in the isopycnal coordinate can be obtained by neglecting the nonlinear terms in (12), deriving an equation for η, and assuming a sinusoidal wave of the form $\eta(x, s, t) = \hat{\eta}^D\hat{\phi}^D(s)\exp(i\kappa(x - c^D t))$, yielding

$${c^D}^2\frac{d}{dZ}\left(\rho\frac{d\hat{\phi}^D}{dZ}\right) + \rho\left(N^2 - {c^D}^2\kappa^2\right)\hat{\phi}^D = 0, \tag{G1}$$

where $(dZ/ds)^{-1}(d/ds) = (d/dZ)$ is used. The boundary conditions are given by (16b) and (16c) but setting $(s, c, \hat{\phi}) = (Z, c^D, \hat{\phi}^D)$. Although the solutions cannot be written in a closed form in general, we can derive algebraic recurrence relationships to calculate the power-series solution, which yields an arbitrary number of the linear dispersion coefficients of KdV-type equations, $\hat{\beta}^{(k)}$, and the associated correction functions, $\hat{\phi}^{D(k)}$ and $\hat{\pi}^{D(k)}$.

We derive the power-series solution using the linear fully nonhydrostatic version of (21) (i.e., $\epsilon = 0$ and $\mu = 1$). Obtaining equations for $\hat{\eta}_n$, writing the equations in a matrix form as in Sec. V B, and assuming a solution of the form $\hat{\boldsymbol{\eta}} = \hat{\eta}_n \hat{\boldsymbol{b}}_n \exp(\mathrm{i}\kappa(x - c_n^D t))$ (no sum on n), we get

$$c_n^{D2}\left(\hat{\mathbf{H}}^{-1} + \kappa^2 \hat{\mathbf{D}}\right)\hat{\boldsymbol{b}}_n = \mathbf{C}^2 \hat{\mathbf{H}}^{-1} \hat{\boldsymbol{b}}_n, \tag{G2}$$

where $\hat{\boldsymbol{\eta}}$, $\mathbf{C}$, and $\hat{\mathbf{H}}$ are defined in (45), and $\hat{\mathbf{D}}$ is a $n_M \times n_M$ matrix whose components are given by

$$(\hat{\mathbf{D}})_{nm} = \hat{D}_{nm}. \tag{G3}$$

Note that the matrices in (G2) are symmetric. The solution in the isopycnal coordinate is given by $\hat{\phi}_n^D = \hat{\boldsymbol{\phi}}^T \hat{\boldsymbol{b}}_n$ [see (48) and (49)]. The orthogonality of $\hat{\boldsymbol{b}}_n$ is given by

$$\hat{\boldsymbol{b}}_n^T \mathbf{C}^2 \hat{\mathbf{H}}^{-1} \hat{\boldsymbol{b}}_m = (c_n^{D2}/\hat{h}_n^D)\delta_{nm}, \tag{G4}$$

where $\hat{h}_n^D(\kappa)$ is the normalization factor. The vertical structure of horizontal velocity is given by $\hat{\pi}_n^D = \hat{\boldsymbol{\pi}}^T \hat{\boldsymbol{a}}_n$ [see (48) and (49)], where

$$\hat{\boldsymbol{a}}_n = \hat{h}_n \hat{\mathbf{H}}^{-1} \hat{\boldsymbol{b}}_n. \tag{G5}$$

To derive a power-series solution, we focus on the pth mode and assume power-series expansion of the relevant variables in κ^2 around $\kappa^{\circ 2}$,

$$c_p^D = c_p^{(0)} + c_p^{(1)}\Delta\kappa^2 + c_p^{(2)}(\Delta\kappa^2)^2 + \cdots, \tag{G6a}$$

$$\hat{\boldsymbol{b}}_p = \hat{\boldsymbol{b}}_p^{(0)} + \hat{\boldsymbol{b}}_p^{(1)}\Delta\kappa^2 + \hat{\boldsymbol{b}}_p^{(2)}(\Delta\kappa^2)^2 + \cdots, \tag{G6b}$$

