Commun. Math. Phys. 370, 1–48 (2019)
Digital Object Identifier (DOI) https://doi.org/10.1007/s00220-019-03483-8

Communications in
Mathematical Physics

Equatorial Wave–Current Interactions

A. Constantin[1], **R. I. Ivanov**[2]

[1] Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria. E-mail: adrian.constantin@univie.ac.at
[2] School of Mathematical Sciences, Technological University Dublin, City Campus, Kevin Street, Dublin D08 NF82, Ireland. E-mail: rossen.ivanov@dit.ie

Received: 21 September 2017 / Accepted: 16 April 2019
Published online: 25 June 2019 – © The Author(s) 2019

Abstract: We study the nonlinear equations of motion for equatorial wave–current interactions in the physically realistic setting of azimuthal two-dimensional inviscid flows with piecewise constant vorticity in a two-layer fluid with a flat bed and a free surface. We derive a Hamiltonian formulation for the nonlinear governing equations that is adequate for structure-preserving perturbations, at the linear and at the nonlinear level. Linear theory reveals some important features of the dynamics, highlighting differences between the short- and long-wave regimes. The fact that ocean energy is concentrated in the long-wave propagation modes motivates the pursuit of in-depth nonlinear analysis in the long-wave regime. In particular, specific weakly nonlinear long-wave regimes capture the wave-breaking phenomenon while others are structure-enhancing since therein the dynamics is described by an integrable Hamiltonian system whose solitary-wave solutions are solitons.

Contents

1. Introduction

The Earth's rotation profoundly affects the dynamics of the atmosphere and of the ocean. However, an in-depth qualitative study of the governing equations for rotating fluids is analytically untractable and prohibitively expensive computationally. For this reason, it is imperative to derive simpler, manageable models which can be studied in detail. A significant feature of equatorial ocean dynamics is that the change of sign of the Coriolis force across the Equator produces an effective waveguide, with the Equator acting as a (fictitious) natural boundary that facilitates azimuthal flow propagation which is symmetric with respect to the Equator: wave–current interactions which propagate in the longitudinal direction and are symmetric with respect to the Equator, featuring so-called Kelvin-waves (large-scale wave motions affected by the Earth's rotation and trapped at the Equator) and a depth-dependent underlying current field. Moreover, due to the pronounced density stratification of the equatorial regions, greater than anywhere else in the ocean, a rather sharp interface—the thermocline—separates a shallow near-surface layer of relatively warm water from a deep layer of colder and denser water, each layer being practically of constant density. Consequently, there are two types of equatorial Kelvin waves: surface and internal waves. These two types of waves have quite different spatial and temporal scales, the internal waves being significantly larger and slower, with much longer wavelengths. Let us also point out that, in the ocean, energy tends to be concentrated in the lower frequencies (see [16]). Internal equatorial Kelvin waves are slow (taking more than two months to cross the equatorial Pacific) and have very long wavelengths (measured in tens/hundreds of kilometres), being typically more strongly excited than other ocean waves. They play an essential role in the "El Niño"-phenomenon (see [22]): a spectacular, naturally occurring anomalous appearance, every few years, of unusually warm water in the central Pacific, transported eastwards by the internal Kelvin waves until a large expanse of the equatorial Pacific becomes much warmer than the average, thus altering the usual heat exchange process with the air above it and resulting in dramatic shifts of weather patterns in a chain reaction around the world. Moreover, the fact that a Kelvin wave is uni-directional facilitates the growth and accumulation of nonlinear perturbations as the wave propagates, so that nonlinear effects can become very large even if the initial amplitude is relatively small—it is not uncommon to record internal wave heights in excess of 40 m. Nonlinearity causes the often observed distortion of some internal equatorial Kelvin waves: the leading edge steepens while the trailing edge becomes flatter. These physical motivations prompted an intensive research activity, with the available studies of equatorial ocean waves falling into one of the following categories:

(I) observational data and/or numerical simulations (see the discussion in [20]);
(II) linear wave perturbations of a depth-dependent underlying current (see the discussion in [7]);
(III) nonlinear studies assuming a passive equatorial current field (see the discussion in [29]).

Category (I) studies convey valuable information but, given the complexity of the encountered flows, by itself it does not suffice to identify clearly the important processes that are at work. Concerning the studies in category (II), the common occurence of nonlinear equatorial phenomena provides the impetus to go beyond the linear theory. As for (III), the importance of the interactions between waves and currents is highlighted by the fact that the underlying current field in the equatorial Pacific, generated by the prevailing westerly ambient wind pattern and the forces created by the Earth's rotation (see the discussion in [3,8]), presents flow-reversal: an eastward jet whose core resides in the thermocline (the Equatorial Undercurrent) is sandwiched between a westward surface flow and an abyssal layer of practically still water. The fact that a flow with piecewise constant vorticity (negative above the thermocline and positive below it) captures this salient feature within a Hamiltonian framework (see Sect. 3) opens up the possibility to develop a rigorous in-depth study of nonlinear wave–current interactions.

A detailed description of the salient physical features of the ocean flow in the equatorial Pacific and of the precise aims of our mathematical analysis is provided in Sect. 2. The analysis of the Hamiltonian structure of the governing equations is pursued in Sect. 3. In Sect. 4 we develop a Hamiltonian perturbation theory. The study of the linearised equations, performed in Sect. 4.1, reveals acute differences between the two most important physical regimes for wave propagation:

- Internal linear short waves propagate slowly eastwards, while the short waves at the surface are fast and can propagate eastwards as well as westwards, with each propagation mode having no noticeable effect at the other interface.
- The effects of linear long-wave propagation are practically confined to the motion of the thermocline, whose oscillations propagate slowly eastwards or westwards. The fact that the westward internal waves are slower is an outcome of the dynamic response of the ocean to the presence of strong underlying currents with flow reversal.

However, there are two major limitations of linear theory. Firstly, the ubiquitous wave-breaking phenomenon is outside the realm of linear waves. Moreover, in the context of equatorial wave–current interactions, the common occurrence of internal solitary waves—localised waves that maintain their coherence, propagating with constant speed and unaltered shape—is not captured by linear theory. In light of this, and given that most ocean energy is concentrated in the long waves, in Sect. 4.2 we lay out the foundations for carrying out a systematic weakly nonlinear analysis in the principal long-wave scaling regimes of geophysical relevance (see Table 1). Of particular interest is the derivation of model equations that capture the most striking geophysical manifestations of nonlinear phenomena:

- Internal solitary waves of permanent form which owe their existence to a balance between wave-steepening effects and wave dispersion—these wave patterns are important because they are often large, very energetic events, and they have therefore a significant role in mass and momentum transport across the ocean. Actually, the corresponding model equation is not merely Hamiltonian, it is completely integrable and therefore the solitary waves present the enhanced structure of a soliton—solitary waves that have an elastic scattering property (after colliding with each other, they eventually emerge unscathed, retaining their shape and speed) and present remarkable stability properties. This opens up the possibility of a detailed analysis of the nonlinear interaction of several such wave patterns by means of an inverse scattering approach, based on an appropriate Riemann–Hilbert problem formulation.

Table 1. Overview of the considered long-wave scaling regimes, expressed in terms of the waves (a similar scaling is performed on the tangential velocities at the two interfaces), and of the main qualitative features that can be accommodated within a specific regime

Long-wave regime	Spatio-temporal scaling of the waves at the two interfaces (internal/surface)	Main qualitative feature of the wave–current interaction
Linear	$\eta(x,t) = \varepsilon\,\eta'(x,t)$, $\eta_1(x,t) = \varepsilon\,\eta_1'(x,t)$, with $\varepsilon \ll 1$, $(x,\ t,\ \eta',\ \eta_1') = \mathrm{O}(1)$	Dispersive character forestalls solitary waves and wave breaking
KdV	$\eta(x,t) = \varepsilon^2\,\eta'(\varepsilon x,t)$, $\eta_1(x,t) = \delta\varepsilon^2\,\eta_1'(\varepsilon x,t/\delta)$, with $\varepsilon \ll 1, \delta = o(\varepsilon^2)$, $(\varepsilon x, t, \eta', \eta_1') = \mathrm{O}(1)$	Dispersion/nonlinearity balance accommodates internal solitons
Inviscid Burgers	$\eta(x,t) = \varepsilon\,\eta'(\varepsilon x,t)$, $\eta_1(x,t) = \delta\varepsilon\,\eta_1'(\varepsilon x,t/\delta)$, with $\varepsilon \ll 1, \delta = \mathrm{O}(\varepsilon^2)$, $(\varepsilon x, t, \eta', \eta_1') = \mathrm{O}(1)$	Strong nonlinearity accommodates internal wave breaking

- Large-amplitude internal waves that break—other than their widespread occurrence, they are perceived as important for the turbulent mixing their death throes produce. This type of solutions correspond to a model similar to the classical nonlinear shallow water equations, modified to account for the presence of underlying currents in a stratified flow. We use the method of characteristics to provide insight into the fascinating process of wave breaking.

In Sect. 5 we overview the obtained results and we offer a perspective on promising future directions for related research.

We conclude this introduction by pointing out that, while we made a conscious effort to avoid being pedantic, there are some aspects of our investigation where a cavalier disregard for proper mathematical rigour is counter-productive. These are best illustrated by some conceptual and technical challenges that contrast to the situation encountered in the periodic setting, being specific for localised wave–current interactions. A central issue is the fact that the Hamiltonian cannot be the total energy of the flow since the kinetic energy contribution of the current is already infinite. Instead, we have to filter out the contribution from the current field and prove that this is achievable at the nonlinear level. Also, defining canonical variables of Schwartz class in the presence of depth-dependent underlying currents is not a foregone result; to the best of our knowledge, this problem has not been addressed in the research literature.

2. Preliminaries

The aim of this section is to briefly discuss the observational record as a physical background for the systematic mathematical approach developed in this paper.

2.1. Key features of equatorial ocean dynamics. Within a band of about 100–150 km of the Equator and extending longitudinally over about 16,000 km, the Pacific Ocean possesses some remarkable features: a significant density stratification (that is greater

than anywhere else in the ocean, see [14]), an underlying current field with flow-reversal, and a wide variety of observed wave propagation phenomena. The hallmark of the pronounced density stratification is the presence of a rather sharp thermocline that separates a shallow near-surface layer of relatively warm water from a deep layer of colder and denser water; the assumption that each layer is of constant density is reasonable (see [31]). Within a band of width about 300 km, centred on the Equator, there is, confined to a depth of about 100 m, a westward current that is driven by the prevailing trade winds, below which lies the Equatorial Undercurrent (EUC)—an eastward jet whose core resides on the thermocline. Below the EUC the flow dies out rapidly so that, at depths in excess of about 500 m there is an abyssal region of almost still water (see [32]). This equatorial background state interacts with various oceanic wave propagation modes, including long waves (with wavelengths exceeding 50 km; see [16]) and short waves (with wavelengths of a few hundreds of metres; see [34]). A comprehensive model of this wave–current interaction must accommodate the density stratification as well as the coupling between the waves at the surface and on the thermocline. Observations provide evidence for highly nonlinear regimes of internal wave motion. While explicit nonlinear solutions can be obtained, they do not cope satisfactorily with the complexity of the equatorial flow due to the limitations on the permissible underlying currents (see the discussion in [7]). For this reason, it is necessary to perform approximations that capture the relevant dynamics. The available approaches rely on linear theory, whether they resort to a numerical treatment by means of finite differences (see [32,39]) or to the method of multiple scales (see [7]). We aim to develop an approach that captures nonlinear effects which, in particular, can explain the propagation of observed equatorial solitary-like waves (see [38]). These are of special interest because they are completely missed by linear theory (see Sect. 4.1.3) and, as pointed out in [2], they are much more easily observable than nonlinear perturbations of an oscillatory wavetrain. Note also that geophysical fluid-dynamical considerations show that the Reynolds number is extremely large (see [30]), so that nonlinear effects typically dominate over viscosity. Field evidence for critical levels (locations where the wave speed equals the mean-flow speed) in the near-surface layer above the thermocline, where the flow-reversal of the underlying current occurs is available (see [34]). It is therefore advisable to rely on the nonlinear alternative to the conventional linear viscous boundary-layer approach for the description of the flow in the neighbourhood of critical levels, where Kelvin 'cat's eye' flow patterns appear (see the discussion in [9] for the somewhat simpler case of gravity water flows with constant non-zero vorticity).

2.2. Basic modelling assumptions. A few realistic simplifying assumptions can be made. Firstly, since the Reynolds numbers are typically very large in geophysical ocean flows, it is reasonable to employ inviscid theory (see [30]). Secondly, because we are considering flows in the neighbourhood of the Equator, it is adequate to use the f-plane approximation in the governing equations (see [26]). Thirdly, since field data shows that the meridional velocities are much smaller than the zonal velocities at the Equator (see [20]), we study flow configurations which are latitude-independent and with a vanishing meridional velocity component. Consequently, we investigate two-dimensional inviscid flows which present no variation in the meridional direction, regarding them as wave–current interactions due to localised wave perturbations of a pure current background state. The presence of underlying depth-dependent currents places us within the framework of flows with non-zero vorticity. Since for wave–current interactions in which the waves are long compared with the mean depth of the effective flow region the importance

of a non-zero mean vorticity preponderates that of its specific distribution (see [13]), the simplest realistic setting is that of flows with constant vorticity above and below the thermocline: negative above, to permit a reversal from the surface westward wind-drift to the eastward-flowing subsurface EUC, and positive below to model a flow that withers with increasing depth. This choice is propitious not just because in two-dimensional flows the vorticity of a particle remains constant as the particle moves about, as one can easily check. Remarkably, for two-dimensional stratified flows with constant vorticity in each layer, a separation of the flow into a pure current and an irrotational wave perturbation can be performed within the framework of the fully nonlinear theory, without recourse to approximations (see Sect. 3.3). This feature permits us to view the nonlinear wave–current interaction as an irrotational wave perturbation of a mean flow representing a pure current.

2.3. The Hamiltonian perspective. The ability to pursue in-depth nonlinear studies is contingent upon a rich structure. Since dissipation is not important, this motivates the quest for a Hamiltonian formulation. The Hamiltonian perspective results in a significant simplification, serving as a guide to the choice of new dependent and independent variables in which the equations take their simplest form. Note that the Hamiltonian formulation of the governing equations for two-dimensional gravity water flows was pioneered in [46] for irrotational flows, and was extended to rotational flows with constant vorticity in [43,44]. A Hamiltonian approach to two-dimensional two-layer gravity water flows with a free surface was developed in [10], and recently the rotational counterpart, with constant vorticity in each layer, was obtained in [6] for periodic flows. This paper establishes the validity of a Hamiltonian perspective in the presence of Coriolis effects in the equatorial f-plane approximation, for zonal flows with no variation in the meridional direction which represent localised perturbations of an underlying pure current background state. A suitable nondimensionalisation reveals that in certain oceanographically relevant regimes the geophysical effects are not merely a small perturbation of the governing equations for gravity water waves: scaling ascertains the relative importance of the different components of the flow, and in this geophysical regime the (linear) contribution due to the Earth's rotation to the longitudinal momentum equation balances the linear part of the material derivative, both being of the same order. We will also show that the Hamiltonian framework is adequate for structure-preserving perturbations. In particular, a specific weakly nonlinear long-wave regime turns out to be structure-enhancing since in this setting the dynamics is described by an integrable Hamiltonian system (with infinitely many degrees of freedom). In other regimes one can derive nonlinear models that capture wave breaking.

3. The Governing Equations

The objective of this section is to present the nonlinear governing equations and to elucidate the basic structure of the associated mean flow. Also, by specifying the associated geophysical scales we can introduce a set of non-dimensional variables which are useful for ascertaining the relative importance of the different components of the flow.

The fundamental model is that of an inviscid two-layer fluid which admits non-zero vorticity. Using the over-bar to represent the physical variables, we choose a coordinate system with its origin at a point on the Earth's surface, with the $\overline{x}$-axis horizontally due East, the $\overline{y}$-axis horizontally due North and the $\overline{z}$-axis upwards (see Fig. 1). Because

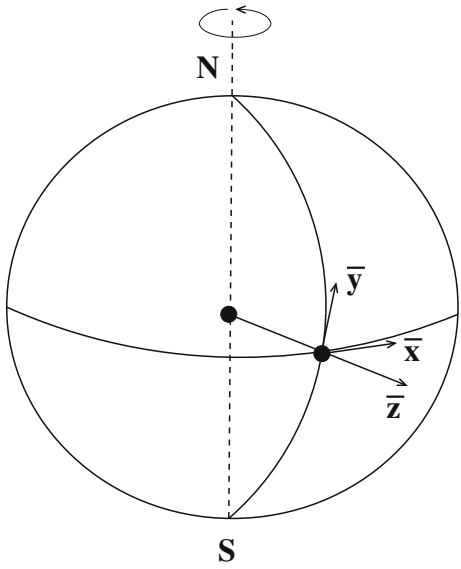

Fig. 1. The rotating frame of reference, with the $\bar{x}$-axis chosen horizontally due east, the $\bar{y}$-axis horizontally due north and the $\bar{z}$-axis upward

we are considering flows in the neighbourhood of the Equator, we use the f-plane approximation (see [26]). Since observations show that the meridional velocities are much smaller than the zonal velocities at the Equator (see [20]) then, by assumption, our flow configuration is latitude-independent with vanishing meridional velocity component, and so we restrict the domain to be at a fixed latitude (see Fig. 2). Let $\overline{z} = \overline{h}_1 + \overline{\eta}_1(\overline{x},\, \overline{t})$ be the free surface, $\overline{z} = \overline{\eta}(\overline{x},\, \overline{t})$ be the thermocline and $\overline{z} = -\overline{h}$ be the flat bed, where $\overline{t}$ stands for time. Here $\overline{h}_1$ and $\overline{h}$ are the mean depths of the near-surface layer and of the abyssal layer, with typical values about 120–250 m and 4 km, respectively (see [14]). The density of the fluid above the thermocline is constant, $\overline{\rho}_1 \approx 10^3$ kg m^{-3}, and is replaced by $\overline{\rho} = \overline{\rho}_1\,(1+r)$ for the fluid below the thermocline, where r is a small positive constant; typically r is about 10^{-3} (see [22]).

Denoting by $(\overline{u}_1(\overline{x},\overline{z},\overline{t}),\ \overline{v}_1(\overline{x},\overline{z},\overline{t}))$ and $(\overline{u}(\overline{x},\overline{z},\overline{t}),\ \overline{v}(\overline{x},\overline{z},\overline{t}))$ the velocity fields in the layers

$$\overline{\mathcal{D}}_1(t) = \{(\overline{x},\overline{z}) : \overline{\eta}(\overline{x},\, \overline{t}) < \overline{z} < \overline{h}_1 + \overline{\eta}_1(\overline{x},\, \overline{t})\}, \quad \overline{\mathcal{D}}(t) = \{(\overline{x},\overline{z}) : -\overline{h} < \overline{z} < \overline{\eta}(\overline{x},\, \overline{t})\},$$

above and below the thermocline, respectively, the equations of motion are the suitably adjusted Euler equations

$$\begin{cases} \overline{u}_{1,\overline{t}} + \overline{u}_1\,\overline{u}_{1,\overline{x}} + \overline{w}_1\,\overline{u}_{1,\overline{z}} + 2\overline{\Omega}\,\overline{w}_1 = -\dfrac{1}{\overline{\rho}_1}\,\overline{P}_{\overline{x}}, \\ \overline{w}_{1,\overline{t}} + \overline{u}_1\,\overline{w}_{1,\overline{x}} + \overline{w}_1\,\overline{w}_{1,\overline{z}} - 2\overline{\Omega}\,\overline{u}_1 = -\dfrac{1}{\overline{\rho}_1}\,\overline{P}_{\overline{z}} - \overline{g}, \end{cases} \quad \text{in} \quad \overline{\mathcal{D}}_1(t), \qquad (3.1)$$

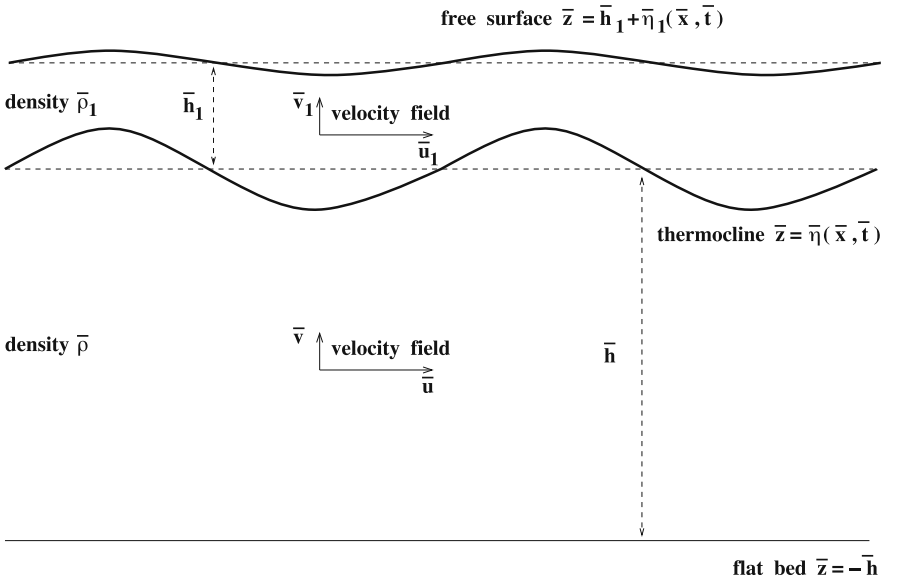

Fig. 2. Sketch of the cross-section of the fluid domain at a fixed latitude: the thermocline $\overline{z} = \overline{\eta}(\overline{x}, \overline{t})$ separates the two layers of different constant densities, the lower boundary is a flat rigid bed, $\overline{z} = -\overline{h}$, while the upper boundary is a free surface of elevation $\overline{z} = \overline{h}_1 + \overline{\eta}_1(\overline{x}, \overline{t})$. The surface and internal waves are coupled, with the amplitude of the oscillations of the thermocline typically considerably larger

$$\begin{cases} \overline{u}_{\overline{t}} + \overline{u}\,\overline{u}_{\overline{x}} + \overline{w}\,\overline{u}_{\overline{z}} + 2\overline{\Omega}\,\overline{w} = -\dfrac{1}{\overline{\rho}}\,\overline{P}_{\overline{x}}, \\ \\ \overline{w}_{\overline{t}} + \overline{u}\,\overline{w}_{\overline{x}} + \overline{w}\,\overline{w}_{\overline{z}} - 2\overline{\Omega}\,\overline{u} = -\dfrac{1}{\overline{\rho}}\,\overline{P}_{\overline{z}} - \overline{g}, \end{cases} \quad \text{in} \quad \overline{\mathcal{D}}(t), \tag{3.2}$$

where $\overline{P} = \overline{P}(\overline{x}, \overline{z}, \overline{t})$ is the pressure, $\overline{g} \approx 9.8\ \mathrm{m\,s^{-2}}$ is the (constant) acceleration of gravity and $\overline{\Omega} \approx 7.29 \times 10^{-5}\ \mathrm{rad\,s^{-1}}$ is the (constant) rotational speed of the Earth about the polar axis (towards the East), supplemented by the equations of mass conservation,

$$\overline{u}_{1,\overline{x}} + \overline{w}_{1,\overline{z}} = 0 \quad \text{in} \quad \overline{\mathcal{D}}_1(t), \tag{3.3}$$

$$\overline{u}_{\overline{x}} + \overline{w}_{\overline{z}} = 0 \quad \text{in} \quad \overline{\mathcal{D}}(t). \tag{3.4}$$

The vorticity distribution is specified by

$$\overline{u}_{1,\overline{z}} - \overline{w}_{1,\overline{x}} = \overline{\gamma}_1 \quad \text{in} \quad \overline{\mathcal{D}}_1(t), \tag{3.5}$$

$$\overline{u}_{\overline{z}} - \overline{w}_{\overline{x}} = \overline{\gamma} \quad \text{in} \quad \overline{\mathcal{D}}(t), \tag{3.6}$$

for suitable (physical) constants $\overline{\gamma}_1$ and $\overline{\gamma}$. The appropriate boundary conditions are, at the free surface, the dynamic and kinematic boundary conditions

$$\overline{P} = \overline{P}_{atm} \quad \text{on} \quad \overline{z} = \overline{h}_1 + \overline{\eta}_1(\overline{x}, \overline{t}), \tag{3.7}$$

$$\overline{w}_1 = \overline{\eta}_{1,\overline{t}} + \overline{u}_1\,\overline{\eta}_{1,\overline{x}} \quad \text{on} \quad \overline{z} = \overline{h}_1 + \overline{\eta}_1(\overline{x}, \overline{t}), \tag{3.8}$$

respectively, where $\overline{P}_{atm}$ is the constant pressure of the atmosphere at the surface of the ocean. At the thermocline we have the kinematic boundary conditions

$$\overline{w}_1 = \overline{\eta}_{\overline{t}} + \overline{u}_1\,\overline{\eta}_{\overline{x}} \quad \text{on} \quad \overline{z} = \overline{\eta}(\overline{x},\,\overline{t}), \tag{3.9}$$

$$\overline{w} = \overline{\eta}_{\overline{t}} + \overline{u}\,\overline{\eta}_{\overline{x}} \quad \text{on} \quad \overline{z} = \overline{\eta}(\overline{x},\,\overline{t}), \tag{3.10}$$

together with the requirement that the pressure is continuous across the thermocline,

$$\overline{P}_+ = \overline{P}_- \quad \text{on} \quad \overline{z} = \overline{\eta}(\overline{x},\overline{t}), \tag{3.11}$$

where $\overline{P}_\pm$ refer to the limits of the pressure $\overline{P}$ from above and below the interface. Finally, at the fixed, horizontal, impermeable bottom, we have the kinematic boundary condition

$$\overline{w} = 0 \quad \text{on} \quad \overline{z} = -\overline{h}. \tag{3.12}$$

Note that (3.9)–(3.10) ensure the continuity of the normal component of the fluid velocity across the thermocline $\overline{z} = \overline{\eta}(\overline{x},\overline{t})$. Since within the inviscid setting the stress at an interface has no tangential component, due to the absence of any interaction derived from friction at the interface, in principle it is permissible for the tangential component of the fluid velocity to present discontinuities at the thermocline. Nevertheless, the available field data for equatorial flows suggests a continuous transition between the two layers (velocity discontinuities across the thermocline would correspond to a delta-sheet of vorticity—see also the discussion at the end of Sect. 3.3), so that we also require a tangential velocity balance:

$$\overline{w}_1\overline{\eta}_{\overline{x}} + \overline{u}_1 = \overline{w}\,\overline{\eta}_{\overline{x}} + \overline{u} \quad \text{on} \quad \overline{z} = \overline{\eta}(\overline{x},\overline{t}). \tag{3.13}$$

Consequently, the velocity field is continuous across the thermocline. However, since the constant vorticities in the upper and lower layer have opposite signs, this continuity property is not valid in the continuously differentiable sense.

