This preprint is the final version published in *J. Fluid Mech.* **881**, 182–217 (2019), apart from a few changes dictated by the journal's house style.

Wave–vortex interactions, remote recoil, the Aharonov–Bohm effect and the Craik–Leibovich equation

Michael Edgeworth McIntyre[1]†,

[1]Department of Applied Mathematics and Theoretical Physics, Cambridge CB3 0WA, UK

(Received 11 May 2018; revised 3 March & 25 July 2019; accepted 2 September 2019)

Three examples of non-dissipative yet cumulative interaction between a single wavetrain and a single vortex are analysed, with a focus on effective recoil forces, local and remote. Local recoil occurs when the wavetrain overlaps the vortex core. All three examples comply with the pseudomomentum rule. The first two examples are two-dimensional and non-rotating (shallow water or gas dynamical). The third is rotating, with deep-water gravity waves inducing an Ursell "anti-Stokes flow". The Froude or Mach number, and the Rossby number in the third example, are assumed small. Remote recoil is all or part of the interaction in all three examples, except in one special limiting case. That case is found only within a severely restricted parameter regime and is the only case in which, exceptionally, the effective recoil force can be regarded as purely local and identifiable with the celebrated Craik–Leibovich vortex force – which corresponds, in the quantum fluids literature, to the Iordanskii force due to a phonon current incident on a vortex. Another peculiarity of that exceptional case is that the only significant wave refraction effect is the Aharonov–Bohm topological phase jump.

1. Introduction

In the vast literature on wave–mean and wave–vortex interactions, there is a tradition of thinking in terms of wave-induced mean forces and the associated wave-induced momentum fluxes or radiation stresses. The tradition goes back many years, to the work of Lord Rayleigh, Léon Brillouin and other great physicists. It continues today in, for instance, work on the fluid dynamics of atmospheres and oceans, as well as on quantum vortices where the wave-induced mean forces are called "Iordanskii forces".

Within the atmosphere–ocean community, the force-oriented viewpoint is important because wave-induced mean forces are recognized as key to solving what used to be three great enigmas – three grand challenges – in atmospheric science. They were to understand the quasi-biennial oscillation of the zonal winds in the equatorial stratosphere, the "antifrictional" self-sharpening of jet streams, and the gyroscopic or Coriolis pumping of global-scale mean circulations in the stratosphere and mesosphere (i.e. between altitudes $\sim$10–100 km) and the consequent water vapour, ozone and pollutant transport and, most dramatically, the refrigeration of the summer mesopause – down to temperatures $\sim 100°$C below radiatively determined temperatures. The history is tortuous and goes back to the 1960s, when all three phenomena were observationally conspicuous but, in terms of mechanism, completely mysterious. See for instance Wallace & Holton (1968), Fritts (1984), Holton *et al.* (1995), Baldwin *et al.* (2001), Dritschel & McIntyre (2008), and Garcia *et al.* (2017).

Recognition of wave-induced mean forces as key to solving all three enigmas, and as

† Email address for correspondence: mem@damtp.cam.ac.uk

essential components of weather and climate models, constituted a gradual, but major, paradigm shift regarding the nature of large-scale momentum transport in atmospheres and oceans. Before the 1960s, such transport tended to be thought of in terms of turbulent eddy viscosities, missing the point that wave-induced momentum transport can be a long-range process more likely to dominate, as in fact it does, over large scales limited not by parcel displacements and mixing lengths but instead by the far greater distances over which waves can propagate.

Much of the atmosphere–ocean literature, especially where it deals with mean forces induced by gravity waves, often takes for granted that the forces can be computed from linearized wave theory alone using what is now called the *pseudomomentum rule* (e.g. Bühler 2014, hereafter B14). That is the accepted basis of gravity-wave "parametrization schemes" in weather and climate models, designed to incorporate the mean forces coming from gravity waves whose wavelengths are too small to be resolved explicitly. It is also the usual basis on which, for instance, Iordanskii forces are computed (e.g. Sonin 1997; Stone 2000*a*). Pseudomomentum, also called quasimomentum or wave momentum, or phonon momentum, is the linear-theoretic wave property whose nondissipative conservation depends, through Noether's theorem, on translational invariance of the mean or background state on which the waves propagate, as distinct from translational invariance of the entire physical system, background plus waves, which implies conservation of momentum.

In a linearized ray-theoretic description the pseudomomentum $\mathbf{p}$ per unit mass is $\mathcal{A}\boldsymbol{k}$, where $\boldsymbol{k}$ is the wavenumber vector and $\mathcal{A}$ the wave-action, i.e. wave-energy divided by intrinsic frequency, per unit mass (e.g. Bretherton & Garrett 1968). The wave-action, wave-energy and pseudomomentum are linear-theoretic wave properties and are $O(a^2)$ in magnitude, where a measures wave amplitude (and will be defined in such a way that $a \ll 1$ validates linearization). The pseudomomentum rule says that $O(a^2)$ wave-induced mean forces can be calculated as if pseudomomentum were momentum and the fluid medium were absent.

As discussed for instance in B14, the rule has been justified mainly for simplified mean flows that are themselves translationally invariant. In such cases it is typical, as is well known, for persistent mean forces with cumulative effects to arise only when the waves break or are otherwise dissipated, leading to a persistent pseudomomentum-flux convergence. However, when the waves are refracted by realistic, three-dimensional backgrounds involving vortices, the situation is fundamentally different. One can get persistent mean forces with cumulative effects in the absence of wave dissipation. Also, it is unclear whether, or when, or in what sense the pseudomomentum rule should hold. The fluid medium is not absent, and it supports a mean pressure field that mediates long-range mean forces of the same order, $O(a^2)$, as those computed from the pseudomomentum rule. Such pressure fields are not wave properties and cannot be computed from linearized wave theory alone. Rather, they require the solution of equations governing the mean or background state correct to $O(a^2)$. Cases in which the rule fails for this reason have long been known, going back as far as Brillouin's pioneering work on acoustic radiation stress (e.g. Brillouin 1936, B14 §12.2.2).

For gravity-wave parametrization, in particular, there are therefore unresolved questions as to how to compute, and indeed how to think about, the wave-induced mean forces for realistic, three-dimensional backgrounds. Current parametrization schemes ignore these questions because they altogether neglect horizontal refraction, giving rise to what is sometimes called the "missing forces" problem for such schemes.

The simplest wave–vortex problems in which these questions arise are to be found in a two-dimensional, non-rotating shallow-water or acoustical setting, with no viscous or other wave dissipation. Two basic examples, the main examples to be studied in this

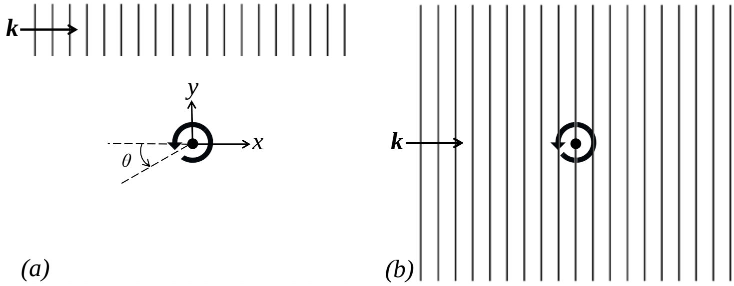

FIGURE 1. Panels (**a**) and (**b**) are schematics of wave–vortex interaction problems (i) and (ii) respectively. Waves of wavenumber $\boldsymbol{k}$ are incident from the left and are weakly refracted by the vortex. Rays are nearly parallel to the x axis. The azimuthal angle θ is defined unconventionally but in a way that will be convenient when discussing the Aharonov–Bohm effect.

paper, are sketched in figures 1*(a)* and 1*(b)*. They will be referred to as problem (i) and problem (ii), respectively. The background flow is a single vortex whose vorticity is confined to a core of radius $r = r_0$, say, with irrotational flow outside. The coordinates are as shown in figure 1*(a)*, with $r^2 = x^2 + y^2$. The vortex weakly refracts a train of gravity waves or sound waves incident from the left. The refraction induces a small difference between incoming and outgoing pseudomomentum fluxes – corresponding to the background not being translationally invariant – and the pseudomomentum rule leads us to expect a persistent $O(a^2)$ mean recoil force to be exerted. That expectation is independent of whether or not the waves overlap the vortex core. One reason for studying the two problems side by side is a desire to understand how overlap or non-overlap affect the way in which the recoil force arises, and where it is exerted, as well as whether it complies with the pseudomomentum rule.

Problem (i), with no overlap, has already been studied in an earlier paper (Bühler & McIntyre 2003, hereafter BM03) but will be revisited here in order to compare it with problem (ii), for which new results will be obtained. Also new will be results for a rapidly-rotating version of problem (i), to be defined below and to be referred to as problem (iii).

Implicit here, as above, is the assumption that the waves can be described by linearized theory for $a \ll 1$ on a background flow of much greater magnitude. Our aim is to obtain precise results by analytical means, in order to gain insight into the questions just noted. To get analytically tractable, precisely soluble problems it turns out that we must also assume, as was done in BM03, that the background flow and the resulting refraction are very weak in the sense that the vortex must be assumed to have small Froude or Mach number

$$\epsilon = U/c_0 \ll 1 , \tag{1.1}$$

where U is a vortex flow speed and c_0 an intrinsic wave speed. Thus the analyses to be presented fall within the asymptotic regime $a \ll \epsilon \ll 1$. For definiteness, U will be taken to be the flow speed at the edge $r = r_0$ of the vortex core, and c_0 the wave speed far from the core. For general $r \geqslant r_0$ the wave speed $c = c(r) = c_0\{1 + O(\epsilon^2 r_0^2/r^2)\}$, from the Bernoulli effect and the r^{-1} dependence of the vortex flow speed.

The regime $a \ll \epsilon \ll 1$ also encompasses the celebrated Lighthill theory of spontaneous sound emission from, and scattering by, unsteady systems of vortices. It can be contrasted with, for instance, the regime $a \sim \epsilon \ll 1$ (e.g. Lelong & Riley 1991; Ward & Dewar 2010; Thomas 2017), in which wave–vortex interactions of the resonant-triad type are possible. The vortical field, if sufficiently complex spatially, can then act as a passive "catalyst" of wave–wave energy transfer very like the Bragg scattering or "elastic scattering" studied in McComas & Bretherton (1977), in a somewhat different context. Yet another regime of interest, one that has been studied very often in past decades, is $a^2 \sim \epsilon \ll 1$, for instance in connection with the generation of Langmuir vortices by the nondissipative Craik–Leibovich instability (e.g. Craik & Leibovich 1976; Leibovich 1980, B14 §11.3). Indeed the regime $a^2 \sim \epsilon \ll 1$ arises naturally in a great variety of problems where mean flows are generated nondissipatively, from rest, entirely through the presence of waves. Then refraction of the waves by the mean flow comes in only at higher order. Further such examples include, among many others, those studied by Longuet-Higgins & Stewart (1964), Bretherton (1969), McIntyre (1981, 1988), Wagner & Young (2015), Haney & Young (2017), Thomas *et al.* (2018), and Thomas & Yamada (2019).

Returning now to problems (i)–(iii), in which $a \ll \epsilon \ll 1$ and refraction takes place at leading order in ϵ, it will be shown in this paper that the pseudomomentum rule is satisfied not only in problem (i) but also in the other two problems, to leading order at least. In all three problems, the background feels a persistent $O(a^2\epsilon^1)$ mean recoil force satisfying the rule.

Problem (ii) is a classical version of the phonon–vortex problem studied in the quantum vortex literature. This classical version is considered for instance by Sonin (1997) and Stone (2000*a*), who take the wavetrain to be infinitely wide, and incident from $x = -\infty$. They argue not only that the pseudomomentum rule holds but also that there is a remarkable simplification, namely that the dominant wave-refraction effect, the sole effect that comes in at leading order, $O(a^2\epsilon^1)$, is the topological phase jump arising from what is called Aharonov–Bohm effect. We will find, however – with cross-checks from two independent methods – that another, quite different refraction effect is also relevant in problem (ii), except in one special limiting case where the Aharonov–Bohm phase jump is indeed dominant. In that special case the length of the wavetrain is taken to infinity first, followed by the width. If the order of limits is reversed, a different answer is obtained and the Aharonov–Bohm phase jump, while still relevant, is no longer the only relevant refraction effect. However, as already emphasized we find that the pseudomomentum rule is still satisfied to leading order, that is, correct to $O(a^2\epsilon^1)$.

The Aharonov–Bohm phase jump is a topological property of the wave field most simply expressed via the following far-field solution to the linearized equations. The solution is well known in the quantum literature and will be verified below, in §2 and Appendix A. For sufficiently large r, and outside a relatively narrow "wake" region surrounding the positive x axis, the wave field has the asymptotic form $A\exp(\mathrm{i}\varPhi)$ where the amplitude $A = O(a)$ is a real constant, with error $O(a\epsilon^2 r_0^2/r^2)$, and the phase $\varPhi$ is given by

$$\varPhi \;=\; k_0(x - c_0 t) \;-\; \alpha\theta \;+\; \text{const.} \;+\; O(\epsilon^2 r_0^2/r^2)\;. \qquad (1.2)$$

The incident wavenumber k_0 is a constant. The azimuthal angle θ is defined as in figure 1*(a)* and ranges from $-\pi$ to π, while α is a constant defined by

$$\alpha \;=\; \varGamma k_0/2\pi c_0 \;=\; U k_0 r_0/c_0 \;=\; k_0 r_0 \epsilon \qquad (1.3)$$

where $\varGamma$ is the Kelvin circulation of the vortex. The phase jump $2\pi\alpha$ across the positive x axis is the Aharonov–Bohm phase jump, a topological defect of (1.2). In a full solution it is smoothed out across the wake region. It measures the effect of the vortex flow in compressing the wavetrain on the positive-y side and stretching it on the negative. The wavecrest shapes $\varPhi =$ const. described by (1.2), with the error term neglected, are plotted in figure 2 for $\alpha = 0.75$, fixing the phase jump at three-quarters of a wavelength. Depending on the value of $k_0 r_0$ this can take us outside the range of validity of our asymptotic regime (see for instance §7), but $\alpha = 0.75$ is chosen here to make the refraction effects visible in the figure. They include the other relevant effect already mentioned,