$$\hat{\boldsymbol{a}}_p = \hat{\boldsymbol{a}}_p^{(0)} + \hat{\boldsymbol{a}}_p^{(1)}\Delta\kappa^2 + \hat{\boldsymbol{a}}_p^{(2)}(\Delta\kappa^2)^2 + \cdots, \tag{G6c}$$

where $\Delta\kappa^2 \equiv \kappa^2 - \kappa^{\circ 2}$. The normalization factor $\hat{h}_n^D$ in general depends on κ but only $\hat{h}_n^{(0)} \equiv \hat{h}_n^D(\kappa^\circ)$ is required to derive the solution. Substituting (G6a) and (G6b) into (G2), and collecting like terms, we get, for $(\Delta\kappa^2)^k (k > 0)$,

$$\sum_{j=0}^{k} \chi_p^{(k-j;0)}\left(\hat{\mathbf{H}}^{-1} + \kappa^{\circ 2}\hat{\mathbf{D}}\right)\hat{\boldsymbol{b}}_p^{(j)} - \mathbf{C}^2 \hat{\mathbf{H}}^{-1}\hat{\boldsymbol{b}}_p^{(k)} = -\sum_{j=1}^{k} \chi_p^{(k-j;0)} \hat{\mathbf{D}} \hat{\boldsymbol{b}}_p^{(j-1)}, \tag{G7}$$

where

$$\chi_p^{(j;m)} = \sum_{l=m}^{j-m} c_p^{(l)} c_p^{(j-l)}. \tag{G8}$$

The result for $k = 0$ is (G2) with $\kappa = \kappa^\circ$, whose eigen solutions are $(c_n^{(0)}, \hat{\boldsymbol{b}}_n^{(0)})$, and $\hat{\boldsymbol{b}}_n^{(0)}$ satisfy the orthogonality (G4) with $\hat{h}_n^D = \hat{h}_n^{(0)}$. Then, we expand $\hat{\boldsymbol{b}}_p^{(k)}$ as

$$\hat{\boldsymbol{b}}_p^{(k)} = \hat{\boldsymbol{b}}_n^{(0)} \hat{\nu}_{np}^{(k)}. \tag{G9}$$

It is apparent that $\hat{\nu}_{np}^{(0)} = \delta_{pn}$. Multiplying (G7) by $\hat{\boldsymbol{b}}_n^{(0)T}$ from the left, substituting the above expansion, and using the orthogonality (G4), we get

$$2c_p^{(0)} c_p^{(k)} \delta_{pn} + (c_p^{(0)2} - c_n^{(0)2})\hat{\nu}_{np}^{(k)} = -\chi_p^{(k;1)}\delta_{pn} - \hat{h}_n^{(0)} \sum_{j=1}^{k} \chi_p^{(k-j;0)} \left(\hat{\boldsymbol{b}}_n^{(0)T} \hat{\mathbf{D}} \hat{\boldsymbol{b}}_m^{(0)}\right) \hat{\nu}_{mp}^{(j-1)} - \sum_{j=1}^{k-1} \chi_p^{(k-j;0)} \hat{\nu}_{np}^{(j)} \quad \text{(no sum on } p, n). \tag{G10}$$

For $k = 1$, it is assumed that the first and last terms on the right-hand-side are 0. Note that there are n_M sets of the above equations because n varies from 0 to $n_M - 1$. For any $k(\geq 1)$, the above algebraic recurrence relationships determine $c_p^{(k)}$ and $\hat{\nu}_{np}^{(k)}$ from the solutions up to $k - 1$. The recurrence relationships do not provide $\hat{\nu}_{pp}^{(k)}$, which affects the vertical structure but not $c_p^{(k)}$. The auxiliary condition to determine $\hat{\nu}_{pp}^{(k)}$ comes from (39). Substituting $\hat{\phi}_p^D = \hat{\boldsymbol{\phi}}^T \hat{\boldsymbol{b}}_p$, (G6b) and (G9) into (39), collecting like terms, and using the orthogonality (G4), we get $\hat{\nu}_{pp}^{(1)} = 0$ and

$$\hat{\nu}_{pp}^{(k)} = -\frac{\hat{h}_p}{2c_p^2} \sum_{j=1}^{k-1} \sum_{q \neq p} \hat{\nu}_{qp}^{(k-j)} \frac{c_q^2}{\hat{h}_q} \hat{\nu}_{qp}^{(j)} \quad \text{for } k > 1 \text{ (no sum on } p). \tag{G11}$$