3.1. The pure current background state. A steady pure current solution of the governing equations, in a flow presenting no variations in the longitudinal direction and with a flat free surface $\overline{z} = \overline{h}_1$ and a flat thermocline $\overline{z} = 0$, is provided by the velocity field $(\overline{U}_1(\overline{z}),\ 0)$ above the thermocline and $(\overline{U}(\overline{z}),\ 0)$ below it, where

$$\overline{U}_1(\overline{z}) = \overline{\gamma}_1\overline{z} + \overline{\gamma}\,\overline{h} \quad \text{for} \quad 0 < \overline{z} < \overline{h}_1, \tag{3.14}$$

$$\overline{U}(\overline{z}) = \overline{\gamma}\,(\overline{z} + \overline{h}) \quad \text{for} \quad -\overline{h} \le \overline{z} \le 0, \tag{3.15}$$

the associated pressure distribution $\overline{\mathfrak{P}}(\overline{z})$ being given by

$$\begin{cases} \overline{P}_{atm} - \overline{\rho}_1\overline{g}\,(\overline{z} - \overline{h}_1) + \overline{\Omega}\,\overline{\rho}_1\Big[\overline{\gamma}_1(\overline{z}^2 - \overline{h}_1^2) + 2\overline{\gamma}\,\overline{h}\,(\overline{z} - \overline{h}_1)\Big] & \text{for} \quad 0 < \overline{z} \le \overline{h}_1, \\ \overline{P}_{atm} - \overline{\rho}\,\overline{g}\,\overline{z} + \overline{\Omega}\,\overline{\rho}\,\overline{\gamma}\,(\overline{z}^2 + 2\overline{h}\overline{z}) - \overline{\Omega}\,\overline{\rho}_1\overline{h}_1(\overline{\gamma}_1\overline{h}_1 + 2\overline{\gamma}\,\overline{h}) + \overline{\rho}_1\overline{g}\,\overline{h}_1 & \text{for} \quad -\overline{h} \le \overline{z} \le 0. \end{cases} \tag{3.16}$$

The current profile (3.14)–(3.15) captures the salient features: a westward surface drift (since $|\overline{\gamma}_1|\overline{h}_1 > \overline{\gamma}\overline{h}$) below which resides an eastward jet (the EUC) which overlies

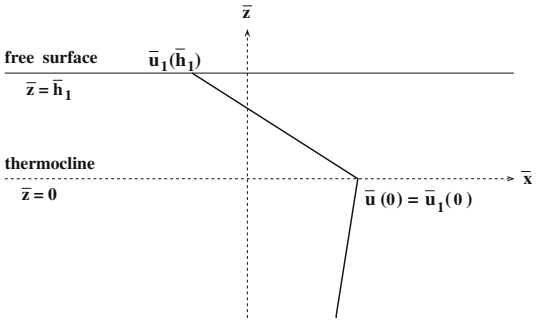

Fig. 3. Depiction of the flow model in the absence of waves (pure current): constant negative vorticity $\overline{\gamma}_1$ in the near-surface homogeneous layer accommodates a flow-reversal from the westward wind-drift at the free surface $\overline{z} = \overline{h}_1$ to an eastward jet whose core resides on the thermocline $\overline{z} = 0$, while a constant positive vorticity in the abyssal homogeneous region below the thermocline permits the adjustment to practically still water at great depths (with no motion on the flat bed $\overline{z} = -\overline{h}$). The corresponding velocity field is given by $\overline{u}_1(\overline{z}) = \overline{\gamma}_1\overline{z} + \overline{\gamma}\,\overline{h}$ and $\overline{w}_1 = 0$ in the near-surface layer $0 < \overline{z} < \overline{h}_1$, with $\overline{u}(\overline{z}) = \overline{\gamma}\,(\overline{z} + \overline{h})$ and $\overline{w} = 0$ in the abyssal layer $-\overline{h} < \overline{z} < 0$; here $\overline{\gamma} > 0 > \overline{\gamma}_1$ and $|\overline{\gamma}_1|\overline{h}_1 > \overline{\gamma}\,\overline{h}$

an abyssal layer where a gradual transition to no motion on the flat bed occurs (see Fig. 3). Note that

$$\int_{-\overline{h}}^{0} \overline{U}_1(\overline{z})\,d\overline{z} + \int_{0}^{\overline{h_1}} \overline{U}(\overline{z})\,d\overline{z} = \frac{1}{2}\,\overline{\gamma}\,\overline{h}^2 + \frac{1}{2}\,\overline{\gamma}_1\,\overline{h}_1^2 + \overline{\gamma}\,\overline{h}\,\overline{h}_1$$

represents the mass transport per unit width (in $\mathrm{m}^2\,\mathrm{s}^{-1}$), so that for

$$(\overline{h}_1/\overline{h}) < -(\overline{\gamma}/\overline{\gamma}_1) + \sqrt{(\overline{\gamma}/\overline{\gamma}_1)^2 - (\overline{\gamma}/\overline{\gamma}_1)}$$

there is a net eastward mass transport. Since typically (see below) $\overline{h}_1/\overline{h}$ and $\overline{\gamma}/\overline{\gamma}_1$ are both of order $O(10^{-2})$, an eastward net flow occurs within 1° from the Equator; this is compensated by a return flow at higher latitudes (see [36]). We seek nonlinear wave–current interactions arising as localised wave perturbations of the background flow (3.14–3.15).

3.2. Nondimensionalisation. The first stage in expressing the governing equations in a useful form is to introduce a suitable non-dimensionalisation of the variables; to this end we write

$$\begin{aligned} &\overline{t} = (\overline{L}/\overline{U}_0)\,t, \quad (\overline{x}, \overline{z}) = \overline{L}\,(x,\ z), \quad (\overline{u}, \overline{w}) = \overline{U}_0\,(u,\ w), \\ &(\overline{u}_1, \overline{w}_1) = \overline{U}_0\,(u_1,\ w_1), \quad \overline{P} = \overline{\rho}_1\overline{U}_0^2 p, \end{aligned} \tag{3.17}$$

where the length scale is $\overline{L} = 500\,\mathrm{m}$ and $\overline{U}_0 = 0.5\,\mathrm{m\,s}^{-1}$ is an appropriate speed scale. (The omission of the overbar now indicates that we are using the non-dimensional version of the corresponding variable.) Set

$$h = \overline{h}/\overline{L}, \qquad h_1 = \overline{h}_1/\overline{L}, \qquad \eta(x,t) = \overline{\eta}(\overline{x},\overline{t})/\overline{L}, \qquad \eta_1(x,t) = \overline{\eta}_1(\overline{x},\overline{t})/\overline{L}. \tag{3.18}$$

Above the thermocline, in the region

$$\mathcal{D}_1(t) = \{(x, z) : \ \eta(x, t) < z < h_1 + \eta_1(x, t)\},$$

Eqs. (3.1), (3.3) and (3.5) therefore become

$$u_{1,t} + u_1 u_{1,x} + w_1 u_{1,z} + \omega\, w_1 = -p_x, \tag{3.19}$$
$$w_{1,t} + u_1 w_{1,x} + w_1 w_{1,z} - \omega\, u_1 = -p_z - g, \tag{3.20}$$
$$u_{1,x} + w_{1,z} = 0, \tag{3.21}$$
$$u_{1,z} - w_{1,x} = \gamma_1, \tag{3.22}$$

while below the thermocline, in the region

$$\mathcal{D}(t) = \{(x, z) : \ -h < z < \eta(x, t)\},$$

Eqs. (3.2), (3.4) and (3.6) become

$$u_t + u u_x + w u_z + \omega\, w = -\frac{1}{1+r}\, p_x, \tag{3.23}$$
$$w_t + u w_x + w w_z - \omega\, u = -\frac{1}{1+r}\, p_z - g, \tag{3.24}$$
$$u_x + w_z = 0, \tag{3.25}$$
$$u_z - w_x = \gamma, \tag{3.26}$$

where

$$\omega = 2\overline{\Omega}\,\overline{L}/\overline{U}_0 \approx 0.15, \ \ g = \overline{g}\,\overline{L}/\overline{U}_0^2 \approx 2 \times 10^4$$
$$\gamma_1 = \overline{L}\,\overline{\gamma}_1/\overline{U}_0 \approx -12.5, \ \ \gamma = \overline{L}\,\overline{\gamma}/\overline{U}_0 \approx 0.25, \tag{3.27}$$

for a flow reversal from a westward surface wind drift of $0.5\,\mathrm{m\,s^{-1}}$ to a maximal eastward EUC speed of $1\,\mathrm{m\,s^{-1}}$ in a layer of mean depth $\overline{h}_1 = 120\,\mathrm{m}$, with an abyssal layer of mean depth $\overline{h} = 4\,\mathrm{km}$; these values, corresponding to $h_1 \approx 0.24$ and $h \approx 8$, are appropriate for the 6000 km stretch between about 140°E and 150°W (see [14]). For a flow that gradually dies out with depth, being motionless at the flat bed, we compute $\overline{\gamma_1} \approx -1.25 \times 10^{-2}\,\mathrm{s^{-1}}$ and $\overline{\gamma} \approx 2.5 \times 10^{-4}\,\mathrm{s^{-1}}$. Written in nondimensional form, the boundary conditions (3.7)–(3.12) become

$$p = \overline{P}_{atm}/(\overline{\rho}_1 \overline{U}_0^2) \quad \text{on} \quad z = h_1 + \eta_1(x, t), \tag{3.28}$$
$$w_1 = \eta_{1,t} + u_1\, \eta_{1,x} \quad \text{on} \quad z = h_1 + \eta_1(x, t), \tag{3.29}$$
$$w_1 = \eta_t + u_1\, \eta_x \quad \text{on} \quad z = \eta(x, t), \tag{3.30}$$
$$w = \eta_t + u\, \eta_x \quad \text{on} \quad z = \eta(x, t), \tag{3.31}$$
$$p_+ = p_- \quad \text{on} \quad z = \eta(x, t), \tag{3.32}$$
$$w = 0 \quad \text{on} \quad z = -h, \tag{3.33}$$
$$w_1 \eta_x + u_1 = w \eta_x + u \quad \text{on} \quad z = \eta(x, t), \tag{3.34}$$

where $p_\pm$ in (3.32) refer to the limits of the function p from above and below the common boundary $z = \eta(x, t)$ of the regions $\mathcal{D}_1(t)$ and $\mathcal{D}(t)$.

The non-dimensional counterpart of the pure current background state (3.14)–(3.15) is

$$U_1(z) = \gamma_1 z + \gamma h \quad \text{for} \quad 0 < z \le h_1, \tag{3.35}$$
$$U(z) = \gamma(z+h) \quad \text{for} \quad -h \le z \le 0, \tag{3.36}$$

with $\mathfrak{p}(z)$, corresponding to the background pressure distribution (3.16), given by

$$\begin{cases} \dfrac{\overline{P}_{atm}}{\overline{\rho}_1 \overline{U}_0^2} - g(z-h_1) + \dfrac{\omega}{2}\left[\gamma_1(z^2-h_1^2) + 2\gamma h(z-h_1)\right] & \text{for } 0 < z \le h_1, \\ \dfrac{\overline{P}_{atm}}{\overline{\rho}_1 \overline{U}_0^2} - (1+r)\left[gz - \dfrac{\omega\gamma}{2}(z^2+2hz)\right] - \dfrac{\omega h_1}{2}(\gamma_1 h_1 + 2\gamma h) + g h_1 & \text{for } -h \le z \le 0. \end{cases} \tag{3.37}$$

3.3. Nonlinear flow separation. For wave–current interactions the "current" component of the flow is defined as the average velocity, and the localised perturbations that fluctuate around this average are ascribed to the wave motion, the direction of wave propagation being along the x-axis. We view the nonlinear equatorial wave–current interaction as a localised wave perturbation of the pure current solution (3.35)–(3.36) in a domain that is infinite in the x-direction. The localised nature of the wave disturbances is captured by assuming the smooth functions $\eta_1(x,t)$ and $\eta(x,t)$ to be of Schwartz class $\mathcal{S}(\mathbb{R})$ in the x-variable at any instant t, with

$$\int_{\mathbb{R}} \eta_1(x,t)\,dx = \int_{\mathbb{R}} \eta(x,t)\,dx = 0, \qquad t \ge 0. \tag{3.38}$$

The corresponding velocity fields (u_1, w_1) and (u, w) are to be smooth in the domains $\mathcal{D}_1(t)$ and $\mathcal{D}(t)$, with both vector functions admitting continuous extensions to the closure of the domains and across their adjacent boundary. It is remarkable that the interpretation of (3.35)–(3.36) as the underlying current can be achieved within the framework of fully nonlinear theory, without performing approximations.

Theorem 1. *If the flow approaches asymptotically the pure current state, then the velocity fields (u_1, w_1) and (u, w) and the pressure $p(x,z,t)$ are, at any instant t, smooth localised perturbations of the background flow (3.35)–(3.36) and of the corresponding background pressure distribution (3.37), in each of the domains $\mathcal{D}_1(t)$ and $\mathcal{D}(t)$, respectively.*

Proof. Consider a solution to the governing equations (3.19)–(3.22), (3.23)–(3.26), (3.28)–(3.34), with no flow along the flat bed $z = -h$. Let us denote by $\mathfrak{U}(z_0,t) = \lim_{\Lambda\to\infty} \dfrac{1}{2\Lambda}\displaystyle\int_{-\Lambda}^{\Lambda} u(x,z_0,t)\,dx$ the underlying current at a depth z_0 below the trough of the thermocline $z = \eta(x,t)$, obtained by averaging. Applying the divergence theorem to the divergence-free vector field $(-w,\, u-\gamma z)$ in the rectangular domain $\{(x,z) : -\Lambda < x < \Lambda,\ -h < z < z_0\}$ yields

$$\begin{aligned} &\int_{-\Lambda}^{\Lambda} \left[u(x,z_0,t) - \gamma z_0\right] dx - \int_{-\Lambda}^{\Lambda} \left[u(x,-h,t) + \gamma h\right] dx \\ &\quad = \int_{-h}^{z_0} w(\Lambda,z,t)\,dz - \int_{-h}^{z_0} w(-\Lambda,z,t)\,dz \to 0 \end{aligned}$$

for $\Lambda \to \infty$. Combining this with the absence of a flow on the flat bed, $u = 0$ on $z = -h$, we get $\mathfrak{U}(z_0,t) = \gamma(z_0+h)$, which coincides with (3.36). Similarly, the

underlying current at a depth z_0 above the crest of the internal wave and below the trough of the surface wave is

$$\mathfrak{U}_1(z_0,t) = \lim_{\Lambda\to\infty} \frac{1}{2\Lambda} \int_{-\Lambda}^{\Lambda} u_1(x,z_0,t)\,dx.$$

Applying the divergence theorem in the domain $\{(x,z) : -\Lambda < x < \Lambda,\ \eta(x,t) < z < z_0\}$ to the vector field $(-w_1,\ u_1 - \gamma_1 z)$, we get

$$\begin{aligned}\int_{-\Lambda}^{\Lambda} \big[u_1(x,z_0,t) - \gamma_1 z_0\big]\,dx &= \int_{\eta(\Lambda,t)}^{z_0} w_1(\Lambda,z,t)\,dz - \int_{\eta(-\Lambda,t)}^{z_0} w_1(-\Lambda,z,t)\,dz\\ &+ \int_{-\Lambda}^{\Lambda} \Big[u_1(x,\eta(x,t),t) - \gamma_1\eta(x,t) + w_1(x,\eta(x,t),t)\,\eta_x(x,\eta(x,t),t)\Big]\,dx,\end{aligned}$$

so that (3.38) and (3.34) yield

$$\begin{aligned}\mathfrak{U}_1(z_0,t) &= \gamma_1 z_0 + \lim_{\Lambda\to\infty} \frac{1}{2\Lambda} \int_{-\Lambda}^{\Lambda} \Big[u_1(x,\eta(x,t),t) + w_1(x,\eta(x,t),t)\,\eta_x(x,\eta(x,t),t)\Big]\,dx\\ &= \gamma_1 z_0 + \lim_{\Lambda\to\infty} \frac{1}{2\Lambda} \int_{-\Lambda}^{\Lambda} \Big[u(x,\eta(x,t),t) + w(x,\eta(x,t),t)\,\eta_x(x,\eta(x,t),t)\Big]\,dx.\end{aligned}$$

This leads to $\mathfrak{U}_1(z_0,t) = \gamma_1 z_0 + \gamma h$, which coincides with (3.35). Indeed, for this it suffices to apply the divergence theorem in the domain $\{(x,z) : -\Lambda < x < \Lambda,\ -h < z < \eta(x,t)\}$ to the vector field $(-w,\ u - \gamma z)$, using (3.34) and (3.38) together with the vanishing of u on the flat bed $z = -h$. □

Let us point out that the velocity field is actually Hölder continuous across the smooth thermocline. Indeed, the velocity field of the pure current background state (3.35)–(3.36) is Lipschitz continuous in the horizontal strip $-h \le z \le h_1$. Since the wave–current interaction is a localized perturbation of this background flow, in an open horizontal strip $\mathbb{O}$ containing the thermocline, the velocity field $\mathbb{V}_{\mathbb{F}}$ of the wave motion is divergence-free, has a piecewise constant curl, decays as it approaches the lateral boundary components of $\mathbb{O}$ at infinity, and is smooth on the upper and lower boundary components $\partial\mathbb{O}_{\pm}$ of $\mathbb{O}$ (since in the fluid domain away from the thermocline, the Laplacian of each velocity component is constant, so that the regularity for Poisson's equation [15] applies). For $s > 3$, the *a priori* elliptic estimate

$$\|\nabla\mathbb{V}_{\mathbb{F}}\|_{L^s(\mathbb{O})} \lesssim \|\mathrm{div}\,\mathbb{V}_{\mathbb{F}}\|_{L^s(\mathbb{O})} + \|\mathrm{curl}\,\mathbb{V}_{\mathbb{F}}\|_{L^s(\mathbb{O})} + \|\mathbb{V}_{\mathbb{F}}\|_{L^\infty(\partial\mathbb{O}_{\pm})}$$

for the div-curl system, interpreted for $\mathbb{V}_{\mathbb{F}} \in L^\infty(\mathbb{O})$ in the sense of distributions (see [4, 42]), ensures, by means of a Sobolev imbedding (see [15]), that $\mathbb{V}_{\mathbb{F}}$ is Hölder continuous with exponent $\alpha < 1/3$ throughout $\mathbb{O}$. Consequently the velocity field of the wave–current interaction is Hölder continuous across the thermocline.

3.4. Hamiltonian formulation. We now present the Hamiltonian formulation of the governing equations (3.19)–(3.22), (3.23)–(3.26), (3.28)–(3.34).

