J. Fluid Mech. (2018), *vol.* 839, *pp.* 408–429.
doi:10.1017/jfm.2018.22

Wave-induced mean flows in rotating shallow water with uniform potential vorticity

Jim Thomas[1], **Oliver Bühler**[1,†] **and K. Shafer Smith**[1]

[1]Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, USA

(Received 25 August 2017; revised 13 December 2017; accepted 1 January 2018)

Theoretical and numerical computations of the wave-induced mean flow in rotating shallow water with uniform potential vorticity are presented, with an eye towards applications in small-scale oceanography where potential-vorticity anomalies are often weak compared to the waves. The asymptotic computations are based on small-amplitude expansions and time averaging over the fast wave scale to define the mean flow. Importantly, we do not assume that the mean flow is balanced, i.e. we compute the full mean-flow response at leading order. Particular attention is paid to the concept of modified diagnostic relations, which link the leading-order Lagrangian-mean velocity field to certain wave properties known from the linear solution. Both steady and unsteady wave fields are considered, with specific examples that include propagating wavepackets and monochromatic standing waves. Very good agreement between the theoretical predictions and direct numerical simulations of the nonlinear system is demonstrated. In particular, we extend previous studies by considering the impact of unsteady wave fields on the mean flow, and by considering the total kinetic energy of the mean flow as a function of the rotation rate. Notably, monochromatic standing waves provide an explicit counterexample to the often observed tendency of the mean flow to decrease monotonically with the background rotation rate.

Key words: quasi-geostrophic flows, rotating flows, shallow water flows

1. Introduction

Nonlinear interactions between waves and suitably defined mean flows have long been studied in fluid mechanics, with early examples focusing on acoustic streaming and related effects (e.g. Lighthill 1978). The field underwent a dramatic extension and deepening in the second half of the last century, when the crucial importance of such interactions for the long-term dynamics of the atmosphere and oceans was realized (as detailed in textbooks such as Andrews, Holton & Leovy 1987; Vallis 2006). By now, wave–mean interaction theory has become a mature field with solid theoretical foundations (e.g. Bühler 2014).

Much of the original wave–mean interaction theory was developed for a situation in which small-amplitude waves are propagating on a non-uniform but laminar basic

† Email address for correspondence: obuhler@cims.nyu.edu

flow of much larger amplitude (Eliassen & Palm 1961). This is a natural model for atmospheric internal waves propagating on a basic state characterized by a strong jet stream, for example. The Eulerian averaging operation that defines the mean flow can then be chosen as large-scale zonal averaging along latitude circles. Mathematically, the situation is described by a perturbation expansion in a small-amplitude parameter $a \ll 1$ in terms of which the basic flow is $O(1)$ and the wave-like disturbances are $O(a)$. Nonlinear effects such as wave-induced changes in the mean flow then accrue at $O(a^2)$ and those are obviously small changes of the $O(1)$ basic state. Significant changes of the basic state are then possible only if there are resonant mean-flow responses that grow as $O(a^2t)$ in time t. For a steady wave field on a laminar flow, the non-acceleration theorem then shows this is only possible if the waves are dissipating or breaking in some location, which led to the well-established theory of dissipative wave drag and critical layers described in the textbooks cited above. In this connection the natural slow time scale for significant mean-flow changes to occur is $t = O(a^{-2})$, as exemplified by the classical theory of the quasi-biennial oscillation in the equatorial lower stratosphere in Plumb (1977).

However, present-day research in atmosphere–ocean fluid dynamics has pushed to smaller spatial scales, where the amplitude of the mean flow is not necessarily larger than that of the waves. A prime example of this is the so-called submesoscale regime in oceanography, which is often characterized by horizontal scales significantly below the deformation radius. At these scales of less than 100 km the relevant mean flow is naturally defined by its slowness when compared to fast wave motions, i.e. the natural Eulerian averaging operation that now defines the mean flow is a slowly varying time average. Of course, such a mean flow is now three-dimensional and itself unsteady, which is a significant departure from the classical theory based on zonal averaging.

For example, a problem of considerable contemporary interest in oceanography concerns the long-term consequences of the interactions between storm-generated near-inertial oscillations and the much slower geostrophically balanced mean flow (e.g. Alford *et al.* 2016). In this physical situation the waves are of comparable magnitude or even larger than the balanced mean flow, which makes the earlier theory based on strong mean flows less relevant. Again, there is no obvious spatial scale separation between waves and mean flows in this situation, so time averaging over the fast wave oscillations is the only option to define the mean flow.

Rigorous general results, i.e. results with rigorous error control, are sparse in this area. With periodic boundary conditions, the particular case where waves and balanced flow are both $O(a)$ is amenable to rigorous two-timing analysis in spectral space (Babin, Mahalov & Nicolaenko 1997; Majda & Embid 1998), with a sub-linear growth solvability condition that determines the evolution on the first nonlinear slow time scale $t = O(a^{-1})$. In the presence of significant Coriolis forces, the main result from these studies has been that the interactions are weak, with no significant energy exchanges between waves and quasi-geostrophic mean flows on this time scale. However, it has not been possible to extend this rigorous multi-scale analysis to the second slow time scale $t = O(a^{-2})$, which would have been of primary interest in classical wave–mean interaction theory. A discussion of the difficulties involved is given in § 6 in Bühler (2000).

In the absence of a rigorous general theory to push to the longer time scale, research in this area has focused on theoretical and numerical case studies of various particular interaction scenarios. An important theme that has emerged in this connection is the notion of a modified diagnostic relation between the streamfunction describing the balanced mean flow and the materially advected

potential vorticity (PV). This modification consists of certain wave-induced terms, such as the vertical curl of the horizontal pseudomomentum vector, that must be added to the standard quasi-geostrophic terms to obtain long-time accuracy for the mean flow. The relevance of such modified diagnostic relations was already implied in the landmark paper of Bretherton (1969), who computed the $O(a^2)$ mean-flow response to compactly supported three-dimensional and two-dimensional wavepackets in the non-rotating Boussinesq equations (see van den Bremer & Sutherland (2014) for a recent extension of this work to higher accuracy and Wagner & Young (2015) for further extension both in wave amplitude and in treating a rotating Boussinesq flow). A general derivation and description of the modified diagnostic relation concept was first presented in Bühler & McIntyre (1998) (hereafter BM98), using rotating shallow-water and three-dimensional Boussinesq flows.

The modified diagnostic relation concept arises naturally in recent studies of the near-inertial wave problem (e.g. Xie & Vanneste 2015; Wagner & Young 2016; Thomas, Smith & Bühler 2017). In this connection, the present authors have recently investigated numerically over very long time periods various cases of interactions between small-amplitude inertia–gravity waves and balanced mean flows in the rotating shallow-water equations. Here the waves and the PV field overlapped in spatial scale, the wave amplitude was $O(a)$ and the PV amplitude was either $O(a)$ or smaller (manuscript in preparation). If the PV amplitude is so weak that it can be taken to be zero one obtains the rigorous constraint that the PV is uniform throughout and equal to its background value always and everywhere. Significant theoretical progress beyond numerical experimentation can be made in this limiting case, and this is the topic of the present paper.

To this end we formulate the equations for the complete $O(a^2)$ mean-flow response induced by $O(a)$ waves on a uniform PV background in rotating shallow water, which extends and in one place corrects previous work on the subject in BM98. We then study the solution to these equations first in the context of slowly varying wave fields, such as a steady wavetrain or unsteady wavepackets, and second in the case of a monochromatic standing wave. The latter case is motivated by the prevalence of nearly monochromatic internal tides in the ocean (e.g. Garrett & Kunze 2007), which are not slowly varying waves in space. Detailed theoretical predictions are obtained for these cases, which are subsequently checked and illustrated by independent direct numerical simulations of the full nonlinear shallow-water equations.

Notably, we do not assume that the mean flow is balanced in our computations, i.e. we allow for a full $O(a^2)$ mean-flow response, which itself can contain wave-like motions. This is different from studies such as Wagner & Young (2015) or Xie & Vanneste (2015), where detailed mean-flow results were obtained but only under the assumption that the mean flow was balanced. In particular, we will investigate dynamical processes that unfold on the intermediate time scale for wave–wave interactions that was explicitly omitted in Wagner & Young (2015), as discussed in their § 4.2 on the set-up for the two-timing asymptotics. Moreover, as we do not assume that the mean flow is balanced, we are able to derive theoretical and numerical results for a wide range of f, from weak to strong rotation.

We pay particular attention to the errors in the mean-flow computation that would result from ignoring the intrinsic time dependence of propagating wavepackets, and we also investigate the overall magnitude of the mean-flow response as a function of the Coriolis parameter f. We find that for very large f the mean-flow response is typically 'choked' and decays quickly with f, but we also show by way of an explicit counterexample using standing waves that the strongest mean-flow response is not always reached in the non-rotating limit $f = 0$.

The plan for the paper is as follows: the governing equations are developed in § 2 and their relevant form for slowly varying waves is discussed in detail in § 3 for both steady wavetrains and propagating wavetrains. The alternative scenario of monochromatic standing waves is developed in § 4, direct numerical simulations are presented in § 5 and concluding remarks are offered in § 6.

Finally, we prefer to use Lagrangian-mean velocity as a natural field with which to describe the mean flow, but we avoid using the full machinery of generalized Lagrangian-mean theory (GLM) by deriving the $O(a^2)$ equations using elementary methods. It would have been shorter to simply refer to the relevant equations in Andrews & McIntyre (1978) (hereafter AM78a), but in this way the present paper is more self-contained.