For horizontal velocity, $\hat{\boldsymbol{a}}^{(k)}$ is calculated by applying (G5) to $\hat{\boldsymbol{b}}_p^{(k)}$ and $\hat{\boldsymbol{a}}_p^{(k)}$. Finally, calculating the power-series solution around $\kappa = 0$, the linear dispersion coefficients $\hat{\beta}^{(k)}$ are given by $\hat{\beta}^{(k)} = (-1)^k c_p^{(k)}$ and the correction functions $\hat{\phi}^{D(k)}$ and $\hat{\pi}^{D(k)}$ by $(-1)^k \hat{\boldsymbol{\phi}}^T \hat{\boldsymbol{b}}_p^{(k)}$ and $(-1)^k \hat{\boldsymbol{\pi}}^T \hat{\boldsymbol{a}}_p^{(k)}$, respectively. Solving the problem analytically up to the 2nd order yields the same $\hat{\beta}_p^{(1)}$, $\hat{\beta}_p^{(2)}$, $\hat{\phi}_p^{D(1)}$, and $\hat{\pi}_p^{D(1)}$ from the 2nd-order KdV theory in Sec. IV, (30b), (33b), (35b), and (35d).

APPENDIX H: POWER-SERIES SOLUTION FOR SIMPLE WAVES IN CONTINUOUSLY STRATIFIED FLUIDS

To derive power-series solution to (50), (51), and (53), we express the relevant variables as power series in $\tilde{R}_{+p}$ around $\tilde{R}_{+p}^\circ$,

$$c_r^N = c_r^{(0)} + \tilde{c}_r^{(1)}\Delta\tilde{R} + \tilde{c}_r^{(2)}\Delta\tilde{R}^2 + \cdots, \tag{H1a}$$

$$\tilde{\boldsymbol{l}}_r = \tilde{\boldsymbol{l}}_r^{(0)} + \tilde{\boldsymbol{l}}_r^{(1)}\Delta\tilde{R} + \tilde{\boldsymbol{l}}_r^{(2)}\Delta\tilde{R}^2 + \cdots, \tag{H1b}$$

$$\hat{\boldsymbol{v}} = \hat{\boldsymbol{v}}^{(0)} + \tilde{\boldsymbol{v}}^{(1)}\Delta\tilde{R} + \tilde{\boldsymbol{v}}^{(2)}\Delta\tilde{R}^2 + \cdots, \tag{H1c}$$

$$\hat{\mathbf{K}} = \hat{\mathbf{K}}^{(0)} + \tilde{\mathbf{K}}^{(1)}\Delta\tilde{R} + \tilde{\mathbf{K}}^{(2)}\Delta\tilde{R}^2 + \cdots, \tag{H1d}$$

where $\Delta\tilde{R} \equiv \tilde{R}_{+p} - \tilde{R}_{+p}^\circ$, $\hat{\boldsymbol{v}}^{(0)} \equiv \hat{\boldsymbol{v}}(\tilde{R}_{+p}^\circ)$, and $c_r^{(0)}$ and $\tilde{\boldsymbol{l}}_r^{(0)}$ are solutions to (50) with $\hat{\mathbf{K}}^{(0)} \equiv \hat{\mathbf{K}}(\hat{\boldsymbol{v}}^{(0)})$. Note that $\tilde{\mathbf{K}}^{(k)}$ depends only on $\tilde{\boldsymbol{v}}^{(k)}$ because $\hat{\mathbf{Y}}$ and $\hat{\mathbf{U}}$ are linear in $\hat{\boldsymbol{\eta}}$ and $\hat{\boldsymbol{u}}$, respectively [see (45)]. Collecting like terms, we get, for $k = 0$,