As a first step towards the choice of new dependent and independent variables that reveal the Hamiltonian structure, we introduce the stream function and the perturbed velocity potential. The equations of mass conservation, (3.21) and (3.25), ensure the

existence of a stream function in each layer, $\psi_1(x,z,t)$ in $\mathcal{D}_1(t)$ and $\psi(x,z,t)$ in $\mathcal{D}(t)$, determined, up to an additive term that depends only on time, by

$$u_1 = \psi_{1,z} \quad \text{and} \quad w_1 = -\psi_{1,x} \quad \text{in} \quad \mathcal{D}_1(t), \qquad u = \psi_z \quad \text{and} \quad w = -\psi_x \quad \text{in} \quad \mathcal{D}(t). \tag{3.39}$$

Relying on (3.21), (3.25), (3.33) and (3.30), we may set

$$\psi(x,z,t) = \int_{-h}^{z} u(x,z',t)\,dz' \quad \text{for} \quad (x,z) \in \mathcal{D}(t), \tag{3.40}$$

$$\psi_1(x,z,t) = \int_{-\infty}^{x} \Big[u_1\big(x',\eta(x',t),t\big)\eta_x(x',t) - w_1(x',\eta(x',t),t)\Big]\,dx' + \int_{\eta(x,t)}^{z} u_1(x,z',t)\,dz' + \frac{\gamma h^2}{2} \quad \text{for} \quad (x,z) \in \mathcal{D}_1(t), \tag{3.41}$$

observing that (3.30) ensures that the first integral in (3.41) is well-defined. Each stream function has a smooth extension to the closure of its respective domain, $\overline{\mathcal{D}(t)}$ and $\overline{\mathcal{D}(t)}$. Actually, since the velocity field is a smooth localised perturbation of the background state (3.35)–(3.36), by evaluating ψ and ψ_1 on $z = \eta(x,t)$ as $x \to -\infty$ we see that (3.40)–(3.41) ensure the existence of a continuous function $\Psi(x,z,t)$ throughout the bulk of the fluid $\mathcal{K}(t) = \{(x,z) : -d \le z \le h_1 + \eta_1(x,t)\}$, with $\Psi = \psi$ on $\overline{\mathcal{D}(t)}$ and $\Psi = \psi_1$ on $\overline{\mathcal{D}(t)}$. The discussion after (3.13) shows that Ψ is continuously differentiable in $\mathcal{K}(t)$ without being twice continuously differentiable (due to discontinuities in the second order partial derivatives across the thermocline). For this reason, rather than using Ψ we prefer to use ψ and ψ_1. With the vorticity distribution defined by (3.22) and (3.26), (3.39) yields

$$\Delta\psi = \gamma \quad \text{in} \quad \mathcal{D}(t), \qquad \Delta\psi_1 = \gamma_1 \quad \text{in} \quad \mathcal{D}_1(t). \tag{3.42}$$

We now introduce harmonic perturbed velocity potentials, φ in $\mathcal{D}(t)$ and φ_1 in $\mathcal{D}_1(t)$, by requiring

$$\begin{cases} u = \varphi_x + \gamma(z+h) & \text{and} \quad w = \varphi_z \quad \text{in} \quad \mathcal{D}(t), \\ u_1 = \varphi_{1,x} + \gamma_1 z + \gamma h & \text{and} \quad w_1 = \varphi_{1,z} \quad \text{in} \quad \mathcal{D}_1(t). \end{cases} \tag{3.43}$$

More precisely, since the velocity field is a smooth localised perturbation of the background state (3.35)–(3.36), with $u = 0$ on the flat bed $z = -h$, we may set

$$\varphi(x,z,t) = \int_{-h}^{z} w(x,z',t)\,dz' \quad \text{for} \quad (x,z) \in \mathcal{D}(t), \tag{3.44}$$

$$\varphi_1(x,z,t) = \int_{-\infty}^{x} \Big[u_1(x',\eta(x',t),t) - \gamma_1\eta(x',t) - \gamma h + w_1(x',\eta(x',t),t)\,\eta_x(x',t)\Big]\,dx' + \int_{\eta(x,t)}^{z} w_1(x,z',t)\,dz', \qquad (x,z) \in \mathcal{D}_1(t)\,; \tag{3.45}$$

to see that the first integral in (3.45) is well-defined, use (3.34) and subsequently the divergence theorem applied to the vector field $(-w,\, u - \gamma z)$ in $\{(x',z) : x^* < x' < x,\ -h < z < \eta(x',t)\}$ with $x^* \to -\infty$. Both perturbed velocity potentials admit continuously differentiable extensions to the closures of their domains, $\varphi \in C^1(\overline{\mathcal{D}(t)})$ and $\varphi_1 \in C^1(\overline{\mathcal{D}_1(t)})$. However, in contrast with the stream functions, any oscillation

of the thermocline impedes a continuous extension of the perturbed velocity potentials to the bulk of the fluid $\mathcal{K}(t)$: while smooth additive functions that depend solely on time may be added to the right sides of (3.44)–(3.45), this process does not lead to an equality of φ and φ_1 on the interface. Indeed, for $\eta \not\equiv 0$, (3.43) and the continuity of the velocity field across the thermocline show that differentiation of a presumed equality $\varphi(x,\, \eta(x,t),\, t) = \varphi_1(x,\, \eta(x,t),\, t) + \varphi_0(t)$ with respect to the x-variable leads to a contradiction since $\gamma \neq \gamma_1$.

The kinematic boundary conditions (3.29)–(3.31) can now be written as

$$\eta_{1,t} = (\varphi_{1,z})_{s_1} - \eta_{1,x}\big[(\varphi_{1,x})_{s_1} + \gamma_1(h_1+\eta_1) + \gamma h\big], \tag{3.46}$$

$$\eta_t = (\varphi_{1,z})_s - \eta_x\big[(\varphi_{1,x})_s + \gamma_1\eta + \gamma h\big], \tag{3.47}$$

$$\eta_t = (\varphi_z)_s - \eta_x\big[(\varphi_x)_s + \gamma\eta + \gamma h\big], \tag{3.48}$$

respectively, where the subscript s_1 means that we consider the traces of the involved functions on the free surface $z = h_1 + \eta_1(x,t)$, while the subscript s denotes traces on the interface $z = \eta(x,t)$.

We can recast the Euler equations (3.19)–(3.20) and (3.23)–(3.24) as

$$\nabla\left[\varphi_{1,t} + \frac{|\nabla\psi_1|^2}{2} + p - (\gamma_1+\omega)\psi_1 + gz\right] = 0 \quad \text{in} \quad \mathcal{D}_1(t),$$

$$\nabla\left[\varphi_t + \frac{|\nabla\psi|^2}{2} + \frac{p}{1+r} - (\gamma+\omega)\psi + gz\right] = 0 \quad \text{in} \quad \mathcal{D}(t),$$

respectively, so that the expressions on which the gradient operates are just time-dependent. Recalling (3.40)–(3.41), (3.44)–(3.45), and taking into account that the flow is a smooth localised perturbation of (3.35)–(3.36) with the background pressure distribution (3.37), we evaluate these expressions as $x \to -\infty$ on $z = \eta(x,t)$ and on $z = -h$, respectively. We get

$$\varphi_{1,t} + \frac{|\nabla\psi_1|^2}{2} + p - (\gamma_1+\omega)\psi_1 + gz = \frac{\overline{P}_{atm}}{\overline{\rho}_1\overline{U}_0^2} + (1+r)\alpha + \beta - \frac{\omega\gamma h^2}{2} \quad \text{in} \quad \mathcal{D}_1(t), \tag{3.49}$$

$$\varphi_t + \frac{|\nabla\psi|^2}{2} + \frac{p}{1+r} - (\gamma+\omega)\psi + gz = \frac{\overline{P}_{atm}}{(1+r)\overline{\rho}_1\overline{U}_0^2} + \alpha - \frac{\omega\gamma h^2}{2} \quad \text{in} \quad \mathcal{D}(t), \tag{3.50}$$

where

$$\alpha = \frac{gh_1}{1+r} - \frac{\omega(\gamma_1 h_1 + 2\gamma h)h_1}{2(1+r)}, \qquad \beta = \frac{\gamma(\gamma-\gamma_1)h^2}{2}.$$

Consequently, we can write the dynamic boundary condition (3.28) as

$$\varphi_{1,t} + \frac{|\nabla\psi_1|^2}{2} - (\gamma_1+\omega)\psi_1 + gz = (1+r)\alpha + \beta \quad \text{on} \quad z = h_1 + \eta_1(x,t), \tag{3.51}$$

while the continuity of the pressure across the thermocline, (3.32), becomes

$$(1+r)\left[(\varphi_t)_s + \frac{|\nabla\psi|_s^2}{2} - (\gamma+\omega)\chi + g\eta\right]$$

$$= \left[(\varphi_{1,t})_s + \frac{|\nabla \psi_1|_s^2}{2} - (\gamma_1 + \omega)\chi + g\eta \right] - \beta - \frac{r\omega\gamma h^2}{2}, \tag{3.52}$$

where

$$\chi(x,t) := \psi(x,\eta(x,t),t) = \psi_1(x,\eta(x,t),t), \quad \chi_1(x,t) := \psi_1(x,\, h_1 + \eta_1(x,t),t). \tag{3.53}$$

Let us now show that the governing equations (3.19)–(3.22), (3.23)–(3.26), (3.28)–(3.34) can be tidied up as a Hamiltonian system by identifying

$$H = E_K + E_P \tag{3.54}$$

as the Hamiltonian, where

$$\begin{aligned} E_K &= (1+r)\int_{\mathbb{R}} \left\{ \int_{-h}^{\eta} \left[\frac{u^2+w^2}{2} \right] dz - \int_{-h}^{0} \frac{\gamma^2 (z+h)^2}{2}\, dz \right\} dx \\ &\quad + \int_{\mathbb{R}} \left\{ \int_{\eta}^{h_1+\eta_1} \left[\frac{u_1^2+w_1^2}{2} \right] dz - \int_{0}^{h_1} \frac{(\gamma_1 z + \gamma h)^2}{2}\, dz \right\} dx, \end{aligned} \tag{3.55}$$

$$\begin{aligned} E_P &= (1+r)\int_{\mathbb{R}} \left\{ \int_{-h}^{\eta} gz\, dz - \int_{-h}^{0} gz\, dz \right\} dx + \int_{\mathbb{R}} \left\{ \int_{\eta}^{h_1+\eta_1} gz\, dz - \int_{0}^{h_1} gz\, dz \right\} dx \\ &= \frac{rg}{2}\int_{\mathbb{R}} \eta^2(x,t)\, dx + \frac{g}{2}\int_{\mathbb{R}} \eta_1^2(x,t)\, dx, \end{aligned} \tag{3.56}$$

are the excess kinetic and potential energies, respectively. Other than the technical spur to avoid infinite energy, the rationale for the above choice is that not all potential energy can be converted into kinetic energy, the portion that is available for this conversion being the difference in the potential energy between the perturbed wave–current state and the pure current background state. In view of (3.17) and (3.27), the nondimensional expression in (3.54) corresponds to the total excess energy per unit width, $\overline{H} = \overline{\rho}_1 \overline{U}_0^2 \overline{L}^2\, H$. To see that $\overline{\rho}_1 \overline{U}_0^2 \overline{L}^2 \approx 625 \times 10^5\ \mathrm{m\,kg\,s^{-2}}$ is a suitable scale, note that, over the 16,000 km length of the EUC, it corresponds to a mean total energy of about $4\ \mathrm{kg\,s^{-2}}$, which is of the order of the mean kinetic energy at the surface of the equatorial Pacific (about $2\ \mathrm{kg\,s^{-2}}$; see [19]). As for the estimate of potential energy, field data gathered at 45 m depth indicates an average ratio between potential and kinetic energy of about 0.76, varying between 0.18 and 3.63 with means of the order of $10^{-2}\ \mathrm{kg\,s^{-2}}$ (see [28]). Moreover, the energy is dominated by the zonal component, which is about one order of magnitude greater than the meridional component, thus corroborating the realistic nature of an approach that invokes the f-plane approximation and the neglect of meridional variations.

Using (3.43) and (3.38), we get

$$\begin{aligned} E_K &= (1+r)\int_{\mathbb{R}} \left\{ \int_{-h}^{\eta} \left[\frac{|\nabla\varphi|^2}{2} + \gamma(z+h)\,\varphi_x \right] dz + \frac{\gamma^2}{6}\,\eta^3 + \frac{\gamma^2 h}{2}\,\eta^2 \right\} dx \\ &\quad + \int_{\mathbb{R}} \left\{ \int_{\eta}^{h_1+\eta_1} \left[\frac{|\nabla\varphi_1|^2}{2} + (\gamma_1 z + \gamma h)\,\varphi_{1,x} \right] dz \right. \\ &\quad \left. + \frac{\gamma_1^2}{6}\,(\eta_1^3 - \eta^3) + \frac{\gamma\gamma_1 h}{2}\,(\eta_1^2 - \eta^2) + \frac{\gamma_1^2 h_1}{2}\,\eta_1^2 \right\} dx. \end{aligned} \tag{3.57}$$

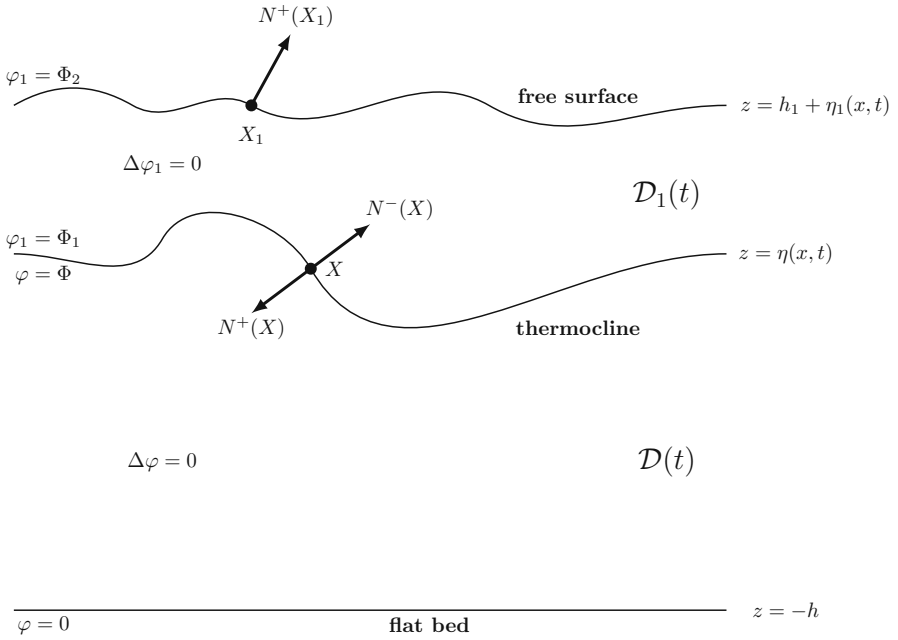

Fig. 4. Sketch of the geometric features that are relevant in the definition of the Dirichlet–Neumann operators associated with the two layers: at the instant t, the outward unit normals for the abyssal layer $\mathcal{D}(t)$ and for the near-surface layer $\mathcal{D}_1(t)$ at a point X on the thermocline are $N^-(X)$ and $N^+(X)$, respectively, while the outward unit normal at a point X_1 on the free surface is $N^+(X_1)$

Since $\Delta\varphi = 0$ in $\mathcal{D}(t)$ with $\varphi = 0$ on $z = -h$, setting

$$\Phi(x,t) = \varphi(x,\eta(x,t),t), \tag{3.58}$$

integration by parts yields

$$\int_{\mathbb{R}}\int_{-h}^{\eta} |\nabla\varphi|^2\, dzdx = \int_{\mathbb{R}} \Phi\left[G(\eta)\Phi\right] dx, \tag{3.59}$$

where $G := G(\eta)$ is the Dirichlet–Neumann operator on $\mathcal{D}(t)$, associating to Φ the normal derivative of φ on the upper boundary $z = \eta(x,t)$ with outward unit normal N; see Fig. 4. Similarly, denoting

$$\Phi_1(x,t) = \varphi_1(x,\eta(x,t),t), \qquad \Phi_2(x,t) = \varphi_1(x,h_1+\eta_1(x,t),t), \tag{3.60}$$

we define on $\mathcal{D}_1(t)$ the Dirichlet–Neumann matrix operator

$$G^+ := G^+(\eta,\eta_1) = \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix}$$

that associates to the boundary values of the harmonic function φ_1 on the lower and upper boundaries, Φ_1 and Φ_2, respectively, the outward normal derivatives $N\cdot\nabla\varphi_1$ at these boundaries. Then

$$\int_{\mathbb{R}}\int_{\eta}^{h_1+\eta_1} |\nabla\varphi_1|^2\, dzdx = \int_{\mathbb{R}} \begin{pmatrix}\Phi_1\\ \Phi_2\end{pmatrix}^{\mathsf{T}} \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix} \begin{pmatrix}\Phi_1\\ \Phi_2\end{pmatrix} dx. \tag{3.61}$$

On the other hand, since $F = \psi - \frac{1}{2}\gamma\,(z+h)^2$ is harmonic in $\mathcal{D}(t)$, with $\nabla F = (-\varphi_z,\ \varphi_x)$ and $F = 0$ on $z = -h$, the divergence theorem applied to $\frac{1}{2}\gamma\,(z+h)^2\,\nabla F$ yields

$$\int_{\mathbb{R}}\int_{-h}^{\eta}\gamma(z+h)\varphi_x\,dzdx = \int_{\mathbb{R}}\frac{\gamma}{2}\left[\eta+h\right]^2(\partial_x\Phi)\,dx = -\int_{\mathbb{R}}\gamma\,(\eta+h)\eta_x\,\Phi\,dx. \tag{3.62}$$

Similarly, the divergence theorem applied to $\frac{1}{2}\,(\gamma_1 z^2 + 2\gamma h z)\nabla F_1$, where $F_1 = \psi_1 - \frac{1}{2}\,(\gamma_1 z^2 + 2\gamma h z)$ is harmonic in $\mathcal{D}_1(t)$ with $\nabla F_1 = (-\varphi_{1,z},\ \varphi_{1,x})$, yields

$$\begin{aligned}&\int_{\mathbb{R}}\int_{\eta}^{h_1+\eta_1}(\gamma_1 z+\gamma h)\varphi_{1,x}dzdx\\ &\quad= -\int_{\mathbb{R}}\Big[\gamma_1(h_1+\eta_1)+\gamma h\Big]\eta_{1,x}\,\Phi_2\,dx + \int_{\mathbb{R}}\Big[\gamma_1\,\eta+\gamma h\Big]\eta_x\,\Phi_1\,dx.\end{aligned} \tag{3.63}$$

From the definition of the Dirichlet–Neumann operators, using (3.43) and (3.29)–(3.31), we see that

$$G_{11}\Phi_1 + G_{12}\Phi_2 = (\varphi_{1,x})_s\,\eta_x - (\varphi_{1,z})_s = -\eta_t - (\gamma_1\eta+\gamma h)\eta_x, \tag{3.64}$$

$$G\Phi = -(\varphi_x)_s\,\eta_x + (\varphi_z)_s = \eta_t + \gamma(\eta+h)\eta_x, \tag{3.65}$$

$$G_{21}\Phi_1 + G_{22}\Phi_2 = -(\varphi_{1,x})_{s_1}\,\eta_{1,x} + (\varphi_{1,z})_{s_1} = \eta_{1,t} + \Big[\gamma_1(h_1+\eta_1)+\gamma h\Big]\eta_{1,x}. \tag{3.66}$$

Adding up the relations (4.9) and (3.65), we obtain

$$G_{11}\Phi_1 + G_{12}\Phi_2 + G\Phi = (\gamma-\gamma_1)\eta\eta_x. \tag{3.67}$$

Let us now define the operator

$$B = B(\eta,\ \eta_1) := G + (1+r)\,G_{11}. \tag{3.68}$$

Introducing the variables

$$\xi = (1+r)\,\Phi - \Phi_1, \qquad \xi_1 = \Phi_2, \tag{3.69}$$

relation (3.67) enables us to express Φ and Φ_1 in terms of ξ and ξ_1:

$$\Phi = B^{-1}\Big(G_{11}\xi - G_{12}\xi_1 + (\gamma-\gamma_1)\eta\eta_x\Big), \tag{3.70}$$

$$\Phi_1 = B^{-1}\Big(-G\xi - (1+r)\,G_{12}\xi_1 + (1+r)(\gamma-\gamma_1)\eta\eta_x\Big). \tag{3.71}$$

Regarding $h_1,\ h,\ \gamma_1,\ \gamma,\ \omega$ as fixed parameters and gathering (3.54)–(3.57), (3.59), (3.61)–(3.63) and (3.70)–(3.71), we express H as a functional depending solely on the variables $\eta,\ \eta_1,\ \xi,\ \xi_1$:

$$H = \mathcal{H}(\eta,\ \eta_1,\ \xi,\ \xi_1). \tag{3.72}$$

Since $x \mapsto \int_{-\infty}^{x}\theta(x')\,dx' \in \mathcal{S}(\mathbb{R})$ if and only if $\theta \in \mathcal{S}(\mathbb{R})$ satisfies $\int_{\mathbb{R}}\theta(x')\,dx' = 0$, we see that the functions φ and φ_1 defined by (3.44)–(3.45) are smooth localised

perturbations. This property is passed on to ξ and ξ_1, while for η and η_1 it is part of our setup.

It is of interest to provide a concise explicit form of the functional $\mathcal{H}$ in (3.72). Since $\Phi_1 = (1+r)\Phi - \xi$ by (3.69), B and G are unbounded self-adjoint operators on $L^2(\mathbb{R})$ while $(G_{12})^* = G_{21}$, and using the relation $(1+r)B^{-1}G_{11} = \mathrm{Id} - B^{-1}G$ that follows from (3.68), we get

$$\begin{aligned}
\mathcal{H} = {} & \frac{\gamma_1^2}{6}\int_{\mathbb{R}} \eta_1^3\,dx + \frac{(1+r)\gamma^2-\gamma_1^2}{6}\int_{\mathbb{R}} \eta^3\,dx + \frac{rg+\gamma h[(1+r)\gamma-\gamma_1]}{2}\int_{\mathbb{R}} \eta^2\,dx \\
& + \frac{g+\gamma_1[\gamma h+\gamma_1 h_1]}{2}\int_{\mathbb{R}} \eta_1^2\,dx - \gamma\int_{\mathbb{R}} (\eta+h)\eta_x\xi\,dx \\
& - \int_{\mathbb{R}} [\gamma_1(h_1+\eta_1)+\gamma h]\eta_{1,x}\xi_1\,dx \\
& - (\gamma-\gamma_1)\int_{\mathbb{R}} \eta\eta_x B^{-1}G\xi dx - (1+r)(\gamma-\gamma_1)\int_{\mathbb{R}} \eta\eta_x B^{-1}G_{12}\xi_1 dx \\
& - \frac{(1+r)(\gamma-\gamma_1)^2}{2}\int_{\mathbb{R}} \eta\eta_x B^{-1}\eta\eta_x dx \\
& + \frac{1}{2}\int_{\mathbb{R}} \Big[\xi_1 G_{22}\xi_1 + \xi G B^{-1}G_{11}\xi - \xi G B^{-1}G_{12}\xi_1 \\
& - \xi_1 G_{21}B^{-1}G\xi - (1+r)\xi_1 G_{21}B^{-1}G_{12}\xi_1\Big]dx. \qquad (3.73)
\end{aligned}$$

We now compute variations of the functional $\mathcal{H}$, with respect to the inner product associated to square-integrable real functions defined on the real line. We consider variations of the wave field, regarding the underlying current field (3.35)–(3.36) as fixed; in particular, the flat bed, the vorticity and the mean depths of the layers do not change. Note that

$$\begin{aligned}
& \delta\Big\{r\int_{\mathbb{R}} \eta^2\,dx + \int_{\mathbb{R}} \eta_1^2\,dx\Big\} \\
& \quad = 2\int_{\mathbb{R}} \Big\{r\,\eta\,\delta\eta + \eta_1\,\delta\eta_1\Big\}\,dx, \qquad (3.74) \\
& \delta\Big\{(1+r)\int_{\mathbb{R}} \Big[\frac{\gamma^2}{6}\eta^3 + \frac{\gamma^2 h}{2}\eta^2\Big]dx + \int_{\mathbb{R}} \Big[\frac{\gamma_1^2}{6}(\eta_1^3-\eta^3) \\
& \qquad + \frac{\gamma\gamma_1 h}{2}(\eta_1^2-\eta^2) + \frac{\gamma_1^2 h_1}{2}\eta_1^2\Big]dx\Big\} \\
& \quad = \int_{\mathbb{R}} \Big\{\Big(\frac{(1+r)\gamma^2-\gamma_1^2}{2}\eta^2 + \gamma h\big[(1+r)\gamma-\gamma_1\big]\eta\Big)\delta\eta \\
& \qquad + \Big(\frac{\gamma_1^2}{2}\eta_1^2 + \gamma_1(\gamma_1 h_1+\gamma h)\eta_1\Big)\delta\eta_1\Big\}\,dx. \qquad (3.75)
\end{aligned}$$