2. Governing equations

We consider the rotating shallow-water equations on flat topography:

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + \boldsymbol{f}\times\boldsymbol{u} + g\boldsymbol{\nabla}h = 0 \quad \text{and} \quad \frac{\partial h}{\partial t} + \boldsymbol{\nabla}\cdot(h\boldsymbol{u}) = 0. \tag{2.1a,b}$$

Here $\boldsymbol{u} = (u, v)$ is the velocity, $\mathrm{D}/\mathrm{D}t = \partial_t + (\boldsymbol{u}\cdot\boldsymbol{\nabla})$ is the material derivative, $\boldsymbol{f} = f\widehat{\boldsymbol{z}}$ is the constant Coriolis vector, h is the layer depth and g is gravity. The associated PV

$$q = \frac{\boldsymbol{\nabla}\times\boldsymbol{u} + f}{h} \quad \text{such that} \quad \frac{\mathrm{D}q}{\mathrm{D}t} = 0 \tag{2.2a,b}$$

holds for smooth flows, i.e. flows without shocks. Here $\boldsymbol{\nabla}\times\boldsymbol{u}$ is shorthand for $\widehat{\boldsymbol{z}}\cdot(\boldsymbol{\nabla}\times\boldsymbol{u})$.

Throughout this paper we assume that q is uniform and equal to its background value f/H, where H is the undisturbed layer depth. The material invariance of q implies that this will remain the case as long as the flow evolves smoothly in time. We therefore have the global vortex stretching constraint

$$q = \frac{f}{H} \quad \Leftrightarrow \quad \boldsymbol{\nabla}\times\boldsymbol{u} = \frac{h-H}{H}f. \tag{2.3a,b}$$

2.1. *Linear wave perturbations*

We assume that the entire flow is a weak perturbation to the rest state $\boldsymbol{u} = 0$ and $h = H$. Specifically, we introduce a small-amplitude parameter $a \ll 1$ and pursue a regular perturbation expansion in powers of a. The linear perturbations $(\boldsymbol{u}', h') = O(a)$ satisfy the usual linear equations

$$\frac{\partial \boldsymbol{u}'}{\partial t} + \boldsymbol{f}\times\boldsymbol{u}' + g\boldsymbol{\nabla}h' = 0 \quad \text{and} \quad \frac{\partial h'}{\partial t} + H\boldsymbol{\nabla}\cdot\boldsymbol{u}' = 0. \tag{2.4a,b}$$

The corresponding linear PV

$$q' = \boldsymbol{\nabla}\times\boldsymbol{u}' - \frac{h'}{H}f = 0 \tag{2.5}$$

by (2.3). Hence the linear dynamics is described entirely by the inertia–gravity wave modes of the system, as the balanced vortical mode is zero. In particular, plane waves with wavenumber $\boldsymbol{k} = (k, l)$ and frequency $\hat{\omega}$ satisfy the dispersion relation

$$\hat{\omega}^2 = f^2 + c^2(k^2 + l^2) \quad \text{where} \quad c^2 = gH. \tag{2.6}$$

The governing equation for the wave energy density

$$E = \frac{1}{2}\left(\overline{u'^2} + \overline{v'^2} + \frac{g}{H}\overline{h'^2}\right) \tag{2.7}$$

follows from (2.4) as

$$\frac{\partial E}{\partial t} + \boldsymbol{\nabla}\cdot(g\overline{h'\boldsymbol{u}'}) = 0. \tag{2.8}$$

Here we introduced a suitable Eulerian averaging operator $\overline{(\ldots)}$ that commutes with ∂_t and $\boldsymbol{\nabla}$ and acts like a linear projection in the sense that $\overline{h'} = 0$ and so on. The precise nature of the averaging operator does not affect the general theory in this section, but specific results will subsequently be based on phase averages over wave oscillations. The wave energy E is an example of a second-order wave property in the sense that $E = O(a^2)$, yet it can be evaluated from the linear, $O(a)$ solution alone. Other wave properties of interest include the $O(a^2)$ Stokes corrections

$$\overline{\boldsymbol{u}}^S = \overline{(\boldsymbol{\xi}'\cdot\boldsymbol{\nabla})\boldsymbol{u}'} \quad \text{and} \quad \overline{h}^S = \overline{(\boldsymbol{\xi}'\cdot\boldsymbol{\nabla})h'}. \tag{2.9a,b}$$

Here $\boldsymbol{\xi}'$ is the linear particle displacement defined by

$$\frac{\partial \boldsymbol{\xi}'}{\partial t} = \boldsymbol{u}' \quad \text{and} \quad \overline{\boldsymbol{\xi}'} = 0. \tag{2.10a,b}$$

The field $\overline{\boldsymbol{u}}^S$ is commonly known as the Stokes drift, and generically $\overline{\phi}^S = \overline{(\boldsymbol{\xi}'\cdot\boldsymbol{\nabla})\phi'}$ denotes the leading-order Stokes correction to any field ϕ. With $\boldsymbol{\xi}'$ in hand the linear continuity equation in (2.4) can be integrated to

$$h' = -H\boldsymbol{\nabla}\cdot\boldsymbol{\xi}'. \tag{2.11}$$

Also, taking the dot product of $\boldsymbol{\xi}'$ with the linear momentum equation in (2.4), averaging and integrating by parts in time yields the $O(a^2)$ virial theorem (cf. AM78a)

$$\frac{1}{2}\frac{\partial^2}{\partial t^2}\overline{|\boldsymbol{\xi}'|^2} + g\overline{h}^S = \overline{\boldsymbol{u}'\cdot(\boldsymbol{u}' + \boldsymbol{f}\times\boldsymbol{\xi}')}. \tag{2.12}$$

A key wave property for wave–mean interaction theory is the $O(a^2)$ Lagrangian-mean pseudomomentum vector $\boldsymbol{p}$ (AM78a) whose Cartesian components are

$$p_i = -\overline{\xi'_{j,i}\left(u'_j + \tfrac{1}{2}\epsilon_{jkm}f_k\xi'_m\right)}. \tag{2.13}$$

Here the summation convention is used, the comma denotes spatial derivatives and ϵ_{jkm} is the Levi-Civita symbol. The evolution law for $\boldsymbol{p}$ follows from contracting the jth component of the linear momentum equation with $-\xi'_{j,i}$ and averaging. This yields

$$\frac{\partial p_i}{\partial t} + \overline{\left(\frac{1}{2}u'_j\left(u'_j + \epsilon_{jkm}f_k\xi'_m\right)\right)}_{,i} = g\overline{\xi'_{j,i}h'}_{,j} = \left(g\overline{\xi'_{j,i}h'} + \frac{g\overline{h'^2}}{2H}\delta_{ji}\right)_{,j}, \tag{2.14}$$

where the latter divergence form makes obvious that this is a conservation law for $\boldsymbol{p}$. The manipulations leading to (2.14) are straightforward except for the treatment of the Coriolis term, which is decomposed as

$$-\xi'_{j,i}\epsilon_{jkm}f_k u'_m = -\frac{1}{2}\frac{\partial}{\partial t}(\xi'_{j,i}\epsilon_{jkm}f_k\xi'_m) + \frac{1}{2}u'_{j,i}\epsilon_{jkm}f_k\xi'_m - \frac{1}{2}\xi'_{j,i}\epsilon_{jkm}f_k u'_m. \tag{2.15}$$

The first term becomes part of $\partial p_i/\partial t$ in (2.14) and by using $\epsilon_{jkm} = -\epsilon_{mkj}$ the remaining two terms can be rearranged as

$$\tfrac{1}{2}\epsilon_{jkm} f_k (u'_{j,i}\xi'_m - \xi'_{j,i}u'_m) = \tfrac{1}{2}\epsilon_{jkm} f_k (u'_{j,i}\xi'_m + \xi'_{m,i}u'_j) = \tfrac{1}{2}\epsilon_{jkm} f_k (u'_j\xi'_m)_{,i} \tag{2.16}$$

and thus become part of the flux term in (2.14). Finally, we note that the virial theorem (2.12) can be used to rewrite (2.14) in the equivalent form

$$\frac{\partial p_i}{\partial t} + \frac{1}{4}\frac{\partial^2}{\partial t^2}\overline{(|\boldsymbol{\xi}'|^2)}_{,i} + \frac{g}{2}\overline{h}^S_{,i} = g\overline{\xi'_{j,i}h'}_{,j}. \tag{2.17}$$

This form will be useful for the mean-flow response equation below.

2.2. *Leading-order mean-flow response*

We now seek to formulate equations for the $O(a^2)$ mean-flow response to the waves. To describe the mean flow we prefer a combination of the Lagrangian-mean velocity $\overline{\boldsymbol{u}}^L$, which is the mean velocity field that moves particles, and the Eulerian-mean height $\overline{h}$, which tracks the mean mass distribution. This is the same approach as in the shallow-water study of Bühler & Jacobson (2001), which however was restricted to slowly varying waves. Of course, the Stokes drift in (2.9) can be used together with the generic formula

$$\overline{\boldsymbol{u}}^L = \overline{\boldsymbol{u}} + \overline{\boldsymbol{u}}^S \tag{2.18}$$

to compute the Eulerian-mean velocity $\overline{\boldsymbol{u}}$ from $\overline{\boldsymbol{u}}^L$ if desired. The derivation of the leading-order mean-flow equations presented here is based on the exact manipulations leading to theorem I in AM78a; see also the exact shallow-water manipulations leading to the corresponding (5.15) in BM98. However, our end result for the mean-flow equations also uses part of the pseudomomentum evolution law in a manner that extends the analysis in Bühler & Jacobson (2001), so we find it most convenient to provide the small-amplitude calculation from scratch.