$$c_r^{(0)} \tilde{\boldsymbol{l}}_r^{(0)H} = \tilde{\boldsymbol{l}}_r^{(0)H} \hat{\mathbf{K}}^{(0)}, \tag{H2a}$$

$$\tilde{\mathbf{L}}^{(0)H} \tilde{\boldsymbol{v}}^{(1)} = \boldsymbol{I}_{+p}, \tag{H2b}$$

and for $k > 0$,

$$\tilde{c}_r^{(k)}\tilde{\boldsymbol{l}}_r^{(0)H} + \tilde{\boldsymbol{l}}_r^{(k)H}\left(c_r^{(0)}\mathbf{I} - \hat{\mathbf{K}}^{(0)}\right) = -\sum_{j=1}^{k-1}\tilde{c}_r^{(k-j)}\tilde{\boldsymbol{l}}_r^{(j)H} + \sum_{j=0}^{k-1}\tilde{\boldsymbol{l}}_r^{(j)H}\tilde{\mathbf{K}}^{(k-j)}, \tag{H3a}$$

$$(k+1)\tilde{\mathbf{L}}^{(0)H}\tilde{\boldsymbol{v}}^{(k+1)} = -\sum_{j=0}^{k-1}(j+1)\tilde{\mathbf{L}}^{(k-j)H}\tilde{\boldsymbol{v}}^{(j+1)}, \tag{H3b}$$

where $\mathbf{I}$ is the unit matrix. For $k = 1$, it is assumed that the first summation term on the right-hand-side (RHS) of (H3a) is 0. Note that no sum is taken on r in (H2a) and (H3a) because there are terms with single r on the RHS. Note also that $\tilde{\boldsymbol{l}}_r^{(k)}$ for all r are required to compute $\tilde{\mathbf{L}}^{(k)}$ from (51), and that the choice of the mode of interest enters the problem through the RHS of (H2b).

We solve this problem by introducing right eigenvectors $\tilde{\boldsymbol{r}}_s^{(0)}$ that satisfy

$$c_s^{(0)}\tilde{\boldsymbol{r}}_s^{(0)} = \hat{\mathbf{K}}^{(0)}\tilde{\boldsymbol{r}}_s^{(0)}. \tag{H4}$$

The left and right eigenvectors satisfy bi-orthogonality,

$$\tilde{\boldsymbol{l}}_r^{(0)H}\tilde{\boldsymbol{r}}_s^{(0)} = \delta_{rs}. \tag{H5}$$

Multiplying (H3a) by $\tilde{\boldsymbol{r}}_s^{(0)}$ from the right, expanding $\tilde{\boldsymbol{l}}_r^{(k)}$ as

$$\tilde{\boldsymbol{l}}_r^{(k)} = \tilde{\boldsymbol{l}}_s^{(0)}\tilde{\nu}_{sr}^{(k)}, \tag{H6}$$

noting $\tilde{\nu}_{sr}^{(0)} = \delta_{rs}$, and using the bi-orthogonality (H5), we get

$$\tilde{c}_r^{(k)}\delta_{rs} + (c_r^{(0)} - c_s^{(0)})\tilde{\nu}_{sr}^{(k)*} = -\sum_{j=1}^{k-1}\tilde{c}_r^{(k-j)}\tilde{\nu}_{sr}^{(j)*} + \sum_{j=0}^{k-1}\tilde{\nu}_{qr}^{(j)*}\left(\tilde{\boldsymbol{l}}_q^{(0)H}\tilde{\mathbf{K}}^{(k-j)}\tilde{\boldsymbol{r}}_s^{(0)}\right) \quad \text{(no sum on } r, s), \tag{H7}$$

where * denotes the complex conjugate. For any $k(\geq 1)$, the algebraic recurrence relationships (H2), (H3b), and (H7) determine $\tilde{c}_{+p}^{(k)}$ and $\tilde{\nu}_{sr}^{(k)}$ from the solutions up to $k - 1$. To express the resulting power series in $\hat{\eta}_p$, we need to invert $\hat{\eta}_p(\Delta\tilde{R})$ and substitute the resulting $\Delta\tilde{R}(\hat{\eta}_p)$ into (H1), yielding