On the other hand,

$$\delta\Big\{\int_{\mathbb{R}}\int_{-h}^{\eta} (z+h)\,\varphi_x\,dzdx\Big\} = \int_{\mathbb{R}} \Big\{\Big((\eta+h)\,\partial_x\Phi\Big)\,\delta\eta - \Big((\eta+h)\,\eta_x\Big)\,\delta\Phi\Big\}\,dx, \qquad (3.76)$$

and

$$\begin{aligned}
&\delta\Big\{\int_{\mathbb{R}}\int_{\eta}^{h_1+\eta_1}(\gamma_1 z+\gamma h)\,\varphi_{1,x}\,dzdx\Big\}\\
&\quad=\int_{\mathbb{R}}\Big\{\Big(-(\gamma_1\eta+\gamma h)\,\partial_x\Phi_1\Big)\,\delta\eta+\Big((\gamma_1\eta+\gamma h)\,\eta_x\Big)\,\delta\Phi_1\\
&\qquad+\Big(\big[\gamma_1(h_1+\eta_1)+\gamma h\big]\partial_x\Phi_2\Big)\,\delta\eta_1-\Big(\big[\gamma_1(h_1+\eta_1)+\gamma h\big]\,\eta_{1,x}\Big)\,\delta\Phi_2\Big\}\,dx,
\end{aligned}\tag{3.77}$$

respectively. For harmonic variations of φ, using the divergence theorem and the identity

$$\delta\Phi=\lim_{\varepsilon\to 0}\frac{(\varphi+\varepsilon\,\delta\varphi)(x,\eta+\varepsilon\delta\eta,t)-\varphi(x,\eta,t)}{\varepsilon}=(\delta\varphi)_s+(\varphi_z)_s\,\delta\eta,$$

we get

$$\begin{aligned}
&\delta\Big\{\int_{\mathbb{R}}\int_{-h}^{\eta}|\nabla\varphi|^2\,dzdx\Big\}\\
&\quad=\int_{\mathbb{R}}|\nabla\varphi|_s^2\,\delta\eta\,dx+2\int_{\mathbb{R}}\Big\{(\varphi_z)_s-(\varphi_x)_s\eta_x\Big\}\Big[\delta\Phi-(\varphi_z)_s\,\delta\eta\Big]dx.
\end{aligned}\tag{3.78}$$

Similarly, for harmonic perturbations of φ_1 we get

$$\begin{aligned}
\delta\Big\{\int_{\mathbb{R}}\int_{\eta}^{h_1+\eta_1}|\nabla\varphi_1|^2\,dzdx\Big\}&=\int_{\mathbb{R}}|\nabla\varphi_1|_{s_1}^2\,\delta\eta_1\,dx-\int_{\mathbb{R}}|\nabla\varphi_1|_s^2\,\delta\eta\,dx\\
&+2\int_{\mathbb{R}}\Big\{(\varphi_{1,z})_{s_1}-(\varphi_{1,x})_{s_1}\eta_{1,x}\Big\}\Big[\delta\Phi_2-(\varphi_{1,z})_{s_1}\delta\eta_1\Big]dx\\
&-2\int_{\mathbb{R}}\Big\{(\varphi_{1,z})_s-(\varphi_{1,x})_s\eta_x\Big\}\Big[\delta\Phi_1-(\varphi_{1,z})_s\delta\eta\Big]dx.
\end{aligned}\tag{3.79}$$

From (3.54)–(3.57) and (3.74)–(3.79), using (3.29)–(3.31), (3.40)–(3.41), (3.44)–(3.45), (3.51)–(3.52), we conclude that

$$\begin{aligned}
\delta H&=\int_{\mathbb{R}}\Big\{\Big((1+r)\eta_t\Big)\delta\Phi+\Big(-\eta_t\Big)\delta\Phi_1+\Big(\eta_{1,t}\Big)\delta\Phi_2\\
&\quad+\Big(-\xi_{1,t}+(\gamma_1+\omega)\chi_1-\frac{(\gamma_1h_1+\gamma h)h_1(\gamma_1+\omega)}{2}\Big)\delta\eta_1\\
&\quad+\Big(-\xi_t+\Big[(1+r)\gamma-\gamma_1+r\omega\Big]\,\chi+\frac{\gamma h^2[\gamma_1-r\omega-(1+r)\gamma]}{2}\Big)\,\delta\eta\Big\}\,dx\\
&=\int_{\mathbb{R}}\Big\{\Big(\eta_t\Big)\,\delta\xi+\Big(\eta_{1,t}\Big)\,\delta\xi_1+\Big(-\xi_{1,t}+(\gamma_1+\omega)\chi_1\Big)\,\delta\eta_1\\
&\quad+\Big(-\xi_t+\Big[(1+r)\gamma-\gamma_1+r\omega\Big]\,\chi\Big)\,\delta\eta\Big\}\,dx,
\end{aligned}\tag{3.80}$$

since the invariance of the total mass in each layer ensures

$$\int_{\mathbb{R}} \delta\eta \, dx = \int_{\mathbb{R}} \delta\eta_1 \, dx = 0. \tag{3.81}$$

Switching from the variables $(\eta,\ \eta_1,\ \xi,\ \xi_1)$ to the new variables $(\eta,\ \eta_1,\ \zeta,\ \zeta_1)$, where

$$\zeta(x,t) = \xi(x,t) + \mu \int_{-\infty}^{x} \eta(x',t)\, dx' \quad \text{where} \quad \mu = \frac{(1+r)\gamma - \gamma_1 + r\omega}{2}, \tag{3.82}$$

$$\zeta_1(x,t) = \xi_1(x,t) + \mu_1 \int_{-\infty}^{x} \eta_1(x',t)\, dx' \quad \text{where} \quad \mu_1 = \frac{\gamma_1 + \omega}{2}, \tag{3.83}$$

using the identities

$$\begin{aligned} 0 &= \int_{\mathbb{R}} \Big(\eta_t \int_{-\infty}^{x} (\delta\eta)\, dx' + (\delta\eta) \int_{-\infty}^{x} \eta_t \, dx' \Big) dx \\ &= \int_{\mathbb{R}} \Big(\eta_{1,t} \int_{-\infty}^{x} (\delta\eta_1)\, dx' + (\delta\eta_1) \int_{-\infty}^{x} \eta_{1,t} \, dx' \Big) dx, \end{aligned}$$

we can write (3.80) as

$$\begin{aligned} \delta H = \int_{\mathbb{R}} \Big\{ &\eta_t \Big(\delta\zeta - \mu \int_{-\infty}^{x} (\delta\eta)\, dx' \Big) + \eta_{1,t} \Big(\delta\zeta_1 - \mu_1 \int_{-\infty}^{x} (\delta\eta_1)\, dx' \Big) \\ &+ \Big(-\zeta_{1,t} + \mu_1 \int_{-\infty}^{x} \eta_{1,t}\, dx' + 2\mu_1 \chi_1 \Big) \delta\eta_1 + \Big(-\zeta_t + \mu \int_{-\infty}^{x} \eta_t \, dx' + 2\mu\chi \Big) \delta\eta \Big\} dx. \end{aligned} \tag{3.84}$$

Recalling (3.53), note that $\eta_t = -\partial_x \chi$ and $\eta_{1,t} = -\partial_x \chi_1$, due to (3.31) and (3.29), while (3.40) and (3.41) yield the asymptotic behaviour $\chi \to \frac{\gamma h^2}{2}$ and $\chi_1 \to \int_0^{h_1} (\gamma_1 z + \gamma h)\, dz + \frac{\gamma h^2}{2} = \frac{\gamma_1 h_1^2 + 2\gamma h h_1 + \gamma h^2}{2}$ for $x \to -\infty$. Consequently

$$\chi(x,t) = \frac{\gamma h^2}{2} - \int_{-\infty}^{x} \eta_t \, dx', \qquad \chi_1(x,t) = \frac{\gamma_1 h_1^2 + 2\gamma h h_1 + \gamma h^2}{2} - \int_{-\infty}^{x} \eta_{1,t}\, dx'. \tag{3.85}$$

Since differentiation of (3.38) with respect to the t-variable yields $\int_{\mathbb{R}} \eta_t \, dx = \int_{\mathbb{R}} \eta_{1,t}\, dx = 0$, using integration by parts, (3.85) and (3.81), we get

$$\begin{aligned} &\int_{\mathbb{R}} \Big(\int_{-\infty}^{x} (\delta\eta) dx' \Big) \eta_t dx = -\int_{\mathbb{R}} (\delta\eta) \Big(\int_{-\infty}^{x} \eta_t dx' \Big) dx = \int_{\mathbb{R}} (\delta\eta) \chi \, dx, \\ &\int_{\mathbb{R}} \Big(\int_{-\infty}^{x} (\delta\eta_1)\, dx' \Big) \eta_{1,t}\, dx = \int_{\mathbb{R}} (\delta\eta_1)\, \chi_1 \, dx. \end{aligned}$$

Cancellations occur in (3.84) due to the above two relations in (3.84), so that

$$\delta H = \int_{\mathbb{R}} \Big\{ (\eta_t)\, \delta\zeta + (\eta_{1,t})\, \delta\zeta_1 - (\zeta_{1,t})\, \delta\eta_1 - (\zeta_t)\, \delta\eta \Big\} dx.$$

Therefore we proved the following result.

Theorem 2. *Let* $\mathfrak{u} = (\eta,\ \eta_1,\ \zeta,\ \zeta_1)^{\mathsf{T}}$. *With* $\mathcal{H}$ *given by the explicit expression* (3.73), *the system*

$$\begin{cases} \dfrac{\delta\mathcal{H}}{\delta\zeta} = \eta_t\,, & \dfrac{\delta\mathcal{H}}{\delta\eta} = -\zeta_t\,, \\ \dfrac{\delta\mathcal{H}}{\delta\zeta_1} = \eta_{1,t}\,, & \dfrac{\delta\mathcal{H}}{\delta\eta_1} = -\zeta_{1,t}\,, \end{cases} \tag{3.86}$$

is the Hamiltonian form $\partial_t\mathfrak{u} = J\,\delta_{\mathfrak{u}}\,\mathcal{H}$ *of the governing equations with respect to the canonical symplectic structure induced by the matrix operator* $J = \left(\begin{smallmatrix} 0 & I_2 \\ -I_2 & 0 \end{smallmatrix}\right)$ *with* $I_2 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, *acting on the phase space* $\mathcal{X} = \mathcal{S}(\mathbb{R}) \times \mathcal{S}(\mathbb{R}) \times \mathcal{S}(\mathbb{R}) \times \mathcal{S}(\mathbb{R})$ *of Schwartz-class functions.*

The above variational derivative of $\mathcal{H} : \mathcal{X} \to \mathbb{R}$ at $\mathfrak{u}_0 \in \mathcal{X}$ is defined by $\langle \mathfrak{u},\ \frac{\delta\mathcal{H}}{\delta\mathfrak{u}}(\mathfrak{u}_0)\rangle = \lim\limits_{\varepsilon\to 0} \frac{\mathcal{H}(\mathfrak{u}_0+\varepsilon\mathfrak{u})-\mathcal{H}(\mathfrak{u}_0)}{\varepsilon}$, with respect to the inner product $\langle\cdot,\cdot\rangle$ in the Hilbert space $L^2(\mathbb{R}) \times L^2(\mathbb{R}) \times L^2(\mathbb{R}) \times L^2(\mathbb{R})$.

4. Hamiltonian Perturbations

We regard the equatorial waves as localised perturbations of the background state, typical examples being solitary waves. A reasonable measure of the "wavelength" of a solitary wave is the spatial extent of the region where the deviation of the profile from its mean level is at least 1% of the maximum height. In the geophysical regime in which the wavelength is long compared to the average depth, both gravity and the rotation of the Earth play a role in the dynamics. In our nondimensional framework $x = 1$ corresponds to 500 m, so that the relevant spatial scaling is

$$x' = \varepsilon x \tag{4.1}$$

with $\varepsilon \ll 10^{-1}$. This defines a physical regime in which the dependent and independent variables are $O(1)$, and the relation to the nondimensional variable x is by means of (4.1); in particular, a perturbation that mainly occurs for values $x' \in (-1,\ 1)$ corresponds to $x \in (-1/\varepsilon,\ 1/\varepsilon)$ and thus represents a physical region spreading over more than 10 km. Our approach to systematic derivations of model equations is from the point of view of Hamiltonian perturbation theory. The parameter ε from (4.1) will also be introduced through choices of scaling of the dependent variables $(\eta,\ \eta_1,\ \zeta,\ \zeta_1)$, corresponding to scaling regimes of interest, in which dispersive and nonlinear effects are brought into play. Since our main interest is the propagation of internal waves, the representative scales are determined by the motion of the thermocline, and we use these scales also for the surface waves. This results in a Hamiltonian that is a function of the small parameter ε. The approximating Hamiltonian systems are obtained by retaining a finite number of terms in the Taylor expansion in ε of the Hamiltonian.

Given a constant $\alpha > 0$ and a smooth local diffeomorphism $f : \mathbb{R}^4 \to \mathbb{R}^4$, the change of variables $[\alpha t' = t,\ \mathfrak{u}' = f(\mathfrak{u})]$ transforms the Hamiltonian system $\partial_t\mathfrak{u} = J\,\delta_{\mathfrak{u}}\mathcal{H}$ to the Hamiltonian system $\partial_{t'}\mathfrak{u}' = J'\,\partial_{\mathfrak{u}'}\mathcal{H}'$ with Hamiltonian $\mathcal{H}'(\mathfrak{u}') = \mathcal{H}(\mathfrak{u})$ if the operator $J' = \alpha(\delta_{\mathfrak{u}} f)J(\delta_{\mathfrak{u}} f)^{\mathsf{T}}$ is symplectic, that is, $(J')^{\mathsf{T}}JJ' = J$ (see [35]); moreover, if the change of variables preserves the Hamiltonian form of all Hamiltonian equations, then J' must be symplectic (see [33]). Due to the physical interpretation in which the

dynamics is governed by the perturbations of the two interfaces, the relevant coordinate transformation in our setting is the spatial scaling

$$\eta' = a\eta, \quad \zeta' = a\zeta, \qquad \eta_1' = b\eta_1, \quad \zeta_1' = b\zeta_1, \tag{4.2}$$

for some positive scalars a and b. Since a $2n \times 2n$ block matrix $\left(\begin{smallmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{smallmatrix}\right)$ is symplectic if and only if the $n \times n$ matrices $M_{11}^{\mathsf{T}} M_{21}$ and $M_{12}^{\mathsf{T}} M_{22}$ are symmetric and $M_{11}^{\mathsf{T}} M_{22} - M_{21}^{\mathsf{T}} M_{12} = I_n$ (see [33]), we see that the scaling transformation (4.2) coupled with the simultaneous temporal scale

$$\alpha t' = t, \tag{4.3}$$

is symplectic if and only if $a = b = \sqrt{\alpha}$. Note for our nondimensionalisation $t = 1$ corresponds to about 17 min. The space-time scale (4.2)–(4.3) specifies the distance and time needed to bring about the balance

$$x' = O(1) \quad \text{and} \quad t' = O(1) \tag{4.4}$$

for the dynamics of the flow. Note that a regime of type (4.2) comes about due to wave perturbations of the pure current background state described in Sect. 3.1. Since these perturbations are harmonic, as shown by the considerations in Sects. 3.3 and 3.4, their overall effect is encoded in the way they operate at the surface and at the interface. Using the relations (3.43), (3.58), (3.60), (3.69), (3.82)–(3.83), one can render (4.2) in terms of the magnitude of the deviation of the surface/interface from the quiescent flat state and of the relative sizes of the wave/current components of the fluid velocity field at these boundaries. We will provide details of this interpretative aspect in each specific case that we discuss.

The linear Dirichlet–Neumann operators G and G^+ are analytic in their dependence on η and (η, η_1), respectively, having convergent Taylor series expansions

$$G(\eta)\zeta = \sum_{j=0}^{\infty} G^{(j)}(\eta)\,\zeta, \tag{4.5}$$

$$G^+(\eta, \eta_1)\begin{pmatrix} \zeta \\ \zeta_1 \end{pmatrix} = \sum_{m_0, m_1 = 0}^{\infty} \begin{pmatrix} G_{11}^{(m_0 m_1)}(\eta, \eta_1) & G_{12}^{(m_0 m_1)}(\eta, \eta_1) \\ G_{21}^{(m_0 m_1)}(\eta, \eta_1) & G_{22}^{(m_0 m_1)}(\eta, \eta_1) \end{pmatrix} \begin{pmatrix} \zeta \\ \zeta_1 \end{pmatrix}, \tag{4.6}$$

with each linear operator $G^{(j)}(\eta)$ homogeneous of degree j in η and each linear operator $G_{ij}^{(m_0 m_1)}(\eta, \eta_1)$ with $i, j = 1, 2$ homogeneous of degree m_0 in η and of degree m_1 in η_1 (see the discussion in [10,11]); moreover, each of the operators $G^{(j)}(\eta)$ and $G_{ii}^{(m_0 m_1)}(\eta, \eta_1)$ with $i = 1, 2$ and $m_0,\ m_1 \geq 0$ is self-adjoint, while $(G_{12}^{(m_0 m_1)}(\eta, \eta_1))^* = G_{21}^{(m_0 m_1)}(\eta, \eta_1)$. Denoting

$$D = -i\,\partial_x$$

and given a smooth function $m : \mathbb{R} \to \mathbb{C}$ whose derivatives of any order have polynomial growth, the Fourier multiplication operator

$$(m(D)f)(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} e^{ik(x-x')} m(k)\, f(x')\, dx'\, dk \tag{4.7}$$

maps $\mathcal{S}(\mathbb{R})$ into $\mathcal{S}(\mathbb{R})$ and extends to a self-adjoint operator in $L^2(\mathbb{R})$ if and only if m is real-valued; see [40]. The operator (4.7) is bounded if and only if $m \in L^\infty(\mathbb{R})$, with $m(D)f$ real-valued or imaginary-valued, for real-valued $f \in \mathcal{S}(\mathbb{R})$, if $m(-x) = \overline{m(x)}$ or if $m(-x) = -\overline{m(x)}$ for all $x \in \mathbb{R}$, respectively. In terms of Fourier multipliers the leading order terms of the Dirichlet–Neumann operators G and G^+ are given by

$$G^{(0)}(\eta) = D\tanh(hD), \tag{4.8}$$

$$G^{(1)}(\eta) = D\eta(x)D - D\tanh(hD)\eta(x)D\tanh(hD), \tag{4.9}$$

while

$$\begin{pmatrix} G_{11}^{(00)}(\eta,\eta_1) & G_{12}^{(00)}(\eta,\eta_1) \\ G_{21}^{(00)}(\eta,\eta_1) & G_{22}^{(00)}(\eta,\eta_1) \end{pmatrix} = \begin{pmatrix} D\coth(h_1D) & -D\operatorname{csch}(h_1D) \\ -D\operatorname{csch}(h_1D) & D\coth(h_1D) \end{pmatrix}, \tag{4.10}$$

$$\begin{pmatrix} G_{11}^{(10)}(\eta,\eta_1) & G_{12}^{(10)}(\eta,\eta_1) \\ G_{21}^{(10)}(\eta,\eta_1) & G_{22}^{(10)}(\eta,\eta_1) \end{pmatrix}$$
$$= \begin{pmatrix} D\coth(h_1D)\eta(x)D\coth(h_1D) - D\eta(x)D & -D\coth(h_1D)\eta(x)D\operatorname{csch}(h_1D) \\ -D\operatorname{csch}(h_1D)\eta(x)D\coth(h_1D) & D\operatorname{csch}(h_1D)\eta(x)D\operatorname{csch}(h_1D) \end{pmatrix}, \tag{4.11}$$

$$\begin{pmatrix} G_{11}^{(01)}(\eta,\eta_1) & G_{12}^{(01)}(\eta,\eta_1) \\ G_{21}^{(01)}(\eta,\eta_1) & G_{22}^{(01)}(\eta,\eta_1) \end{pmatrix}$$
$$= \begin{pmatrix} -D\operatorname{csch}(h_1D)\eta_1(x)D\operatorname{csch}(h_1D) & D\operatorname{csch}(h_1D)\eta_1(x)D\coth(h_1D) \\ D\coth(h_1D)\eta_1(x)D\operatorname{csch}(h_1D) & -D\coth(h_1D)\eta_1(x)D\coth(h_1D) + D\eta_1(x)D \end{pmatrix}. \tag{4.12}$$

The above formulas are indicative of the nonlinear nonlocal dependence of the operators $G(\eta)$ and $G^+(\eta,\eta_1)$ on η and (η,η_1), respectively. Concerning the effect of scaling, note that a simple change of variables confirms that the transformation (4.1) replaces the Fourier multiplier operator $m(D)$ by $m(\varepsilon D')$, with $D' = -i\partial_{x'}$. More precisely, for $f \in \mathcal{S}(\mathbb{R})$ we have

$$m(D)f(x) = m(\varepsilon D')F(x'), \tag{4.13}$$

where $F(x) = f(x/\varepsilon)$. On the other hand, the effect of the scaling (4.2) is expressed by

$$G^{(j)}(\eta')\zeta' = a^{j+1}G^{(j)}(\eta)\zeta, \qquad j \geq 0, \tag{4.14}$$

with more intricate but workable scaling formulas expressing $G_{ij}^{(m_0m_1)}(\eta',\eta_1')$ in terms of $G_{ij}^{(m_0m_1)}(\eta,\eta_1)$ for $1 \leq i, j \leq 2$ and $m_0,\ m_1 \geq 0$.