First, applying the Eulerian averaging operator to the continuity equation in (2.1) and collecting terms at $O(a^2)$ leads to

$$\frac{\partial \overline{h}}{\partial t} + H\boldsymbol{\nabla}\cdot\overline{\boldsymbol{u}}^L = H\boldsymbol{\nabla}\cdot\left(\overline{\boldsymbol{u}}^S - \frac{\overline{h'\boldsymbol{u}'}}{H}\right). \tag{2.19}$$

Using (2.9) and (2.10) the right-hand side can be rewritten as

$$H(\overline{\xi'_j u'_{i,j}} + \overline{\xi'_{j,j}u'_i})_{,i} = H(\overline{\xi'_j u'_i})_{,ij} = \frac{H}{2}(\overline{\xi'_j u'_i} + \overline{\xi'_i u'_j})_{,ij} = \frac{H}{2}\frac{\partial}{\partial t}(\overline{\xi'_i \xi'_j})_{,ij}. \tag{2.20}$$

This leads to the $O(a^2)$ continuity equation in the final form

$$\frac{\partial \overline{h}}{\partial t} + H\boldsymbol{\nabla}\cdot\overline{\boldsymbol{u}}^L = \frac{H}{2}\frac{\partial}{\partial t}(\overline{\xi'_i\xi'_j})_{,ij}. \tag{2.21}$$

The right-hand side corresponds to the divergence effect in GLM theory (e.g. McIntyre 1988; Bühler 2014). In particular, if the mean layer depth $\tilde{h}$ is defined in finite-amplitude GLM theory by (In general, the mean layer depth $\tilde{h}$ defined by (2.22) equals neither $\overline{h}$ nor $\overline{h}^L$.)

$$\frac{\partial \tilde{h}}{\partial t} + \boldsymbol{\nabla}\cdot(\tilde{h}\overline{\boldsymbol{u}}^L) = 0 \tag{2.22}$$

then (2.21) implies the $O(a^2)$ result

$$\tilde{h} = \overline{h} - \frac{H}{2}(\overline{\xi_i' \xi_j'})_{,ij}. \tag{2.23}$$

This small-amplitude GLM result is well known for incompressible flows, for which $\overline{h}$ is constant, but it also holds in the present compressible case. Finite-amplitude GLM theory yields the exact PV relation first derived in (5.12-17) of BM98:

$$\overline{q}^L = \frac{\boldsymbol{\nabla} \times (\overline{\boldsymbol{u}}^L - \boldsymbol{p}) + f}{\tilde{h}}, \tag{2.24}$$

where $\boldsymbol{p}$ is the exact GLM pseudomomentum vector. To $O(a^2)$ this corresponds to

$$H\overline{q}^L = \boldsymbol{\nabla} \times (\overline{\boldsymbol{u}}^L - \boldsymbol{p}) - f\frac{\tilde{h} - H}{H} + f, \tag{2.25}$$

where the approximate pseudomomentum in (2.13) is sufficient. This expression holds for general PV fields, but under the global constraint (2.3) we have that $\overline{q}^L = f/H$ holds exactly and therefore

$$\boldsymbol{\nabla} \times \overline{\boldsymbol{u}}^L = \boldsymbol{\nabla} \times \boldsymbol{p} + f\frac{\tilde{h} - H}{H} = \boldsymbol{\nabla} \times \boldsymbol{p} + f\left(\frac{\overline{h} - H}{H} - \frac{1}{2}(\overline{\xi_i' \xi_j'})_{,ij}\right) \tag{2.26}$$

holds to $O(a^2)$. This helpful diagnostic relation shows how $\boldsymbol{\nabla} \times \overline{\boldsymbol{u}}^L$ is controlled by $\boldsymbol{\nabla} \times \boldsymbol{p}$ and the background vortex stretching term. Of course, even for a given right-hand side of (2.26) this yields only $\boldsymbol{\nabla} \times \overline{\boldsymbol{u}}^L$ but not $\boldsymbol{\nabla} \cdot \overline{\boldsymbol{u}}^L$, so $\overline{\boldsymbol{u}}^L$ is only partially determined by it. To go beyond this it is necessary to consider the full mean momentum equation.

Applying the Eulerian averaging operator to the momentum equation yields

$$\frac{\partial \overline{\boldsymbol{u}}}{\partial t} + \boldsymbol{f} \times \overline{\boldsymbol{u}} + g\boldsymbol{\nabla}\overline{h} = -\overline{(\boldsymbol{u}' \cdot \boldsymbol{\nabla})\boldsymbol{u}'}. \tag{2.27}$$

Adding Stokes drift terms results in

$$\frac{\partial \overline{\boldsymbol{u}}^L}{\partial t} + \boldsymbol{f} \times \overline{\boldsymbol{u}}^L + g\boldsymbol{\nabla}\overline{h} = \frac{\partial \overline{\boldsymbol{u}}^S}{\partial t} + \boldsymbol{f} \times \overline{\boldsymbol{u}}^S - \overline{(\boldsymbol{u}' \cdot \boldsymbol{\nabla})\boldsymbol{u}'}. \tag{2.28}$$

The time derivative of the $O(a^2)$ Stokes drift is

$$\frac{\partial \overline{\boldsymbol{u}}^S}{\partial t} = \frac{\partial}{\partial t}\overline{(\boldsymbol{\xi}' \cdot \boldsymbol{\nabla})\boldsymbol{u}'} = \overline{(\boldsymbol{u}' \cdot \boldsymbol{\nabla})\boldsymbol{u}'} - \overline{(\boldsymbol{\xi}' \cdot \boldsymbol{\nabla})(\boldsymbol{f} \times \boldsymbol{u}' + g\boldsymbol{\nabla}h')}. \tag{2.29}$$

The leads to cancellations in (2.28), which now takes the form

$$\frac{\partial \overline{\boldsymbol{u}}^L}{\partial t} + \boldsymbol{f} \times \overline{\boldsymbol{u}}^L + g\boldsymbol{\nabla}\overline{h} = -\overline{(\boldsymbol{\xi}' \cdot \boldsymbol{\nabla})(g\boldsymbol{\nabla}h')} \equiv -g\overline{\boldsymbol{\nabla}h}^S. \tag{2.30}$$

Here $\overline{\boldsymbol{\nabla}h}^S$ is the Stokes correction to the height field gradient, in analogy with (2.9). Equation (2.30) could have been expected from GLM theory because $\overline{\boldsymbol{\nabla}h}^L = \boldsymbol{\nabla}\overline{h} + \overline{\boldsymbol{\nabla}h}^S$. Notably, $\overline{\boldsymbol{\nabla}h}^S$ is not irrotational. Indeed, in component form $-g\overline{\boldsymbol{\nabla}h}^S$ is

$$-g\overline{\xi_j' h'_{,ij}} = -g(\overline{\xi_j' h'_{,j}})_{,i} + g\overline{\xi'_{j,i} h'_{,j}}. \tag{2.31}$$

The first term is a gradient term $-g\nabla\overline{h}^S$ but the second term is not and it connects to the pseudomomentum evolution law (2.17). The outcome is the Lagrangian-mean momentum equation in the final form

$$\frac{\partial\overline{\boldsymbol{u}}^L}{\partial t}+\boldsymbol{f}\times\overline{\boldsymbol{u}}^L+g\nabla\overline{h}=\frac{\partial\boldsymbol{p}}{\partial t}-\frac{g}{2}\nabla\overline{h}^S+\frac{1}{4}\frac{\partial^2}{\partial t^2}\nabla\overline{|\boldsymbol{\xi}'|^2}. \tag{2.32}$$

Together, equations (2.21) and (2.32) determine the complete $O(a^2)$ mean-flow response $(\overline{\boldsymbol{u}}^L, \overline{h})$. The specific form of these equations has been chosen to highlight the impact of time-dependent wave fields.

3. Slowly varying waves

The classical prototype of a slowly varying wave is a freely propagating compact wavepacket, i.e. a linear solution of the form

$$h'(x, y, t)=aH\,A(\mu\kappa[\boldsymbol{x}-\boldsymbol{c}_g t])\exp(\mathrm{i}[\boldsymbol{k}\cdot x-\omega t-\alpha]), \tag{3.1}$$

with real part understood. Here the wavenumber vector $\boldsymbol{k}=(k, l)$ and the frequency $\omega>0$ satisfy (2.6), $\kappa=|\boldsymbol{k}|$, $\boldsymbol{c}_g$ is the group velocity associated with $\boldsymbol{k}$, and we retained an explicit phase shift parameter $\alpha\in[0, 2\pi]$ that allows the Eulerian averaging operator to be based on phase averaging. Finally, $A(\cdot)$ is a generic smooth real amplitude envelope function, with $A(0)=1$ and smooth decay to zero away from the wavepacket. In the numerical simulations described in § 5, A is a Gaussian. The small parameter $\mu\ll 1$ controls the slow envelope variation such that $(A_t, |\nabla A|)=O(\mu A)(\omega, \kappa)$. For example, the spatial envelope scale is $L=1/(\mu\kappa)$. The linear velocity $\boldsymbol{u}'$ follows from (3.1) and the plane-wave relations

$$\boldsymbol{u}'=(u', v')=\frac{h'}{\kappa^2 H}\,(\omega\boldsymbol{k}-\mathrm{i}\boldsymbol{f}\times\boldsymbol{k})\,. \tag{3.2}$$

Our plan is to derive the $O(a^2)$ mean-flow response to such slowly varying wave fields.

3.1. *Second-order equations*

Phase-averaged $O(a^2)$ wave properties follow from the real parts of (3.1) and (3.2) in the usual way. In particular, at leading order we can use the plane-wave equalities

$$\overline{h}^S=\frac{\overline{h'^2}}{H}=a^2H\frac{A^2}{2}\quad\text{and}\quad\frac{g\overline{h'^2}}{2H}=\frac{g\overline{h}^S}{2}=\left(1-\frac{f^2}{\omega^2}\right)\frac{E}{2}. \tag{3.3a,b}$$

This shows how A is related to the phase-averaged wave energy E, and that the potential part of the wave energy is generally less than half of the full wave energy for rotating shallow-water waves. We also find that

$$\overline{h}=\tilde{h}\quad\text{and}\quad\boldsymbol{p}=\overline{\boldsymbol{u}}^S=\frac{\overline{h'\boldsymbol{u}'}}{H}=\frac{\boldsymbol{k}}{\omega}E \tag{3.4a,b}$$

all hold to leading order. The equalities in (3.4) are valid for arbitrary values of the Coriolis parameter f, and deviations from these equalities arise only at $O(\mu)$ in the slowly varying parameter.