$$c_r^N = c_r^{(0)} + \hat{c}_r^{(1)}\Delta\hat{\eta}_p + \hat{c}_r^{(2)}\Delta\hat{\eta}_p^2 + \cdots, \tag{H8a}$$

$$\hat{\boldsymbol{v}} = \hat{\boldsymbol{v}}^{(0)} + \hat{\boldsymbol{v}}^{(1)}\Delta\hat{\eta}_p + \hat{\boldsymbol{v}}^{(2)}\Delta\hat{\eta}_p^2 + \cdots, \tag{H8b}$$

where $\Delta\hat{\eta}_p \equiv \hat{\eta}_p - \hat{\eta}_p^\circ$ and $\hat{\eta}_p^\circ \equiv \hat{\eta}_p(\tilde{R}_{+p}^\circ)$. Calculating the power-series solution around $\hat{\eta}_p = 0$ assuming $\hat{\boldsymbol{v}}^{(0)} = \boldsymbol{O}$, the nonlinear coefficients $\hat{\alpha}^{(k)}$ are given by $\hat{\alpha}^{(k)} = \hat{c}_p^{(k)}$, and the correction functions $\hat{\phi}^{N(k)}$ and $\hat{\pi}^{N(k)}$ by $\hat{\boldsymbol{\phi}}^T\hat{\boldsymbol{\eta}}^{(k+1)}$ and $\hat{\boldsymbol{\pi}}^T\hat{\boldsymbol{u}}^{(k+1)}$, respectively.

Note that the recurrence relationships (H2), (H3b), and (H7) do not determine $\tilde{\nu}_{rr}^{(k)}$. However, it is important to note that the power-series solution (H8) has already been determined because $\tilde{\nu}_{rr}^{(k)}$ affects the solution only when $\Delta\tilde{R}$ appears explicitly.

Although numerical computation is required to calculate the power-series solution in general, we can derive analytic solution to (H2), (H3b), and (H7) around $\hat{\eta}_p = 0$ up to $k = 2$ as follows. To deal with weakly nonlinear solution, it is convenient to use the index for each vertical mode n but with positive and negative signs for the right- and left-propagating solutions, respectively. It is also convenient to work with block vectors and matrices with the sizes $n_M \times 1$ and $n_M \times n_M$, respectively. Writing $\tilde{\boldsymbol{l}}_{\pm n}^{(0)} = [\tilde{\boldsymbol{l}}_{\pm n}^{\eta T}, \tilde{\boldsymbol{l}}_{\pm n}^{uT}]^T$ and $\tilde{\boldsymbol{r}}_{\pm n}^{(0)} = [\tilde{\boldsymbol{r}}_{\pm n}^{\eta T}, \tilde{\boldsymbol{r}}_{\pm n}^{uT}]^T$, and normalizing the left eigenvectors such that the component corresponding $\hat{u}_p$ is 1, the solutions for $k = 0$ are

$$c_{\pm n}^{(0)} = \pm c_n, \tag{H9a}$$

$$(\tilde{\boldsymbol{l}}_{\pm n}^\eta)_m = \pm\frac{c_n}{\hat{h}_n}\delta_{mn}, \quad (\tilde{\boldsymbol{l}}_{\pm n}^u)_m = \delta_{mn}, \tag{H9b}$$

$$(\tilde{\boldsymbol{r}}_{\pm n}^\eta)_m = \pm\frac{\hat{h}_n}{2c_n}\delta_{mn}, \quad (\tilde{\boldsymbol{r}}_{\pm n}^u)_m = \frac{1}{2}\delta_{mn}, \tag{H9c}$$

$$\tilde{\eta}_p^{(1)} = \frac{\hat{h}_p}{2c_p}, \quad \tilde{u}_p^{(1)} = \frac{1}{2}, \quad \tilde{\eta}_q^{(1)} = \tilde{u}_q^{(1)} = 0, \tag{H9d}$$