4.1. The linearised equations. The quiescent state (no waves) corresponds to the trivial solution $\eta = \eta_1 = \zeta = \zeta_1 = 0$. Since most practical predictions of water waves use linear approximations, it is natural to investigate the linearised equations about the state of rest. An elegant way to derive them is to truncate the Taylor expansion of the Hamiltonian at its quadratic term. The quadratic part of the Hamiltonian (3.73) is

$$H^{(2)} = \frac{rg + \gamma h[(1+r)\gamma - \gamma_1]}{2}\int_{\mathbb{R}} \eta^2\,dx + \frac{g + \gamma_1(\gamma_1 h_1 + \gamma h)}{2}\int_{\mathbb{R}} \eta_1^2\,dx$$

$$
\begin{aligned}
&-\gamma\, h \int_{\mathbb{R}} \eta_x \xi\, dx - (\gamma_1 h_1 + \gamma h) \int_{\mathbb{R}} \eta_{1,x} \xi_1\, dx \\
&+ \frac{1}{2} \int_{\mathbb{R}} \xi\, \frac{D \tanh(hD)}{1 + r + \tanh(hD)\tanh(h_1 D)} \xi\, dx \\
&+ \int_{\mathbb{R}} \xi\, \frac{D \tanh(hD)\operatorname{sech}(h_1 D)}{1 + r + \tanh(hD)\tanh(h_1 D)} \xi_1\, dx \\
&+ \frac{1}{2} \int_{\mathbb{R}} \xi_1\, \frac{D \tanh(hD) + (1+r) D \tanh(h_1 D)}{1 + r + \tanh(hD)\tanh(h_1 D)} \xi_1\, dx
\end{aligned} \tag{4.15}
$$

since $B \approx D\tanh(hD) + (1+r)D\coth(h_1 D)$, $G \approx D\tanh(hD)$ and

$$
\begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix} \approx \begin{pmatrix} D\coth(h_1 D) & -D\operatorname{csch}(h_1 D) \\ -D\operatorname{csch}(h_1 D) & D\coth(h_1 D) \end{pmatrix}.
$$

To obtain the linearised form of (3.86) we have to exhibit in (4.15) a functional dependence on $\mathfrak{u} = (\eta,\ \eta_1,\ \zeta,\ \zeta_1)^{\mathsf{T}}$. This can be achieved as follows. Note that (3.82)–(3.83) yield

$$
\partial_x \xi = \partial_x \zeta - \mu\eta, \qquad \partial_x \xi_1 = \partial_x \zeta_1 - \mu_1 \eta_1, \tag{4.16}
$$

$$
D\xi = D\zeta + i\mu\eta = D\Big(\zeta - \mu \int_{-\infty}^{x} \eta_x\, dx'\Big),
$$

$$
D\xi_1 = D\zeta_1 + i\mu_1\eta_1 = D\Big(\zeta_1 - \mu_1 \int_{-\infty}^{x} \eta_{1,x}\, dx'\Big). \tag{4.17}
$$

Using the calculus of pseudodifferential operators with smooth symbols, by means of (4.17) we can handle the last three terms in (4.15) by relying on the fact that although the functions and the operators of interest are all real-valued, the natural set-up for Fourier multipliers is the complex inner product on $L^2(\mathbb{R})$, which is also adequate for taking advantage of the self-adjointness of operators of the type (4.7). We get

$$
\begin{aligned}
H^{(2)} = &\Big(\frac{rg + \gamma h[(1+r)\gamma - \gamma_1]}{2} - \mu\gamma h\Big) \int_{\mathbb{R}} \eta^2\, dx \\
&+ \Big(\frac{g + \gamma_1(\gamma_1 h_1 + \gamma h)}{2} - \mu_1(\gamma_1 h_1 + \gamma h)\Big) \int_{\mathbb{R}} \eta_1^2\, dx \\
&+ \gamma h \int_{\mathbb{R}} \eta \zeta_x\, dx + (\gamma_1 h_1 + \gamma h) \int_{\mathbb{R}} \eta_1 \zeta_{1,x}\, dx \\
&+ \mu \int_{\mathbb{R}} \eta_x\, \frac{\tanh(hD)}{(1+r)D + D\tanh(hD)\tanh(h_1 D)} \zeta\, dx \\
&+ \frac{\mu^2}{2} \int_{\mathbb{R}} \eta\, \frac{\tanh(hD)}{(1+r)D + D\tanh(hD)\tanh(h_1 D)} \eta\, dx \\
&+ \frac{1}{2} \int_{\mathbb{R}} \zeta\, \frac{D\tanh(hD)}{1 + r + \tanh(hD)\tanh(h_1 D)} \zeta\, dx \\
&+ \int_{\mathbb{R}} \zeta\, \frac{D\tanh(hD)\operatorname{sech}(h_1 D)}{1 + r + \tanh(hD)\tanh(h_1 D)} \zeta_1\, dx \\
&+ \mu\mu_1 \int_{\mathbb{R}} \eta\, \frac{\tanh(hD)\operatorname{sech}(h_1 D)}{(1+r)D + D\tanh(hD)\tanh(h_1 D)} \eta_1\, dx
\end{aligned}
$$

$$
\begin{aligned}
&+\mu\int_{\mathbb{R}}\eta_x\,\frac{\tanh(hD)\mathrm{sech}(h_1D)}{(1+r)D+D\tanh(hD)\tanh(h_1D)}\zeta_1\,dx\\
&+\mu_1\int_{\mathbb{R}}\eta_{1,x}\,\frac{\tanh(hD)\mathrm{sech}(h_1D)}{(1+r)D+D\tanh(hD)\tanh(h_1D)}\zeta\,dx\\
&+\frac{1}{2}\int_{\mathbb{R}}\zeta_1\,\frac{D\tanh(hD)+(1+r)D\tanh(h_1D)}{1+r+\tanh(hD)\tanh(h_1D)}\zeta_1\,dx\\
&+\frac{\mu_1^2}{2}\int_{\mathbb{R}}\eta_1\,\frac{\tanh(hD)+(1+r)\tanh(h_1D)}{(1+r)D+D\tanh(hD)\tanh(h_1D)}\eta_1\,dx\\
&+\mu_1\int_{\mathbb{R}}\eta_{1,x}\,\frac{\tanh(hD)+(1+r)\tanh(h_1D)}{(1+r)D+D\tanh(hD)\tanh(h_1D)}\zeta_1\,dx.
\end{aligned}
$$

The linearised equations of motion are

$$
\begin{cases}
\eta_t=\dfrac{\delta H^{(2)}}{\delta\zeta},\quad \xi_t=-\dfrac{\delta H^{(2)}}{\delta\eta},\\
\eta_{1,t}=\dfrac{\delta H^{(2)}}{\delta\zeta_1},\ \xi_{1,t}=-\dfrac{\delta H^{(2)}}{\delta\eta_1}.
\end{cases}
\tag{4.18}
$$

Given an initial data $\mathfrak{u}_0$, representing a localised wave perturbation of the pure current background state described in Sect. 3.1, we can solve the linear system (4.18) using the Fourier transform. Indeed, setting $\hat{f}(k)=\int_{\mathbb{R}}f(x)e^{-ikx}dx$ for $f\in\mathcal{S}(\mathbb{R})$ in each component of $\mathfrak{u}$, the system (4.18) is transformed for any fixed $k\in\mathbb{R}$ into the linear autonomous system of ordinary differential equations

$$
\partial_t\,\hat{\mathfrak{u}}(k,t)=M(k)\,\hat{\mathfrak{u}}(k,t),
\tag{4.19}
$$

where

$$
M(k)=\begin{pmatrix}
-i\gamma hk+i\mu k\,\Theta(k) & i\mu_1 k\,\mathrm{sech}(h_1k)\,\Theta(k) & k^2\Theta(k) & k^2\mathrm{sech}(h_1k)\,\Theta(k)\\
i\mu k\,\mathrm{sech}(h_1k)\,\Theta(k) & -i\Gamma_1 k+i\mu_1 k\,\Theta_1(k) & k^2\mathrm{sech}(h_1k)\,\Theta(k) & k^2\Theta_1(k)\\
\Gamma-\mu^2\,\Theta(k) & -\mu\mu_1\mathrm{sech}(h_1k)\,\Theta(k) & -i\gamma hk+i\mu k\Theta(k) & i\mu k\,\mathrm{sech}(h_1k)\,\Theta(k)\\
-\mu\mu_1\mathrm{sech}(h_1k)\,\Theta(k) & \omega\Gamma_1-g-\mu_1^2\Theta_1(k) & i\mu_1 k\,\mathrm{sech}(h_1k)\,\Theta(k) & -i\Gamma_1k+i\mu_1k\Theta_1(k)
\end{pmatrix},
$$

and

$$
\begin{cases}
\Gamma\;=r(\omega\gamma h-g), & \Theta(k)=\dfrac{\tanh(hk)}{k[1+r+\tanh(hk)\tanh(h_1k)]},\\
\Gamma_1\;=\gamma_1h_1+\gamma h, & \Theta_1(k)\;=\dfrac{\tanh(hk)+(1+r)\tanh(h_1k)}{k[1+r+\tanh(hk)\tanh(h_1k)]}.
\end{cases}
\tag{4.20}
$$

Note that in the absence of vorticity ($\gamma=\gamma_1=0$) the matrix $M(k)$ is real but the context of equatorial wave–current interactions (for which $\gamma\neq0$) brings about purely imaginary entries on the main diagonal. The complex system (4.19) can be cast in the form of a real Hamiltonian system with double degree of freedom (of dimension eight) by separating the real and complex part of each of the four components of the vector $\hat{\mathfrak{u}}\in\mathbb{C}^4$. However, in our context it is advisable to work with the complex linear Hamiltonian system to avoid going beyond the treshold of Galois theory for roots of polynomials. In particular, finding the spectrum of $M(k)$ is computationally within reach

for a fixed value of ξ since explicit formulas are available for its quartic characteristic polynomial. On the other hand, since the available theoretical studies of Hamiltonian systems deal with systems in $\mathbb{R}^{2n}$, there is a need for theoretical adjustment but with guaranteed structural features. For example, the matrix $M(k)$ being Hamiltonian means that $JM(k)$ is self-adjoint, so $M(k) = J^{-1}(-\overline{M(k)}^{\mathsf{T}})J$. Therefore the matrices $M(k)$ and $\overline{M(k)}^{\mathsf{T}}$ are similar. Thus if $\Lambda(k)$ is an eigenvalue of $M(k)$, then so is $-\overline{\Lambda(k)}$ and the whole Jordan block structure will be the same for $\Lambda(k)$ and $-\overline{\Lambda(k)}$.

The unique solution to (4.19) with initial data $\hat{\mathfrak{u}}_0(k) = \int_{\mathbb{R}} \mathfrak{u}_0(x)\, e^{ikx}\, dx$, given by $\hat{\mathfrak{u}}(k,t) = e^{M(k)t}\, \hat{\mathfrak{u}}_0(k)$ for $t \geq 0$, corresponds, by means of the inverse Fourier transform, to the solution

$$\mathfrak{u}(x,t) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{M(k)t}\, \hat{\mathfrak{u}}_0(k)\, e^{ikx} dk, \qquad t \geq 0 \tag{4.21}$$

of (4.18) with initial data $\mathfrak{u}_0 \in \mathcal{S}(\mathbb{R})$. To a purely imaginary eigenvalue $\Lambda(k) = -ikc$ with $c \in \mathbb{R}\backslash\{0\}$ of $M(k)$, with corresponding eigenvector $\mathfrak{v}(k) \neq 0$, we can associate the oscillatory mode solution $e^{M(k)t}\mathfrak{v}(k) = e^{-ikct}\, \mathfrak{v}(k)$ of (4.19), and for the initial data $\mathfrak{u}_0 \in \mathcal{S}(\mathbb{R})$ with $\hat{\mathfrak{u}}_0(k) = \mathfrak{v}(k)$, we can interpret the solution

$$\frac{1}{2\pi} \int_{\mathbb{R}} e^{ik(x-ct)}\, \mathfrak{v}(k)\, dk \tag{4.22}$$

as a linear superposition of such modes. This way a purely imaginary eigenvalue $\Lambda(k) = -ikc$ of $M(k)$ corresponds to the fundamental oscillation mode $e^{ik(x-ct)}\, \mathfrak{v}(k)$ with frequency $\frac{|k|}{2\pi}$, propagating at the constant speed c. Due to (3.17), a fixed nondimensional value of $k \neq 0$ corresponds in physical variables to a harmonic oscillation of wavelength $\frac{2\pi}{|k|}\, \overline{L}$. With our choice $\overline{L} = 500\,\mathrm{m}$, since ocean waves have wavelengths in excess of 50 m, the qualitative behaviour of the solutions to the linear system (4.19) for $|k| < 64$ is physically relevant. Note that an eigenvalue $\Lambda(k)$ with nonzero real part is the symptom of an oscillation mode whose amplitude grows exponentially in time since $\Lambda(k)$ or $-\overline{\Lambda(k)}$ leads to a coefficient of the type $e^{\omega(k)t}$ with $\omega(k) > 0$. This Kelvin-Helmholtz instability corresponds to a situation when a wave is capable of extracting energy from the background pure current state, by drawing either kinetic energy from the pure current motion or potential energy from the stratification.[1] Since these instability phenomena are a prelude to mixing and turbulence, it is important to understand whether or not they are inherent to the equatorial wave–current interactions. For considerations of spectral type the explicit formulas for the roots of the quartic characteristic polynomial are too unwieldy. We advocate an approach that takes advantage of the fact that each entry of the matrix $M(k)$ depends smoothly (actually, the dependence is real-analytic) on the parameter k. This ensures that the eigenvalues depend continuously on k, even if splitting and coalescence of eigenvalues may take place (see the discussion in [21], Chapter VII) and the eigenvalues need not be differentiable, which makes it difficult to translate this feature quantitatively. Also, note that the eigenspaces can behave quite singularly—they are not necessarily continuous. However, for simple eigenvalues the eigenvalue itself as well as the (unique) normalised eigenvector will exhibit a real-analytic dependence on k (see [25]). Note that the multiparameter version of this last result fails spectacularly: for simple eigenvalues, a locally Lipschitz-dependence holds but differentiability might

[1] The wave motion involves lifting of denser fluid and sinking of lighter fluid, both of which cause a change in the centre of gravity which brings about a change of potential energy. In the absence of dissipation, the supply/loss of potential energy can only be provided by conversion of/into kinetic energy.

fail and it may not even be possible to choose eigenvectors in a continuous way (see the discussion in [24]).

Since our main interest is the propagation of localised perturbations, it is advantageous to use the Paley-Wiener theorem to translate the feature that $x \mapsto \mathrm{u}(x,t)$ has compact support at every instant into the requirement that $k \mapsto \hat{\mathrm{u}}(k,t)$ is an analytic function of exponential type. The general solution of (4.19) is a linear combination of solutions of the form $t^n e^{\Lambda(k)t}\mathfrak{U}$, where n is a nonnegative integer, $\mathfrak{U}$ is a constant vector and $\Lambda(k)$ is an eigenvalue of $M(k)$. Since the fact that no entry of the matrix $M(k)$ exhibits a superlinear growth in k towards infinity ensures, by inspection of Ferrari's classical algebraic solution to the quartic, the existence of a constant $m > 0$ such that $|\Lambda(k)| \leq m(1+|k|)$ for all $k \in \mathbb{R}$, we deduce that $k \mapsto \hat{\mathrm{u}}(k,t)$ remains forever of exponential type if it is initially so. Consequently the linearised system captures the propagation of localised wave perturbations of the pure current background state.

A simple calculation shows that $M(0)$ has $\Lambda(0) = 0$ as an eigenvalue of algebraic multiplicity four and geometric multiplicity two, with the eigenspace spanned by the eigenvectors $(1\,0\,0\,0)^{\mathsf{T}}$ and $(0\,1\,0\,0)^{\mathsf{T}}$. This behaviour is quite different from that encountered for waves in the physically relevant regime $0 < |k| < 64$.

Lemma. *The matrix $M(k)$ has four distinct purely imaginary eigenvalues for $0 < |k| < 64$.*

Proof. Note that $\Lambda(k) \in \mathbb{C}$ is an eigenvalue of $M(k)$ with eigenvector $(v_1\ v_2\ v_3\ v_4)^{\mathsf{T}}$ if and only if $\lambda(k) = \frac{i\Lambda(k)}{k}$ is an eigenvalue with corresponding eigenvector $(v_1\ v_2\ ikv_3\ ikv_4)^{\mathsf{T}}$ of the real matrix

$$R(k) = \begin{pmatrix} \gamma h - \mu\,\Theta(k) & -\mu_1\,\mathrm{sech}(h_1k)\,\Theta(k) & \Theta(k) & \mathrm{sech}(h_1k)\,\Theta(k) \\ -\mu\,\mathrm{sech}(h_1k)\,\Theta(k) & \Gamma_1 - \mu_1\Theta_1(k) & \mathrm{sech}(h_1k)\,\Theta(k) & \Theta_1(k) \\ -\Gamma + \mu^2\,\Theta(k) & \mu\mu_1\mathrm{sech}(h_1k)\,\Theta(k) & \gamma h - \mu\,\Theta(k) & -\mu\,\mathrm{sech}(h_1k)\,\Theta(k) \\ \mu\mu_1\mathrm{sech}(h_1k)\,\Theta(k) & -\omega\Gamma_1 + g + \mu_1^2\,\Theta_1(k) & -\mu_1\,\mathrm{sech}(h_1k)\,\Theta(k) & \Gamma_1 - \mu_1\Theta_1(k) \end{pmatrix}.$$

To investigate the roots of the characteristic polynomial $p(\lambda) = \det(R - \lambda I_4)$, we first perform two sets of operations to simplify its structure. Add the first row multiplied by μ to the third row, and the second row multiplied by μ_1 to the fourth row, and in the outcome add the third column multiplied by μ to the first column and the fourth column multiplied by μ_1 to the second column to obtain a determinant edxpressed in terms of $X = \lambda - \gamma h$, corresponding to the wave speed relative to the maximum speed of the EUC, as

$$\hat{p}(X) = \begin{vmatrix} -X & 0 & \Theta & \mathfrak{s}\Theta \\ 0 & \gamma_1 h_1 - X & \mathfrak{s}\Theta & \Theta_1 \\ -\Gamma - 2\mu X & 0 & -X & 0 \\ 0 & \hat{g} - 2\mu_1 X & 0 & \gamma_1 h_1 - X \end{vmatrix}, \tag{4.23}$$

where $\hat{g} = g + \gamma_1^2 h_1 - \omega\gamma h$ and $\mathfrak{s} = \mathfrak{s}(k) = \mathrm{sech}(h_1 k)$. Expanding the above determinant by minors along the first column yields

$$\hat{p}(0) = -\Gamma\Theta\left\{\hat{g}(\Theta_1 - \mathfrak{s}^2\Theta) - \gamma_1^2 h_1^2\right\} > 0 \tag{4.24}$$

since $\Theta > 0 > \Gamma$ and the inequality $s < (1+s)\tanh(s)$ for $s > 0$, in combination with the relation

$$\Theta_1 - \mathfrak{s}^2\Theta = \frac{\tanh(h_1 k)}{k}, \tag{4.25}$$

yields $\gamma_1^2 h_1^2/(\Theta_1 - \mathfrak{s}^2\Theta) = \gamma_1^2 h_1 h_1 k/\tanh(h_1 k) < \gamma_1^2 h_1 (1 + h_1|k|) < \hat{g}$ for $|k| < 64$. On the other hand,

$$\hat{p}\Big(-\frac{\Gamma}{2\mu}\Big) = \frac{\Gamma^2}{4\mu^2}\left\{\Big(\gamma_1 h_1 + \frac{\Gamma}{2\mu}\Big)^2 - \Theta_1\Big(\hat{g} + \frac{\mu_1\Gamma}{\mu}\Big)\right\} < 0 \tag{4.26}$$

since the inequality $\tanh(s) < s < (1+s)\tanh(s)$ for $s > 0$ yields

$$h + h_1 > \frac{\tanh(hk)}{k} + \frac{\tanh(h_1 k)}{k} > \Theta_1(k) > \frac{\tanh(h_1 k)}{k} > \frac{h_1}{1 + h_1|k|}, \tag{4.27}$$

while $\mu > 0 > \mu_1$, $\Gamma < 0$, and

$$\Theta_1\Big(\hat{g} + \frac{\mu_1\Gamma}{\mu}\Big) > \Theta_1\hat{g} > \frac{h_1\hat{g}}{1 + |k|h_1} > \gamma_1^2 h_1^2 > \Big(\gamma_1 h_1 + \frac{\Gamma}{2\mu}\Big)^2 \quad \text{for} \quad |k| < 64.$$

Note that

$$h > \frac{\tanh(kh)}{k} > \Theta(k) > \frac{\tanh(hk)}{k(2+r)} = \frac{h}{2+r}\,\frac{\tanh(hk)}{hk} > \frac{h}{(2+r)(1+h|k|)}. \tag{4.28}$$

An expansion of (4.23) by minors along the first column yields

$$\hat{p}(-\mu\Theta B) = -(\hat{g} + 2\mu\mu_1\Theta B)\mathfrak{s}^2\mu^2\Theta^3 B^2 < 0, \tag{4.29}$$

since $B = 1 + \sqrt{1 - \Gamma/(\mu^2\Theta)} > 2$ satisfies $B(2-B)\mu^2\Theta = \Gamma$ and $\hat{g} + 2\mu\mu_1\Theta B > 0$.

We conclude from $-\mu\Theta B < 0 < -\Gamma/(2\mu)$ and (4.24), (4.26), (4.29), that the quartic (4.23), with leading term X^4, has four distinct real roots, two positive and two negative. □

4.1.1. Long waves. We now describe the main features of the linear wave propagation under the over-arching assumption of long waves, characterised by wavelengths in excess of 16 km and corresponding roughly to $|k| < 0.2$.

The lemma in the preamble of this section validates the applicability of perturbation theory. Using the approximations

$$\mathfrak{s}(k) \approx 1, \qquad \Theta(k) \approx \frac{h}{1+r}, \qquad \Theta_1(k) \approx \frac{h}{1+r} + h_1, \tag{4.30}$$

for $k \to 0$, in combination with

$$-\Gamma \approx rg, \qquad \hat{g} \approx g + \gamma_1^2 h_1, \qquad 2\mu \approx (1+r)\gamma - \gamma_1, \qquad 2\mu_1 \approx \gamma_1, \tag{4.31}$$

we can approximate the determinant (4.23) by

$$p_0(X) = \begin{vmatrix} -X & 0 & \frac{h}{1+r} & \frac{h}{1+r} \\ 0 & \gamma_1 h_1 - X & \frac{h}{1+r} & \frac{h}{1+r} + h_1 \\ rg + [\gamma_1 - (1+r)\gamma]X & 0 & -X & 0 \\ 0 & g + \gamma_1^2 h_1 - \gamma_1 X & 0 & \gamma_1 h_1 - X \end{vmatrix}.$$

An expansion of the above determinant by minors along the first column yields the polynomial

$$p_0(X) = X^4 - (\gamma_1 h_1 - \gamma h)X^3 - [g(h+h_1) + \gamma\gamma_1 hh_1]X^2 + ghh_1(\gamma_1 - \gamma)X + \frac{rg^2 hh_1}{1+r}. \tag{4.32}$$

While the intricacy of the coefficients prevents us from using the explicit quartic formulas to gain insight into the location of the roots of p_0, we can nevertheless take advantage of the approach used to reduce the solution of the quartic to solving the resolvent cubic. More precisely, we seek real numbers A, α and β such that

$$p_0(X) = \left(X^2 - \frac{\gamma_1 h_1 - \gamma h}{2} X + A\right)^2 - (\alpha X + \beta)^2, \tag{4.33}$$

a process factorising the quartic into a product of two quadratics whose roots are readily available. Expanding the brackets on the above right side and comparing the coefficients with (4.32), we get

$$\beta^2 = A^2 - \frac{rg^2 hh_1}{1+r}, \tag{4.34}$$

$$\alpha^2 = 2A + g(h+h_1) + \frac{(\gamma_1 h_1 + \gamma h)^2}{4}, \tag{4.35}$$

$$2\alpha\beta = -A(\gamma_1 h_1 - \gamma h) - ghh_1(\gamma_1 - \gamma), \tag{4.36}$$

and the two available expressions for $4\alpha^2\beta^2$ yield a cubic equation for A,

$$4\left\{2A + g(h+h_1) + \frac{(\gamma_1 h_1 + \gamma h)^2}{4}\right\}\left\{A^2 - \frac{rg^2 hh_1}{1+r}\right\} - \left\{A(\gamma_1 h_1 - \gamma h) + ghh_1(\gamma_1 - \gamma)\right\}^2 = 0. \tag{4.37}$$

It is a daunting task to use Cardano's formulas to identify a real root A of the cubic (4.37) so that α^2 and β^2, defined in (4.34)–(4.35), are positive. It turns out that a computational approach can be avoided by using the available structure. Indeed, in terms of $a = A + g(h+h_1)/2$, we can write the cubic $\mathcal{A}(a)$ on the left side of (4.37) in the form

$$4\left\{2a + \frac{(\gamma_1 h_1 + \gamma h)^2}{4}\right\}\left\{\left[a - \frac{g(h+h_1)}{2}\right]^2 - \frac{rg^2 hh_1}{1+r}\right\} - \left\{a(\gamma_1 h_1 - \gamma h) + \frac{g(h-h_1)}{2}(\gamma_1 h_1 + \gamma h)\right\}^2.$$

Since

$$\mathcal{A}(0) = g^2(\gamma_1 h_1 + \gamma h)^2\left\{\frac{(h+h_1)^2}{4} - \frac{rhh_1}{1+r} - \frac{(h-h_1)^2}{4}\right\} = g^2(\gamma_1 h_1 + \gamma h)^2 \frac{hh_1}{1+r} > 0,$$

we see that $\mathcal{A}(-[\gamma_1 h_1 + \gamma h]^2/8) < 0 < \mathcal{A}(0)$. Thus $\mathcal{A}$ has a negative root $a_0 > -(\gamma_1 h_1 + \gamma h)^2/8$ and the cubic (4.37) has a root A_0 such that

$$-\frac{g(h+h_1)}{2} - \frac{(\gamma_1 h_1 + \gamma h)^2}{8} < A_0 = a_0 - \frac{g(h+h_1)}{2} < -\frac{g(h+h_1)}{2}. \tag{4.38}$$

Since $\gamma_1 h_1 + \gamma h < 0$, corresponding to A_0 we get from (4.35) a value $\alpha_0 < 0$ with

$$\frac{\gamma_1 h_1 + \gamma h}{2} < \alpha_0 < 0, \tag{4.39}$$

while (4.34) yields a value $\beta_0 > 0$ with

$$g\sqrt{\Big(\frac{h+h_1}{2}\Big)^2 - \frac{rhh_1}{1+r}} < \beta_0 < g\sqrt{\Big(\frac{h+h_1}{2} + \frac{(\gamma_1 h_1 + \gamma h)^2}{8g}\Big)^2 - \frac{rhh_1}{1+r}}. \tag{4.40}$$

Writing now (4.33) in the form

$$p_0(X) = \Big\{X^2 - \Big(\frac{\gamma_1 h_1 - \gamma h}{2} + \alpha_0\Big) X + A_0 - \beta_0\Big\}\Big\{X^2 - \Big(\frac{\gamma_1 h_1 - \gamma h}{2} - \alpha_0\Big) X + A_0 + \beta_0\Big\},$$

we see that the four real roots of the quartic p_0 are given by

$$X_{1,2} = \frac{\gamma_1 h_1 - \gamma h + 2\alpha_0}{4} \pm \sqrt{\frac{(\gamma_1 h_1 - \gamma h + 2\alpha_0)^2}{16} + \beta_0 - A_0}, \tag{4.41}$$

$$X_{3,4} = \frac{\gamma_1 h_1 - \gamma h - 2\alpha_0}{4} \pm \sqrt{\frac{(\gamma_1 h_1 - \gamma h - 2\alpha_0)^2}{16} - \beta_0 + A_0}. \tag{4.42}$$

By (3.17) and (3.27), the numerical range of the relevant physical constants is

$$h \approx 8, \qquad h_1 \approx 0.24, \qquad \gamma h \approx 2, \qquad \gamma_1 h_1 \approx -3, \qquad g \approx 2 \times 10^4, \qquad r \approx 10^{-3}. \tag{4.43}$$

We will first obtain approximations of the roots (4.41)–(4.42) taking only into account the relative sizes of the above parameters, more precisely, the fact that

$$g \gg 1 \quad \text{while} \quad h + h_1,\ \gamma h,\ \gamma_1 h_1,\ rg \quad \text{are all} \quad O(1). \tag{4.44}$$

The specification (4.43) singles out the details for the ocean dynamics in the equatorial Pacific but other values, subject to (4.44), are relevant for flows in the equatorial regions of the Atlantic and Indian ocean, respectively.