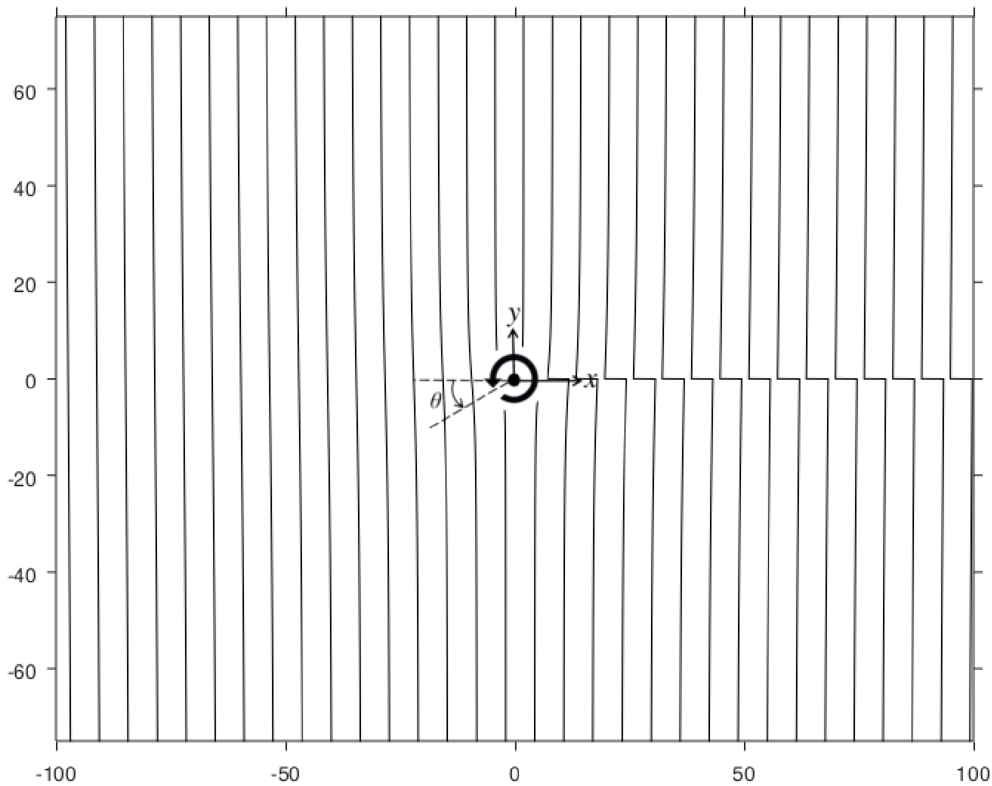

FIGURE 2. Wavecrests plotted from the far-field solution (1.2), with $\alpha = 0.75$. The unit of length is taken as k_0^{-1} so that the unrefracted wavelength is 2π. The Aharonov–Bohm phase jump appears as a phase discontinuity on the positive x-axis. In a full solution this discontinuity is smoothed out across a relatively narrow "wake" region.

which is that, except on the y axis, the wavecrests are slightly rotated away from the y direction, by $O(\epsilon r_0/r)$. This latter effect is also important in problem (i), as noted in BM03 and in B14 §14.2.

There remains the question of *where* the wave-induced recoil force is exerted. The question is ambiguous as it stands, and can be answered in more than one way, but there is one and only one way that avoids bringing in the $O(a^2)$ mean pressure field. It has the further advantage of being the only way that is relevant to gravity-wave parametrization. It is to ask, then answer, the question thus: if the waves were removed and the recoil force exerted artificially, as an external applied force, where should it be exerted in order to have the same effect on the mean flow? As shown in BM03 and B14 §14.2, in the case of problem (i), the answer is not at locations where the waves are refracted – as a naive invocation of the pseudomomentum rule might suggest – but, rather, solely at the location of the vortex core. BM03 therefore called the recoil "remote". In problem (i), the vortex core can be arbitrarily distant from locations where the waves are refracted. There is of course nothing mysterious about this remoteness – no violation of Newton's Third Law – because pressure fields can mediate actions and reactions continuously, over substantial distances, just as they do in ordinary vortex–vortex interactions.

To arrive at this picture BM03 relied mainly on a thought-experiment in which an artificial "holding force" was applied to the vortex core, in such a way as to cancel the recoil due to the waves. It was shown by careful analysis that, by applying this holding force, the mean flow and wave field can be kept exactly steady, with exactly constant total momentum. Here, following B14 §14.2, we ask instead how the mean flow responds to the net pseudomomentum flux in the absence of a holding force. The answer is then that the vortex translates, and keeps on translating persistently, in a direction perpendicular to the recoil force – a classic Magnus-force-like scenario. It translates because it is advected by an $O(a^2)$ "Bretherton flow" induced by the wave field (Bretherton 1969). With no holding force, therefore, the *Kelvin impulse* of the vortex (e.g. Batchelor 1967, eq. (7.3.7), and eq. (3.3) below) changes in just the same way as if the waves were removed and the recoil force artificially applied to the vortex core. Because $a^2 \ll a \ll \epsilon$, the wave field can still be treated as steady. And in each problem studied here and in BM03, the Bretherton flow organizes itself such that the rate of change of impulse corresponds to a recoil force that is consistent with the pseudomomentum rule.

By way of illustration, figure 3 depicts schematically the Bretherton flow in a version of problem (i) solved in §5.1 of BM03, *q.v.* for the analytical details. The heavy curve represents a narrow wavetrain that is slightly deflected as it goes past the vortex, in such a way that the net pseudomomentum flux into the region points in the positive y direction. Ray theory is used to describe the waves, as throughout BM03, assuming $k_0 r_0 \gg 1$. The $O(a^2)$ mean flow within the wavetrain is dominated by the Stokes drift. A small portion of its mass flux leaks sideways as a consequence of wave refraction, forming the Bretherton flow, which is irrotational outside the wavetrain. In the example shown it advects the vortex core leftward, in the negative x direction. The corresponding recoil force – a force that would move the vortex core leftward in the absence of waves – is therefore a force in the positive y direction, like the net pseudomomentum flux. Its magnitude is shown by BM03's analysis to be consistent with the pseudomomentum rule.

In BM03 and B14, as in the present work, it is assumed that the vortex core size $r = r_0$ is small enough to allow the core to be carried bodily along by the Bretherton flow, whose scale is much larger, with strain rate much less than vorticity since $a^2 \ll a \ll \epsilon$ (cf. Kida 1981). That in turn makes the results independent of detailed core structure, i.e. of the function $\omega_0(r)$ where ω_0 is the vorticity, but dependent only on the Kelvin circulation $\mathit{\Gamma} = \iint \omega_0 \, dxdy$.

In the case of problem (ii), the same remote-recoil effects will be found to occur. In addition, because of overlap, there is a *local recoil* corresponding to advection of the vortex core by the Stokes drift of the wavetrain. This local contribution is given by the celebrated Craik–Leibovich vortex force, ω_0 times the Stokes drift, equation (2.12) below, and is directed toward negative y in the case of figure 1*(b)*. Its quantum vortex counterpart is the Iordanskii force, with the Stokes drift corresponding to the phonon current. The special limiting case where the Aharonov–Bohm effect is dominant is also special in another way, namely that the local contribution is the only contribution. The remote contribution vanishes, in that particular limit. Generically, however, both contributions are important, as will be shown.

The plan of the paper is as follows. §2 introduces the equations to be used, and verifies the far-field solution (1.2). §3 recalls how the Kelvin impulse $\boldsymbol{I}$ of a vortex responds to a force applied to its core, and proves a general theorem relating the pseudomomentum field to the rate of change of $\boldsymbol{I}$. This is a variant of the impulse–pseudomomentum theorem first proved in Bühler & McIntyre (2005). It provides one way of seeing that the pseudomomentum rule holds in all three of our problems, independently of our explicit calculations of wave refraction and net pseudomomentum flux. The theorem does, however, depend heavily on the smallness of ϵ and leaves open some challenging questions about the wider validity of the rule. §4 briefly revisits problem (i), in preparation for its

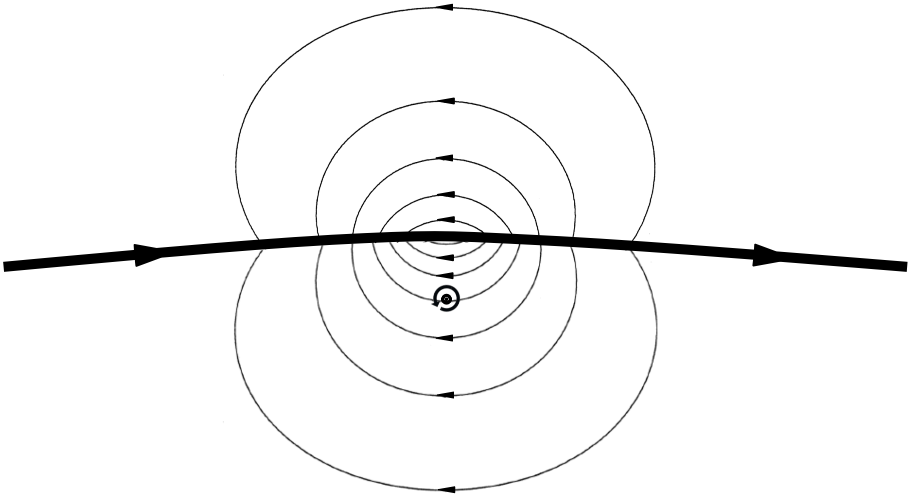

FIGURE 3. Schematic of the Bretherton flow arising in a version of problem (i) studied in BM03. The $O(a^2)$ mean flow within a narrow wavetrain, whose ray path is shown by the heavy curve, is dominated by the Stokes drift. A small portion of its mass flux, $O(a^2\epsilon^1)$ in this case, leaks sideways as a consequence of wave refraction. To describe this situation the refraction problem must be considered correct to two orders in ϵ, as was done in §5.1 of BM03. Refraction effects enter at both orders, not only the $O(\epsilon)$ effects illustrated in figure 2 but also an $O(\epsilon^2)$ change in the direction of the absolute group velocity, exaggerated in this schematic.

extension to problem (ii) in §§5–7. In §8 we formulate and solve problem (iii), the rapidly-rotating version of problem (i). In that version, the waves are high-frequency deep-water surface gravity waves, and the mean flow obeys quasigeostrophic shallow-water dynamics in a fluid layer whose depth H is sufficiently large by comparison with k_0^{-1}. The mean flow feels rotation strongly but the waves feel it only weakly. The $O(a)$ wavemotion can be treated as irrotational to sufficient accuracy. A point of interest is that the rotation produces a tendency for the Stokes drift to be cancelled by the well-known Eulerian-mean "anti-Stokes flow" (Ursell 1950; Hasselmann 1970; Pollard 1970; Lane *et al.* 2007). It might be thought that the cancellation suppresses the Bretherton flow and hence the remote recoil, but the analysis will show otherwise. §9 offers some concluding remarks, emphasizing challenges for future work.

2. Equations used

To verify the far-field wave solution (1.2) we need the linearized equations outside the vortex core. The irrotational background velocity field $\boldsymbol{u}_0$ is $\boldsymbol{u}_0(r) = U r_0 r^{-1}\hat{\boldsymbol{\theta}} = \epsilon c_0 r_0 r^{-1}\hat{\boldsymbol{\theta}}$ where $\hat{\boldsymbol{\theta}}$ is the unit vector in the θ direction. The equations are most succinctly written in their Bernoulli form

$$\left(\frac{\partial}{\partial t} + \boldsymbol{u}_0\cdot\boldsymbol{\nabla}\right)\phi' \;=\; -\,c^2\eta' + \chi' \ , \tag{2.1}$$

$$\left(\frac{\partial}{\partial t} + \boldsymbol{u}_0\cdot\boldsymbol{\nabla}\right)\eta' \;=\; -\,\nabla^2\phi' \ , \tag{2.2}$$

where ϕ' is the velocity potential for the irrotational wavemotion, $\boldsymbol{u}' = \boldsymbol{\nabla}\phi'$, say, while η' is the fractional layer-thickness or density disturbance, in the shallow water or acoustical interpretation respectively, and χ' is a prescribed oscillatory forcing potential. Such forcing is a convenient way of representing wave sources and sinks, as used in BM03 and B14. In the limiting cases of problem (ii) these sources and sinks will recede to infinity, leaving $\chi' = 0$ for all finite (x, y). Equation (2.2) is the linearized mass-conservation equation. Eliminating η' and noting that $(\partial_t + \boldsymbol{u}_0\cdot\boldsymbol{\nabla})\, c = 0$, because c is a function of r alone, we have

$$\left(\frac{\partial}{\partial t} + \boldsymbol{u}_0\cdot\boldsymbol{\nabla}\right)^2 \phi' \;-\; c^2\nabla^2\phi' \;=\; \left(\frac{\partial}{\partial t} + \boldsymbol{u}_0\cdot\boldsymbol{\nabla}\right)\chi' \;. \tag{2.3}$$

Now if $\phi' \propto \exp(\mathrm{i}\varPhi)$ with $\varPhi$ as in (1.2)–(1.3), we have a local wavenumber vector

$$\boldsymbol{k} \;=\; \boldsymbol{\nabla}\varPhi \;=\; k_0\hat{\boldsymbol{x}} - \alpha r^{-1}\hat{\boldsymbol{\theta}} + O(\epsilon^2 k_0 r_0^2 r^{-2}) \;=\; k_0\{\hat{\boldsymbol{x}} - \epsilon r_0 r^{-1}\hat{\boldsymbol{\theta}} + O(\epsilon^2 r_0^2 r^{-2})\} \tag{2.4}$$

where $\hat{\boldsymbol{x}}$ is the unit vector in the x direction and where the error term has been assumed to have length-scale $\gtrsim k_0^{-1}$. We note that $\boldsymbol{k}$ is slightly rotated away from the x direction, pointing slightly into the background flow, as already seen in figure 2 where the wavecrests are rotated away from the y direction. So

$$\boldsymbol{\nabla}\phi' \;=\; \mathrm{i}\boldsymbol{k}\phi' \;=\; \mathrm{i}k_0\left\{\hat{\boldsymbol{x}} - \epsilon r_0 r^{-1}\hat{\boldsymbol{\theta}} + O(\epsilon^2 r_0^2 r^{-2})\right\}\phi' \;, \tag{2.5}$$

and with $\boldsymbol{u}_0(r) \;=\; \epsilon c_0 r_0 r^{-1}\hat{\boldsymbol{\theta}}$ we have, noting that $\hat{\boldsymbol{\theta}}\cdot\hat{\boldsymbol{x}} \;=\; \sin\theta\,$,

$$\left(\frac{\partial}{\partial t} + \boldsymbol{u}_0\cdot\boldsymbol{\nabla}\right)\phi' \;=\; \mathrm{i}k_0 c_0\left\{-1 + \epsilon r_0 r^{-1}\sin\theta + O(\epsilon^2 r_0^2 r^{-2})\right\}\phi' \tag{2.6}$$

and

$$\left(\frac{\partial}{\partial t} + \boldsymbol{u}_0\cdot\boldsymbol{\nabla}\right)^2\phi' \;=\; -k_0^2 c_0^2\left\{1 - 2\epsilon r_0 r^{-1}\sin\theta + O(\epsilon^2 r_0^2 r^{-2})\right\}\phi' \;, \tag{2.7}$$

which equals $c_0^2\nabla^2\phi' \;=\; -|\boldsymbol{k}|^2 c_0^2\phi'$ to the same accuracy and therefore satisfies (2.3) with $\chi' \;=\; 0$, to that accuracy. The next order $O(\epsilon^2 r_0^2 r^{-2})$ fails to satisfy (2.3), because a contribution $-2k_0^2 c_0^2\epsilon^2 r_0^2 r^{-2}\phi'$ on the right of (2.7) disagrees with a contribution $-k_0^2 c_0^2\epsilon^2 r_0^2 r^{-2}\phi'$ to $c_0^2\nabla^2\phi'$, with no coefficient 2. At higher orders there are contributions from $\boldsymbol{\nabla}(\sin\theta)$ that also disagree. We note in passing that, by contrast, $\exp(\mathrm{i}\varPhi)$ with no error term is an exact, and not merely a far-field asymptotic, solution to the Schrödinger equation of the original Aharonov–Bohm problem (details in Appendix A). The Schrödinger equation (A 1) differs from (2.3) except in the limit $\epsilon \to 0$ (e.g. Stone 2000*a*).