and those for $k = 1$ are

$$\tilde{c}_{+p}^{(1)} = \frac{3\hat{N}_{ppp}^A}{4}, \tag{H10a}$$

$$\tilde{\nu}_{+n,+n}^{(1)} = \tilde{\nu}_{-n,-n}^{(1)} = 0, \tag{H10b}$$

$$\tilde{\nu}_{+n,-n}^{(1)} = \tilde{\nu}_{-n,+n}^{(1)} = \frac{\hat{N}_{nnp}^A}{8c_p}, \tag{H10c}$$

$$\tilde{\nu}_{\pm m,+n}^{(1)} = \frac{\hat{N}_{mnp}^A}{4(c_n \mp c_m)}\left(\pm\frac{c_n}{c_m} + \frac{c_n}{c_p} + 1\right)(m \neq n), \tag{H10d}$$

$$\tilde{\nu}_{\pm m,-n}^{(1)} = \frac{\hat{N}_{mnp}^A}{4(-c_n \mp c_m)}\left(\mp\frac{c_n}{c_m} - \frac{c_n}{c_p} + 1\right)(m \neq n), \tag{H10e}$$

$$\tilde{\eta}_p^{(2)} = \frac{\hat{N}_{ppp}^A\hat{h}_p}{32c_p^2}, \quad \tilde{u}_p^{(2)} = -\frac{\hat{N}_{ppp}^A}{32c_p}, \tag{H10f}$$

$$\tilde{\eta}_q^{(2)} = \frac{3\hat{N}_{qpp}^A\hat{h}_p}{8(c_p^2 - c_q^2)}, \quad \tilde{u}_q^{(2)} = \frac{\hat{N}_{qpp}^A\hat{h}_p}{8c_p\hat{h}_q}\frac{c_p^2 + 2c_q^2}{c_p^2 - c_q^2}, \tag{H10g}$$

where $q \neq p$ and no sum is taken on p because it is the single mode of interest. The solution for the phase speed for $k = 2$ is

$$\tilde{c}_{+p}^{(2)} = -\frac{3(\hat{N}_{ppp}^A)^2}{64c_p} + \frac{3\hat{h}_p}{16c_p}\sum_{q\neq p}\frac{(\hat{N}_{ppq}^A)^2}{\hat{h}_q}\frac{5c_p^2 + 4c_q^2}{c_p^2 - c_q^2}. \tag{H11}$$

The above solution is written as a power series in $\Delta\tilde{R}$. To derive the solution in $\hat{\eta}_p$, we need to invert $\hat{\eta}_p(\Delta\tilde{R})$. Using the series reversion formula (e.g., Sec. 5.7 in Ref. 21 or Sec. 3.6 in Ref. 48), we get

$$\Delta\tilde{R} \sim \frac{2c_p}{\hat{h}_p}\hat{\eta}_p - \frac{\hat{N}_{ppp}^A c_p}{4\hat{h}_p^2}\hat{\eta}_p^2. \tag{H12}$$

Substituting (H9)–(H12) into (H1a) and (H1c) yields (30a), (33a), (35a), and (35c). Note that, if $\tilde{\boldsymbol{l}}_{\pm n}^{(0)}$ is multiplied by an arbitrary scaling factor $a_{\pm n}$, the solution in $\tilde{R}_{+p}$ changes because $\tilde{\boldsymbol{r}}_{\pm n}^{(0)}$ and $\tilde{R}_{+p}$ are multiplied by $a_{\pm n}^{-1}$ and a_{+p}, respectively. However, the power-series solution in $\hat{\eta}_p$ is independent of the arbitrary scaling factors hence requires no auxiliary condition.

[1] A. R. Osborne and T. L. Burch, "Internal solitons in the Andaman Sea," Science **208**, 451–460 (1980).

[2] P. E. Holloway, E. Pelinovsky, T. Talipova, and B. Barnes, "A nonlinear model of internal tide transformation on the Australian North West Shelf," J. Phys. Oceanogr. **27**, 871–896 (1997).