The estimates (4.38), (4.39) and (4.40) yield

$$X_{1,2} \approx \pm\sqrt{g(h+h_1)}, \tag{4.45}$$

while Viète's relations enable us to infer from (4.32) that

$$\sum_{n=1}^{4} X_n^2 = \Big(\sum_{n=1}^{4} X_n\Big)^2 - 2 \sum_{m,n=1, n\neq m}^{4} X_m X_n = 2g(h+h_1) + \gamma_1^2 h_1 + \gamma^2 h^2,$$

so that $X_3^2 + X_4^2 \approx \gamma_1^2 h_1^2 + \gamma^2 h^2$. Using now (4.32) to write $p_0(X) = 0$ for $X = X_3$ and $X = X_4$ as

$$\frac{X^4 - (\gamma_1 h_1 - \gamma h)X^3 - \gamma\gamma_1 hh_1 X^2}{g} - \Big\{(h+h_1)X^2 - hh_1(\gamma_1 - \gamma)X - \frac{rghh_1}{1+r}\Big\} = 0,$$

we can approximate $X_{3,4}$ by the roots of the quadratic $\left\{(h+h_1)X^2 - hh_1(\gamma_1-\gamma)X - \frac{rghh_1}{1+r}\right\}$, that is,

$$X_{3,4} \approx \frac{(\gamma_1-\gamma)hh_1}{2(h+h_1)} \pm \sqrt{\frac{(\gamma_1-\gamma)^2h^2h_1^2}{4(h+h_1)^2} + \frac{rghh_1}{(1+r)(h+h_1)}}. \tag{4.46}$$

Since $c = \lambda = \gamma h + X$ and $\overline{c} = c\,\overline{U}_0$, we obtain from (3.17)–(3.18) that the dispersion relations corresponding to (4.45) and (4.46) are, respectively,

$$\overline{c_{1,2}} \approx \overline{\gamma}\,\overline{h} \pm \sqrt{\overline{g}(\overline{h}+\overline{h}_1)}, \tag{4.47}$$

and

$$\overline{c_{3,4}} \approx \overline{\gamma}\,\overline{h} + \frac{(\overline{\gamma}_1-\overline{\gamma})\overline{h}\,\overline{h}_1}{2(\overline{h}+\overline{h}_1)} \pm \sqrt{\frac{(\overline{\gamma}_1-\overline{\gamma})^2\overline{h}^2\overline{h}_1^2}{4(\overline{h}+\overline{h}_1)^2} + \frac{r\overline{g}\,\overline{h}\,\overline{h}_1}{(1+r)(\overline{h}+\overline{h}_1)}}. \tag{4.48}$$

Note that a root X of the quartic $\hat{p}(X)$ is related by means of $\lambda = \gamma h + X$ to an eigenvalue λ of the matrix $R(k)$, and the corresponding eigenvector $\mathbb{E} = (E_1\ E_2\ E_3\ E_4)^{\intercal}$ corresponds to the eigenvector $(E_1\ E_2\ \frac{E_3}{ik}\ \frac{E_4}{ik})^{\intercal}$ for the eigenvalue $\Lambda(k) = -ik\lambda$ of the matrix $M(k)$. If $\mathbb{E}$ is an eigenvector of $R(k)$, then, multiplying the first equation of the system $R(k)\mathbb{E} = \lambda\mathbb{E}$ by μ and adding it to the second equation yields

$$E_3 = -\left(\mu + \frac{\Gamma}{X}\right)E_1,$$

while the second equation multiplied by μ_1 and added to the fourth equation yields

$$E_4 = \frac{\mu_1(X-\gamma_1h_1) + \omega\Gamma_1 - g}{\gamma_1h_1 - X}\,E_2,$$

and the first equation reads

$$-(\mu\Theta + X)E_1 - \mu_1\mathfrak{s}\Theta E_2 + \Theta E_3 + \mathfrak{s}\Theta E_4 = 0.$$

Using the first two relations in the third yields

$$E_1 = \frac{\mathfrak{s}\Theta\{g - \omega\Gamma_1 - 2\mu_1(X-\gamma_1h_1)\}X}{(X-\gamma_1h_1)[X^2 + \Theta(2\mu X + \Gamma)]}\,E_2. \tag{4.49}$$

Taking (4.30) and (4.31) into account, (4.49) becomes

$$E_1 \approx \frac{h\{g - \gamma_1(X-\gamma_1h_1)\}X}{(X-\gamma_1h_1)\{X^2 + (\gamma-\gamma_1)hX - rgh\}}\,E_2. \tag{4.50}$$

For $X = X_{1,2}$, (4.45) and (4.44) yield $E_1 \approx \frac{ghX}{X^3}\,E_2 \approx \frac{h}{h+h_1}\,E_2$, that is,

$$\eta \approx \eta_1. \tag{4.51}$$

while for $X = X_{3,4}$ the quadratic polynomial providing (4.46) yields $[(\gamma - \gamma_1)X - rg]hh_1 \approx -(h + h_1)X^2$, so that (4.50) simplifies to $E_1 \approx \frac{h_1\{g-\gamma_1(X-\gamma_1 h_1)\}}{X(X-\gamma_1 h_1)} E_2 \approx \frac{h_1 g}{X(X-\gamma_1 h_1)} E_2$, and in this case

$$\eta \gg \eta_1 ; \tag{4.52}$$

for example, for the values specified in (4.43), the ratio η/η_1 is about 500 (see Fig. 5). The above considerations prove the validity of the following result.

Theorem 3. *Within the framework of linear theory, long waves in slow-mode propagate at speeds given by the formula* (4.48)*, with the effects confined to the motion of the thermocline, due to* (4.52)*, while the propagation speed of the fast long-waves is given by* (4.47)*.*

We showed that the Paley-Wiener theorem applied to the linear system (4.19) ensures that a localised disturbance occurring in the equatorial mid-Pacific will generate a unique solution in the form of a localised wave perturbation of the underlying pure current background state, that we can regard as a superposition of periodic modes by means of (4.22). Since in the ocean, energy tends to be concentrated in the lower frequencies, the physically most relevant regime is that of long waves, in which the asymptotic procedure implemented above supresses high frequencies by acting like a low-pass filter. From (4.47) and (4.48), using the values in (4.43), we obtain four possible propagation speeds in this regime:

$$\overline{c_1} \approx 202 \text{ m s}^{-1}, \qquad \overline{c_2} \approx -200 \text{ m s}^{-1}, \tag{4.53}$$

$$\overline{c_3} \approx 1.56 \text{ m s}^{-1}, \qquad \overline{c_4} \approx -1.05 \text{ m s}^{-1}. \tag{4.54}$$

The speed values (4.53) correspond to the extremely rare event[2] of tsunami waves: if triggered by a submarine earthquake or by a meteorite impact, these would consist, in view of (4.51), of surface waves coupled with internal waves of roughly the same amplitude, propagating along the Equator eastwards at 727 km h^{-1} and westwards at 720 km h^{-1}. While these fast waves are exceptional, note that on Sunday, 22 May 1960, several powerful earthquakes in rapid succession, occurred along 1000 km of fault parallel to the Chilean coastline, with the epicentre at about 200 km off the coast of Central Chile (see the discussion in [5]). Tsunami waves were generated which propagated across the Pacific Ocean and records of the travel time of the tsunami indicate a speed of about 720 km h^{-1}, which matches well our predictions. On the other hand, the two values (4.54) represent ubiquituous slow internal waves (known as eastward Kelvin waves and westward Rossby waves): (4.52) yields that the corresponding surface waves are insignificant, given that the typical amplitude of an internal wave is about 20 m. Note that the predicted speed values (4.54) for the slow internal waves fit reasonably well with reported field data (see the discussion in [7]). They highlight an important feature of the equatorial ocean dynamics in the Pacific: the eastward internal wave speed exceeds the

[2] While four out of every five tsunamis happen in the Pacific Ocean, initiated by earthquakes along the boundary of one of the Earth's main geologically active tectonic plates, those in the tropical Pacific tend to be modest in size. However, a few times every century, tsunami waves generated by great earthquakes in the North Pacific or along the Pacific coast of South America sweep across the entire Pacific. As for large asteriod impacts into the deep ocean, these are very rare—the only currently known deep-ocean impact event on Earth being the Eltanin asteroid impact and tsunami, ocurring about 2.5 million years ago when a asteriod exceeding 1 km in diameter did strike the ocean about 1500 km SSW of Chile (see [17]).

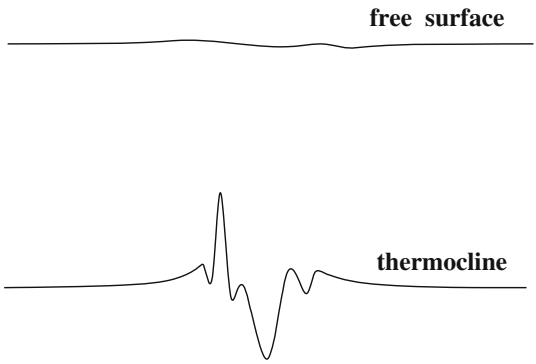

Fig. 5. Linear long-wave theory: coupled localised wave perturbations over a large spatial scale exist but the amplitude of the internal wave exceeds more than 500 times that of the surface wave. Typically detected oscillations of the thermocline from its mean level are about 10 m

maximum speed of the EUC, while the westward internal wave is slower than the maximum speed of the surface wind-driven current. Consequently, the dynamic response of the equatorial Pacific presents an east–west asymmetry due to the interaction between waves and the depth-dependent underlying currents.

4.1.2. Short waves. High-frequency ocean waves, with wavelengths of 50–100 m, correspond to the range $31 < |k| < 64$, in which the approximations $\tanh(h_1 k) \approx 1$ and $\tanh(kh) \approx 1$ are reasonable, so that $\Theta(k) \approx \frac{1}{(2+r)|k|}$ and $\Theta_1(k) \approx \frac{1}{|k|}$. On the other hand, the fact that $s \mapsto s/\sinh(s)$ is decreasing for $s > 0$ yields

$$\begin{aligned} 0 &< \max\{|\mu\mu_1|,\ |\mu k|,\ |\mu_1 k|,\ k^2\}\operatorname{sech}(h_1 k)\,\Theta(k) \\ &< k^2 \operatorname{sech}(h_1 k)\,\Theta(k) < \frac{k}{(2+r)\sinh(h_1 k)}; \end{aligned}$$

in particular, $\sinh(kh_1) > 817$ for $k > 31$ yields the upper bound 0.0189 above. Retaining only the leading order of each entry, we approximate the matrix $M(k)$ in (4.19) by

$$M_\infty(k) = \begin{pmatrix} -i\gamma hk & 0 & |k|/(2+r) & 0 \\ 0 & -i\Gamma_1 k & 0 & |k| \\ \Gamma & 0 & -i\gamma hk & 0 \\ 0 & \omega\Gamma_1 - g & 0 & -i\Gamma_1 k \end{pmatrix}, \tag{4.55}$$

with $\Gamma,\ \Gamma_1 < 0$. The characteristic polynomial

$$p_\infty(k, \Lambda) = \Big\{(\Lambda + i\gamma hk)^2 - \Gamma|k|/(2+r)\Big\}\Big\{(\Lambda + i\Gamma_1 k)^2 + (g - \omega\Gamma_1)|k|\Big\}$$

has the four disjoint purely imaginary eigenvalues

$$\Lambda_\pm(k) = -i\gamma hk \pm i\sqrt{|\Gamma k|/(2+r)}, \quad \Lambda^\pm(k) = -i\Gamma_1 k \pm i\sqrt{(g - \omega\Gamma_1)|k|} \tag{4.56}$$

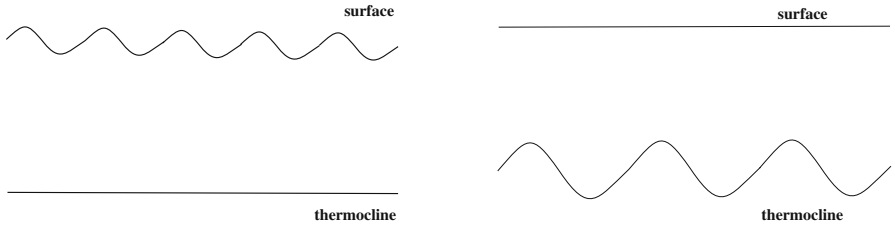

Fig. 6. Linear short-wave theory: there exist travelling periodic surface waves with speeds (4.58), with practically no recognisable effect at the thermocline (left), as well as travelling periodic internal waves with speeds (4.57), which are not detectable at the surface (right)

with eigenvectors $\mathfrak{u}_{\pm}(\xi) = \left(1\ 0\ \pm i\sqrt{\frac{(2+r)|\Gamma|}{|k|}}\ 0\right)^{\mathsf{T}}$ and $\mathfrak{u}^{\pm}(\xi) = \left(0\ 1\ 0\ \pm i\sqrt{\frac{g-\omega\Gamma_1}{|k|}}\right)^{\mathsf{T}}$, respectively. Consequently we have four types of high-frequency oscillations: two modes in which periodic sinusoidal disturbances of the thermocline propagate at the speeds

$$c_{\pm}(k) = \gamma h \mp \frac{|k|}{k}\sqrt{\frac{|\Gamma|}{(2+r)\,|k|}} \tag{4.57}$$

without any effect on the (flat) free surface, and two modes in which periodic harmonic oscillations of the free surface propagate at the speeds

$$c^{\pm}(k) = \Gamma_1 \mp \frac{|k|}{k}\sqrt{\frac{g-\omega\Gamma_1}{|k|}} \tag{4.58}$$

with an undisturbed (flat) thermocline (see Fig. 6). The appearance of dispersive effects is in marked contrast with the long waves: the initial disturbance can be thought of as being composed of the sum of a great many modes and, as time passes by, due to (4.57)–(4.58), the longer components (that is, those for which $|k|$ is relatively smaller) will propagate faster than the shorter modes.

Let us point out succinctly the practical meaning of the nondimensional formulas (4.57)–(4.58). In view of (3.17), the nondimensional travelling mode $e^{ik(x-ct)}$, of wavelength $\frac{2\pi}{|k|}$, corresponds in physical variables to a harmonic oscillation of wavelength $\frac{2\pi}{|k|}\,\overline{L}$, propagating at the speed $\overline{c} = c\,\overline{U_0}$. More precisely, due to (3.27), (3.18), (3.17), (3.82), (3.83) and (4.20), the dimensional counterpart of (4.57) for the wavelength $\mathfrak{L} = \frac{2\pi}{|k|}\,\overline{L}$ is

$$\overline{c_{\pm}}(\mathfrak{L}) = \overline{\gamma}\,\overline{h} \pm \sqrt{\frac{r\,(\overline{g} - 2\overline{\Omega}\,\overline{\gamma}\,\overline{h})}{2+r}\,\frac{\mathfrak{L}}{2\pi}} \approx \overline{\gamma}\,\overline{h} \pm \sqrt{\frac{r\overline{g}\mathfrak{L}}{2\pi(2+r)}}, \tag{4.59}$$

while that of (4.58) is

$$\overline{c^{\pm}}(\mathfrak{L}) = (\overline{\gamma}_1\overline{h}_1 + \overline{\gamma}\,\overline{h}) \pm \sqrt{\left(\overline{g} - 2\overline{\Omega}\,[\overline{\gamma}_1\overline{h}_1 + \overline{\gamma}\,\overline{h}]\right)\frac{\mathfrak{L}}{2\pi}} \approx (\overline{\gamma}_1\overline{h}_1 + \overline{\gamma}\,\overline{h}) \pm \sqrt{\frac{\overline{g}\mathfrak{L}}{2\pi}}. \tag{4.60}$$

The previous considerations prove the following result.

Theorem 4. *High-frequency linear waves can propagate as oscillations of the free surface and of the thermocline, with the corresponding speeds, provided by the formulas* (4.60) *and* (4.59)*, respectively, imparting a dispersive character to the flow. The propagation modes (at the surface or at the termocline) are decoupled, with no noticeable effect occurring at the other interface.*

Note that formula (4.60) predicts that sinusoidal surface waves with a wavelength $\mathfrak{L} = 100$ m (corresponding to $k = 10\pi$) propagate at speeds

$$\overline{c^+}(100) \approx 12\ \mathrm{m\,s^{-1}}, \quad \overline{c^-}(100) \approx -13\ \mathrm{m\,s^{-1}}.$$

These values of the wavelength and of the wave speeds are realistic for wind-generated surface waves in the Pacific (see [23]). On the other hand, by (4.59), the predicted propagation speeds of sinusoidal internal waves with a wavelength of 100 m are

$$\overline{c_+}(100) \approx 1.28\ \mathrm{m\,s^{-1}}, \quad \overline{c_-}(100) \approx 0.72\ \mathrm{m\,s^{-1}}.$$

Such values, close to the maximal speed of the Equatorial Undercurrent, are reported in field data (see the discussion in [7]).

4.1.3. Solitary waves. For the linear system (4.18) the existence of a smooth localised wave solution $\mathfrak{u}(x - c_0 t)$ which propagates without change of shape at the constant speed c_0 is equivalent, upon applying the Fourier transform, to $\Lambda(k) = -ikc_0$ being an eigenvalue of $M(k)$ with corresponding eigenvector $\hat{\mathfrak{u}}(k)$ for every $k \in \mathbb{R}$ for which $\hat{\mathfrak{u}}(k) \neq 0$. In physical terms this means that the internal solitary wave profile $\eta \in \mathcal{S}(\mathbb{R})$ is expressed in Fourier integral form $\eta(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{\eta}(k)\, e^{ikx} dk$ as consisting of a superposition of harmonic wave components $\frac{1}{2\pi}\hat{\eta}(\xi)\, e^{ikx}$ of wavelength $2\pi/|k|$ and amplitude $\frac{1}{2\pi}|\hat{\eta}(k)|$, each travelling out at the same velocity so that the shape of the initial localised disturbance is not altered as time passes by, the only effect being a mere translation in the direction of wave propagation. We now prove that linear theory does not capture the solitary wave phenomenon.

Theorem 5. *The linearised problem does not admit solitary wave solutions.*

Proof. The existence of a solitary wave corresponding to some non-trivial $\mathfrak{u} \in \mathcal{S}(\mathbb{R}^4)$ and some propagation speed $c_0 \neq 0$ would ensure that c_0 is a root of the characteristic polynomial $p(\lambda)$ associated to the real 4×4 matrix $R(k)$ for all k in some open interval (a, b) where $\hat{\mathfrak{u}}(k) \neq 0$. But since all entries of $R(k)$ are real-analytic functions of the variable k, this would mean that the real-analytic function $k \mapsto p(c_0) = \det[R(k) - c_0 I]$ vanishes for $k \in (a, b)$. Consequently it must be identically zero, so that $p(c_0) = 0$ for all $k \in \mathbb{R}$. Since $\lim_{k\to\infty} \operatorname{sech}(h_1 k) = \lim_{k\to\infty} \Theta(k) = \lim_{k\to\infty} \Theta_1(k) = 0$, letting $k \to \infty$ in the matrix $[R(k) - c_0 I]$ yields the singular matrix

$$\begin{pmatrix} \gamma h - c_0 & 0 & 0 & 0 \\ 0 & \Gamma_1 - c_0 & 0 & 0 \\ -\Gamma & 0 & \gamma h - c_0 & 0 \\ 0 & -\omega\Gamma_1 + g & 0 & \Gamma_1 - c_0 \end{pmatrix}.$$

Consequently we should have $c_0 = \gamma h$ or $c_0 = \Gamma_1$. Letting $k \to 0$ in the identity (4.24) yields

$$\lim_{k\to 0} p(\gamma h) = \frac{\Gamma h h_1(\gamma_1^2 h_1 - \hat{g})}{1 + r} \neq 0,$$

which rules out the possibility $c_0 = \gamma h$. On the other hand, from (4.23) we can express $\lim_{k\to 0} p(\Gamma_1)$ expanding by minors along the last row yields to obtain the value

$$(\omega\Gamma_1 - g)\left\{\frac{(\Gamma + 2\mu\gamma_1 h_1)hh_1}{1+r} + \Big(\frac{h}{1+r} + h_1\Big)\gamma_1^2 h_1^2\right\} \neq 0,$$

so that $c_0 \neq \Gamma_1$. The obtained contradiction completes the proof. □

4.2. Weakly nonlinear models. Even the most powerful computers cannot handle the full range of space and time scales that must be resolved to successfully simulate numerically the real-world equatorial flows. Model equations and conceptual simplifications can be brought to bear on the problem, thereby providing us with a considerable level of understanding of various physically relevant regimes for equatorial wave–current interactions. The results of the previous subsection show that linear models are not sufficient to capture the solitary wave phenomenon. Recall that linear theory corresponds to retaining only quadratic terms in the expansion of the Hamiltonian, with the expansion parameter describing the amplitude of the wave perturbation of a background pure current state. Weakly nonlinear models can be derived by expanding the Hamiltonian up to cubic terms in the wave amplitude, thereby permitting nonlinear interactions to become relevant.