The expression (2.4) for $\boldsymbol{k}$ is consistent with ray (JWKB) theory, as can be checked from BM03 (4.12) or B14 (14.5). Also of interest is the direction of the absolute group velocity

$$\boldsymbol{C}^{\mathrm{abs}} = \frac{c\boldsymbol{k}}{|\boldsymbol{k}|} + \boldsymbol{u}_0(r) = \frac{c_0\boldsymbol{k}}{|\boldsymbol{k}|} + \boldsymbol{u}_0(r) + O(\epsilon^2 c_0 r_0^2 r^{-2}) \;. \tag{2.8}$$

Correct to $O(\epsilon c_0 r_0 r^{-1})$, $\boldsymbol{C}^{\mathrm{abs}}$ is parallel to $\hat{\boldsymbol{x}}$, as can be checked by taking the y components of (2.4) and of $\boldsymbol{u}_0(r) \;=\; \epsilon c_0 r_0 r^{-1}\hat{\boldsymbol{\theta}}$. Propagation due to the y component of $\boldsymbol{k}$ cancels advection due to the y component of $\boldsymbol{u}_0$. The cancellation follows also from the irrotationality of the background flow outside the vortex core, in virtue of the curl-curvature formula of ray theory, B14 p. 86.

We avoided relying on ray theory here, when verifying (1.2), because integrating the ray equations over large distances might, conceivably, accumulate significant errors in $\varPhi$, giving incorrect values for the Aharonov–Bohm phase jump, whereas the error $O(\epsilon^2 r_0^2/r^2)$ in (1.2) is small enough to rule out any such accumulation. A convenient corollary is that (1.2) can be used in problem (i) as well as in problem (ii), because when ray theory is valid it is permissible, correct to $O(\epsilon)$, to replace the constant amplitude A by a y-dependent amplitude envelope that restricts the wavetrain appropriately, as sketched in figure 1*(a)*, again using the $O(\epsilon)$ property $\boldsymbol{C}^{\mathrm{abs}} \,||\, \hat{\boldsymbol{x}}$ just noted (as contrasted with the $O(\epsilon^2)$ bending of ray paths in figure 3).

Ray theory will, on the other hand, be sufficient for our treatment of problem (iii), in which the Aharonov–Bohm phase jump has no role. Details are postponed until §8.

For the mean flow, a natural and efficient framework for solving problems (i)–(iii) is that of generalized Lagrangian-mean (GLM) theory, as laid out for instance in B14. However, except where stated otherwise the reader unfamiliar with GLM theory can read the equations as involving, to sufficient accuracy, only the Eulerian-mean velocity $\overline{\boldsymbol{u}}$ and the Stokes drift $\overline{\boldsymbol{u}}^{\mathrm{S}}$, or phonon current per unit mass. Whenever the $O(a)$ wavemotions are irrotational and describable by ray theory, the exact GLM pseudomomentum $\mathbf{p}$ per unit mass can be replaced by $\overline{\boldsymbol{u}}^{\mathrm{S}}$ and the exact Lagrangian-mean velocity $\overline{\boldsymbol{u}}^{\mathrm{L}}$ by $\overline{\boldsymbol{u}}+\overline{\boldsymbol{u}}^{\mathrm{S}}$, with error $O(a^2\epsilon^2)$; see e.g. B14, equations (10.15) and (10.17). Then the combination $\overline{\boldsymbol{u}}^{\mathrm{L}}-\mathbf{p}$, which occurs frequently in the exact theory, can be read simply as $\overline{\boldsymbol{u}}$, and the exact mean vorticity $\widetilde{\boldsymbol{\omega}}$ defined by

$$\widetilde{\boldsymbol{\omega}} \;=\; \boldsymbol{\nabla}\times(\overline{\boldsymbol{u}}^{\mathrm{L}}-\mathbf{p}) \tag{2.9}$$

can be read simply as $\overline{\boldsymbol{\omega}}$, the Eulerian-mean vorticity. (The quantity $\widetilde{\boldsymbol{\omega}}$ is the simplest exact measure of mean vorticity. It arises from frozen-field distortions of the three-dimensional vorticity field by the wave-induced displacement field.) The relation $\mathbf{p}=\overline{\boldsymbol{u}}^{\mathrm{S}}$ is always valid sufficiently far from the vortex core, in all three problems, where ray theory is always valid. In problems (i) and (ii) we need only the vertical or z component of (2.9).

The power and economy of the GLM formalism comes from Kelvin's circulation theorem and its consequence, e.g. B14 §10.2.9, that $\overline{\boldsymbol{u}}^{\mathrm{L}}$ exactly advects mean vorticities $\widetilde{\boldsymbol{\omega}}$, or $\widetilde{\boldsymbol{\omega}}+\boldsymbol{f}$ in problem (iii), with $\boldsymbol{f}$ the vector Coriolis parameter. This will prove useful throughout our analyses. In problem (iii) it expresses Ursell's insight into the anti-Stokes flow, in a succinct and natural way to be pointed out in §8. The advection property is neatly summarized by the exact three-dimensional form of the nondissipative equation for $\widetilde{\boldsymbol{\omega}}$; see e.g. B14, equations (10.99) and (10.153). It is

$$\frac{\partial\widetilde{\boldsymbol{\omega}}}{\partial t} \;-\; \boldsymbol{\nabla}\times\{\overline{\boldsymbol{u}}^{\mathrm{L}}\times(\boldsymbol{f}+\widetilde{\boldsymbol{\omega}})\} \;=\; 0 \tag{2.10}$$

or alternatively

$$\frac{\overline{D}^{\mathrm{L}}\widetilde{\boldsymbol{\omega}}}{Dt} \;+\; (\boldsymbol{f}+\widetilde{\boldsymbol{\omega}})\,\boldsymbol{\nabla}\cdot\overline{\boldsymbol{u}}^{\mathrm{L}} \;=\; (\boldsymbol{f}+\widetilde{\boldsymbol{\omega}})\cdot\boldsymbol{\nabla}\overline{\boldsymbol{u}}^{\mathrm{L}}, \tag{2.11}$$

where $\overline{D}^{\mathrm{L}}/Dt=\partial/\partial t+\overline{\boldsymbol{u}}^{\mathrm{L}}\cdot\boldsymbol{\nabla}$. If we now set $\boldsymbol{f}=0$ and apply the foregoing recipe to (2.10), replacing $\widetilde{\boldsymbol{\omega}}$ by $\overline{\boldsymbol{\omega}}$ and $\overline{\boldsymbol{u}}^{\mathrm{L}}$ by $\overline{\boldsymbol{u}}+\overline{\boldsymbol{u}}^{\mathrm{S}}$, then we get the approximate version of (2.10) known as the Craik–Leibovich equation:

$$\frac{\partial\overline{\boldsymbol{\omega}}}{\partial t} - \boldsymbol{\nabla}\times(\overline{\boldsymbol{u}}\times\overline{\boldsymbol{\omega}}) \;=\; \boldsymbol{\nabla}\times(\overline{\boldsymbol{u}}^{\mathrm{S}}\times\overline{\boldsymbol{\omega}})\;. \tag{2.12}$$

The right-hand side of (2.12) is the curl of the Craik–Leibovich vortex force $\overline{\boldsymbol{u}}^{\mathrm{S}}\times\overline{\boldsymbol{\omega}}$, which as mentioned earlier makes a local contribution $\overline{\boldsymbol{u}}^{\mathrm{S}}\times\boldsymbol{\omega}_0$ to the effective force on the vortex in problem (ii), where $\boldsymbol{\omega}_0=\boldsymbol{\nabla}\times\boldsymbol{u}_0$. Equation (2.12) was originally derived by Craik & Leibovich (1976), via a much longer route, to study another problem – the dynamics of Langmuir vortices – assuming incompressible flow $\boldsymbol{\nabla}\cdot\boldsymbol{u}=0$ and steady wave fields, and under the strong parameter restriction $a^2\sim\epsilon\ll 1$, i.e. that all mean flows, whether wave-induced or pre-existing, have order of magnitude $O(a^2)$. The route via GLM just recalled, which is not only much shorter but also has wider validity, was first pointed out by Leibovich (1980).

To complete the mean-flow equations we need a mass-conservation equation. As usual in GLM, we define a mean two-dimensional density or layer depth $\widetilde{h}$ such that the areal mass element $\propto\widetilde{h}\,dxdy$ exactly; see equations (10.42)–(10.47) of B14. Then mass conservation is expressed by

$$\frac{\partial \tilde{h}}{\partial t} + \boldsymbol{\nabla}_{\mathrm{H}} \cdot (\tilde{h}\,\overline{\boldsymbol{u}}^{\mathrm{L}}_{\mathrm{H}}) = 0\,, \tag{2.13}$$

where suffix H denotes horizontal projection, on to the xy plane, superfluous in the two-dimensional acoustical setting but needed in the shallow-water setting and in problem (iii). In all three problems, however, we shall find that (2.13) can be simplified to

$$\boldsymbol{\nabla}_{\mathrm{H}} \cdot \overline{\boldsymbol{u}}^{\mathrm{L}}_{\mathrm{H}} = 0 \tag{2.14}$$

if we are willing to work to the lowest significant accuracy – the lowest that captures the remote-recoil effects to be discussed. This will keep the analysis extraordinarily simple yet able to illustrate the main points. Equation (2.14) will be justified shortly for problems (i) and (ii) and in §8 for problem (iii), with error estimates. We can then define a streamfunction, $\tilde{\psi}$ say, such that the horizontal velocity components can be written as

$$\overline{u}^{\mathrm{L}}_{\mathrm{H}} = -\frac{\partial \tilde{\psi}}{\partial y} \qquad \text{and} \qquad \overline{v}^{\mathrm{L}}_{\mathrm{H}} = \frac{\partial \tilde{\psi}}{\partial x}\,. \tag{2.15}$$

The Bretherton flow has streamfunction

$$\tilde{\psi}_{\mathrm{B}} = \tilde{\psi} - \tilde{\psi}_0 \tag{2.16}$$

where $\tilde{\psi}_0$ is the streamfunction for the nondivergent velocity field $\boldsymbol{u}_0$ of the vortex flow.

To compute the Bretherton flows in problems (i) and (ii) to sufficient accuracy (see §4), we need only (2.15)–(2.16) and the vertical component of (2.9). We have $\widetilde{\boldsymbol{\omega}} = \boldsymbol{\nabla} \times (\overline{\boldsymbol{u}}^{\mathrm{L}} - \mathbf{p}) = \boldsymbol{\nabla} \times \boldsymbol{u}_0 = \boldsymbol{\omega}_0\,,$ expressing irrotationality outside the vortex core. The vertical component of $\boldsymbol{\nabla} \times (\overline{\boldsymbol{u}}^{\mathrm{L}} - \boldsymbol{u}_0)$ is just $\nabla^2_{\mathrm{H}} \tilde{\psi}_{\mathrm{B}}$. We therefore have

$$\nabla^2_{\mathrm{H}} \tilde{\psi}_{\mathrm{B}} = \hat{\boldsymbol{z}} \cdot \boldsymbol{\nabla} \times \mathbf{p} \tag{2.17}$$

where $\hat{\boldsymbol{z}}$ is the vertical unit vector. The right-hand side of (2.17) is known as soon as we know the wave pseudomomentum field $\mathbf{p}$, which as mentioned earlier can be identified with the $\overline{\boldsymbol{u}}^{\mathrm{S}}$ field whenever the wavemotion is irrotational and ray theory applies.

For problem (iii) it will be shown in §8 that we need only two modifications. First, the vorticity $\nabla^2_{\mathrm{H}} \tilde{\psi}_{\mathrm{B}}$ must be replaced in the standard way by $(\nabla^2_{\mathrm{H}} - L_{\mathrm{D}}^{-2})\tilde{\psi}_{\mathrm{B}}$, the quasigeostrophic potential vorticity (PV), of the Bretherton flow, where L_{D} is the Rossby deformation length-scale $L_{\mathrm{D}} = f^{-1}(gH)^{1/2}$ where g is gravity and $f = |\boldsymbol{f}|$. Second, we must replace $\mathbf{p}$, which for deep-water surface gravity waves is strongly z-dependent, by its vertical average $\langle \mathbf{p} \rangle$. So in place of (2.17) we have simply

$$(\nabla^2_{\mathrm{H}} - L_{\mathrm{D}}^{-2})\tilde{\psi}_{\mathrm{B}} = \hat{\boldsymbol{z}} \cdot \boldsymbol{\nabla} \times \langle \mathbf{p} \rangle\,. \tag{2.18}$$

The derivation of (2.18) involves some delicate arguments about the asymptotics and will be postponed until §8 and Appendix C. The elliptic operators in (2.17) and (2.18) illustrate, by implication, a generic property of Bretherton flows, that they can extend well outside the wavetrain where $\mathbf{p} \neq 0$. That is one way of seeing the generic nature of remote recoil.

We now justify replacing (2.13) by (2.14) for problems (i) and (ii). Among the errors thus incurred, the largest is $O(a^2\epsilon^1)$. It arises in problem (i), from the variation of wave amplitude A across the wavetrain and illustrating, incidentally, a need to avoid textbook arguments for the near-incompressibility of low-Mach-number or low-Froude-number flows. Those arguments do not take into account the kinds of spatial heterogeneity that are possible here, especially in problem (i).