[3] C. Jackson, "Internal wave detection using the moderate resolution imaging spectroradiometer (MODIS)," J. Geophys. Res. **112**, C11012, doi:10.1029/2007JC004220 (2007).

[4]D. Farmer and L. Armi, "The generation and trapping of solitary waves over topography," Science **283**, 188–190 (1999).
[5]J. Pan and D. A. Jay, "Dynamic characteristics and horizontal transports of internal solitons generated at the Columbia River plume front," Cont. Shelf Res. **29**, 252–262 (2009).
[6]S. A. Thorpe, "Asymmetry of internal seiche in Loch Ness," Nature **231**, 306–308 (1971).
[7]L. Boegman, J. Imberger, G. N. Ivey, and J. P. Antenucci, "High-frequency internal waves in large stratified lakes," Limnol. Oceanogr. **48**, 895–919 (2003).
[8]M. Preusse, H. Freistühler, and F. Peeters, "Seasonal variation of solitary wave properties in Lake Constance," J. Geophys. Res. **117**, C04026, doi: 10.1029/2011JC007403 (2012).
[9]R. Grimshaw, "Korteweg-de Vries equation," in *Nonlinear Waves in Fluids: Recent Advances and Modern Applications*, edited by R. Grimshaw (Springer, 2005), pp. 1–28.
[10]E. Pelinovsky, O. Polukhina, A. Slunyaev, and T. Talipova, "Internal solitary waves," in *Solitary Waves in Fluids*, edited by R. H. J. Grimshaw (WIT Press, 2007), pp. 85–110.
[11]T. Kakutani and N. Yamasaki, "Solitary waves on a two-layer fluid," J. Phys. Soc. Jpn. **45**, 674–679 (1978).
[12]K. R. Helfrich and W. K. Melville, "Long nonlinear internal waves," Annu. Rev. Fluid Mech. **38**, 395–425 (2006).
[13]K. G. Lamb and L. Yan, "The evolution of internal wave undular bores: Comparisons of a fully nonlinear numerical model with weakly nonlinear theory," J. Phys. Oceanogr. **26**, 2712–2734 (1996).
[14]R. Grimshaw, E. Pelinovsky, and T. Talipova, "Solitary wave transformation in a medium with sign-variable quadratic nonlinearity and cubic nonlinearity," Phys. D **132**, 40–62 (1999).
[15]C.-Y. Lee and R. Beardsley, "The generation of long nonlinear internal waves in a weakly stratified shear flow," J. Geophys. Res. **79**, 453–462, doi: 10.1029/jc079i003p00453 (1974).
[16]P. E. Holloway, E. Pelinovsky, and T. Talipova, "A generalized Korteweg-de Vries model of internal tide transformation in the coastal zone," J. Geophys. Res. **104**, 18333–18350, doi:10.1029/1999jc900144 (1999).
[17]E. N. Pelinovskii, O. E. Polukhina, and K. Lamb, "Nonlinear internal waves in the ocean stratified in density and current," Oceanology **40**, 757–766 (2000).
[18]R. Grimshaw, E. Pelinovsky, and O. Poloukhina, "Higher-order Korteweg-de Vries models for internal solitary waves in a stratified shear flow with a free surface," Nonlinear Processes Geophys. **9**, 221–235 (2002).
[19]D. J. Benney, "Long nonlinear waves in fluid flows," J. Math. Phys. **45**, 52–63 (1966).
[20]G. B. Whitham, *Linear and Nonlinear Waves* (Wiley, 1974).
[21]G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists*, 6th ed. (Elsevier, 2005).
[22]W. Choi and R. Camassa, "Fully nonlinear internal waves in a two-fluid system," J. Fluid Mech. **396**, 1–36 (1999).
[23]A. E. Gill, *Atmosphere-Ocean Dynamics* (Academic Press, 1982).
[24]S. G. Llewellyn-Smith and W. R. Young, "Conversion of the barotropic tide," J. Phys. Oceanogr. **32**, 1554–1566 (2002).
[25]R. Maugé and T. Gerkema, "Generation of weakly nonlinear nonhydrostatic internal tides over large topography: A multi-modal approach," Nonlinear Processes Geophys. **15**, 233–244 (2008).