The mean-flow response is governed by (2.21) and (2.32), but these equations simplify considerably in the wavepacket regime. We assume that the time and space derivatives of the mean flow are $O(\mu a^2)$ in response to the envelope variation of the wavepacket. Higher-order contributions in $\mu \ll 1$ can be neglected. For example, second derivatives of wave properties would yield $O(\mu^2 a^2)$ terms that are negligible. This yields the simpler $O(\mu a^2)$ mean-flow equations

$$\frac{\partial \overline{h}}{\partial t} + H\nabla \cdot \overline{\boldsymbol{u}}^L = 0 \tag{3.5}$$

replacing (2.21), and

$$\frac{\partial \overline{\boldsymbol{u}}^L}{\partial t} + \boldsymbol{f} \times \overline{\boldsymbol{u}}^L + g\nabla \overline{h} = \frac{\partial \boldsymbol{p}}{\partial t} - \frac{g}{2}\nabla \overline{h}^S \tag{3.6}$$

replacing (2.32). Hence, the mean-flow response is governed by the linear rotating shallow-water equations subject to the time-dependent wavepacket-related body force in (3.6). The PV constraint yields

$$\nabla \times \overline{\boldsymbol{u}}^L = \nabla \times \boldsymbol{p} + f\frac{\overline{h} - H}{H}. \tag{3.7}$$

We consider a standard Helmholtz decomposition of $\overline{\boldsymbol{u}}^L$ in terms of a rotational streamfunction ψ^L and a divergent potential ϕ^L, say, such that

$$\overline{\boldsymbol{u}}^L = \widehat{\boldsymbol{z}} \times \nabla \psi^L + \nabla \phi^L \quad \Rightarrow \quad \nabla \times \overline{\boldsymbol{u}}^L = \nabla^2 \psi^L \quad \text{and} \quad \nabla \cdot \overline{\boldsymbol{u}}^L = \nabla^2 \phi^L. \tag{3.8a–c}$$

The PV constraint (3.7) then links ψ^L and $\overline{h}$, but it contains no information on ϕ^L, which depends on the time-dependent solution of the mean-flow response equations.

3.2. *Steady mean-flow response and rotational regimes*

In previously studied cases the slowly varying wave field was augmented with suitable sources and sinks such that a steady wavetrain situation can be reached (BM98, Bühler & McIntyre (2003)). Then all time derivatives can be neglected in (3.5) and (3.6), which yields

$$\nabla \cdot \overline{\boldsymbol{u}}^L = 0 \quad \text{and} \quad \boldsymbol{f} \times \overline{\boldsymbol{u}}^L + g\nabla \overline{h} = -\frac{g}{2}\nabla \overline{h}^S. \tag{3.9a,b}$$

Basically, this has the same effect here as the assumption of a balanced mean flow in § 3.3 of Xie & Vanneste (2015) and throughout the analysis of Wagner & Young (2015). The steady mean-flow response satisfies

$$\phi^L = 0 \quad \text{and} \quad f\psi^L = g\left(\overline{h} + \tfrac{1}{2}\overline{h}^S - H\right). \tag{3.10a,b}$$

Final substitution into the PV constraint (3.7) then yields the modified diagnostic relation for steady wavetrains (its counterpart for general steady wave fields is given in (4.10) below)

$$\left(\nabla^2 - \frac{1}{L_D^2}\right)\psi^L = \nabla \times \boldsymbol{p} - \frac{f\overline{h}^S}{2H}, \quad \text{where} \quad L_D = \frac{c}{f} \tag{3.11}$$

is the Rossby deformation length of the shallow-water system. We note in passing that the term $-f\overline{h}^S/2H$ differs from the corresponding term given in BM98, which is due to an error in BM98 (see appendix A).

It is instructive to consider the typical solution of (3.11) for a fixed wavetrain shape (i.e. fixed $\boldsymbol{k}$ and envelope shape) as a function of the Coriolis parameter f. There are basically three rotational regimes. The first regime is that of weak rotation, in which $f \ll \mu c\kappa$ and hence the deformation scale L_D is much larger than the wavetrain envelope scale $L = 1/(\mu\kappa)$. In this regime (3.11) reduces to $\nabla^2\psi^L = \boldsymbol{\nabla}\times\boldsymbol{p}$, which means that $\overline{\boldsymbol{u}}^L$ is the least-squares projection of $\boldsymbol{p}$ onto non-divergent vector fields. This yields the familiar dipolar Bretherton return flow around the wavepacket (e.g. Bühler & McIntyre 2003). The maximum of $\overline{\boldsymbol{u}}^L$ occurs at the core of the wavetrain and its precise value depends on the aspect ratio of the wavetrain (e.g. § 14.3.3 in Bühler 2014).

In the second regime of moderate rotation $f \approx \mu c\kappa$ and hence the deformation scale L_D is comparable to the envelope scale L, but still much larger than the wavenumber scale $1/\kappa$, so $\omega \gg f$ holds. Clearly, as f increases the magnitude of the pseudomomentum decreases, because $\boldsymbol{p} = \boldsymbol{k}E/\omega$ and ω increases with f. In other words, with increasing f less pseudomomentum is available for a given amount of wave energy E, which tends to reduce the magnitude of the Bretherton flow. To elucidate this regime further we use (3.3) to rewrite (3.11) as

$$\left(\nabla^2 - \frac{1}{L_D^2}\right)\psi^L = -\frac{\boldsymbol{k}\times\boldsymbol{\nabla}E}{\omega} - \frac{f}{c^2}\left(1 - \frac{f^2}{\omega^2}\right)\frac{E}{2}. \tag{3.12}$$

The last term is approximately $-fE/(2c^2)$ in this regime, and we also have the scaling $\boldsymbol{\nabla}E \sim E/L = E/L_D$. It is then easy to check that all terms in (3.12) are of equal magnitude. Hence the mean flow consists of a superposition of a dipolar Bretherton flow (modified by the finite deformation length) together with an anticyclonic mean-flow pattern centred at the wavetrain, which is induced by the negative monopole in the last term. Both flow components decay exponentially with distance from the wavetrain.

Finally, in the regime of strong rotation $f \gg \mu c\kappa$ the deformation scale L_D is much smaller than the envelope scale L. In this regime the primary wave is strongly affected by the rotation (i.e. it is an inertia–gravity wave) and therefore ω and f are comparable. In this regime the undifferentiated rotational terms dominate and hence (3.12) reduces to the algebraic relation

$$f \gg \mu c\kappa\text{:}\qquad -\frac{1}{L_D^2}\psi^L \approx -\frac{f}{c^2}\left(1 - \frac{f^2}{\omega^2}\right)\frac{E}{2} \quad\Rightarrow\quad \psi^L \approx \frac{E}{2f}\left(\frac{c^2\kappa^2}{f^2 + c^2\kappa^2}\right). \tag{3.13a,b}$$

This is a weak anticyclonic flow compactly supported on the footprint of the wavetrain; it is weak because $|\boldsymbol{\nabla}\psi^L| = O(\mu)|\boldsymbol{p}|$, so the mean flow is necessarily small compared to the pseudomomentum (or Stokes drift) magnitude. Ultimately, if $f \gg \mu c\kappa$ then the primary waves are inertial oscillations with $\omega \approx f$ and then (3.13) shows that the mean flow decays as f^{-3}. In summary, for weak rotation $f \ll \mu c\kappa$ we obtain a Bretherton return flow, for moderate rotation $f \approx \mu c\kappa$ we obtain a superposition of the Bretherton flow and an anticyclonic flow pattern centred at the wavetrain, and for strong rotation $f \gg \mu c\kappa$ that anticyclonic flow pattern dominates, but its amplitude decays to zero as $1/f^3$. Overall, increasing f is seen to diminish the mean-flow response to a steady wavetrain.

The anticyclonic flow pattern for strong rotation, with its characteristic sharp confinement to the wavetrain footprint, has already been pointed out qualitatively for the shallow-water model in BM98. Interestingly, it has also been found recently in the three-dimensional rotating Boussinesq flow study in Wagner & Young (2015). As illustrated by their figure 3(*c*,*d*), they observed precisely this kind of flow pattern on a central stratification surface in the case of a vertically ducted mode-one internal wave between two rigid planes. This does not seem to be a coincidence, because on the one hand they assumed a very slowly evolving wavepacket (such that $\mu = a^2$ in the present notation), which makes the steady wavetrain analysis relevant, and on the other hand a vertically ducted mode-one internal wave is known to behave similarly to a shallow-water inertia–gravity wave, at least at the linear level. What appears to be the case, then, is that this similarity extends to the slowly varying weakly nonlinear wave–mean interactions considered here.