For our asymptotic regime we need to let $a \to 0$ and $\epsilon \to 0$ keeping $a \ll \epsilon$, for a given geometry of the background flow and incident wave field. Where the vortex flow $\boldsymbol{u}_0(r)$ crosses the wavetrain, in problem (i), it encounters $\tilde{h}$ values that are reduced by $O(a^2\epsilon^0)$

because of the Brillouin radiation stress in the wavetrain. And since $\boldsymbol{\nabla}_{\mathrm{H}} \cdot \boldsymbol{u}_0 = 0$, the resulting contribution to $\boldsymbol{\nabla}_{\mathrm{H}} \cdot (\tilde{h}\overline{\boldsymbol{u}}^{\mathrm{L}}_{\mathrm{H}})$ in (2.13), which is neglected in going to (2.14), is just $\boldsymbol{u}_0 \cdot \boldsymbol{\nabla}_{\mathrm{H}} \tilde{h}$ to leading order, with magnitude $O(a^2\epsilon^1)$ since $\boldsymbol{u}_0 = O(\epsilon)$. (This contribution is significant, however, at the greater accuracy required in the case of figure 3, as can be seen from (5.1) of BM03, even though it will not be required in the present analyses.)

The $O(a^2\epsilon^0)$ local reduction in $\tilde{h}$ within the wavetrain (set-down, in the shallow-water setting) is necessary to ensure that the $O(a^2)$ sideways mean fluid acceleration vanishes. The sideways gradients of Brillouin radiation stress and $O(a^2)$ mean pressure must cancel. Only the isotropic part of the Brillouin radiation stress is involved, the so-called "hard-spring" contribution, unrelated to pseudomomentum fluxes, and equal to $\partial \ln c/\partial \ln h$ times wave-energy per unit area, where h is layer depth or two-dimensional mass density (e.g. Brillouin 1936, B14 §10.5.1).

3. Impulse and pseudomomentum

The impulse–pseudomomentum theorem applies to problems (i)–(iii) as well as to a more general set of problems involving multiple vortices and more complicated wave fields, with arbitrary wave source and sink regions. The theorem provides an elegant way of showing that the pseudomomentum rule is satisfied in all these problems. There is, however, a severe limitation. The theorem relies crucially on horizontal nondivergence, (2.14). So it applies only at the lowest significant order of accuracy. There is a challenge here since the limitation puts the case of figure 3 outside the scope of the theorem. As just pointed out, (5.1) of BM03 shows that (2.14) is not accurate enough in that case; in fact (2.14) must then be replaced by the anelastic equation

$$\boldsymbol{\nabla}_{\mathrm{H}} \cdot (\tilde{h}\overline{\boldsymbol{u}}^{\mathrm{L}}_{\mathrm{H}}) \;=\; 0\ . \tag{3.1}$$

Yet BM03's analysis shows that the pseudomomentum rule still holds, a point to which we will return.

Before proceeding, we revisit the thought-experiment in which the waves are removed and an artificial external force field $\boldsymbol{F}$ exerted on the vortex core $\omega_0(r)$, producing a rate of change of Kelvin impulse. In order to make the vortex translate bodily without change of shape, with velocity $\boldsymbol{u}_{\mathrm{tr}}$, say, we need $\boldsymbol{F} = -\omega_0\,\hat{\boldsymbol{z}} \times \boldsymbol{u}_{\mathrm{tr}}$. (The curl of this force field is just that required to move the vortex core through the fluid at velocity $\boldsymbol{u}_{\mathrm{tr}}$, while the divergence sets up the dipolar pressure field required to produce the corresponding changes outside the core, where the velocity field is irrotational.) Being transverse to the vortex motion, the resultant force $\boldsymbol{R}$ has the character of a Magnus force,

$$\boldsymbol{R} \;=\; \iint \boldsymbol{F}\,dxdy \;=\; -\hat{\boldsymbol{z}} \times \boldsymbol{u}_{\mathrm{tr}} \iint \omega_0\,dxdy \;=\; -\mathit{\Gamma}\,\hat{\boldsymbol{z}} \times \boldsymbol{u}_{\mathrm{tr}}\ . \tag{3.2}$$

The Kelvin impulse is defined for our two-dimensional shallow water or acoustical domain as

$$\boldsymbol{I} \;=\; \iint (y,\,-x)\,Q\,dxdy \;=\; \iint -\hat{\boldsymbol{z}} \times \boldsymbol{x}\,Q\,dxdy \tag{3.3}$$

(e.g. Batchelor 1967, equation (7.3.7)) where $\boldsymbol{x} = (x, y)$ and where Q is the vorticity, $Q = \omega_0$ in this case. When the vortex translates in response to $\boldsymbol{F}$, the rate of change of $\boldsymbol{I}$ is just $\boldsymbol{R}$; cf. (3.7) below. In the corresponding thought-experiment for problem (iii), the same statements hold if Q is redefined as the quasigeostrophic PV.

The impulse–pseudomomentum theorem makes the following assumptions, in addition to (2.14) and its consequences (2.15)–(2.18). The wave field together with its sources and sinks is taken to have finite extent, prior to taking any infinite-wavetrain limits that might be of interest, while the domain of integration is taken infinite so as to enclose within it the vortex core, or cores, as well as all the waves and their source and sink regions. It is assumed that the pseudomomentum field satisfies a two-dimensional equation of the form (see Appendix D)

$$\frac{\partial \mathsf{p}}{\partial t} + \boldsymbol{\nabla}_{\mathrm{H}} \cdot \mathsf{B} = -(\boldsymbol{\nabla}_{\mathrm{H}} \overline{\boldsymbol{u}}^{\mathrm{L}}) \cdot \mathsf{p} + \boldsymbol{\mathcal{F}} , \tag{3.4}$$

with vertical averaging understood in problem (iii). The first term on the right comes from wave refraction and scattering by the mean flow, and $\boldsymbol{\mathcal{F}}$ is the rate of generation or absorption of pseudomomentum in the wave source and sink regions, per unit area. In the refraction term, p contracts with $\overline{\boldsymbol{u}}^{\mathrm{L}}$ and not with $\boldsymbol{\nabla}_{\mathrm{H}}$. On the left, the precise form of the pseudomomentum flux tensor B is immaterial but we note for later reference that, wherever ray theory holds, we shall have, in Cartesian tensor notation, with i and j running from 1 to 2, the standard group-velocity property

$$\mathsf{B}_{ij} = \mathsf{p}_i \, C_j^{\mathrm{abs}} \tag{3.5}$$

where $\boldsymbol{C}^{\mathrm{abs}}$ is the absolute group velocity, $\boldsymbol{u}_0$ plus the intrinsic group velocity, $c\boldsymbol{k}/|\boldsymbol{k}|$ in problems (i) and (ii) as in (2.8), and $c\boldsymbol{k}/2|\boldsymbol{k}|$ in problem (iii). The divergence operator contracts with $\boldsymbol{C}^{\mathrm{abs}}$ so that, in Cartesians, the ith component of $\boldsymbol{\nabla}_{\mathrm{H}} \cdot \mathsf{B}$ is $\mathsf{B}_{ij,j}$. The theorem states that

$$\boxed{\frac{d}{dt}(\boldsymbol{I} + \mathsf{P}) = \iint \boldsymbol{\mathcal{F}} \, dxdy} \tag{3.6}$$

where $\mathsf{P} = \iint \mathsf{p} \, dxdy$, the total pseudomomentum, again with vertical averaging understood in problem (iii).

The proof begins by noting that $Q\,dxdy$ is materially invariant, so that

$$\frac{d\boldsymbol{I}}{dt} = \iint \frac{\overline{D}_{\mathrm{H}}^{\mathrm{L}}(y, -x)}{Dt} Q \, dxdy = \iint (\overline{v}_{\mathrm{H}}^{\mathrm{L}}, -\overline{u}_{\mathrm{H}}^{\mathrm{L}}) \, Q \, dxdy . \tag{3.7}$$

From here on, with everything in two dimensions (x, y), we drop the suffixes H so that, for instance, $\boldsymbol{\nabla}_{\mathrm{H}}$ will be denoted by $\boldsymbol{\nabla}$. Then, recalling (2.9) and (2.15)–(2.18), we have

$$\frac{d\boldsymbol{I}}{dt} = \iint Q \, \boldsymbol{\nabla}\tilde{\psi} \, dxdy = \iint \Big\{ (\nabla^2 - L_{\mathrm{D}}^{-2})\tilde{\psi} - \hat{\boldsymbol{z}} \cdot \boldsymbol{\nabla} \times \mathsf{p} \Big\} \boldsymbol{\nabla}\tilde{\psi} \, dxdy , \tag{3.8}$$

with L_{D} finite in problem (iii) but infinite in problems (i)–(ii). Now $\nabla^2\tilde{\psi}\boldsymbol{\nabla}\tilde{\psi}$ contributes nothing because, in Cartesians, its ith component is $\tilde{\psi}_{,jj}\tilde{\psi}_{,i} = (\tilde{\psi}_{,j}\tilde{\psi}_{,i})_{,j} - \frac{1}{2}(\tilde{\psi}_{,j}\tilde{\psi}_{,j})_{,i}$, which integrates to zero. The integrated terms at infinity vanish because, if we consider a domain of integration having radius $r \to \infty$, the integrated terms have integrands $O(r^{-2})$ in problems (i) and (ii) and $O(\exp(-2r/L_{\mathrm{D}}))$ in problem (iii), from the vortex-only contributions. The Bretherton flows, being dipolar because of the $\boldsymbol{\nabla}$ on the right-hand sides of (2.17) and (2.18), decay at the same rate or faster as $r \to \infty$. In problem (iii) we have the additional contribution $-L_{\mathrm{D}}^{-2}\tilde{\psi}\boldsymbol{\nabla}\tilde{\psi} \propto \frac{1}{2}\boldsymbol{\nabla}(\tilde{\psi}^2)$, which also integrates to zero because $\tilde{\psi}^2 = O(\exp(-2r/L_{\mathrm{D}}))$. Therefore (3.8) reduces, in all three problems, to

$$\frac{d\boldsymbol{I}}{dt} = -\iint (\hat{\boldsymbol{z}} \cdot \boldsymbol{\nabla} \times \mathsf{p}) \, \boldsymbol{\nabla}\tilde{\psi} \, dxdy . \tag{3.9}$$

Upon exchanging the dot with the cross and then integrating by parts, using the finite extent of the wave field, we see that the right-hand side is equal to

$$-\iint \{(\hat{\boldsymbol{z}} \times \boldsymbol{\nabla}) \cdot \mathsf{p}\} \boldsymbol{\nabla}\tilde{\psi} \, dxdy = \iint \{\mathsf{p} \cdot (\hat{\boldsymbol{z}} \times \boldsymbol{\nabla})\} \boldsymbol{\nabla}\tilde{\psi} \, dxdy = \iint (\boldsymbol{\nabla}\overline{\boldsymbol{u}}^{\mathrm{L}}) \cdot \mathsf{p} \, dxdy \tag{3.10}$$

since $\hat{\boldsymbol{z}} \times \boldsymbol{\nabla}$ commutes with $\boldsymbol{\nabla}$, and $\hat{\boldsymbol{z}} \times \boldsymbol{\nabla}\tilde{\psi} = \overline{\boldsymbol{u}}^{\mathrm{L}}$ by (2.15), so that the integrand on the right is minus the refraction term in (3.4). On eliminating that term between (3.4) and

(3.10), and noting that $\boldsymbol{\nabla}\cdot\mathbf{B}$ integrates to zero, again because of the finite extent of the wave field, we arrive at (3.6).

The proof has no dependence on whether or not ray theory applies. It depends only on (2.14)–(2.18) and on the form of (3.4), not on any particular formulae for $\mathbf{p}$, $\mathbf{B}$ and $\boldsymbol{\mathcal{F}}$. The group-velocity property (3.5) will, however, be useful when considering pseudomomentum budgets in detail, for problems (i)–(iii), because along with ray theory it always holds far from the vortex core, together with the approximation $\mathbf{p} = \overline{\boldsymbol{u}}^{\mathrm{S}}$.

The case of figure 3 prompts the question, can we replace (2.14) by (3.1) and still prove the theorem? After considerable effort, the author has been forced to the conclusion that we cannot. At higher orders in ϵ there is an incompatibility between the per-unit-mass basis of vorticity – being the curl of velocity rather than of mass transport – and the per-unit-volume, or per-unit-area, basis of conservation relations, and their refractive extensions such as (3.4), in which $\mathbf{p}$ would need to be replaced by $\tilde{h}\mathbf{p}$ in order to attain enough accuracy to be compatible with (3.1) (again see Appendix D). But one cannot simply insert a factor $\tilde{h}$ into the integrand of (3.3), because $\tilde{h}Q\,dxdy$ is not materially invariant and the subsequent steps from (3.7) onward are invalidated.

It seems likely that this limitation is not a limitation of the pseudomomentum rule as such, but only a limitation of the Kelvin impulse concept. As is well known, the ability to replace momentum budgets by impulse budgets depends on incompressibility. Incompressibility is needed in order to banish to infinity the large-scale $O(a^2)$ pressure-field adjustments in transient situations, including transient versions of thought-experiments like that associated with the Magnus relation (3.2).

Indeed, there is a variant of the case of figure 3 in which the pseudomomentum rule holds to all orders in ϵ, as was shown in BM03 §5.2. However, that result depended on keeping the mean flow exactly steady – over an infinite domain – by applying an artificial "holding force" to the vortex core as described in §1, and then taking account of the full momentum budget in the far field correct to $O(a^2)$. Conditions in the far field are greatly simplified by assuming exact steadiness everywhere. The result depends on the generic relation between the fluxes of momentum and pseudomomentum. As shown in GLM theory they differ only by an isotropic term (see Appendix D), which can be balanced by changes in mean pressure. Whether that can lead to further generalization remains to be explored.

4. Bretherton flows and recoil forces for problem (i)

In this section we review BM03's leading-order results on recoil forces in problem (i), as a preliminary to the subsequent analyses of problem (ii). We confine attention to recoil forces computed from Bretherton flows, omitting BM03's ray-theoretic calculations of the $O(a^2\epsilon^1)$ net pseudomomentum flux. The impulse–pseudomomentum theorem tells us that such calculations must give the same results, at this lowest order, as indeed they were found to do. However, as pointed out in BM03, the Bretherton flows provide the simplest route to the results. Just as $\boldsymbol{R}$ in (3.2) can be computed correct to $O(a^2\epsilon^1)$ from $\boldsymbol{u}_{\mathrm{tr}}$ correct to $O(a^2\epsilon^0)$, because Γ is an $O(\epsilon)$ quantity, a wave-induced recoil can be computed correct to $O(a^2\epsilon^1)$ from a Bretherton flow correct to $O(a^2\epsilon^0)$, that is, from a Bretherton flow computed for an *unrefracted* wavetrain.