It is more convenient to write the governing Hamiltonian system (3.86) in terms of the variables $(q,\ q_1,\ \eta,\ \eta_1)$, where

$$q = \zeta_x, \qquad q_1 = \zeta_{1,x}. \tag{4.61}$$

Using (3.82)–(3.83), (3.69), (3.58), (3.60), (3.43) and (3.13) we get

$$\frac{q - (\mu + \gamma - \gamma_1)\eta}{r} = \nabla\varphi\Big|_{z=\eta(x,t)} \cdot (1, \eta_x), \qquad q_1 - \mu_1\eta_1 = \nabla\varphi_1\Big|_{z=h_1+\eta_1(x,t)} \cdot (1, \eta_x),$$

so that $[q - (\mu + \gamma - \gamma_1)\eta|/r$ and $q_1 - \mu_1\eta_1$ are the tangential velocities of the wave perturbations at the interface and free surface, respectively. Therefore the variables $(q,\ q_1,\ \eta,\ \eta_1)$ basically encode the profiles and the propagation speeds of the interface and of the free surface, whilst taking into account the vorticity values above and below the thermocline as well as the effect of the Earth's rotation. Note that (3.38) in combination with (3.82)–(3.83) ensure $q,\ q_1 \in \mathcal{S}(\mathbb{R})$. Since for $q = \zeta_x$ and $\tilde{q} = \tilde{\zeta}_x$, using integration by parts, we see that

$$\frac{\partial\mathcal{H}}{\partial\zeta} = -\partial_x\frac{\partial\mathcal{H}}{\partial q}, \qquad \frac{\partial\mathcal{H}}{\partial\zeta_1} = -\partial_x\frac{\partial\mathcal{H}}{\partial q_1}.$$

Similarly, since (3.82)–(3.83) yield $\xi_x = q - \mu\eta$ and $\xi_{1,x} = q_1 - \mu_1\eta_1$, we get

$$\frac{\partial\mathcal{H}}{\partial\xi} = -\partial_x\frac{\partial\mathcal{H}}{\partial\xi_x} = -\partial_x\frac{\partial\mathcal{H}}{\partial q} + \mu\,\partial_x\frac{\partial\mathcal{H}}{\partial\eta}, \qquad \frac{\partial\mathcal{H}}{\partial\xi_1} = -\partial_x\frac{\partial\mathcal{H}}{\partial\xi_{1,x}} = -\partial_x\frac{\partial\mathcal{H}}{\partial q_1} + \mu_1\,\partial_x\frac{\partial\mathcal{H}}{\partial\eta_1}. \tag{4.62}$$

Therefore, in terms of the new variables, the Hamiltonian system (3.86) takes on the form

$$\frac{d}{dt}\begin{pmatrix} q \\ q_1 \\ \eta \\ \eta_1 \end{pmatrix} = -\partial_x \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \partial\mathcal{H}/\partial q \\ \partial\mathcal{H}/\partial q_1 \\ \partial\mathcal{H}/\partial\eta \\ \partial\mathcal{H}/\partial\eta_1 \end{pmatrix}. \tag{4.63}$$

4.2.1. KdV equation describing the evolution of the thermocline. Taking $\varepsilon \ll 1$ and $\delta \ll 1$ to be small parameters, we consider the shallow-water (long wave) regime defined by the spatial scale

$$x' = \varepsilon x, \tag{4.64}$$

and the wave perturbation scales

$$\begin{cases} \eta(x,t) = \varepsilon^2 \eta'(x',t), & q(x,t) = \varepsilon^2 q'(x',t), \\ \eta_1(x,t) = \delta\varepsilon^2 \eta_1'\Big(x', \dfrac{t}{\delta}\Big), & q_1(x,t) = \delta\varepsilon^2 q_1'\Big(x', \dfrac{t}{\delta}\Big), \end{cases} \tag{4.65}$$

in which η and q, as well as η_1 and q_1, are considered of the same order of magnitude, with the surface wave having a smaller amplitude than the internal wave—an impulse that generates wave perturbations of the background pure current state will typically have a pronounced oscillatory effect at the thermocline[3] whereas the sea surface is hardly affected at all.[4] The interpretation of the scaling (4.65), in which the typical internal wave amplitude ε^2 and wavelength ε^{-1} are in quadratic balance, is that x', t, η', q', η_1' and q_1' are all of order $O(1)$, with (4.64) specifying the distance needed to bring about this balance for the dynamics of the flow. Note that the nondimensionalisation performed in Sect. 3.2 fixed the reference scales of 500 m for length and 0.5 m s^{-1} for speed, while typical amplitudes and propagation speeds of equatorial internal waves do not exceed 50 m and 0.1 m s^{-1}, respectively. This is consistent with the choice $\varepsilon \ll 1$ in (4.65). The time scale at the surface, introduced in (4.65), captures the fact that changes in the surface wave profile occur faster than at the thermocline; note also that the discussion of Hamiltonian perturbations at the beginning of Sect. 4 shows that to use different spatial scales for the two interfaces requires a different adjustment of the two time scales. The scaling (4.64)–(4.65) introduces the small parameters ε and δ into the Hamiltonian (3.73), transforming (4.63) accordingly into the Hamiltonian system

$$\frac{d}{dt}\begin{pmatrix} q' \\ q_1' \\ \eta' \\ \eta_1' \end{pmatrix} = -\partial_x \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \partial\mathcal{H}'/\partial q' \\ \partial\mathcal{H}'/\partial q_1' \\ \partial\mathcal{H}'/\partial \eta' \\ \partial\mathcal{H}'/\partial \eta_1' \end{pmatrix}, \tag{4.66}$$

with Hamiltonian $\mathcal{H}' = \mathcal{H}\varepsilon^{-3}$.

Due to (4.8), (4.9), (4.13), (4.14), under the scaling (4.65) the Taylor expansion for the Dirichlet–Neumann operator $G(\eta)$ is

$$\begin{aligned} G(\eta') = {} & \varepsilon^2 h (D')^2 + \varepsilon^4 \Big\{ -\tfrac{1}{3} h^3 (D')^4 + D'\eta'(x') D' \Big\} \\ & + \varepsilon^6 \Big\{ -\tfrac{2}{15} h^5 (D')^6 - h^2 (D')^2 \eta'(x') (D')^2 \Big\} + O(\varepsilon^8) \end{aligned} \tag{4.67}$$

[3] The impulse causes the denser water to rise above the flat equilibrium thermocline level, but gravity pulls it back down because the water beneath the thermocline is heavier than that above it, and the inertia acquired during the falling movement causes the water to penetrate below the level of equilibrium of the thermocline, thus setting up an oscillatory motion. This wave disturbance propagates horizontally since the local increase in pressure that occurs where a parcel of water rises above the equilibrium level pushes the fluid below it away from that place, generating in this process an acceleration (and thus also a force) which also comprises a horizontal component.

[4] Since the internal restoring buoyancy force is much smaller than that at the surface, the density variation across the thermocline being about one-thousandth of that across the air-water interface.

since $\tanh(s) = s - \frac{1}{3}s^3 + \frac{2}{15}s^5 + O(s^7)$ and for $j \geq 2$ the terms $G^{(j)}(\eta')$ are of order $O(\varepsilon^8)$, while

$$\begin{aligned} G_{11}(\eta',\,\eta_1') = {} & \tfrac{1}{h_1} + \varepsilon^2\Big\{\tfrac{1}{3}h_1(D')^2 + \tfrac{1}{h_1^2}\eta'(x')\Big\} + \varepsilon^4\Big\{-\tfrac{1}{45}h_1^3(D')^4 \\ & + \tfrac{1}{3}\eta'(x')(D')^2 + \tfrac{1}{3}(D')^2\eta'(x')\Big\} \\ & - \varepsilon^2\delta\,\tfrac{1}{h_1^2}\eta_1'(x') + \varepsilon^4\delta\Big\{\tfrac{1}{6}\eta_1'(x')(D')^2 + \tfrac{1}{6}(D')^2\eta_1'(x')\Big\} + O(\varepsilon^6), \end{aligned} \tag{4.68}$$

since $\coth(s) = s^{-1} + \frac{1}{3}s - \frac{1}{45}s^3 + O(s^5)$ and $\operatorname{csch}(s) = s^{-1} - \frac{1}{6}s + \frac{7}{360}s^3 + O(s^5)$. Using (4.67)–(4.68) for the asymptotic expansion of the operator B, defined in (3.68), gives

$$\begin{aligned} B(\eta',\,\eta_1') = {} & \tfrac{1+r}{h_1} + \varepsilon^2\Big\{h(D')^2 + \tfrac{1+r}{3}h_1(D')^2 + \tfrac{1+r}{h_1^2}\eta'(x')\Big\} - \varepsilon^2\delta\,\tfrac{1+r}{h_1^2}\eta_1'(x') \\ & - \varepsilon^4\Big\{\tfrac{1}{3}h^3(D')^4 - D'\eta'(x')D' \\ & + \tfrac{1+r}{45}h_1^3(D')^4 - \tfrac{1+r}{3}\eta'(x')(D')^2 - \tfrac{1+r}{3}(D')^2\eta'(x')\Big\} \\ & + \varepsilon^4\delta\Big\{\tfrac{1+r}{6}\eta_1'(x')(D')^2 + \tfrac{1+r}{6}(D')^2\eta_1'(x')\Big\} + O(\varepsilon^6), \end{aligned} \tag{4.69}$$

and, with respect to this expansion, the inverse operator is

$$\begin{aligned} [B(\eta',\,\eta_1')]^{-1} = {} & \tfrac{h_1}{1+r} - \tfrac{\varepsilon^2 h_1^2}{(1+r)^2}\Big\{h(D')^2 + \tfrac{1+r}{3}h_1(D')^2 + \tfrac{1+r}{h_1^2}\eta'(x')\Big\} \\ & + \tfrac{\varepsilon^2\delta}{1+r}\eta_1'(x') + O(\varepsilon^4). \end{aligned} \tag{4.70}$$

Similarly we see that

$$G_{12}(\eta',\,\eta_1') = -\tfrac{1}{h_1} + \varepsilon^2\Big\{-\tfrac{1}{h_1^2}\eta' + \tfrac{h_1}{6}(D')^2\Big\} + O(\varepsilon^3), \tag{4.71}$$

$$G_{21}(\eta',\,\eta_1') = -\tfrac{1}{h_1} + \varepsilon^2\Big\{-\tfrac{1}{h_1^2}\eta' + \tfrac{h_1}{6}(D')^2\Big\} + O(\varepsilon^3), \tag{4.72}$$

$$G_{22}(\eta',\,\eta_1') = \tfrac{1}{h_1} + \varepsilon^2\Big\{\tfrac{1}{h_1^2}\eta' + \tfrac{h_1}{3}(D')^2\Big\} + O(\varepsilon^3). \tag{4.73}$$

Let us now choose $\delta = O(\varepsilon)$ in (4.64). Writing the Hamiltonian (3.73) as functionally dependent on η', η_1', ξ', ξ_1', we retain terms up to order $O(\varepsilon^5)$ and thus reduce the Hamiltonian (3.73) to

$$\begin{aligned} H^{(5)} = {} & \varepsilon^6\,\tfrac{(1+r)\gamma^2-\gamma_1^2}{6}\int_{\mathbb{R}}(\eta')^3\,\frac{dx'}{\varepsilon} + \varepsilon^4\,\tfrac{rg+\gamma h[(1+r)\gamma-\gamma_1]}{2}\int_{\mathbb{R}}(\eta')^2\,\frac{dx'}{\varepsilon} \\ & + \varepsilon^4\delta^2\,\tfrac{g+\gamma_1(\gamma h+\gamma_1 h_1)}{2}\int_{\mathbb{R}}(\eta_1')^2\,\frac{dx'}{\varepsilon} \\ & - \varepsilon^4\,\gamma\int_{\mathbb{R}}(\varepsilon^2\eta'+h)\eta_{x'}'\xi'\,\frac{dx'}{\varepsilon} - \varepsilon^4\delta^2(\gamma h+\gamma_1 h_1)\int_{\mathbb{R}}\eta_{1,x'}'\xi_1'\,\frac{dx'}{\varepsilon} \\ & + \varepsilon^4\delta\,\tfrac{h}{1+r}\int_{\mathbb{R}}\xi'(D')^2\xi_1'\,\frac{dx'}{\varepsilon} \end{aligned}$$

$$+\varepsilon^4 \tfrac{h}{2(1+r)} \int_{\mathbb{R}} \xi'(D')^2\xi' \frac{dx'}{\varepsilon} - \varepsilon^6 \tfrac{h^3}{3(1+r)} \int_{\mathbb{R}} \xi'(D')^4\xi' \frac{dx'}{\varepsilon}$$
$$+\varepsilon^6 \tfrac{1}{2(1+r)} \int_{\mathbb{R}} \xi' D'\eta'(x')D'\,\xi' \frac{dx'}{\varepsilon}$$
$$-\varepsilon^6 \tfrac{h^2h_1}{(1+r)^2} \int_{\mathbb{R}} \xi'(D')^4\,\xi' \frac{dx'}{\varepsilon} + \varepsilon^4\delta^2(\tfrac{h_1}{2} + \tfrac{h}{2(1+r)}) \int_{\mathbb{R}} \xi_1'(D')^2\,\xi_1' \frac{dx'}{\varepsilon}$$

since (4.64) and (3.82)–(3.83) yield $\xi = \varepsilon\xi'$ and $\xi_1 = \varepsilon^2\xi_1'$, and the only pseudo-differential operators for which we have to take into account one term beyond leading order are G_{22} and

$$GB^{-1}G_{11} = \varepsilon^2 \tfrac{h}{1+r}(D')^2 + \varepsilon^4\Big\{ -\tfrac{h^3}{3(1+r)}(D')^4$$
$$+ \tfrac{1}{1+r} D'\eta'(x')D' - \tfrac{h^2h_1}{(1+r)^2}(D')^4\Big\} + O(\varepsilon^5),$$
$$G_{21}B^{-1}G_{12} = \tfrac{1}{(1+r)h_1} + \varepsilon^2\Big\{\tfrac{1}{(1+r)h_1^2}\eta' - \tfrac{2h_1}{3(1+r)}(D')^2 - \tfrac{h}{(1+r)^2}(D')^2\Big\} + O(\varepsilon^3).$$

On the other hand, in analogy to (4.62), the relations

$$\partial_{x'}\xi' = q' - \mu\eta', \quad \partial_{x'}\xi_1' = q_1' - \mu_1\eta_1', \tag{4.74}$$

yield

$$\begin{aligned} H^{(5)}(q', q_1', \eta', \eta_1') = {} & \varepsilon^3 \frac{1}{2}\int_{\mathbb{R}} \left(\frac{h}{r+1}q'^2 - 2\frac{\mu h}{r+1}q'\eta' + A_1\eta'^2 + \delta^2(g-\omega\Gamma_1)(\eta_1')^2\right) dx' \\ & + \varepsilon^3\delta\,\frac{h}{r+1}\int_{\mathbb{R}} \left(q'q_1' - \mu\eta'q_1' - \mu_1\eta_1'q' + \mu\mu_1'\eta'\eta_1'\right) dx' \\ & + \varepsilon^3\gamma h\int_{\mathbb{R}} q'\eta'\,dx' + \varepsilon^3\delta^2\Gamma_1\int_{\mathbb{R}} q_1'\eta_1'\,dx' \\ & - \varepsilon^5\frac{A_2}{2}\int_{\mathbb{R}} \left((q_x')^2 - 2\mu q_{x'}'\eta_{x'}' + \mu^2(\eta_{x'}')^2\right) dx' \\ & + \varepsilon^5\frac{1}{2(1+r)}\int_{\mathbb{R}} \eta'(q')^2\,dx' + \varepsilon^5\left(\frac{\gamma}{2} - \frac{\mu}{r+1}\right)\int_{\mathbb{R}} q'(\eta')^2\,dx' \\ & + \varepsilon^5\frac{A_3}{6}\int_{\mathbb{R}} (\eta')^3\,dx' \\ & + \varepsilon^3\delta^2\frac{1}{2}\left(\frac{h}{r+1} + h_1\right)\int_{\mathbb{R}} \left((q_1')^2 - 2\mu_1 q_1'\eta_1' + \mu_1^2(\eta_1')^2\right) dx' \end{aligned} \tag{4.75}$$

where

$$A_1 = rg - r\omega\gamma h + \frac{h\mu^2}{r+1}, \quad A_2 = \frac{h^2[(1+r)h + 3h_1]}{3(1+r)^2},$$
$$A_3 = (1+r)\gamma^2 - \gamma_1^2 - 3\gamma\mu + \frac{3\mu^2}{r+1}. \tag{4.76}$$

We now neglect terms of order $O(\varepsilon^6)$ in the Hamiltonian $\mathcal{H}$, thus replacing (4.66) by a truncated Hamiltonian system. The canonical equations for the variables related to the free surface are linear

$$\eta'_{1,t} = -\partial_{x'}\left[\delta\Gamma_1\eta'_1 + \frac{h}{r+1}(q' - \mu\eta') + \delta\left(\frac{h}{r+1} + h_1\right)(q'_1 - \mu'_1\eta'_1)\right], \quad (4.77)$$

$$q'_{1,t} = -\partial_{x'}\left[\delta\Gamma_1 q'_1 - \delta\mu_1\left(\frac{h}{r+1} + h_1\right)(q'_1 - \mu_1\eta'_1) - \frac{\mu_1 h}{r+1}(q' - \mu\eta') + \delta(g - \omega\Gamma_1)\eta'_1\right], \quad (4.78)$$

while those for the variables at the thermocline are nonlinear

$$\eta'_t = -\partial_{x'}\left[\gamma h\eta' + \frac{h}{r+1}\Big(q' - \mu\eta' + \delta(q'_1 - \mu_1\eta'_1)\Big) + \varepsilon^2 A_2(q'_{x'x'} - \mu\eta'_{x'x'}) + \varepsilon^2\frac{1}{r+1}q'\eta' + \varepsilon^2\left(\frac{\gamma}{2} - \frac{\mu}{r+1}\right)\eta'^2\right], \quad (4.79)$$

$$q'_t = -\partial_{x'}\left[\gamma h q' - \frac{\mu h}{r+1}q' + A_1\eta' - \delta\frac{h\mu}{r+1}(q'_1 - \mu_1\eta'_1) - \varepsilon^2 A_2\mu(q'_{x'x'} - \mu\eta'_{x'x'}) + \varepsilon^2\frac{1}{r+1}\frac{q'^2}{2} + \varepsilon^2\left(\gamma - \frac{2\mu}{r+1}\right)\eta' q' + \varepsilon^2 A_3\frac{\eta'^2}{2}\right]. \quad (4.80)$$

For $\delta \ll \varepsilon^2$ we can neglect the δ-terms above and decouple the internal wave motion from the oscillations of the free surface: the system (4.77)–(4.80) simplifies to

$$\eta'_{1,t} = -\frac{h}{r+1}(q' - \mu\eta')_{x'}, \quad (4.81)$$

$$q'_{1,t} = -\frac{\mu_1 h}{r+1}(q' - \mu\eta')_{x'}, \quad (4.82)$$

$$\eta'_t = -\partial_{x'}\left[\gamma h\eta' + \frac{h}{r+1}(q' - \mu\eta') + \varepsilon^2 A_2(q'_{x'x'} - \mu\eta'_{x'x'}) + \varepsilon^2\frac{1}{r+1}q'\eta' + \varepsilon^2\left(\frac{\gamma}{2} - \frac{\mu}{r+1}\right)\eta'^2\right], \quad (4.83)$$

$$q'_t = -\partial_{x'}\left[\gamma h q' - \frac{\mu h}{r+1}q' + A_1\eta' - \varepsilon^2 A_2\mu(q'_{x'x'} - \mu\eta'_{x'x'}) + \varepsilon^2\frac{1}{r+1}\frac{q'^2}{2} + \varepsilon^2\left(\gamma - \frac{2\mu}{r+1}\right)\eta' q' + \varepsilon^2 A_3\frac{\eta'^2}{2}\right]. \quad (4.84)$$

Note that (4.81)–(4.82) yield $(q'_1 - \mu_1\eta'_1)_t = 0$; these equations also show that the motion of the free surface is determined by the initial data and the characteristics (η', q') of the displacements of the thermocline. On the other hand, from (4.83)–(4.84) we infer that the leading order linear equations for η' and q' are

$$\eta'_t = -\frac{h}{r+1}q'_{x'} + \left(\frac{h\mu}{r+1} - \gamma h\right)\eta'_{x'}, \quad (4.85)$$

$$q'_t = \left(\frac{h\mu}{r+1} - \gamma h\right)q'_{x'} - A_1\,\eta'_{x'}. \quad (4.86)$$

Therefore the wave speed c of a linear travelling wave, in which the (x', t)-dependence is solely in terms of $(x' - ct)$, satisfies the quadratic equation

$$\left(c - \gamma h + \frac{h\mu}{1+r}\right)^2 = \frac{h}{r+1} A_1 \tag{4.87}$$

which, in view of (4.76), has the solutions

$$c = \gamma h - \frac{h\mu}{r+1} \pm \sqrt{\left(\frac{h\mu}{r+1}\right)^2 + \frac{hr(g - \gamma h\omega)}{r+1}}. \tag{4.88}$$

Note that $g \gg \gamma h\omega$, so that the solutions are real. Recalling from (3.36) that γh is the speed of the underlying current at the mean level $z = 0$ of the thermocline, we see that in (4.88) the plus sign corresponds to the speed of waves outrunning the current (downstream linear waves), while the minus sign corresponds to the speed of waves running counter to the current (upstream linear waves). At this (leading) order c is a constant and dispersion effects are not observable. To deal with nonlinear effects, setting

$$c_0 = c - \gamma h, \qquad c_1 = c - \gamma h + \frac{\mu h}{r+1} = \pm\sqrt{\left(\frac{h\mu}{r+1}\right)^2 + \frac{hr(g - \gamma h\omega)}{r+1}}, \tag{4.89}$$

with c given by (4.88), we observe that (4.85)–(4.86) show that $q' = ((r+1)c_1\eta')/h$ at the leading order. We therefore expect that

$$q' = \frac{(r+1)c_1}{h}\eta' + \varepsilon^2 b_1 \eta'^2 + \varepsilon^2 b_2 \eta'_{x'x'} + O(\varepsilon^3) \tag{4.90}$$

for some constants $b_{1,2}$ that are yet to be determined. With the Ansatz (4.90), we substitute q' in (4.83)–(4.84) and write both two equations in terms of η' only up to the order $O(\varepsilon^2)$, thus obtaining two evolutionary equations for η', which should coincide up to $O(\varepsilon)$. The equality of their coefficients allows us to find the constants b_1 and b_2, as follows:

$$b_1 = \frac{r+1}{2h}\left[\frac{\gamma}{2} - \frac{\mu}{r+1} + \frac{h}{2(r+1)c_1} A_3 - \frac{c_1}{2h}\right], \tag{4.91}$$

$$b_2 = \frac{A_2}{2h}\left(\frac{\mu^2 h}{c_1} - \frac{(r+1)^2 c_1}{h}\right). \tag{4.92}$$

The resulting evolution equation for η is

$$\eta'_t + c\eta'_{x'} + \varepsilon^2 \frac{A_2(r+1)c_0^2}{2hc_1}\eta'_{x'x'x'} + \varepsilon^2 \frac{1}{2c_1}\left(\frac{3c_0^2}{h} + 3\gamma c_0 + h\gamma^2 - \frac{h\gamma_1^2}{1+r}\right)\eta'\eta'_{x'} = 0, \tag{4.93}$$

while the corresponding q' can be recovered from (4.90). Although we can write (4.93) in the standard form of the KdV-equation in a moving frame of reference by means of a scaling transformation (see [35]), we prefer to work with the form (4.93) to facilitate the physical interpretation of the mathematical results. One appealing feature of the KdV equation (4.93) is its bi-Hamiltonian structure: the equation has two expressions

as a Hamiltonian evolutionary equation, with the two Hamiltonian operators compatible, that is, any linear combination is again a Hamiltonian operator (see [35]). The bi-Hamiltonian structure ensures the existence of infinitely many integrals of motion that are functionally independent. Moreover, the equation can be solved exactly using inverse scattering theory: starting with arbitrary initial data in the Schwartz class, the solution that evolves from this data develops into a finite number of localised solitary waves (solitons) with speeds proportional to the amplitude, plus an oscillatory tail (see the discussion in [5]). Each solitary wave recovers its localised identity after interacting with other solitary waves (re-emerging unscated after such interaction, the only hallmark of the interaction being a phase shift), while the oscillatory tail disperses and spreads out in space. Therefore the solution evolves in to an ordered set of solitons, with the tallest in front, followed by an oscillatory tail. The details of this general picture can be predicted fairly explicitly from detailed knowledge of the initial data. For example, the internal solitary wave solution is a wave of depression, given in the original variables by

$$\eta'(x,t) = -A\,\mathrm{sech}^2\left\{\frac{1}{2}\sqrt{\frac{-A\mathcal{A}}{3\mathcal{B}}}\left[(x-x_0)-\left(c-\frac{A\mathcal{A}}{3}\right)t\right]\right\},$$

with $A > 0$ and $\mathcal{A} = \frac{1}{2c_1}\left(\frac{3c_0^2}{h}+3\gamma c_0+h\gamma^2-\frac{h\gamma_1^2}{r+1}\right) < 0$, $\mathcal{B} = \frac{A_2(r+1)c_0^2}{2hc_1} > 0$. We see that the propagation speed of the solitary wave, $c+\frac{A|\mathcal{A}|}{3}$, exceeds the speed c of the leading-order linear wave by an amount that is proportional to the wave amplitude A. Therefore larger solitary waves travel faster. With regard to the soliton interaction of two solitary waves that are initially separated, with the larger (and faster) one overtaking the smaller (and slower) one, they emerge unscated from the interaction, but each is shifted, this being the hallmark of the nonlinear nature of the interaction—in contrast to mere superposition. More precisely, the faster/slower solitary wave is shifted forward/backward by $\sqrt{\frac{12\mathcal{B}}{a_1|\mathcal{A}|}}\ln\frac{\sqrt{a_1}+\sqrt{a_2}}{\sqrt{a_1}-\sqrt{a_2}}$ and $\sqrt{\frac{12\mathcal{B}}{a_2|\mathcal{A}|}}\ln\frac{\sqrt{a_1}+\sqrt{a_2}}{\sqrt{a_1}-\sqrt{a_2}}$, respectively, where $a_1 > a_2 > 0$ are the amplitudes of the solitary waves. We point out that case studies for such wave phenomena in equatorial regions are available in the literature (see the interpretation of space shuttle photographs in [47], which show the relevance of soliton theory to the propagation of equatorial internal waves, and the discussion in [2], where it is pointed out that internal solitons are likely to be generated by the relaxation of the trade winds at the western boundary of the equatorial Pacific that triggers the "El Niño" event), although in general the presence of underlying currents is neglected. We see that the effect of the currents is encoded in the coefficients of the equation, so that these currents alter the shape and speed of these solitons (in comparison to the classical irrotational setting).