3.3. *Freely propagating wavepacket*

The governing equation for $\boldsymbol{\nabla}\cdot\overline{\boldsymbol{u}}^L$ follows from (3.5) and (3.6) as

$$\left(\frac{\partial^2}{\partial t^2}+f^2-c^2\nabla^2\right)\boldsymbol{\nabla}\cdot\overline{\boldsymbol{u}}^L=\frac{\partial}{\partial t}\left(\frac{\partial}{\partial t}\boldsymbol{\nabla}\cdot\boldsymbol{p}+f\boldsymbol{\nabla}\times\boldsymbol{p}-\frac{g}{2}\nabla^2\overline{h}^S\right). \tag{3.14}$$

Similarly, the equations for $\overline{h}$ and $\boldsymbol{\nabla}\times\overline{\boldsymbol{u}}^L$ are

$$\left(\frac{\partial^2}{\partial t^2}+f^2-c^2\nabla^2\right)\frac{\overline{h}-H}{H}=-\left(\frac{\partial}{\partial t}\boldsymbol{\nabla}\cdot\boldsymbol{p}+f\boldsymbol{\nabla}\times\boldsymbol{p}-\frac{g}{2}\nabla^2\overline{h}^S\right) \tag{3.15}$$

and

$$\left(\frac{\partial^2}{\partial t^2}+f^2-c^2\nabla^2\right)\boldsymbol{\nabla}\times\overline{\boldsymbol{u}}^L=\left(\frac{\partial^2}{\partial t^2}-c^2\nabla^2\right)\boldsymbol{\nabla}\times\boldsymbol{p}-f\frac{\partial}{\partial t}\boldsymbol{\nabla}\cdot\boldsymbol{p}+\frac{fg}{2}\nabla^2\overline{h}^S. \tag{3.16}$$

Subtracting f times (3.15) from (3.16) shows that the PV constraint (3.7) is satisfied. It is helpful to apply a Helmholtz decomposition also to $\boldsymbol{p}$ such that

$$\boldsymbol{p}=\widehat{\boldsymbol{z}}\times\boldsymbol{\nabla}\psi^P+\boldsymbol{\nabla}\phi^P\quad\Rightarrow\quad\boldsymbol{\nabla}\times\boldsymbol{p}=\nabla^2\psi^P\quad\text{and}\quad\boldsymbol{\nabla}\cdot\boldsymbol{p}=\nabla^2\phi^P. \tag{3.17a–c}$$

This simplifies the equations for ϕ^L and ψ^L to

$$\left(\frac{\partial^2}{\partial t^2}+f^2-c^2\nabla^2\right)\phi^L=\frac{\partial}{\partial t}\left(\frac{\partial}{\partial t}\phi^P+f\psi^P-\frac{g}{2}\overline{h}^S\right) \tag{3.18}$$

and

$$\left(\frac{\partial^2}{\partial t^2}+f^2-c^2\nabla^2\right)\psi^L=\left(\frac{\partial^2}{\partial t^2}-c^2\nabla^2\right)\psi^P-f\frac{\partial}{\partial t}\phi^P+\frac{fg}{2}\overline{h}^S. \tag{3.19}$$

These are forced shallow-water wave equation with Coriolis forces. Now, equation (3.18) quantifies how a time-dependent wavepacket produces a non-zero ϕ^L as a response. The most important part of that response is the pattern that travels with a freely propagating wavepacket, and in order to compute that pattern it is convenient to move into frame of reference moving with the group velocity $\boldsymbol{c}_g$ of the wavepacket, because in that co-moving frame the pattern is steady. This is achieved by replacing

$\partial/\partial t$ by $-\boldsymbol{c}_g \cdot \boldsymbol{\nabla}$. For definiteness, we consider a wavepacket aligned with the x-axis such that $\boldsymbol{k} = (k, 0)$ and therefore $\partial/\partial t$ is replaced by $-c_g \partial/\partial x$. This leads to the time-independent co-moving equation

$$\left((c^2 - c_g^2)\frac{\partial^2}{\partial x^2} + c^2\frac{\partial^2}{\partial y^2} - f^2\right)\phi^L = c_g\frac{\partial}{\partial x}\left(-c_g\frac{\partial}{\partial x}\phi^P + f\psi^P - \frac{g}{2}\overline{h}^S\right). \tag{3.20}$$

The question is whether ϕ^L decays to zero away from the wavepacket, which is necessary to obtain a well-posed solution. This is true if $f \neq 0$, but not if $f = 0$. Specifically, the group velocity magnitude according to the dispersion relation (2.6) is

$$c_g = \frac{c^2 k}{\omega} \quad \Rightarrow \quad \frac{c_g^2}{c^2} = \frac{c^2 k^2}{f^2 + c^2 k^2} \leqslant 1. \tag{3.21a,b}$$

The inequality is strict if $f \neq 0$, i.e. in the rotating case. This shows that if $f \neq 0$ then the left-hand side in (3.20) is a scaled modified Helmholtz operator in x and y, and hence the solution ϕ^L indeed decays away from the wavepacket, albeit more slowly in y than in x. However, if $f = 0$ then (3.20) reduces to

$$c^2\frac{\partial^2}{\partial y^2}\phi^L = c\frac{\partial}{\partial x}\left(-c\frac{\partial}{\partial x}\phi^P - \frac{g}{2}\overline{h}^S\right), \tag{3.22}$$

the solution of which does not decay in y away from the wavepacket. This means that if $f = 0$ then the true solution to the wavepacket forcing does not settle down to a co-moving pattern, but instead remains time-dependent even in the wavepacket frame. This is a familiar resonance phenomenon in non-rotating compressible flows, where localized forcing moving with the sound speed c produces a time-dependent bow wave that spreads to larger and larger distances in the direction transversely to the moving forcing region. The non-rotating shallow-water equations are equivalent to two-dimensional homentropic compressible fluid equations and hence this analogy is relevant. We therefore conclude that a compactly supported co-moving mean-flow response to a freely propagating wavepacket in shallow water is only possible in the rotating case $f \neq 0$.

3.4. *Comparison of steady and unsteady wavepacket response*

It is instructive to compare the mean-flow response $\overline{\boldsymbol{u}}^L$ based on a steady wavetrain to that corresponding to a freely moving, unsteady wavepacket. For example, this illustrates the error that can be expected if the steady wavetrain results are used as an approximation to the co-moving case, as would be the case if the mean flow were treated as balanced. This comparison is straightforward for the potential part ϕ^L of $\overline{\boldsymbol{u}}^L$, which is simply zero in the steady approximation, but given by the solution of (3.20) in the co-moving case. For the rotational part ψ^L the situations is more interesting because ψ^L is non-zero in both cases. We again consider a wavepacket propagating along the x-axis and denote the steady approximation to the streamfunction by ψ_s^L and its unsteady, co-moving counterpart by ψ_u^L. It then follows from (3.11), (3.16) and (3.21) that

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - \frac{1}{L_D^2}\right)\psi_s^L = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\psi^P - \frac{f}{2c^2}g\overline{h}^S \tag{3.23}$$

whilst

$$\left(\frac{f^2}{\omega^2}\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}-\frac{1}{L_D^2}\right)\psi_u^L=\left(\frac{f^2}{\omega^2}\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)\psi^P-\frac{f}{2c^2}g\bar{h}^S-\kappa\frac{f}{\omega}\frac{\partial}{\partial x}\phi^P. \tag{3.24}$$

These two equations differ by the scaled x-derivatives, which is the direction of propagation of the wavepacket, as well as by an additional ϕ^P source term in (3.24). For a wavepacket with uniform aspect ratio $\boldsymbol{\nabla}\cdot\boldsymbol{p}$ is comparable to $\boldsymbol{\nabla}\times\boldsymbol{p}$, so this term cannot be expected to be small in general. Still, in the weak rotation regime ($f\ll\mu c\kappa$) both equations simply reduce to the Bretherton flow, i.e. $\psi_s^L=\psi_u^L=\psi^P$. However, in the moderate rotation regime ($f\sim\mu c\kappa$) their solutions can already be expected to differ noticeably, as all the different terms play a role. Finally, in the strong rotation regime ($f\geqslant c\kappa$) both equations can be inverted analytically by ignoring the second derivative terms, which yields

$$f\geqslant c\kappa\text{:}\qquad \psi_s^L=\frac{g\bar{h}^S}{2f}\quad\text{and}\quad \psi_u^L=\frac{g\bar{h}^S}{2f}+\frac{c^2\kappa}{f\omega}\frac{\partial}{\partial x}\phi^P. \tag{3.25a,b}$$

Hence the co-moving solution for ψ^L has an additional term, which is of the same order of magnitude as the preceding anticyclonic monopole term. Moreover, the additional term is non-local, i.e. it is not confined to the wavepacket footprint. Specifically, in the present example $\nabla^2\phi^P=\boldsymbol{\nabla}\cdot\boldsymbol{p}\propto\partial E/\partial x$, which implies that outside the wavepacket $\partial\phi^P/\partial x\propto-\cos(2\theta)/r^2$ in polar coordinates (r,θ) centred at the wavepacket. This predicts that for strong rotation the co-moving ψ^L develops negative dips fore and aft of the wavepacket, whereas no such dips are present in the steady approximation. This will be confirmed by the simulations in § 5.1 below. In summary, for moderate and strong rotation we can expect the co-moving ψ^L for a freely propagating wavepacket to differ markedly from its steady wavetrain approximation.

4. Monochromatic standing waves

The results so far suggest that increasing the background rotation f leads to a reduction in the mean-flow response $\bar{\boldsymbol{u}}^L$, such that the Lagrangian-mean flow for strong rotation is found to be substantially weaker than the Stokes drift or pseudomomentum. Similar findings have been established in BM98 and more recently in Wagner & Young (2015). Moreover, the same broad tendency was also found by Bühler & Holmes-Cerfon (2009) in the context of particle diffusion induced by random waves in the rotating shallow-water system, where it turned out that the effective diffusivity induced by the waves was a decreasing function of f (see their figure 2*a*). It is therefore instructive to display a simple counterexample, i.e. a situation in which the mean-flow response $\bar{\boldsymbol{u}}^L$ is not maximized in the non-rotating limit. Such behaviour turns out to arise in the case of a monochromatic standing wave, which could be relevant for forced wave motion such as tides in the ocean, for example.