To take advantage of this simplification, in problem (i), we need to consider a wavetrain of finite length. That is because of the curl-curvature formula mentioned below (2.8), with its implication that the absolute group velocity $\boldsymbol{C}^{\mathrm{abs}}$ remains parallel to $\hat{\boldsymbol{x}}$ when wave refraction is computed correct to $O(\epsilon)$. It remains parallel despite the $O(\epsilon)$ refraction effects illustrated in figure 2; recall the cancellation noted below (2.8). The $O(a^2\epsilon^1)$ net pseudomomentum flux and recoil force, which depend entirely on $O(\epsilon)$ wavecrest rotations of the kind illustrated in figure 2, therefore vanish when the wave source and sink are allowed to recede to infinity.

So to obtain the simplest, leading-order version of problem (i), following BM03, we let the unrefracted wavetrain extend between wave source and sink regions centred at finite locations $(x,\, y) = (-X, Y)$ and $(+X, Y)$, say, as sketched in figure 4 (heavy straight line). For consistency with BM03's use of ray theory we take $X, Y \gg W$, and $W \gg k_0^{-1}$, where W is a width-scale for the wavetrain. In order to generate and absorb an approximately monochromatic wavetrain, the wave source and sink are prescribed, just as in BM03, by taking the oscillatory forcing potential χ' to be $\exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x} - \mathrm{i}kct)$ times a slowly-varying forcing envelope whose length-scale is at least of the same order of magnitude as W, and where the real part is understood. However, the forcing envelope scale is kept $\ll X, Y$, allowing us to think of the source and sink regions as approximately localized. It is convenient to take $\boldsymbol{k}\cdot\boldsymbol{x} = k_0 x$ in the source along with $kct = k_0 ct$, with constant k_0.

The $O(a^2\epsilon^0)$ Bretherton flow for the unrefracted wavetrain satisfies (2.17), with evanescence at infinity. It consists of the Stokes drift $\overline{\boldsymbol{u}}^{\mathrm{S}}$ straight along the wavetrain, parallel to $\hat{\boldsymbol{x}}$, together with irrotational return flows symmetrically on both sides, as sketched in figure 4. All the vorticity of this Bretherton flow comes from $\hat{\boldsymbol{z}}\cdot\boldsymbol{\nabla}\times\mathsf{p}$ on the right-hand side of (2.17). Within the wavetrain but outside the wave source and sink, $\hat{\boldsymbol{z}}\cdot\boldsymbol{\nabla}\times\mathsf{p}$ is simply minus the horizontal shear of the Stokes drift.

In virtue of the incompressibility expressed by (2.14)–(2.16), the irrotational return flow outside the wavetrain, which advects the vortex core, is the same as if it were induced by a two-dimensional mass sink at the left end of the wavetrain and a mass source at the right end whose strengths are equal to the mass flow in the Stokes drift within the wavetrain. Omitting factors ρ, where ρ is fluid mass density, we define the source and sink strengths $\pm S$ as the volume fluxes in a layer of unit depth; thus

$$S = \int \overline{u}^{\mathrm{S}}(y)\,dy = \int \mathsf{p}_1(y)\,dy\ , \tag{4.1}$$

with the integral taken across the wavetrain. The pseudomomentum within the wavetrain has been written correct to $O(a^2\epsilon^0)$ as $\mathsf{p}_1(y)\hat{\boldsymbol{x}}$. Using $W \ll X$ and $W \ll Y$, again following BM03, we can approximate the mass-source flow in problem (i) as radially outward from $(x,\, y) = (X, Y)$ at speed $S/[2\pi\{(x-X)^2 + (y-Y)^2\}^{1/2}]$, and similarly the mass-sink flow as radially inward toward $(x,\, y) = (-X, Y)$. When these flows are added vectorially we obtain the flow pattern sketched in figure 4; and the net velocity advecting the vortex core at $(x,\, y) = (0, 0)$ is

$$\overline{\boldsymbol{u}}^{\mathrm{L}}(0,0) \;=\; \frac{S}{\pi}\,\frac{X}{X^2+Y^2}\,(-\hat{\boldsymbol{x}}) \tag{4.2}$$

correct to $O(a^2\epsilon^0)$. Because of the Magnus relation this corresponds to a resultant recoil force $\boldsymbol{R} = -\mathit{\Gamma}\,\hat{\boldsymbol{z}}\times\overline{\boldsymbol{u}}^{\mathrm{L}}(0,0)$, i.e.

$$\boldsymbol{R} \;=\; \frac{\mathit{\Gamma} S}{\pi}\,\frac{X}{X^2+Y^2}\,(+\hat{\boldsymbol{y}}) \tag{4.3}$$

correct to $O(a^2\epsilon^1)$, where $\hat{\boldsymbol{y}}$ the unit vector in the y direction.

Notice that $\overline{\boldsymbol{u}}^{\mathrm{L}}(0,0)$ and $\boldsymbol{R}$ tend toward zero in the formal limit $X \to \infty$, cross-checking what was deduced from the direction of $\boldsymbol{C}^{\mathrm{abs}}$ and consequent vanishing of the $O(a^2\epsilon^1)$ net pseudomomentum flux and recoil force in that limit. As the wavetrain gets longer, the irrotational return part of the Bretherton flow becomes increasingly spread out in the y direction, diluting its effect at $(x,\, y) = (0, 0)$.

5. Bretherton flows and recoil forces for problem (ii)

The dilution effect just pointed out is the easiest way of seeing the noninterchangeability of limits in problem (ii). The same dilution effect occurs for any wavetrain whose

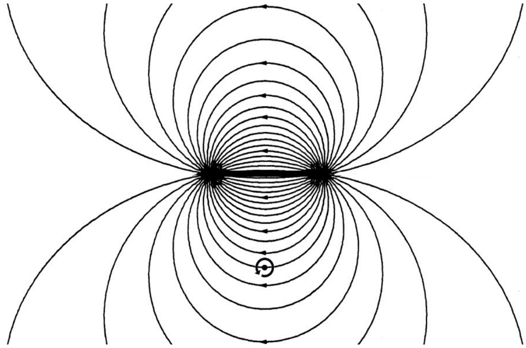

FIGURE 4. Schematic of Bretherton-flow streamlines in problem (i), as analysed in BM03 correct to lowest order $O(a^2\epsilon^0)$ for the finite wavetrain whose ray path is shown by the heavy straight line. At this order the Stokes drift is nondivergent except within the wave source and sink regions. The waves propagate from a source on the left to a sink on the right.

width W is held fixed while its length $L = 2X \to \infty$. In that formal limit the magnitude of the return flow at any fixed point tends toward zero; and it remains zero if the formal limit $W \to \infty$ is taken subsequently. The vortex core is then advected solely by the Stokes drift $\overline{\boldsymbol{u}}^{\mathrm{S}}(0,0) = \overline{\boldsymbol{u}}^{\mathrm{S}} = \mathsf{p}_1\hat{\boldsymbol{x}}$, which is now constant across the wavetrain, so that with $L \to \infty$ first in problem (ii), (4.2) and (4.3) are replaced by

$$\overline{\boldsymbol{u}}^{\mathrm{L}}(0,0) \;=\; \overline{\boldsymbol{u}}^{\mathrm{S}} \;=\; \mathsf{p}_1\hat{\boldsymbol{x}} \tag{5.1}$$

and

$$\boldsymbol{R} \;=\; -\varGamma\mathsf{p}_1\hat{\boldsymbol{y}}\;. \tag{5.2}$$

Not only the magnitudes but also the signs have changed. These results hold for arbitrary k_0r_0, because in the unrefracted wavetrain we always have $\overline{\boldsymbol{u}}^{\mathrm{S}} = \mathbf{p} = \mathsf{p}_1\hat{\boldsymbol{x}}$. Notice that (5.2) is equal to the Craik–Leibovich vortex force $\overline{\boldsymbol{u}}^{\mathrm{S}} \times \overline{\boldsymbol{\varpi}} = \overline{\boldsymbol{u}}^{\mathrm{S}} \times \boldsymbol{\omega}_0$ integrated over the vortex core, corresponding to the Iordanskii force in the quantum fluids literature (e.g. Sonin 1997), with $\overline{\boldsymbol{u}}^{\mathrm{S}} = \mathbf{p}$ corresponding to the phonon current per unit mass. The relation to the Aharonov–Bohm effect is discussed in the next two sections.

In any other version of problem (ii) there will be two contributions, one from the Craik–Leibovich vortex force and the other from the return part of the Bretherton flow. In the opposite formal limit, with the width W of the wavetrain going to infinity first, the two contributions cancel. The return flow is then uniform, and equal and opposite to the Stokes drift, being diluted only near the extremities $y \sim \pm\frac{1}{2}W$. In that formal limit, therefore, the recoil force $\boldsymbol{R}$ vanishes.

In all other versions of problem (ii) there is always some dilution, making the return flow weaker than the Stokes drift and keeping the sign of the recoil opposite to that in problem (i). For instance, consider the "square" limit $L = W \to \infty$. Then it is readily shown that

$$\overline{\boldsymbol{u}}^{\mathrm{L}}(0,0) \;=\; \tfrac{1}{2}\overline{\boldsymbol{u}}^{\mathrm{S}}(0,0) \;=\; \tfrac{1}{2}\mathsf{p}_1\hat{\boldsymbol{x}} \tag{5.3}$$

so that in place of (5.2) we have

$$\boldsymbol{R} \;=\; -\tfrac{1}{2}\varGamma\mathsf{p}_1\hat{\boldsymbol{y}}\;. \tag{5.4}$$

More generally, the factor $\frac{1}{2}$ is replaced by $\{1 - 2\pi^{-1}\lim\,\arctan(W/L)\}$.

To derive this last result, one can regard the wide wavetrain as made up of narrow wavetrains each with $S = \mathsf{p}_1\,dy$, for constant p_1, and then integrate (4.3), with Y replaced by y, over the whole wavetrain to get the contribution to $\boldsymbol{R}$ from the return flow. That contribution, for $L \to \infty$ and $W \to \infty$ in various ways, is therefore

$$\frac{\mathit{\Gamma}\mathsf{p}_1}{\pi}\,\lim\int_{-\frac{1}{2}W}^{\frac{1}{2}W}\frac{X\,dy}{X^2+y^2}\,(+\hat{\boldsymbol{y}})\;=\;\frac{2\mathit{\Gamma}\mathsf{p}_1}{\pi}\,\lim\,\arctan(W/L)\;\hat{\boldsymbol{y}}\;. \tag{5.5}$$

In the next two sections, the foregoing results for problem (ii) will be cross-checked against computations of far-field pseudomomentum fluxes correct to $O(a^2\epsilon^1)$, taking account of the refracted wavecrest shapes illustrated in figure 2 and their mathematical description (1.2). The impulse–pseudomomentum theorem tells us that the results must agree; but it is interesting, nevertheless, not only to carry out the cross-checks but also to see how the refraction works in more detail, thereby gaining mechanistic insight, and another view of the noninterchangeability of limits.†

6. Wave refraction in problem (ii): the far field outside the wake

The impulse-pseudomomentum theorem implies that $\boldsymbol{R}$ can be computed as $-\oint \mathsf{B}\cdot\hat{\boldsymbol{n}}\,ds$, correct to $O(a^2\epsilon^1)$, for a steady wave field whose sources and sinks lie outside the contour of integration. The unit normal $\hat{\boldsymbol{n}}$ is directed outward, and B is the pseudomomentum flux tensor appearing in (3.4). We take advantage of the simplicity of the refracted far-field solution (1.2), and its compatibility with ray theory and the group-velocity property (3.5), by letting the contour expand appropriately as $L\to\infty$ and $W\to\infty$. It is convenient to take the contour to be a rectangle with dimensions L by W, where L is now to be read as the length of the wavetrain excluding its source and sink regions, as they recede to infinity. Then, correct to $O(a^2\epsilon^1)$,

$$\boldsymbol{R}\;=\;-\lim\oint\mathsf{B}\cdot\hat{\boldsymbol{n}}\,ds\;=\;\lim\left(\int_{-\frac{1}{2}W}^{\frac{1}{2}W}\mathsf{B}\cdot\hat{\boldsymbol{x}}\,dy\bigg|_{x=-\frac{1}{2}L}-\int_{-\frac{1}{2}W}^{\frac{1}{2}W}\mathsf{B}\cdot\hat{\boldsymbol{x}}\,dy\bigg|_{x=\frac{1}{2}L}\right). \tag{6.1}$$

We have used (1.2) and (3.5) to neglect the contributions from the sides of the rectangle parallel to $\hat{\boldsymbol{x}}$, as follows. For the transverse, y component, the only relevant component of (3.5) on the sides parallel to $\hat{\boldsymbol{x}}$ is $\mathsf{B}_{22}=\hat{\boldsymbol{y}}\cdot\mathsf{p}\,\boldsymbol{C}^{\rm abs}\cdot\hat{\boldsymbol{y}}$. In the far field, again thanks to ray theory, we have $\mathsf{p}=\overline{\boldsymbol{u}}^{\rm S}=\mathcal{A}\boldsymbol{k}$ where $\mathcal{A}$ is the wave-action per unit mass and where $\boldsymbol{k}=\boldsymbol{\nabla}\mathit{\Phi}=k_0\{\hat{\boldsymbol{x}}-\epsilon r_0 r^{-1}\hat{\boldsymbol{\theta}}+O(\epsilon^2 r_0^2 r^{-2})\}$, from (1.2) and (1.3) or from (2.4). We also have $c=c(r)=c_0\{1+O(\epsilon^2 r_0^2 r^{-2})\}$, as noted in §1. Denoting $\mathsf{p}\cdot\hat{\boldsymbol{x}}$ by p_1 as before, and $\mathsf{p}\cdot\hat{\boldsymbol{y}}$ by p_2, we have $\mathsf{p}_2/\mathsf{p}_1=\boldsymbol{k}\cdot\hat{\boldsymbol{y}}/\boldsymbol{k}\cdot\hat{\boldsymbol{x}}=O(\epsilon r_0 r^{-1})$. From the formula (2.8), again noting the cancellation of leading-order y components, we have $\boldsymbol{C}^{\rm abs}\cdot\hat{\boldsymbol{y}}=O(\epsilon^2 c_0 r_0^2 r^{-2})$. Hence $\mathsf{B}_{22}=\mathsf{p}_1 c_0$ times $O(\epsilon^3 r_0^3 r^{-3})$. For the longitudinal, x component, we have $\mathsf{B}_{12}=\mathsf{p}_1\boldsymbol{C}^{\rm abs}\cdot\hat{\boldsymbol{y}}=\mathsf{p}_1 c_0$ times $O(\epsilon^2 r_0^2 r^{-2})$. Both make negligible contributions as the rectangle expands to infinity.