[26]S. D. Griffiths and R. H. J. Grimshaw, "Internal tide generation at the continental shelf modeled using a modal decomposition: Two-dimensional results," J. Phys. Oceanogr. **37**, 428–451 (2007).
[27]K. Shimizu, "A theory of vertical modes in multilayer stratified fluids," J. Phys. Oceanogr. **41**, 1694–1707 (2011).
[28]S. M. Kelly, J. D. Nash, K. I. Martini, M. H. Alford, and E. Kunze, "The cascade of tidal energy from low to high modes on a continental slope," J. Phys. Oceanogr. **42**, 1217–1232 (2012).
[29]R. Long, "Some aspects of the flow of stratified fluids. Part I. A theoretical investigation," Tellus **5**, 42–58 (1953).
[30]B. Benjamin, "Internal waves of finite amplitude and permanent form," J. Fluid Mech. **25**, 241–270 (1966).
[31]J. W. Miles, "On solitary waves," Tellus **31**, 456–462 (1979).
[32]J. A. Gear and R. Grimshaw, "A second-order theory for solitary waves in shallow fluids," Phys. Fluids **26**, 14–29 (1983).
[33]K. Lamb, "Theoretical descriptions of shallow-water solitary internal waves: Comparisons with fully nonlinear waves," in *The 1998 WHOI/IOS/ONR Internal Solitary Wave Workshop: Contributed Papers*, edited by T. Duda and D. Farmer (Woods Hole Oceanographic Institution, 1999), pp. 209–217.
[34]A. Kasahara, "Various vertical coordinate systems used for numerical weather prediction," Mon. Weather Rev. **102**, 509–522 (1974).
[35]K. G. Lamb and B. Wan, "Conjugate flows and flat solitary waves for a continuously stratified fluid," Phys. Fluids **10**, 2061–2079 (1998).
[36]V. D. Djordjevic and L. G. Redekopp, "The fission and disintegration of internal solitary waves moving over two-dimensional topography," J. Phys. Oceanogr. **8**, 1016–1024 (1978).
[37]C. G. Koop and G. Butler, "An investigation of internal solitary waves in a two-fluid system," J. Fluid Mech. **112**, 225–251 (1981).
[38]J. Lighthill, *Waves in Fluids* (Cambridge University Press, 1978).
[39]K. Shimizu, J. Imberger, and M. Kumagai, "Horizontal structure and excitation of primary motions in a strongly stratified lake," Limonol. Oceanogr. **52**, 2641–2655 (2007).
[40]A. de la Fuente, K. Shimizu, Y. Niño, and J. Imberger, "Nonlinear and weakly nonhydrostatic inviscid evolution of internal gravitational basin-scale waves in a large, deep lake: Lake Constance," J. Geophys. Res. **115**, C12045, doi:10.1029/2009JC005839 (2010).
[41]L. A. Ostrovsky and J. Grue, "Evolution equations for strongly nonlinear internal waves," Phys. Fluids **15**, 2934–2948 (2003).
[42]J. W. Miles, "Solitary waves," Annu. Rev. Fluid Mech. **12**, 11–43 (1980).
[43]M. van Dyke, *Perturbation Methods in Fluid Mechanics* (Parabolic Press, 1975).
[44]M. Miyata, "An internal solitary wave of large amplitude," La Mer **23**, 43–48 (1985).
[45]C. H. Su and C. S. Gardner, "Korteweg-de Vries equation and generalization. III: Derivation of the Korteweg-de Vries equation and Burgers equation," J. Math. Phys. **10**, 536–539 (1969).
[46]A. E. Green and P. M. Naghdi, "A derivation of equations for wave propagation in water of variable depth," J. Fluid Mech. **78**, 237–246 (1976).
[47]K. Shimizu, "Application of modal analysis to strongly stratified lakes," Ph.D. thesis, University of Western Australia, Crawley, Western Australia, Australia, 2009.
[48]M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables*, Applied Mathematics Series Vol. 55 (National Bureau of Standards, Washington, D.C., 1972).