4.1. *Linear standing wave fields*

A monochromatic standing wave is defined by a height field disturbance of the form

$$h'(x,y,t)=\cos\omega t\,\hat{h}(x,y), \tag{4.1}$$

where $\hat{h}(x, y)$ is a real-valued eigenfunction of the two-dimensional Laplacian such that

$$\nabla^2\hat{h} = -\frac{\omega^2 - f^2}{c^2}\hat{h}. \tag{4.2}$$

We disregard the resonant case $\omega^2 = f^2$, but both propagating waves with $\omega^2 > f^2$ and evanescent waves with $\omega^2 < f^2$ are included. The corresponding linear velocity field is readily computed as

$$\boldsymbol{u}' = \frac{g\omega}{\omega^2 - f^2}\left(-\sin\omega t\,\nabla\hat{h} - \cos\omega t\frac{\boldsymbol{f}\times\nabla\hat{h}}{\omega}\right) \tag{4.3}$$

and the particle displacements are

$$\boldsymbol{\xi}' = \frac{g}{\omega^2 - f^2}\left(\cos\omega t\,\nabla\hat{h} - \sin\omega t\frac{\boldsymbol{f}\times\nabla\hat{h}}{\omega}\right). \tag{4.4}$$

We use time averaging over the monochromatic wave period $2\pi/\omega$ such that

$$\overline{\cos^2\omega t} = \overline{\sin^2\omega t} = \tfrac{1}{2} \quad \text{and} \quad \overline{\cos\omega t\sin\omega t} = 0. \tag{4.5a,b}$$

The Stokes correction for h is

$$\overline{h}^S = \frac{g}{\omega^2 - f^2}\frac{1}{2}|\nabla\hat{h}|^2. \tag{4.6}$$

Notably, this implies $\overline{h}^L \geqslant \overline{h}$ for propagating waves ($\omega^2 > f^2$) and *vice versa* for evanescent waves. Straightforward if lengthy computations also show that

$$\overline{\boldsymbol{u}}^S = -\left(\frac{g}{\omega^2 - f^2}\right)^2 \boldsymbol{f}\times\nabla\left(\frac{1}{2}|\nabla\hat{h}|^2 + \frac{\omega^2 - f^2}{4c^2}\hat{h}^2\right) \tag{4.7}$$

and that the Lagrangian-mean pseudomomentum (2.13) is

$$\boldsymbol{p} = \left(\frac{g}{\omega^2 - f^2}\right)^2\left(3 - \frac{f^2}{\omega^2}\right)\frac{1}{4}\left[(\boldsymbol{f}\times\nabla\hat{h})\cdot\nabla\right]\nabla\hat{h}. \tag{4.8}$$

Clearly, $\overline{\boldsymbol{u}}^S$ and $\boldsymbol{p}$ are now given by very different expressions, in contrast with their near equality in the slowly varying case. Importantly, both fields are identically zero if $f = 0$.

4.2. *Mean-flow response*

The monochromatic wave field has steady mean $O(a^2)$ fields by construction and hence the steady mean-flow response according to (2.21) and (2.32) leads again to (3.10), i.e.

$$\phi^L = 0 \quad \text{and} \quad f\psi^L = g\left(\overline{h} + \tfrac{1}{2}\overline{h}^S - H\right). \tag{4.9a,b}$$

The solution is completed by using the full expression for the PV constraint in (2.26), whereas previously we ignored its last term due to the assumption of slowly varying waves. This results in the modified diagnostic relation

$$\left(\nabla^2 - \frac{1}{L_D^2}\right)\psi^L = \boldsymbol{\nabla}\times\boldsymbol{p} - \frac{f}{2}\left(\frac{\overline{h}^S}{H} + \overline{(\xi_i'\xi_j')}_{,ij}\right). \tag{4.10}$$

Clearly, if $f=0$ then the right-hand side of (4.10) simply vanishes, because $\boldsymbol{p}$ as given by (4.8) is zero in that limit. This means that ψ^L and therefore $\overline{\boldsymbol{u}}^L$ are both zero if $f=0$. There is no Bretherton flow in this limit, in fact there is no mean motion at all, although the mean height field $\overline{h}=H-(\overline{h}^S)/2$ is undulated according to (4.6). This establishes that non-zero mean motion in response to a monochromatic standing wave requires non-zero values of f, so this counterexample illustrates that $\overline{\boldsymbol{u}}^L$ is not necessarily maximized in the non-rotating limit. This will be illustrated by numerical simulations below.

5. Direct numerical simulations

We undertook direct numerical simulations of the fully nonlinear shallow-water equations (2.1) in order to the test and evaluate the predictions for the mean-flow response derived in the previous sections. The simulation tool was a standard pseudo-spectral code for the rotating shallow-water equations with doubly periodic boundary conditions. The size of the square domain was $(150\pi)^2$, the grid resolution was 512^2, and a hyper-diffusion operator $\nu\nabla^8$ with $\nu = 10^{-13}$ was added to for numerical stability.

5.1. *Wavepacket simulations*

The aim of these simulations is to check the validity of the mean-flow analysis in § 3.4. A wavepacket was chosen according to (3.1) and (3.2) with parameters

$$a=0.001, \quad g=H=c=1, \quad \boldsymbol{k}=(1,1), \quad \kappa=\sqrt{2}. \tag{5.1a–d}$$

This choice of $\boldsymbol{k}$ implies that the group velocity is at a 45° angle in the xy-plane. The very low wave amplitude was necessary in order to avoid complications due to shock formation in the non-rotating case. As discussed in Bühler (1998), without rotation a plane shallow-water wave with amplitude $a=0.1$ forms a shock already within a single wavelength of propagation. With the present choice of $a=0.001$ the shock formation is delayed to a hundred wavelengths, which was enough for our purpose. The wavepacket envelope was the slowly varying Gaussian

$$A = \exp\left(-\frac{\mu^2\kappa^2(x^2+y^2)}{2}\right) \quad \text{with } \mu = 0.0316. \tag{5.2}$$

Three different values of the Coriolis parameter

$$f=\{0,\,0.02,\,2\} \quad \Rightarrow \quad \frac{f}{c\kappa}=\{0,\,0.014,\,1.4\} \quad \Rightarrow \quad \frac{f}{\mu c\kappa}=\{0,\,0.45,\,4.5\} \tag{5.3a–c}$$

were chosen to illustrate the role of rotation. They broadly correspond to the regimes of weak, moderate and strong rotation discussed before.

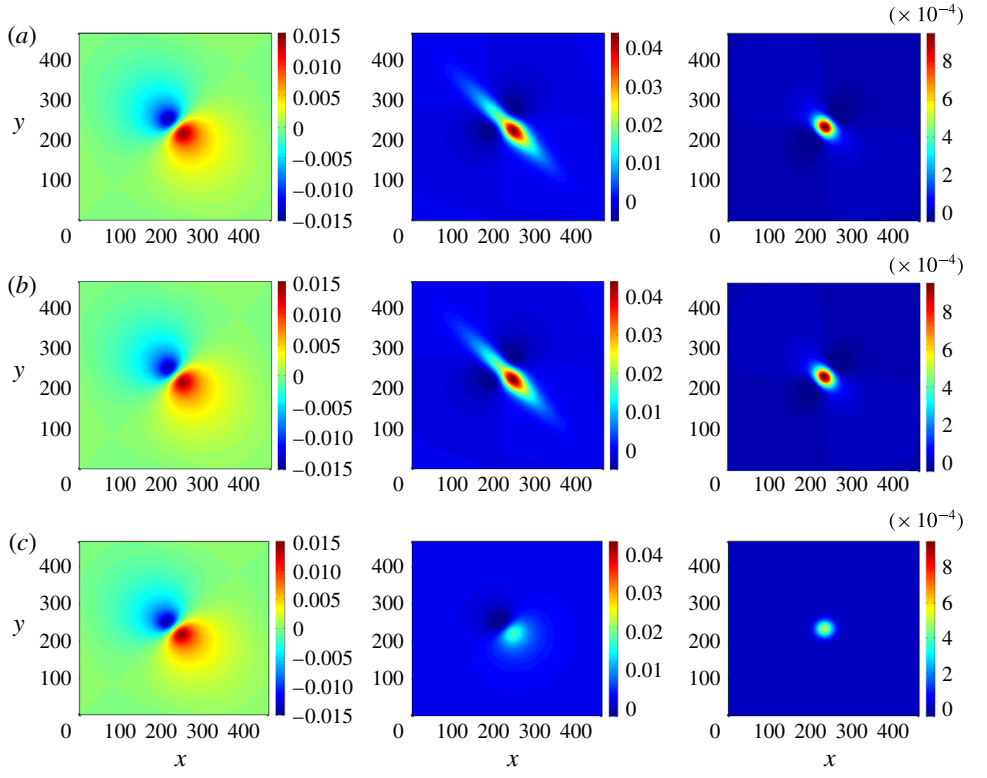

FIGURE 1. Snapshots of ψ^L from wavepacket simulations. The columns from left to right correspond to the cases of weak, moderate, and strong rotation in (5.3). (*a*) Shows analytic results based on the co-moving equations in § 3.4, (*b*) is from the direct numerical simulations and (*c*) shows analytic results based on the steady approximation in § 3.4.

The Eulerian averaging operator was implemented as an average over the phase α in the wavepacket fields based on (3.1). This meant that we ran several wavepacket simulations with different values of α and then averaged over these runs. This led to a very clear and satisfactory elimination of the fast spatial scales.

Now, the nonlinear simulations were initialized with a superposition of the $O(a)$ wavepacket fields and the anticipated co-moving $O(a^2)$ Eulerian-mean-flow response. However, despite our best efforts, the mean flow in the simulations with non-zero f exhibited strong start-up inertial oscillations centred at the initial wavepacket position. We hypothesize that these strong start-up inertial oscillations are somehow related to the exact higher-derivative terms in the mean-flow equations (2.21) and (2.32), which were omitted in the wavepacket analysis. In order to obtain clear numerical results in this challenging situation, we had to perform another crucial averaging step: for runs with non-zero f we time averaged the already phase-averaged mean flow over the inertial period $2\pi/f$, which filtered the start-up inertial oscillations. After we eventually settled on this non-trivial numerical procedure very clear numerical results were obtained, which are illustrated in the figures.