Now it is clear from §5 that $\boldsymbol{R}\parallel\hat{\boldsymbol{y}}$, so that $\boldsymbol{R}=R\hat{\boldsymbol{y}}$, say, correct to $O(a^2\epsilon^1)$. For the sake of brevity, therefore, we restrict attention from now on to evaluating R from the y component of (6.1). We do this in two stages, to be described in this and the

† Perhaps surprisingly, the quantum fluids literature – going back over the past fifty years or so – tends to ignore many of the points under discussion, including the $O(a^2)$ mean flow problem, the distinction between momentum and pseudomomentum, and the noninterchangeability of limits. The author has, however, found one big quantum fluids paper (Sonin 1997) in which the non-interchangeability of limits is mentioned toward the end of the paper, almost as an afterthought; see below equation (83) therein. Another paper (Wexler & Thouless 1998) takes a different path but flags up the dangers of manipulating divergent infinite series. Some but not all of the $O(a^2)$ effects are discussed in Stone (2000*b*), while all are consistently dealt with in Guo & Bühler (2014), within the Gross–Pitaevskii superfluid model, but only for problem (i). The issues still seem to be surrounded by controversy, perhaps involving unconscious assumptions (e.g. McIntyre 2017) about, for instance, the distinction between particles and quasiparticles.

next section. The first stage is to compute the contribution R_o from outside the wake. The noninterchangeability of limits comes from that contribution. Outside the wake, we can again use (1.2), (2.4), (2.8) and (3.5). At the second stage, in the next section, we compute $R = R_\mathrm{o} + R_\mathrm{w}$, where R_w is the contribution from within the wake. It will be found that R_w agrees with (5.2) and that it is proportional to the Aharonov–Bohm phase jump. The first contribution R_o will be found to agree with (5.5) and to depend solely on the other relevant refraction effect, the $O(\epsilon)$ rotation of wavecrests seen in figure 2 and expressed by the $\hat{\boldsymbol{\theta}}$ term in $\boldsymbol{k}$. From here on we denote B evaluated from (1.2), (2.4), (2.8) and (3.5) by B_o, so that

$$R_\mathrm{o} = \lim \hat{\boldsymbol{y}} \cdot \left(\int_{-\frac{1}{2}W}^{\frac{1}{2}W} \mathsf{B}_\mathrm{o} \cdot \hat{\boldsymbol{x}}\, dy \bigg|_{x=-\frac{1}{2}L} - \int_{-\frac{1}{2}W}^{0-} \mathsf{B}_\mathrm{o} \cdot \hat{\boldsymbol{x}}\, dy \bigg|_{x=\frac{1}{2}L} - \int_{0+}^{\frac{1}{2}W} \mathsf{B}_\mathrm{o} \cdot \hat{\boldsymbol{x}}\, dy \bigg|_{x=\frac{1}{2}L} \right) \tag{6.2}$$

which, in the limit, correctly excludes the wake contribution because of the relative narrowness of the wake, whose width $w \ll W$, as detailed in the next section.

We can evaluate $\hat{\boldsymbol{y}} \cdot \mathsf{B}_\mathrm{o} \cdot \hat{\boldsymbol{x}} = \mathsf{p}_2 \boldsymbol{C}^{\mathrm{abs}} \cdot \hat{\boldsymbol{x}}$ from $\boldsymbol{C}^{\mathrm{abs}} \cdot \hat{\boldsymbol{x}} = c_0\{1 + O(\epsilon r_0 r^{-1})\}$ and $\mathsf{p}_2 = \mathsf{p}_1 \boldsymbol{k} \cdot \hat{\boldsymbol{y}} / \boldsymbol{k} \cdot \hat{\boldsymbol{x}} = -\mathsf{p}_1\{\epsilon r_0 r^{-1} \hat{\boldsymbol{\theta}} \cdot \hat{\boldsymbol{y}} + O(\epsilon^2 r_0^2 r^{-2})\} = \mathsf{p}_1 k_0^{-1}\{\partial \Phi / \partial y + O(\epsilon^2 r_0^2 r^{-2})\}$ with Φ as in (1.2), and where p_1 can now be read as the incident pseudomomentum, neglecting refraction, that is, $\mathsf{p}_1 =$ const. as in §§4–5. Correct to $O(a^2\epsilon^1)$, therefore, the first integral on the right of (6.2) reduces to $c_0 \mathsf{p}_1 k_0^{-1} \lim \int_{-\alpha\tilde{\theta}}^{+\alpha\tilde{\theta}} d\Phi = c_0 \mathsf{p}_1 k_0^{-1} \lim\,(2\alpha\tilde{\theta}) = (\Gamma \mathsf{p}_1/\pi) \lim \tilde{\theta}$, where $\tilde{\theta} = \arctan(W/L) > 0$. The noninterchangeability of the limits $L \to \infty$ and $W \to \infty$ is now evident.

At each fixed y, and correct to $O(a^2\epsilon^1)$, p_2 is an odd function of x, and $\hat{\boldsymbol{y}} \cdot \mathsf{B}_\mathrm{o} \cdot \hat{\boldsymbol{x}}$ also. Therefore the second and third integrals in (6.2) add up to a contribution equal to that from the first integral, so that altogether

$$R_\mathrm{o} \;=\; \frac{2\Gamma \mathsf{p}_1}{\pi} \lim\, \arctan(W/L) \tag{6.3}$$

correct to $O(a^2\epsilon^1)$, in agreement with (5.5). Notice incidentally that problem (i) now appears as a trivial variant of the above, obtained by selecting appropriate subsets of rays in cases where ray theory is valid all along the wavetrain.

7. Wave refraction in problem (ii): the far field within the wake

To complete the work on problem (ii) we need to evaluate the remaining contribution to the y component of (6.1),

$$R_\mathrm{w} \;=\; -\lim \int_{\mathrm{wake}} \hat{\boldsymbol{y}} \cdot \mathsf{B}_\mathrm{w} \cdot \hat{\boldsymbol{x}}\, dy \bigg|_{x=\frac{1}{2}L}, \tag{7.1}$$

and to verify that it agrees with (5.2). Here B_w stands for B within the wake.

We evaluate (7.1) in two cases that are analytically tractable, $k_0 r_0 \ll 1$ and $k_0 r_0 \gg 1$. In the second case we use ray tracing across the vortex core, and in the first we draw on the work of Ford & Llewellyn Smith (1999, hereafter FLS), who carried out a careful and thorough asymptotic analysis of weak refraction and scattering in that case, building on earlier contributions including that of Sakov (1993); see also Belyaev & Kopiev (2008).

7.1. *The long-wave case $k_0 r_0 \ll 1$*

FLS's results are subject to a severe restriction on the range of α values for which they are valid. From (1.3) we see that α is now the product of two small quantities $k_0 r_0$ and ϵ. The results are nevertheless attractive for our purposes because, when valid, they show that the wake has a simple Fresnel-diffractive structure with width-scale $w \sim k_0^{-1/2} L^{1/2} \gg k_0^{-1}$, and angular scale asymptotically zero, as $L \to \infty$. Across the wake there is a smooth phase transition such that the ray-theoretic formulae used in §6 still hold, including $\mathsf{p} \propto \boldsymbol{k}$ and $\mathsf{p}_2 = \mathsf{p}_1 k_0^{-1} \partial\Phi/\partial y$, where Φ now denotes the phase within the wake, as distinct from the phase given by (1.2). The wave field within the wake still has the form $A\exp(\mathrm{i}\Phi)$, where the amplitude A is still real and constant to sufficient accuracy across the wake, with relative error $O(\epsilon)$, and where the phase Φ increases by $2\pi\alpha$ going anticlockwise across the wake. Therefore we can evaluate $\int \hat{\boldsymbol{y}}\cdot \mathsf{B}_\mathrm{w}\cdot \hat{\boldsymbol{x}}\, dy$ across the wake as $\int c_0 \mathsf{p}_2\, dy = c_0 \mathsf{p}_1 k_0^{-1} \int d\Phi = 2\pi\alpha c_0 \mathsf{p}_1 k_0^{-1} = \Gamma \mathsf{p}_1$, verifying that (7.1) with its minus sign does agree with (5.2).

FLS's solution also contains a Born-scattering term with amplitude $O(r^{-1/2})$, which however contributes nothing. The magnitude $O(r^{-1})$ of its pseudomomentum flux makes it potentially able to contribute to $-\oint \mathsf{B}\cdot\hat{\boldsymbol{n}}\, ds$. But for our rectangular integration contour the contribution to $\hat{\boldsymbol{y}}\cdot \mathsf{B}\cdot \hat{\boldsymbol{n}}$ is an odd function of y, which integrates to zero.

The reader who wishes to check the foregoing against FLS in more detail may find the following notes useful. Outside the wake, the leading far-field term in FLS's solution agrees with the foregoing for any $\alpha \ll 1$ because we then have, from (1.2),

$$\begin{aligned} \exp(\mathrm{i}\Phi) &= \exp(-\mathrm{i}\alpha\theta)\exp\{\mathrm{i}(k_0(x-c_0t)+\text{const.})\} \\ &= (1-\mathrm{i}\alpha\theta)\exp\{\mathrm{i}(k_0(x-c_0t)+\text{const.})\} + O(\alpha^2)\ . \end{aligned} \tag{7.2}$$

This agrees with FLS's (2.14), (2.20) and the leading term in their (5.7), after allowing for their different definition of θ and remembering that our θ jumps from $+\pi$ to $-\pi$, going anticlockwise across the positive x axis. In their dimensionless notation, our α is written as $M^2\Gamma\omega/2\pi$, where their M is our ϵ and their $\Gamma\omega/2\pi$ is our $k_0 r_0/\epsilon$, all taken positive. The Born scattering term is the next, $O(r^{-1/2})$ term in their (5.7), with outgoing waves $\propto r^{-1/2}\exp\{\mathrm{i}k_0(r-c_0t)\}$. The Fresnel wake is described by their (5.12). The phase transition at fixed x is given by the sum $C+S$ of two real-valued Fresnel integrals, suitably scaled, and is therefore an odd function of y/w, with Φ asymptoting toward $\pm\pi\alpha + \text{const.}$ with gentle oscillations on the scale w. The factor $\exp(-\mathrm{i}\eta^2)$ in FSL's (5.11), where $\eta^2 = \frac{1}{2}(y/w)^2$, converts a Born-like factor $\exp\{\mathrm{i}k_0(r-c_0t)\}$ into a plane-wave factor $\exp\{\mathrm{i}k_0(x-c_0t)\}$, to sufficient accuracy within the wake, matching up with our (7.2).

Belyaev & Kopiev (2008) reconsider FLS using a different solution technique, that of Aharonov & Bohm (1959) and Berry *et al.* (1980). They also discuss the conceptual issue of whether (1.2) above can usefully be regarded as a plane wave outside the Fresnel wake, in the limit $r \to \infty$. However, the wave field properties required by the foregoing analysis of the pseudomomentum budget are unaffected. Those properties are, most crucially, the validity of (1.2) outside the wake, the phase continuity across the wake, and the y-antisymmetry of the Born contribution to $\hat{\boldsymbol{y}}\cdot \mathsf{B}\cdot \hat{\boldsymbol{n}}$. And we note again that the analysis is independently confirmed by the end-to-end cross-check from §§3–5.

7.2. *The short-wave case $k_0 r_0 \gg 1$*

We now evaluate (7.1) in the case $kr_0 \gg 1$, using ray theory. Attention is restricted to the simplest case, the Rankine vortex model in which the core is in solid rotation, with constant angular velocity $\Omega = \frac{1}{2}|\boldsymbol{\omega}_0| = |\boldsymbol{u}_0(r)|/r = \Gamma/2\pi r_0^2$, taken positive, i.e. anticlockwise, for definiteness, as in the figures.

Correct to $O(\epsilon)$ the rays outside the core and its lee are straight, as already remarked, with absolute group velocities $\boldsymbol{C}^\mathrm{abs}$ parallel to $\hat{\boldsymbol{x}}$, but slightly rotated wavecrests, as seen in figure 2. Inside the core, the ray-tracing equations, e.g. (2.14) of BM03, verify what is obvious from rotating the reference frame while keeping the same intrinsic phase and group velocity $c = c_0\{1+O(\epsilon^2)\}$, namely that the wavenumber vector $\boldsymbol{k}$ rotates with angular velocity Ω as a ray point crosses the core at velocity $c_0\hat{\boldsymbol{x}} + O(\epsilon)$. Thus the rays

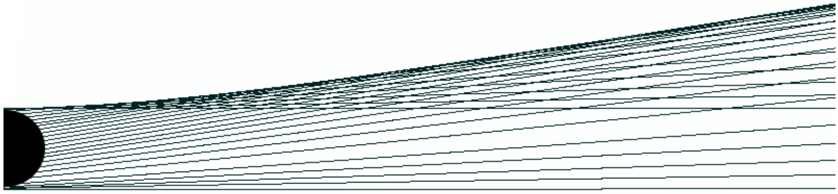

FIGURE 5. Rays emerging from the vortex core and forming a wake with a caustic, computed correct to $O(\epsilon)$ from the ray-tracing equations. The $O(\epsilon)$ deflections are exaggerated in the plot. The rays entering the core, not shown, are initially parallel to the x axis at $y/r_0 = 0$, ± 0.1, ± 0.2, ± 0.3, ± 0.4, ± 0.5, ± 0.6, ± 0.7, ± 0.8, ± 0.9, ± 0.95, ± 0.98, ± 0.995, ± 1.

bend slightly to the left as they cross the core into its lee, where they are straight again but no longer quite parallel to the x axis. In fact the group velocity $\boldsymbol{C}^{\mathrm{abs}}$ rotates twice as fast as $\boldsymbol{k}$, with angular velocity 2Ω, because of the changing $O(\epsilon)$ contribution from the y component of $\boldsymbol{u}_0 = \Omega(-y, x)$ in (2.8) as the ray point crosses the core. This and the straightness of the rays outside the core are special cases of the curl-curvature formula mentioned below (2.8). The formula implies generally that, correct to $O(\epsilon)$, the group velocity vector rotates with angular velocity $\boldsymbol{\omega}_0 = \boldsymbol{\nabla} \times \boldsymbol{u}_0$.