4.2.2. Inviscid Burgers regime. Equatorial ocean dynamics present peculiar features when compared to off-equatorial regions. A remarkable anomaly is that, despite high levels of internal wave energy and shear due to the presence of strong, zonal, basin-scale currents, study observations show that internal wave breaking is less frequent than one might anticipate, with rates near the Equator less than 10% of those typical at mid-latitudes for internal waves of comparable energy (see [18]). Nevertheless, internal wave breaking does occur in the equatorial Pacific: observations indicate that, in a process in which the interactions with the large shear caused by the EUC play a significant rôle (with the EUC enhancing the internal wave activity—see the discussion in [27]), a strong wind stress propagates downward and generates quite significant internal waves

(see [45]) which often break. The breaking is preceded by a steepening of the internal wave profile which results in finite-time gradient blow-up. We will now indicate how the Hamiltonian perturbation approach yields in a specified physical regime model equations that capture this type of behaviour.

Taking $\varepsilon \ll 1$ and $\delta \ll 1$ to be small parameters, we consider the shallow-water regime defined by the spatial scale

$$x' = \varepsilon x, \tag{4.94}$$

and the wave perturbation scales

$$\begin{cases} \eta(x,t) = \varepsilon \eta'(x',t), & q(x,t) = \varepsilon\, q'(x',t), \\ \eta_1(x,t) = \delta\varepsilon\, \eta_1'\Big(x',\ \dfrac{t}{\delta}\Big), & q_1(x,t) = \delta\varepsilon\, q_1'\Big(x',\ \dfrac{t}{\delta}\Big), \end{cases} \tag{4.95}$$

in which x', t, η', q', η_1', and q_1' are of the order of magnitude $O(1)$, with the surface perturbations faster but less significant than those at the thermocline (in view of the fact that $\delta \ll 1$). The regime (4.94)–(4.95) captures internal waves of larger amplitude than the long-wave regime (4.64)–(4.65), transforming (4.63) into the Hamiltonian system

$$\frac{d}{dt}\begin{pmatrix} q' \\ q_1' \\ \eta' \\ \eta_1' \end{pmatrix} = -\partial_x \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \partial\tilde{\mathcal{H}}/\partial q' \\ \partial\tilde{\mathcal{H}}/\partial q_1' \\ \partial\tilde{\mathcal{H}}/\partial \eta' \\ \partial\tilde{\mathcal{H}}/\partial \eta_1' \end{pmatrix}, \tag{4.96}$$

with Hamiltonian $\tilde{\mathcal{H}} = \mathcal{H}/\varepsilon$. For $\delta = O(\varepsilon)$, we follow the approach pursued in Sect. 4.1 and expand the Hamiltonian (3.73) up to order $O(\varepsilon^3)$, obtaining

$$\begin{aligned} H^{(3)}(q', q_1', \eta', \eta_1') = {} & \varepsilon \frac{1}{2}\int_{\mathbb{R}} \left(\frac{h}{r+1} q'^2 - 2\frac{\mu h}{r+1} q'\eta' + A_1 \eta'^2 + \delta^2 (g - \omega\Gamma_1)(\eta_1')^2 \right) dx' \\ & + \varepsilon\delta \frac{h}{r+1}\int_{\mathbb{R}} \left(q'q_1' - \mu\eta' q_1' - \mu_1 \eta_1' q' + \mu\mu_1' \eta'\eta_1' \right) dx' \\ & + \varepsilon\gamma h \int_{\mathbb{R}} q'\eta'\, dx' + \varepsilon\delta^2\Gamma_1 \int_{\mathbb{R}} q_1'\eta_1'\, dx' \\ & - \varepsilon^3 \frac{A_2}{2}\int_{\mathbb{R}} \left((q_x')^2 - 2\mu q_{x'}'\eta_{x'}' + \mu^2 (\eta_{x'}')^2 \right) dx' \\ & + \varepsilon^3 \frac{1}{2(1+r)}\int_{\mathbb{R}} \eta'(q')^2\, dx' + \varepsilon^5 \left(\frac{\gamma}{2} - \frac{\mu}{r+1} \right) \int_{\mathbb{R}} q'(\eta')^2\, dx' \\ & + \varepsilon^3 \frac{A_3}{6}\int_{\mathbb{R}} (\eta')^3\, dx' \\ & + \varepsilon\delta^2 \frac{1}{2}\left(\frac{h}{r+1} + h_1 \right) \int_{\mathbb{R}} \left((q_1')^2 - 2\mu_1 q_1'\eta_1' + \mu_1^2 (\eta_1')^2 \right) dx', \end{aligned} \tag{4.97}$$

with the notation in (4.76). We now neglect terms of order $O(\varepsilon^4)$ in the Hamiltonian $\mathcal{H}$, thus replacing (4.96) by a truncated Hamiltonian system with Hamiltonian $H^{(3)}/\varepsilon$. The canonical equations for the variables related to the free surface are linear

$$\eta'_{1,t} = -\partial_{x'}\left[\delta\Gamma_1\eta'_1 + \frac{h}{r+1}(q' - \mu\eta') + \delta\left(\frac{h}{r+1} + h_1\right)(q'_1 - \mu'_1\eta'_1)\right], \quad (4.98)$$

$$q'_{1,t} = -\partial_{x'}\left[\delta\Gamma_1 q'_1 - \delta\mu_1\left(\frac{h}{r+1} + h_1\right)(q'_1 - \mu_1\eta'_1) - \frac{\mu_1 h}{r+1}(q' - \mu\eta') + \delta(g - \omega\Gamma_1)\eta'_1\right], \quad (4.99)$$

while those for the variables at the thermocline are nonlinear

$$\eta'_t = -\partial_{x'}\left[\gamma h\eta' + \frac{h}{r+1}\Big(q' - \mu\eta' + \delta(q'_1 - \mu_1\eta'_1)\Big) + \varepsilon^2 A_2(q'_{x'x'} - \mu\eta'_{x'x'}) + \varepsilon\frac{1}{r+1}q'\eta' + \varepsilon\left(\frac{\gamma}{2} - \frac{\mu}{r+1}\right)\eta'^2\right], \quad (4.100)$$

$$q'_t = -\partial_{x'}\left[\gamma hq' - \frac{\mu h}{r+1}q' + A_1\eta' - \delta\frac{h\mu}{r+1}(q'_1 - \mu_1\eta'_1) - \varepsilon^2 A_2\mu(q'_{x'x'} - \mu\eta'_{x'x'}) + \varepsilon\frac{1}{r+1}\frac{q'^2}{2} + \varepsilon\left(\gamma - \frac{2\mu}{r+1}\right)\eta'q' + \varepsilon A_3\frac{\eta'^2}{2}\right]. \quad (4.101)$$

For $\delta = O(\varepsilon^2)$ we can neglect the δ-terms and the ε^2-terms above and decouple the motion of the thermocline from that of the free surface: the system (4.98)–(4.101) simplifies to

$$\eta'_{1,t} = -\frac{h}{r+1}(q' - \mu\eta')_{x'}, \quad (4.102)$$

$$q'_{1,t} = -\frac{\mu_1 h}{r+1}(q' - \mu\eta')_{x'}, \quad (4.103)$$

$$\eta'_t = -\partial_{x'}\left[\gamma h\eta' + \frac{h}{r+1}(q' - \mu\eta') + \varepsilon\frac{1}{r+1}q'\eta' + \varepsilon\left(\frac{\gamma}{2} - \frac{\mu}{r+1}\right)\eta'^2\right], \quad (4.104)$$

$$q'_t = -\partial_{x'}\left[\gamma hq' - \frac{\mu h}{r+1}q' + A_1\eta' + \varepsilon\frac{1}{r+1}\frac{q'^2}{2} + \varepsilon\left(\gamma - \frac{2\mu}{r+1}\right)\eta'q' + \varepsilon A_3\frac{\eta'^2}{2}\right]. \quad (4.105)$$

The Eqs. (4.102)–(4.103) show that the motion of the free surface is determined by the initial data and the characteristics (η', q') of the displacements of the thermocline, with $(q'_1 - \mu_1\eta'_1)_t = 0$. Note that (4.104)–(4.105) show that the leading order linear equations for η' and q' are (4.85)–(4.86), so that the wave speed c of a linear travelling wave, in which the (x', t)-dependence is solely in terms of $(x' - ct)$, satisfies the quadratic equation (4.87) whose solutions are given by (4.88). The choice of the plus sign in (4.88) yields the speed of linear waves outrunning the current, while (3.82) shows that the minus sign corresponds to the linear waves propagating westwards. To investigate nonlinear effects, using the notation (4.89), we infer from (4.85)–(4.86) that $q' = ((r+1)c_1\eta')/h$ at leading order. We therefore expect that

$$q' = \frac{(r+1)c_1}{h}\eta' + \varepsilon b_1\eta'^2 + O(\varepsilon^2) \quad (4.106)$$

for some constant b_1 to be determined. With the Ansatz (4.106), we substitute q in (4.104)–(4.105) and write both two equations in terms of η' up to the order of $O(\varepsilon^2)$,

thus obtaining two evolutionary equations for η', which should coincide up to $O(\varepsilon)$. From the equality of their coefficients we obtain

$$b_1 = \frac{r+1}{2h}\left[\frac{\gamma}{2} - \frac{\mu}{r+1} + \frac{hA_3}{2(r+1)c_1} - \frac{c_1}{2h}\right]. \tag{4.107}$$

The resulting evolution equation for η is the inviscid Burgers equation

$$\eta'_t + c\eta'_x + \varepsilon^2 \frac{1}{2c_1}\left(\frac{3c_0^2}{h} + 3\gamma c_0 + h\gamma^2 - \frac{h\gamma_1^2}{1+r}\right)\eta'\eta'_x = 0, \tag{4.108}$$

while q' can be recovered from (4.106). It is well-known (see, for example, the discussion in [12]) that any initial data in the Schwartz class will lead to finite-time blow-up for the solution to (4.108) in the form of wave breaking: the solution remains bounded but its slope becomes unbounded in finite time.

5. Concluding Discussion

Within the framework of two-dimensional equatorial flows in the f-plane, with no meridional variations, we presented a Hamiltonian formulation of the nonlinear governing equations for wave–current interactions in a two-layer inviscid fluid with a flat bed and a free surface, for flows with constant vorticity in each layer. A key step was to prove, without recourse to approximations, that the flow can be viewed as an irrotational perturbation—possibly of the same order—of a mean flow that represents the underlying current field. We have also derived several simplified models for the equatorial wave–current interaction across the Pacific through systematic structure-preserving perturbation theory. These have genuine potential as simplified diagnostic and prognostic models since they accommodate all the salient features of equatorial ocean dynamics in the Pacific (strong stratification, as well as a current field with flow-reversal) and are able to capture nonlinear effects. In particular, at some specific geophysical scales the weakly nonlinear long-wave regime turns out to be structure-enhancing since in this setting the nonlinear dynamics is described by a KdV equation—an integrable infinite-dimensional Hamiltonian system. The derived models can discriminate between different physical effects in observations and simulations, dictated by different scales. For example, the variety of oceanographically relevant scales permits us to single out various types of interplay between dispersion and nonlinearity, ranging from regimes in which dispersion and nonlinearity balance each other and allow wave solutions that propagate without change of form to dispersionless regimes that favour wave breaking.

The general systematic approach developed in this paper is also useful for developing accurate numerical procedures. Note that, even within the framework of irrotational two-layer inviscid flow with a free surface, the problem of treating computationally the nonlinear interactions between the surface and internal modes of oscillation is challenging (see the discussion in [37]). For example, for fixed densities $\overline{\rho} > \overline{\rho}_1$ of the layers but allowing variations of their mean depths $\overline{h}$ and $\overline{h}_1$, in the rigid-lid approximation the polarity of the internal solitary waves (that is, whether they are waves of elevation or depression) changes once in the range $\overline{h}_1/\overline{h} \in (0, 1)$, while no such change occurs in the coupled configuration (see [1,10]). Furthermore, the presence of non-zero vorticity in each layer complicates considerably the dynamics. Showing that within the nonlinear framework it is possible to split the velocity field into an underlying steady current component and a harmonic wave velocity field opens up the possibility to find numerically

solitary wave solutions to the fully-nonlinear problem using a boundary integral method based on Cauchy integral formula (see the discussion in [41]).

The Hamiltonian approach developed in this paper and the above mentioned topics open up exciting perspectives for future interdisciplinary research on equatorial wave–current interactions.

Acknowledgements. Open access funding provided by University of Vienna. This research was supported by the WWTF research Grant MA16-009 and by the EPSRC Grant No. EP/K032208/1. The authors would also like to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for support and hospitality during the programme "Nonlinear water waves", where the final stages of the work on this paper were undertaken. The authors are very grateful to the referee for comments and suggestions.

References

1. Benjamin, T.B.: Internal waves of finite amplitude and permanent form. J. Fluid Mech. **25**, 241–270 (1966)
2. Boyd, J .P.: Equatorial solitary waves. Part I: Rossby solitons. J. Phys. Oceanogr. **10**, 1699–1717 (1980)
3. Boyd, J.P.: Dynamics of the Equatorial Ocean. Springer, Berlin (2018)
4. Cheng, C.H., Shkoller, S.: Solvability and regularity for an elliptic system prescribing the curl, divergence, and partial trace of a vector field on Sobolev-class domains. J. Math. Fluid Mech. **19**, 375–422 (2017)
5. Constantin, A.: Nonlinear Water Waves with Applications to Wave–Current Interactions and Tsunamis. SIAM, Philadelphia (2011)
6. Constantin, A., Ivanov, R.I., Martin, C.-I.: Hamiltonian formulation for wave–current interactions in stratified rotational flows. Arch. Ration. Mech. Anal. **221**, 1417–1447 (2016)
7. Constantin, A., Johnson, R.S.: The dynamics of waves interacting with the equatorial undercurrent. Geophys. Astrophys. Fluid Dyn. **109**, 311–358 (2015)
8. Constantin, A., Johnson, R.S.: Ekman-type solutions for shallow-water flows on a rotating sphere: a new perspective on a classical problem. Phys. Fluids **31**, 021401 (2019)
9. Constantin, A., Strauss, W., Varvaruca, E.: Global bifurcation of steady gravity water waves with critical layers. Acta Math. **195–262**, 2016 (2017)
10. Craig, W., Guyenne, P., Kalisch, H.: Hamiltonian long-wave expansions for free surfaces and interfaces. Commun. Pure Appl. Math. **58**, 1587–1641 (2005)
11. Craig, W., Guyenne, P., Sulem, C.: The surface signature of internal waves. J. Fluid Mech. **710**, 277–303 (2012)
12. Dafermos, C.M.: Hyperbolic Conservation Laws in Continuum Physics. Springer, Berlin (2010)
13. DaSilva, A.F.T., Peregrine, D.H.: Steep, steady surface waves on water of finite depth with constant vorticity. J. Fluid Mech. **125**, 281–302 (1988)
14. Fedorov, A.V., Brown, J.N.: Equatorial waves. In: Steele, J. (ed.) Encyclopedia of Ocean Sciences, pp. 3679–3695. Academic Press, Cambridge (2009)
15. Gilbarg, D., Trudinger, N.S.: Elliptic Partial Differential Equations of Second Order. Springer, Berlin (2001)
16. Gill, A.: Atmosphere-Ocean Dynamics. Academic Press, Cambridge (1982)
17. Goff, J., Chagué-Goff, C., Archer, M., Dominey-Howes, D., Turney, C.: The Eltanin asteroid impact: possible South Pacific palaeomegatsunami footprint and potential implications for the Pliocene–Pleistocene transition. J. Quat. Sci. **27**, 660–670 (2012)
18. Gregg, M.C., Sanford, T.B., Winkel, D.P.: Reduced mixing from the breaking of internal waves in equatorial waters. Nature **422**, 513–515 (2003)
19. Ishikawa, Y., Awaji, T., Akimoto, K.: Global surface circulation and its kinetic energy distribution derived from drifting buoys. J. Oceanogr. **53**, 489–516 (1997)
20. Johnson, G.C., McPhaden, M.J.: Equatorial Pacific ocean horizontal velocity, divergence, and upwelling. J. Phys. Oceanogr. **31**, 839–849 (2001)

21. Kato, T.: Perturbation Theory for Linear Operators. Springer, Berlin (1980)
22. Kessler, W.S., McPhaden, M.J.: Oceanic equatorial waves and the 1991–1993 El Niño. J. Clim. **8**, 1757–1774 (1995)
23. Kinsman, B.: Wind Waves. Prentice-Hall, Englewood Cliffs (1965)
24. Kurdyka, K., Păunescu, L.: Hyperbolic polynomials and multiparameter real-analytic perturbation theory. Duke Math. J. **141**, 123–149 (2008)
25. Lax, P.D.: Linear Algebra. Wiley, New York (1997)
26. LeBlond, P.H., Mysak, L.A.: Waves in the Ocean. Elsevier, Amsterdam (1978)
27. Lien, R.-C., D'Asaro, E.A., McPhaden, M.J.: Internal waves and turbulence in the upper central equatorial Pacific: Lagrangian and Eulerian observations. J. Phys. Oceanogr. **32**, 2619–2639 (2002)
28. Lien, R.-C., McPhaden, M.J., Gregg, M.C.: High-frequency internal waves at 0°, 140°W and their possible relationship to deep-cycle turbulence. J. Phys. Oceanogr. **26**, 581–600 (1996)
29. Majda, A.J.: Introduction to PDEs and Waves for the Atmosphere and Ocean. Courant Lecture Notes in Mathematics. AMS, Providence (2003)
30. Maslowe, S.A.: Critical layers in shear flows. Ann. Rev. Fluid Mech. **18**, 405–432 (1986)
31. McCreary, J.P.: Modeling equatorial ocean circulation. Ann. Rev. Fluid Mech. **17**, 359–409 (1985)
32. McPhaden, M.J., Proehl, J.A., Rothstein, L.M.: The interaction of equatorial Kelvin waves with realistically sheared zonal currents. J. Phys. Oceanogr. **17**, 1499–1515 (1986)
33. Meyer, K.H., Hall, G.R.: Introduction to Hamiltonian Dynamical Systems and the n-Body Problem. Springer, New York (1992)
34. Moum, J.N., Nash, J.D., Smyth, W.D.: Narrowband oscillations in the upper equatorial ocean. Part i: interpretation as shear instability. J. Phys. Oceanogr. **41**, 397–411 (2011)
35. Olver, P.J.: Applications of Lie Groups to Differential Equations. Springer, New York (1993)
36. Pacanowski, R.C., Philander, S.G.H.: The generation of equatorial currents. J. Geophys. Res. **85**, 1123–1136 (1980)
37. Parau, E.I., Woolfenden, H.C.: Numerical computation of solitary waves in a two-layer fluid. J. Fluid Mech. **688**, 528–550 (2011)
38. Pinkel, R., Merrifield, M., McPhaden, M., Picaut, J., Rutledge, S., Siegel, D., Washburn, L.: Solitary waves in the western equatorial Pacific Ocean. Geophys. Res. Lett. **24**, 1603–1606 (1997)
39. Proehl, J.A., McPhaden, M.J., Rothstein, L.M.: A numerical approach to equatorial oceanic wave-mean flow interactions. In: O'Brien, J.J. (ed.) Advanced Physical Oceanographic Modelling, pp. 111–126. D. Reidel Publishing Company, Dordrecht (1986)
40. Reed, P., Simon, B.: Methods of Modern Mathematical Physics. Fourier Analysis, Self-Adjointness, vol. II. Academic Press, New York (1975)
41. Vanden-Broeck, J.-M.: Solitary waves in water: numerical methods and results. In: Grimshaw, R.H.J. (ed.) Solitary Waves in Fluids, pp. 55–84. WIT Press, Southampton (2007)
42. von Wahl, W.: Estimating ∇u by div u and curl u. Math. Methods Appl. Sci. **15**, 123–143 (1992)
43. Wahlén, E.: A Hamiltonian formulation of water waves with constant vorticity. Lett. Math. Phys. **79**, 303–315 (2007)
44. Wahlén, E.: Hamiltonian long-wave approximations of water waves with constant vorticity. Phys. Lett. A **372**, 2597–2602 (2008)
45. Wijesekera, H.W., Dillon, T.M.: Internal waves and mixing in the upper equatorial Pacific Ocean. J. Geophys. Res. **96**, 7115–7125 (1991)
46. Zakharov, V.E.: Stability of periodic waves of finite amplitude on the surface of a deep fluid. J. Appl. Mech. Tech. Phys. **9**, 1990–1994 (1968)
47. Zheng, Q., Klemas, V., Yan, X.-H.: Dynamic interpretation of space shuttle photographs: deepwater internal waves in the western equatorial Indian Ocean. J. Geophys. Res. **100**, 2579–2589 (1995)

Communicated by W. Schlag