Figure 1 shows results for ψ^L at a moment in the simulations when the wavepacket is in the centre of the domain. It illustrates excellent agreement between the analytic results obtained in the co-moving frame and the direct numerical simulations. On the

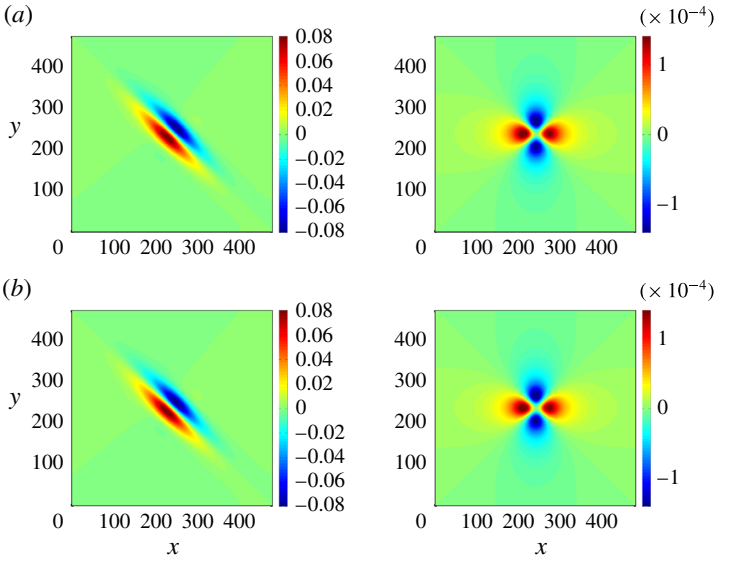

FIGURE 2. Snapshots of ϕ^L from wavepacket simulations with moderate rotation on the left and strong rotation on the right. (*a*) The analytic solution for ϕ^L in the co-moving frame and (*b*) shows results from the direct numerical simulation. No results are shown for the steady approximation, in which $\phi^L = 0$ identically. Figure 3 below shows simulation results for the $f = 0$ case, in which a steady co-moving solution for ϕ^L does not exist (cf. § 3.3).

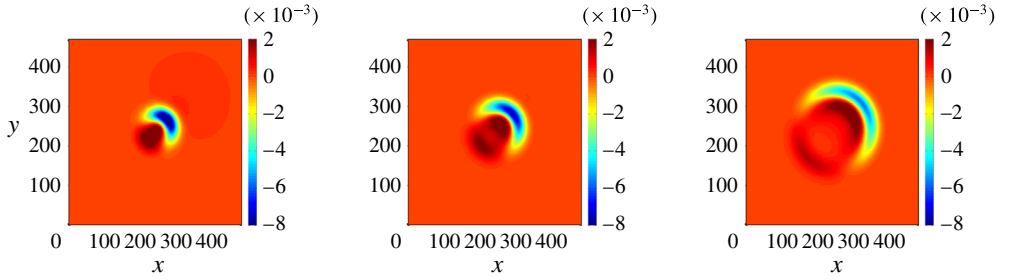

FIGURE 3. Direct wavepacket simulation time sequence of ϕ^L in the non-rotating case $f = 0$ at three different times ($t = 40$, 60 and 100), illustrating that ϕ^L does not reach a steady state in this case.

other hand, the steady approximation is only really accurate in the non-rotating case, when all three methods reduce to the Bretherton flow. As expected from § 3.4, in the strong rotation case the steady approximation captures the monopole supported on the wavepacket footprint, but it misses the non-local part of ψ^L outside the wavepacket.

Figure 2 shows analogous results for ϕ^L in the case of moderate and strong rotation, omitting the non-rotating case for which no co-moving result exists. Again, there is excellent agreement between the analytic results in the co-moving frame and the direct numerical simulations. Notably, the magnitude of ϕ^L is comparable to the magnitude of ψ^L in these cases, so both components are needed for an accurate prediction of $\overline{\boldsymbol{u}}^L$. Finally, in the non-rotating case a time sequence of ϕ^L centred at the wavepacket shown in figure 3 illustrates the growing and widening bow wave in this resonant case.

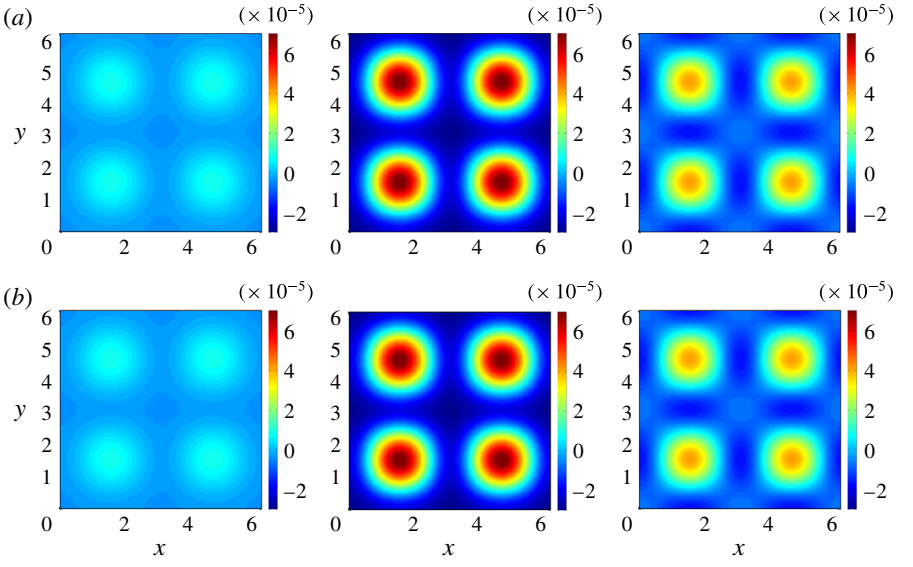

FIGURE 4. Snapshots of ψ^L for standing wave case with $\hat{h} = \cos x \cos y$. The potential $\phi^L = 0$ for standing waves. Asymptotic results are in (a) and direct simulation results are in (b). $f = 0.1$, 1 and 10 from left to right.

5.2. *Standing wave simulations*

For the standing wave field we chose the simple pattern $\hat{h} = \cos x \cos y$ and this field together with the corresponding linear velocity fields was then used to initialize the numerical simulations. Results are presented in figure 4. As predicted by inspecting (4.8) and (4.10), the mean flow vanishes at $f = 0$ and on increasing f the Lagrangian flow increases in magnitude, reaches a peak, and then begins to decay. Although the strength of the mean flow varies with f, unlike the case of a wavepacket, the spatial structure of the mean flow remains the same, at least for this very simple choice of $\hat{h}$.

5.3. *Kinetic mean-flow energy dependence on f*

We computed the dependence of the total, domain-integrated mean-flow kinetic energy on the rotation rate f for both wavepackets and standing waves. As f was varied the wave amplitudes were adjusted as needed in order to keep the total energy of the wave field fixed. Results are shown in figure 5(a) corresponding to standing waves and the lower panel to wavepackets.

In both cases we find that for very large f, the mean flow is 'choked' and its kinetic energy decreases to zero, as noted before. However, for low f the behaviour is markedly different in the two cases: for standing waves the mean-flow kinetic energy goes to zero also at $f = 0$, and hence there is an energy maximum for some intermediate value of f, whilst for wavepackets the mean-flow kinetic energy is monotonically decreasing with f. Furthermore, from figure 5(b), we see that at low rotation rates the potential part contributes more with negligible contribution from streamfunction part, a scenario that reverses at high rotation rates. Hence for wavepackets we observe a non-trivial cross-over of mean-flow kinetic energy from its divergent to its rotational form as f is increased. This is not the case for standing waves, of course, for which the divergent part of $\overline{\boldsymbol{u}}^L$ is always zero.

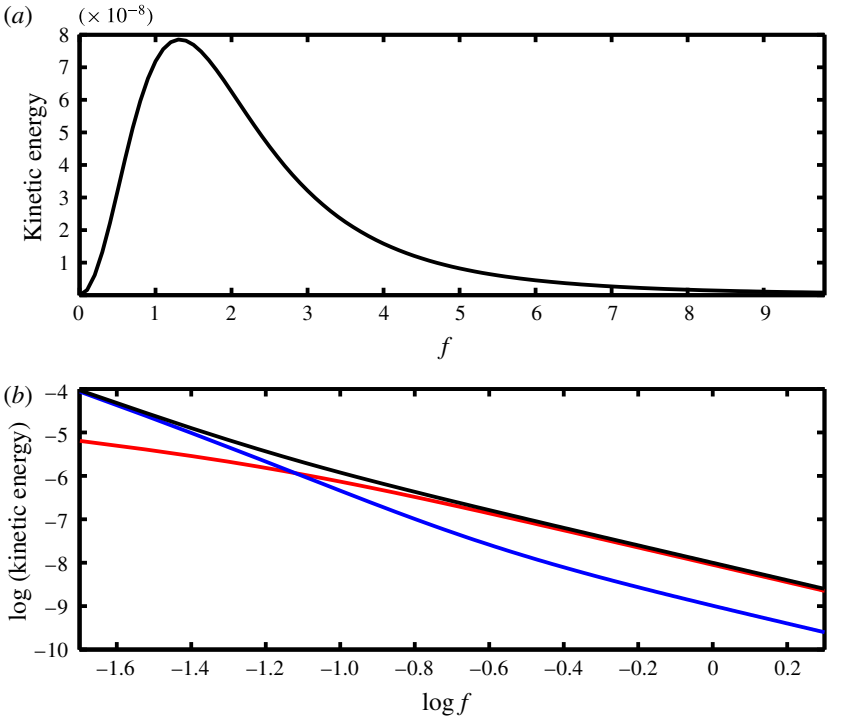

FIGURE 5. Domain-integrated kinetic mean-flow energy versus f for standing wave (a) and wavepacket (b) cases. The total energy of the wave field was fixed whilst varying f. The black curve is the total mean-flow kinetic energy and the blue and red curves indicate its divergent and rotational components, respectively. For the standing wave the divergent, ϕ^L-related part of the energy is identically zero and hence the black and red curves coincide. Note the logarithmic scales on the second plot.