For weak refraction the ray undergoing the greatest deflection is that crossing the widest part of the core, at $y = 0$. The rays from $-r_0 < y < 0$ therefore splay out slightly, while those from $0 < y < r_0$ cross one another and form a caustic, extending slightly outside the tangent line $y = +r_0$, as shown in figure 5 with the deflections exaggerated. A full analysis is beyond our scope here; to evaluate (7.1) we will simply add up the leading-order pseudomomentum fluxes $\hat{\boldsymbol{y}} \cdot \mathsf{B}_{\mathrm{w}} \cdot \hat{\boldsymbol{x}} = c_0 \mathsf{p}_2$ as if carried by each ray independently. This assumes that the refraction term in the $O(a^2\epsilon^1)$ pseudomomentum law (3.4) works in the same way, to leading order at least, whether or not the rays go through a caustic.

The treatment of the ray deflections as small of order ϵ, and the neglect of diffractive effects, becomes delicate and very restrictive when combined with the formal limit $L \to \infty$. It will nevertheless yield the correct result, in agreement with (5.2), as will now be shown. The agreement will also lend support to our assumption about (3.4) and caustics.

The following is a shortcut to the results from the ray-tracing equations shown in figure 5, using the rotation-rate of $\boldsymbol{k}$ already mentioned. A ray point entering the core at $y = -r_0 \sin\theta$, say, for some fixed θ with $|\theta| < \pi/2$, already has a wavenumber with nonvanishing y component $\boldsymbol{k}\cdot\hat{\boldsymbol{y}} = -\alpha r_0^{-1}\hat{\boldsymbol{\theta}}\cdot\hat{\boldsymbol{y}} = +\alpha r_0^{-1}\cos\theta$, from (1.2) or (2.4) giving, at that location, $\hat{\boldsymbol{y}} \cdot \mathsf{B}_{\mathrm{w}} \cdot \hat{\boldsymbol{x}} = c_0 \mathsf{p}_2 = c_0 \mathsf{p}_1 k_0^{-1} \alpha r_0^{-1} \cos\theta = \mathsf{p}_1 \Gamma (2\pi r_0)^{-1} \cos\theta = \mathsf{p}_1 \Omega r_0 \cos\theta$. As the ray point crosses the core, over a distance $2r_0\cos\theta$ and taking a time $2c_0^{-1} r_0 \cos\theta$, the vectors $\boldsymbol{k}$ and p rotate with angular velocity Ω, so that $c_0\mathsf{p}_2$ increases by a further small amount $2\mathsf{p}_1 \Omega r_0 \cos\theta$ (equal to $c_0 \mathsf{p}_1$ times the net rotation angle $2c_0^{-1}\Omega r_0 \cos\theta$), and again by a further small amount $\mathsf{p}_1 \Omega r_0 \cos\theta$ after exiting the core and reaching sufficiently large $x > 0$. This last increment is the same increment as in (1.2) between the far edge of the core and $x \to \infty$. However, it is to be added to the new far-core-edge value $3\mathsf{p}_1 \Omega r_0 \cos\theta$ rather than to the original value $-\mathsf{p}_1 \Omega r_0 \cos\theta$ implied by (1.2) and (2.4). With our ray still nearly parallel to the x axis after exiting the core, we are using the fact that the $O(\epsilon)$ *rates of change* of $\boldsymbol{k} \cdot \hat{\boldsymbol{y}}$ outside the core are the same as those implied by (1.2), (2.4) and (2.8).

So, adding all the contributions just noted, we have that the ray has a total end-to-end change $\mathsf{p}_1 \Omega r_0 \cos\theta + 2\mathsf{p}_1 \Omega r_0 \cos\theta + \mathsf{p}_1 \Omega r_0 \cos\theta = 4\mathsf{p}_1 \Omega r_0 \cos\theta$, in $c_0 \mathsf{p}_2$, corresponding to an end-to-end deflection angle β, say, $= \mathsf{p}_2/\mathsf{p}_1|_{x\to\infty} = 4c_0^{-1}\Omega r_0 \cos\theta = 4\epsilon\cos\theta$, with maximum value 4ϵ, in an anticlockwise sense. Integrating the change in $c_0 \mathsf{p}_2$ over all the rays that cross the core, from $y = -r_0$ to $y = r_0$, noting that $dy = -r_0 \cos\theta\, d\theta$ and that $\int_{-\pi/2}^{\pi/2} \cos^2\theta\, d\theta = \pi/2$, we find from (7.1) with due care over signs that $R_{\rm w} = -2\pi \mathsf{p}_1 \Omega r_0^2 = -\Gamma \mathsf{p}_1$, agreeing with (5.2) as expected.

To summarize, (6.1) gives us, both for $k_0 r_0 \ll 1$ and for $k_0 r_0 \gg 1$,

$$R \;=\; R_{\rm w} + R_{\rm o} \;=\; -\left\{1 - \frac{2}{\pi}\lim\, \arctan\left(\frac{W}{L}\right)\right\} \Gamma \mathsf{p}_1 \tag{7.3}$$

for arbitrary limiting values of W/L, agreeing with the independent derivations in §§3–5. Furthermore, recalling that those independent derivations are valid for arbitrary $k_0 r_0$, we see also that (7.3) must be a result far more robust than is suggested by the delicacy of the flux computations for $k_0 r_0 \ll 1$ and $k_0 r_0 \gg 1$. On the other hand, everything still depends on the smallness of ϵ. Numerical solutions for cases of stronger refraction show wakes oriented at substantial angles away from the x axis; see for instance figures 2–3 of Coste *et al.* (1999). And it is still an open question as to what might or might not replace the impulse–pseudomomentum theorem for arbitrary ϵ.

8. Problem (iii)

The main reason for being interested in this rapidly-rotating version of problem (i), in which the waves are deep-water gravity waves, is the existence of the Ursell anti-Stokes flow. This is an Eulerian-mean flow $\overline{\boldsymbol{u}}$ that largely cancels the strongly z-dependent Stokes drift $\overline{\boldsymbol{u}}^{\rm S}$ of the waves. Indeed the cancellation is exact, for finite-amplitude waves, when the wave field is exactly steady and exactly homogeneous across an infinite xy domain (Ursell 1950; Pollard 1970). (In a thought-experiment starting with irrotational waves, in such a domain, the mean flow undergoes a free inertial oscillation about the anti-Stokes state. It is sometimes forgotten that this thought-experiment was clearly analysed and understood in Ursell's pioneering work.) In problem (iii), however, contrary to what might at first be thought, the anti-Stokes flow fails to suppress remote recoil.

We consider an unstratified rapidly-rotating system of finite depth H under gravity $-g\hat{\boldsymbol{z}}$, so that the z direction is vertically upward. The vector Coriolis parameter $\boldsymbol{f}$ is parallel to $\hat{\boldsymbol{z}}$. The waves have a wavenumber $k = |\boldsymbol{k}|$ that is large enough to make $\exp(-kH)$ negligible. The intrinsic wave frequency $kc = (gk)^{1/2} \gg f = |\boldsymbol{f}|$, so that rotation affects the wave dynamics only weakly.

In addition to a and ϵ we now have another small parameter, the mean-flow Rossby number

$$\mathrm{Ro} \;=\; U/fr_0 \;\ll\; 1\;, \tag{8.1}$$

whose smallness will bring in the Taylor–Proudman effect and give us quasigeostrophic mean-flow dynamics. As before, the velocity scale U will be taken as the velocity of the vortex flow at the edge of the core, $r = r_0$. The core will be defined by nonvanishing quasigeostrophic potential vorticity, $(\nabla_{\rm H}^2 - L_{\rm D}^{-2})\tilde{\psi}_0$ in the notation of §2.

The anti-Stokes flow can be regarded as a consequence of the Taylor–Proudman effect together with the exact advection property expressed by the mean vorticity equations (2.10) and (2.11). This is no more than a rephrasing of Ursell's original argument, putting it within the GLM framework. Focusing on the present case $\mathrm{Ro} \ll 1$, we can regard inertia waves as fast waves with a strong Coriolis restoring force. The Taylor–Proudman effect arises from the corresponding stiffness of the lines of absolute vorticity $\boldsymbol{f} + \widetilde{\boldsymbol{\omega}}$, which must tend to stay vertical, on average at least. In particular, they cannot be continually sheared over by the mean flow; and for this purpose the mean flow is the Lagrangian-mean flow $\overline{\boldsymbol{u}}^{\rm L}$, as equations (2.10) and (2.11) make clear. A Lagrangian-mean flow without vertical shear is a Stokes drift plus an Eulerian-mean anti-Stokes flow, plus an additional contribution that is independent of z – in this case the vortex flow plus the Bretherton flow that mediates remote recoil.

To tackle problem (iii) we must first derive (2.18). The starting point is the vertical component of (2.11). Writing $f+\widetilde{\omega}$ for the vertical component of $\boldsymbol{f}+\widetilde{\boldsymbol{\omega}}$, and $\overline{w}^{\mathrm{L}}$ for the vertical component of $\overline{\boldsymbol{u}}^{\mathrm{L}}$, we have

$$\frac{\overline{D}^{\mathrm{L}}\widetilde{\omega}}{Dt} \;+\; (f+\widetilde{\omega})\boldsymbol{\nabla}\cdot\overline{\boldsymbol{u}}^{\mathrm{L}} \;=\; (\boldsymbol{f}+\widetilde{\boldsymbol{\omega}})\cdot\boldsymbol{\nabla}\overline{w}^{\mathrm{L}} \tag{8.2}$$

exactly. Upon cancelling a pair of terms in $\partial\overline{w}^{\mathrm{L}}/\partial z$, this reduces to

$$\frac{\overline{D}^{\mathrm{L}}\widetilde{\omega}}{Dt} \;+\; (f+\widetilde{\omega})\boldsymbol{\nabla}_{\mathrm{H}}\cdot\overline{\boldsymbol{u}}^{\mathrm{L}}_{\mathrm{H}} \;=\; \widetilde{\boldsymbol{\omega}}_{\mathrm{H}}\cdot\boldsymbol{\nabla}_{\mathrm{H}}\overline{w}^{\mathrm{L}}. \tag{8.3}$$

As before, suffix H denotes horizontal projection. Writing $\widetilde{\omega}=\omega_0+\widetilde{\omega}_{\mathrm{B}}$ and $\overline{\boldsymbol{u}}^{\mathrm{L}}=\boldsymbol{u}_0+\overline{\boldsymbol{u}}^{\mathrm{L}}_{\mathrm{B}}$ where the vortex-only contributions ω_0 and $\boldsymbol{u}_0$ are z-independent, with $\boldsymbol{u}_0$ horizontal, and the wave-induced contributions $\widetilde{\omega}_{\mathrm{B}}$ and $\overline{\boldsymbol{u}}^{\mathrm{L}}_{\mathrm{B}}$ are $O(a^2)$, we note that in the first term on the left the contribution $\overline{\boldsymbol{u}}^{\mathrm{L}}_{\mathrm{B}}\cdot\boldsymbol{\nabla}\,\widetilde{\omega}_{\mathrm{B}} \;=\; \overline{\boldsymbol{u}}^{\mathrm{L}}_{\mathrm{B}}\cdot\boldsymbol{\nabla}_{\mathrm{H}}\,\widetilde{\omega}_{\mathrm{B}} \;+\; \overline{w}^{\mathrm{L}}\partial\widetilde{\omega}_{\mathrm{B}}/\partial z \;=\; O(a^4)$ and is therefore negligible. (We need not restrict ϵ at this stage.) There are two further such $O(a^4)$ contributions, namely the right-hand side, and on the left $\widetilde{\omega}_{\mathrm{B}}\,\boldsymbol{\nabla}_{\mathrm{H}}\cdot\overline{\boldsymbol{u}}^{\mathrm{L}}_{\mathrm{H}}=\widetilde{\omega}_{\mathrm{B}}\,\boldsymbol{\nabla}_{\mathrm{H}}\cdot\overline{\boldsymbol{u}}^{\mathrm{L}}_{\mathrm{B}}$ since $\boldsymbol{\nabla}_{\mathrm{H}}\cdot\boldsymbol{u}_0=0$. The $O(a^2)$ contribution $\omega_0\,\boldsymbol{\nabla}_{\mathrm{H}}\cdot\overline{\boldsymbol{u}}^{\mathrm{L}}_{\mathrm{H}}=\omega_0\,\boldsymbol{\nabla}_{\mathrm{H}}\cdot\overline{\boldsymbol{u}}^{\mathrm{L}}_{\mathrm{B}}$ is also negligible, against $f\,\boldsymbol{\nabla}_{\mathrm{H}}\cdot\overline{\boldsymbol{u}}^{\mathrm{L}}_{\mathrm{H}}$, because of the smallness of Ro. After neglecting all these contributions we can take the vertical average of (8.3), using the z-independence of $\boldsymbol{u}_0$ and $\boldsymbol{\omega}_0$. Denoting vertical averages by angle brackets as before and noting that $\boldsymbol{u}_0\cdot\boldsymbol{\nabla}_{\mathrm{H}}\,\omega_0=0$ we get

$$\left(\frac{\partial}{\partial t}+\boldsymbol{u}_0\cdot\boldsymbol{\nabla}_{\mathrm{H}}\right)\langle\widetilde{\omega}_{\mathrm{B}}\rangle \;+\; \langle\overline{\boldsymbol{u}}^{\mathrm{L}}_{\mathrm{B}}\rangle\cdot\boldsymbol{\nabla}_{\mathrm{H}}\,\omega_0 \;+\; f\,\boldsymbol{\nabla}_{\mathrm{H}}\cdot\langle\overline{\boldsymbol{u}}^{\mathrm{L}}_{\mathrm{H}}\rangle \;=\; 0 \tag{8.4}$$

or, written more compactly, again with negligible error $O(a^4)$,

$$\frac{\overline{D}^{\mathrm{L}}_{\mathrm{H}}\langle\widetilde{\omega}\rangle}{Dt} \;+\; f\,\boldsymbol{\nabla}_{\mathrm{H}}\cdot\langle\overline{\boldsymbol{u}}^{\mathrm{L}}_{\mathrm{H}}\rangle \;=\; 0\;, \tag{8.5}$$

where we have defined $\overline{D}^{\mathrm{L}}_{\mathrm{H}}/Dt=\partial/\partial t+\langle\overline{\boldsymbol{u}}^{\mathrm{L}}_{\mathrm{H}}\rangle\cdot\boldsymbol{\nabla}_{\mathrm{H}}$. In a closely similar way, the vertical average of the three-dimensional mass-conservation equation, B14 equation (10.47), simplifies to a vertically-averaged version of (2.13),

$$\frac{\overline{D}^{\mathrm{L}}_{\mathrm{H}}\tilde{h}}{Dt} \;+\; \tilde{h}\,\boldsymbol{\nabla}_{\mathrm{H}}\cdot\langle\overline{\boldsymbol{u}}^{\mathrm{L}}_{\mathrm{H}}\rangle \;=\; 0\;, \tag{8.6}$$

again with negligible error $O(a^4)$. As before, the mean layer depth $\tilde{h}=\tilde{h}(x,y,t)$ is defined such that $\rho\tilde{h}\,dxdy$ is the areal mass element, where ρ is the constant mass density. Elimination of $\boldsymbol{\nabla}_{\mathrm{H}}\cdot\langle\overline{\boldsymbol{u}}^{\mathrm{L}}_{\mathrm{H}}\rangle$ between (8.5) and (8.6) gives us that $\langle\widetilde{\omega}\rangle-f\ln\tilde{h}$, plus an arbitrary additive constant, is a material invariant under advection by $\langle\overline{\boldsymbol{u}}^{\mathrm{L}}_{\mathrm{H}}\rangle$. Fractional changes in $\tilde{h}$ are small, $(\tilde{h}-H)/H=O(\mathrm{Ro})$, and so taking the additive constant to be $f\ln H$ and using $\ln\tilde{h}-\ln H=\ln(\tilde{h}/H)=\ln\{(\tilde{h}-H+H)/H\}=(\tilde{h}-H)/H+O(\mathrm{Ro}^2)$, we get

$$\frac{\overline{D}^{\mathrm{L}}_{\mathrm{H}}\tilde{q}}{Dt} \;=\; 0 \tag{8.7}$$

where

$$\tilde{q} \;=\; \tilde{q}(x,y,t) \;=\; \langle\widetilde{\omega}\rangle-\frac{f}{H}(\tilde{h}-H)\;, \tag{8.8}$$

which is the appropriate form of the quasigeostrophic potential vorticity of the mean flow. In any thought-experiment in which the waves are switched on after the vortex is established, (8.7) implies that the $\tilde{q}$ field is unchanged by the presence of the waves, apart from the advection of the vortex core by the Bretherton flow. See also Appendix B. So in problem (iii) we have $\tilde{q}=q_0$ where q_0 is the potential vorticity of the vortex alone.