6. Concluding remarks

We have computed the full $O(a^2)$ wave-induced mean-flow response in rotating shallow water under the key assumption of uniform PV in a number of illustrative examples. As we did not assume at the outset that the mean flow had to be balanced, we were able to compute the full mean-flow response, which includes wave-like contributions to $\overline{\boldsymbol{u}}^L$ mediated by the potential ϕ^L. For steady waves these contributions are zero and we recovered the modified diagnostic relations between ψ^L and certain wave properties that are usually associated with balanced mean flows. For unsteady wavepackets it is less obvious how to define a useful modified diagnostic relation, but this was achieved by focusing on the co-moving velocity field $\overline{\boldsymbol{u}}^L$ travelling with the wavepacket. Notably, for rotating flows this co-moving field contained comparable contributions from both ψ^L and ϕ^L, so in this sense the concept of a modified diagnostic relation was here extended to include ϕ^L as well. *Mutatis mutandis*, we expect these shallow-water results to generalize to three-dimensional rotating Boussinesq flows, including the peculiar finding that for standing waves the mean-flow response is maximized at an intermediate value of the Coriolis parameter f. Likewise, the extension of the present examples to weak PV perturbations would presumably lead to the usual quasi-geostrophic terms being

added to the modified diagnostic relations. Of course, the nature and consequence of the long-time interactions between non-trivial PV signals and wave fields remain open questions.

We note in closing a heuristic application of our wavepacket analysis to the effective horizontal diffusion of particles in rotating shallow water. This problem was solved in Bühler & Holmes-Cerfon (2009) in the context of wave fields realized as small-amplitude Gaussian random fields. Alternatively, one could realize a prescribed wave power spectrum as an uncorrelated ensemble of compactly supported wavepackets, to each of which the mean-flow analysis of the present paper applies. (The phase-averaging method deployed in § 5.1 yields a smooth $\overline{\boldsymbol{u}}^L$, but we believe that a single wavepacket experiment would produce the same irreversible particle displacements augmented by further reversible oscillations on the scale of the primary wave crests.) Specifically, any particle would then experience the time history of the co-moving velocity field $\overline{\boldsymbol{u}}^L$ due to a passing wavepacket. At $O(a^2)$, the net displacement suffered by the particle due to such a velocity field is therefore proportional to the line integral of the co-moving velocity field along a line parallel to the group velocity and intersecting the starting position of the particle. For example, these are lines at a 45° angle in the figures in § 5.1. The net displacement therefore only depends on the distance between the original particle position and the wavepacket track. Moreover, it appears from inspection of the spatial patterns of ψ^L and ϕ^L in figures 1 and 2 that only the rotational, streamfunction pattern makes a non-zero contribution to the net displacement of the particle, because the integral of ϕ^L contributions cancels along lines parallel to the group velocity whereas the integral of ψ^L does not. If this heuristic argument is true it would indicate that the non-zero ϕ^L fields in the wavepacket case, which are comparable in magnitude to the ψ^L fields, nevertheless do not lead to lasting particle displacements, and are therefore of secondary importance for transport and diffusion.

Acknowledgements

We gratefully acknowledge constructive and thought-provoking comments by the referees, which improved our paper. O.B. acknowledges financial support from the United States National Science Foundation grant DMS-1312159 and Office of Naval Research grant N00014-15-1-2355. K.S.S. and J.T. acknowledge support from the Center for Prototype Climate Modeling, a unit of the New York University Abu Dhabi Institute.

Appendix A. Error in BM98

The second term in the diagnostic relation (3.11) for the steady wavetrain response can be written as

$$-\frac{f\overline{h}^S}{2H} = -\frac{f}{c^2}\frac{g\overline{h}^S}{2} = -\frac{f}{c^2}\frac{g\overline{h'^2}}{2H} = -\frac{f}{c^2}\left(1-\frac{f^2}{\omega^2}\right)\frac{E}{2}. \tag{A 1}$$

On the other hand, combining (6.6) and (7.1) from BM98 in the present shallow-water case ($\gamma = 2$) yields a different term, namely

$$-\frac{f}{c^2}\frac{1}{2}\overline{|\boldsymbol{u}'|^2} = -\frac{f}{c^2}\left(1+\frac{f^2}{\omega^2}\right)\frac{E}{2}. \tag{A 2}$$

Hence in the present paper the term is proportional to the potential part of the wave energy whereas in BM98 it was proportional to the kinetic part of the wave energy. Those are not equal and there is a quantitative discrepancy with BM98.

The reason for the discrepancy is a subtle error in the manipulation of the mean momentum equation in BM98. This involved using (2.5) to write

$$-\overline{(\boldsymbol{u}'\cdot\boldsymbol{\nabla})\boldsymbol{u}'}=-\frac{1}{2}\boldsymbol{\nabla}\overline{|\boldsymbol{u}'|^2}-\overline{(\boldsymbol{\nabla}\times\boldsymbol{u}')\times\boldsymbol{u}'}=-\frac{1}{2}\boldsymbol{\nabla}\overline{|\boldsymbol{u}'|^2}-\boldsymbol{f}\times\frac{\overline{h'\boldsymbol{u}'}}{H} \tag{A 3}$$

and then replacing the last term with $-\boldsymbol{f}\times\overline{\boldsymbol{u}}^S$ based on the plane-wave relations (3.4). However, this introduces an $O(\mu a^2)$ error term, as the plane-wave relations are not exact for a slowly varying wavetrain. The omitted error term is

$$-\boldsymbol{f}\times\left(\frac{\overline{h'\boldsymbol{u}'}}{H}-\overline{\boldsymbol{u}}^S\right)=\boldsymbol{f}\times\left(\overline{(\xi_j'\boldsymbol{u}')}_j\right)=-\boldsymbol{\nabla}f\overline{\xi'v'}=-\frac{1}{2}\boldsymbol{\nabla}\overline{\boldsymbol{u}'\cdot\boldsymbol{f}\times\boldsymbol{\xi}'}. \tag{A 4}$$

This uses $\overline{\xi'u'}=\overline{\eta'v'}=0$ and $\overline{\eta'u'}=-\overline{\xi'v'}$ in the present situation of a steady wavetrain. Combining this with the first term in (A 3) and applying the steady virial theorem (2.12) finally yields

$$-\frac{1}{2}\boldsymbol{\nabla}\left(\overline{|\boldsymbol{u}'|^2}+\overline{\boldsymbol{u}'\cdot\boldsymbol{f}\times\boldsymbol{\xi}'}\right)=-\frac{g\overline{h}^S}{2}, \tag{A 5}$$

which eventually recovers the correct term, as in the present paper.

REFERENCES

ALFORD, M. H., MACKINNON, J. A., SIMMONS, H. L. & NASH, J. D. 2016 Near-inertial internal gravity waves in the ocean. *Ann. Rev. Marine Sci.* **8**, 95–123.

ANDREWS, D. G., HOLTON, J. R. & LEOVY, C. B. 1987 *Middle Atmosphere Dynamics*. Academic Press.

ANDREWS, D. G. & MCINTYRE, M. E. 1978 An exact theory of nonlinear waves on a Lagrangian-mean flow. *J. Fluid Mech.* **89**, 609–646.

BABIN, A., MAHALOV, A. & NICOLAENKO, B. 1997 Global splitting and regularity of rotating shallow-water equations. *Eur. J. Mech.* B **16** (5), 725–754.

VAN DEN BREMER, T. S. & SUTHERLAND, B. R. 2014 The mean flow and long waves induced by two-dimensional internal gravity wavepackets. *Phys. Fluids* **26** (10), 106601.

BRETHERTON, F. P. 1969 On the mean motion induced by internal gravity waves. *J. Fluid Mech.* **36**, 785–803.

BÜHLER, O. 1998 A shallow-water model that prevents nonlinear steepening of gravity waves. *J. Atmos. Sci.* **55**, 2884–2891.

BÜHLER, O. 2000 On the vorticity transport due to dissipating or breaking waves in shallow-water flow. *J. Fluid Mech.* **407**, 235–263.

BÜHLER, O. 2014 *Waves and Mean Flows*. Cambridge University Press.

BÜHLER, O. & HOLMES-CERFON, M. 2009 Particle dispersion by random waves in rotating shallow water. *J. Fluid Mech.* **638**, 5–26.

BÜHLER, O. & JACOBSON, T. E. 2001 Wave-driven currents and vortex dynamics on barred beaches. *J. Fluid Mech.* **449**, 313–339.

BÜHLER, O. & MCINTYRE, M. E. 1998 On non-dissipative wave–mean interactions in the atmosphere or oceans. *J. Fluid Mech.* **354**, 301–343.

BÜHLER, O. & MCINTYRE, M. E. 2003 Remote recoil: a new wave–mean interaction effect. *J. Fluid Mech.* **492**, 207–230.

ELIASSEN, A. & PALM, E. 1961 On the transfer of energy in stationary mountain waves. *Geofys Publ.* **22** (3), 1–23.

GARRETT, C. & KUNZE, E. 2007 Internal tide generation in the deep ocean. *Annu. Rev. Fluid Mech.* **39**, 57–87.

LIGHTHILL, J. 1978 *Waves in Fluids*. Cambridge University Press.

MAJDA, A. J. & EMBID, P. 1998 Averaging over fast gravity waves for geophysical flows with unbalanced initial data. *Theoret. Comput. Fluid Dynam.* **11** (3), 155–169.

MCINTYRE, M. E. 1988 A note on the divergence effect and the Lagrangian-mean surface elevation in periodic water waves. *J. Fluid Mech.* **189**, 235–242.

PLUMB, R. A. 1977 The interaction of two internal waves with the mean flow: implications for the theory of the quasi-biennial oscillation. *J. Atmos. Sci.* **34**, 1847–1858.

THOMAS, J., SMITH, K. S. & BÜHLER, O. 2017 Near-inertial wave dispersion by geostrophic flows. *J. Fluid Mech.* **817**, 406–438.

VALLIS, G. K. 2006 *Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-scale Circulation*. Cambridge University Press.

WAGNER, G. L. & YOUNG, W. R. 2015 Available potential vorticity and wave-averaged quasi-geostrophic flow. *J. Fluid Mech.* **785**, 401–424.

WAGNER, G. L. & YOUNG, W. R. 2016 A three-component model for the coupled evolution of near-inertial waves, quasi-geostrophic flow and the near-inertial second harmonic. *J. Fluid Mech.* **802**, 806–837.

XIE, J.-H. & VANNESTE, J. 2015 A generalised-lagrangian-mean model of the interactions between near-inertial waves and mean flow. *J. Fluid Mech.* **774**, 143–169.