The final step in deriving (2.18) is to make explicit use of hydrostatic and geostrophic balance. Some delicate scale analysis is involved at this stage. The full details are given in Appendix C, in which the key points are as follows. Hydrostatic balance, meaning the overall balance for a complete fluid column, implies that horizontal pressure gradients on the bottom, underneath the wavetrain, are given by $\rho g \boldsymbol{\nabla}_{\rm H} \tilde{h}$, again because $\rho \tilde{h}\, dxdy$ is the areal mass element. Geostrophic balance then gives (2.15) with $\tilde{\psi} = g(\tilde{h} - H)/f$. Then $\tilde{q} = q_0$ together with (2.9) and (2.16) gives (2.18). The Taylor–Proudman effect extends the geostrophic relation upward into the wavetrain; $\langle \overline{\boldsymbol{u}}^{\rm L}_{\rm H} \rangle = \overline{\boldsymbol{u}}^{\rm L}_{\rm H}$. Radiation stresses within the wavetrain cannot break the overall hydrostatic balance because such stresses have no foothold on the bottom boundary, in virtue of our assumption that $\exp(-kH)$ is negligibly small. That allows us to neglect the net vertical, radiation-stress-induced external force on the fluid column – in contrast, it should be noted, with the situation of figure 3. For further comments see Appendix C. In Appendix C we also note that the exact wave solution of Pollard (1970) provides some useful cross-checks.

With (2.18) in place, we can now invoke the impulse-pseudomomentum theorem to assert that recoil forces can be computed either from Bretherton flows correct to $O(a^2\epsilon^0)$ or from net pseudomomentum fluxes correct to $O(a^2\epsilon^1)$. In the remainder of this section we carry out both computations, in the case of a small vortex core with $r_0 \ll L_{\rm D}$, providing mechanistic insight as well as an end-to-end cross-check on our derivation of (2.18).

First consider the Bretherton flow. Because it satisfies (2.18), it decays sideways like $\exp(-|y|/L_{\rm D})$, on the fixed length-scale $L_{\rm D}$. Therefore there is no dilution effect like that in problem (i). With $L \to \infty$, and with a narrow wavetrain for which $W \ll L_{\rm D}$ and $W \ll Y$, in the notation of §4, we have, for $|y - Y| > W$, outside the unrefracted wavetrain, with Y the distance to the vortex core,

$$\overline{\boldsymbol{u}}^{\rm L}_{\rm B}(x,\, y) \;=\; (S/2L_{\rm D})\exp(-|y-Y|/L_{\rm D})(-\hat{\boldsymbol{x}}) \tag{8.9}$$

where S is still defined by (4.1) but with vertical averaging understood. So, for our small vortex core with $r_0 \ll L_{\rm D}$, carried bodily by the z-independent Bretherton flow, we take $y = 0$ in (8.9) to get

$$\boldsymbol{R} \;=\; (\mathit{\Gamma} S/2L_{\rm D})\exp(-|Y|/L_{\rm D})(+\hat{\boldsymbol{y}})\;, \tag{8.10}$$

with $\mathit{\Gamma}$ evaluated at the edge of the core. The signs are the same as those in problem (i).

Second, we compute $\boldsymbol{R}$ from the $O(a^2\epsilon^1)$ pseudomomentum flux $\mathsf{B}_{21} = \hat{\boldsymbol{y}} \cdot \mathsf{B} \cdot \hat{\boldsymbol{x}}$, using ray theory. As in §7.2, the rays start exactly parallel to the x axis, with $\mathsf{p}_2 \to 0$ as $x \to -\infty$, and finish after bending slightly, through an $O(\epsilon)$ end-to-end deflection angle $\beta = \mathsf{p}_2/\mathsf{p}_1|_{x\to\infty}$. The vortex flow has velocity $\boldsymbol{u}_0(r) = \hat{\boldsymbol{\theta}}\,\partial\psi_0/\partial r$, say, where the quasigeostrophic streamfunction ψ_0 satisfies $(\nabla^2 - L_{\rm D}^{-2})\psi_0 = 0$ outside the core. Defining $r' = r/L_{\rm D}$, we have

$$\psi_0 \;=\; -\frac{\mathit{\Gamma}}{2\pi} K_0(r') \qquad \text{outside the core,} \tag{8.11}$$

where $K_0(r')$ is the modified Bessel function asymptoting to $(\pi/2r')^{1/2}\exp(-r')$ for $r' \gg 1$ and to $-\ln(r')$ for $r' \ll 1$, near the core. The Kelvin circulation $\mathit{\Gamma}$ is again defined to be the circulation at the core edge $r = r_0$, namely $\pm 2\pi r_0|\boldsymbol{u}_0(r_0)| = \pm 2\pi r_0 U$, with positive sign when the vortex is cyclonic as in the figures. For $r > r_0$ the circulation is not constant, but decays exponentially like $r'^{1/2}\exp(-r')$.

To verify agreement with (8.10) we need only calculate β. The curl-curvature formula tells us that β is nonzero at $O(\epsilon)$, because the relative vorticity $\nabla^2\psi_0 = L_{\rm D}^{-2}\psi_0$ is nonzero outside the core. A cyclonic vortex core is surrounded by anticyclonic vorticity and the rays therefore bend to the right, rather than to the left as in §7, so that $\operatorname{sgn}\beta = -\operatorname{sgn}\mathit{\Gamma}$. Notice incidentally that there will no longer be any far-field subtleties, or issues with noninterchangeable limits, thanks to the exponential decay of ψ_0. Another effect of that decay is that the right-bending rays must splay out slightly when they pass to the left of the vortex, but cross one another and form a caustic when to the right. The presence or absence of a caustic makes no difference to the results.

The deep water waves in problem (iii) have intrinsic frequency $kc = (gk)^{1/2}$ and intrinsic group velocity $\boldsymbol{C} = C\boldsymbol{k}/k$ where $C = \frac{1}{2}c = \frac{1}{2}(g/k)^{1/2}$. The absolute group velocity $\boldsymbol{C}^{\text{abs}} = \frac{1}{2}c_0\hat{\boldsymbol{x}} + O(\epsilon)$, with $O(\epsilon)$ contributions coming both from $\boldsymbol{u}_0$ and from refractive changes in wavenumber $\boldsymbol{k}$. Following a ray point moving at speed $\frac{1}{2}c_0 + O(\epsilon)$, the curl-curvature formula says that the direction of $\boldsymbol{C}^{\text{abs}}$ rotates clockwise away from the x direction at an angular velocity equal to the (negative) relative vorticity $\nabla^2\psi_0 = L_{\text{D}}^{-2}\psi_0$. So for weak refraction we have

$$\beta \;=\; (\tfrac{1}{2}c_0)^{-1}\int_{-\infty}^{\infty} L_{\text{D}}^{-2}\psi_0(x,Y)\,dx \;=\; -\frac{\Gamma}{\pi c_0 L_{\text{D}}^2}\int_{-\infty}^{\infty} K_0\{(x^2+Y^2)^{1/2}/L_{\text{D}}\}\,dx\;. \tag{8.12}$$

The integral on the right is exactly equal to $L_{\text{D}}\,\pi\exp(-|Y|/L_{\text{D}})$, as will be shown shortly. Hence $\beta = -(\Gamma/c_0 L_{\text{D}})\exp(-|Y|/L_{\text{D}})$. Remembering that $C = \frac{1}{2}c_0$, we see that there is an end-to-end difference in pseudomomentum fluxes representing a rate of import $-\frac{1}{2}c_0\,\mathsf{p}_2\big|_{x\to\infty} = -\frac{1}{2}c_0\beta\mathsf{p}_1 = +(\Gamma\mathsf{p}_1/2L_{\text{D}})\exp(-|Y|/L_{\text{D}})$ of y-pseudomomentum per unit y-distance, correct to $O(a^2\epsilon^1)$. Recalling the definition of S in (4.1), with vertical averaging understood, we sum over all the rays to find the total recoil force in the y direction as

$$\boldsymbol{R} \;=\; (\Gamma S/2L_{\text{D}})\exp(-|Y|/L_{\text{D}})(+\hat{\boldsymbol{y}})\;, \tag{8.13}$$

in agreement with (8.10).

The integral on the right of (8.12) is equal to L_{D} times the value at $y' = Y/L_{\text{D}}$ of the function $I(y')$ defined by $I(y') = \int_{-\infty}^{\infty} K_0(r')dx'$ where $x' = x/L_{\text{D}}$ and $y' = y/L_{\text{D}}$ so that $r'^2 = x'^2 + y'^2$. Now $K_0(r')$ is equal to its Laplacian in the x', y' plane, except at the origin where the Laplacian has a delta function $-2\pi\delta(x')\delta(y')$ in place of the integrable logarithmic singularity in K_0 itself. For any $y' \neq 0$ we therefore have

$$I(y') \;=\; \int_{-\infty}^{\infty}\left(\frac{\partial^2}{\partial x'^2} + \frac{\partial^2}{\partial y'^2}\right)K_0(r')dx' \;=\; \frac{d^2}{dy'^2}\int_{-\infty}^{\infty} K_0(r')dx' \;=\; \frac{d^2}{dy'^2}I(y') \tag{8.14}$$

and, taking the delta function into account, we have for all y' from $-\infty$ to $+\infty$

$$\frac{d^2}{dy'^2}I(y') - I(y') \;=\; -2\pi\delta(y')\;, \tag{8.15}$$

whose solution evanescent at infinity is $I(y') = \pi\exp(-|y'|)$, corresponding to the result asserted.

9. Concluding remarks

Despite their restricted parameter range, the problems studied here are enough to remind us that remote recoil, as such, is generic and ubiquitous. Remote recoil will occur whenever wave-induced mean flows extend outside wavetrains or wave packets and advect coherent vortices. Remote recoil is excluded, or made subdominant, by asymptotic theories of wave–current interactions that assume slowly-varying mean currents with a single large length-scale and correspondingly weak vorticity or PV anomalies.

The main question left open by this work concerns the scope of the pseudomomentum rule. As remarked at the end of §3, the rule is known to be valid in a wider range of cases than those considered here, even though in a still wider context there are known exceptions including the case of one-dimensional sound waves in a rigid tube, as noted long ago in Brillouin's classic works on radiation stress. Further exceptions include the internal-gravity-wave problem studied in McIntyre (1973) and the rotating problems studied in Thomas *et al.* (2018, & refs.), in some ways similar to our problem (iii). The failure of the rule in these latter cases, and in Brillouin's, is related to $O(a^2)$ mean pressure reactions from confining boundaries (more detail in Appendix C below). We may similarly expect failure of the rule in laboratory experiments such as those of Humbert *et al.* (2017), conducted in tanks or channels with confining walls that can support $O(a^2)$ mean pressures. Section 3 reminds us that the impulse–pseudomomentum theorem depends on having a sufficiently large fluid domain enclosing the regions occupied by waves and vortices.

For the reasons indicated at the end of section 3, even in a large domain the scope of the pseudomomentum rule in wave–vortex interactions is very much a nontrivial question calling for further research, probably involving numerical experimentation along the lines of the strong-refraction experiments of Coste *et al.* (1999). Even though the Kelvin impulse concept depends on banishing large-scale pressure-field adjustments to infinity, the basic thought-experiment associated with the Magnus relation (3.2), that of applying a force to move a vortex core, works, by contrast, in a relatively local way. This poses not only a technical but also a nontrivial conceptual challenge.

Regarding quantum vortices, it would be interesting to see how the present analysis of problem (ii) extends to the Gross–Pitaevskii superfluid model, a context in which problem (i) was studied in Guo & Bühler (2014). In the corresponding version of problem (ii) we can expect to find the same noninterchangeability of limits and the same caveats regarding the Aharonov–Bohm effect, pointing to a remote-recoil contribution in addition to the Iordanskii force. The Gross–Pitaevskii model provides a simple representation of quantum vortex cores (Berloff 2004), whose supersonic flow velocities might vitiate any attempt at a weak-refraction theory, even though the small core size might, on the other hand, imply that bodily advection of the core – back and forth by a larger-scale wavemotion as well as persistently by the mean flow – could still be a useful simplifying feature.

Acknowledgements: Pavel Berloff provided the first stimulus to embark on this study. I thank him and Natalia Berloff, Oliver Bühler, Victor Kopiev, Hayder Salman, Mike Stone, Jim Thomas, Jacques Vanneste and Bill Young for their interest and for their very useful comments – some of them on substantial technical points – during the writing and revision of the paper. I should also like to thank the three referees, whose comments were extremely challenging and of great value in helping me to sharpen the presentation, in the course of two major revisions.

Appendix A. The Schrödinger equation and the phase function (1.2)

In the quantum problem originally studied by Aharonov & Bohm (1959), the wave field $\phi = \exp(\mathrm{i}\mathit{\Phi})$ with $\mathit{\Phi}$ defined by (1.2) is not only a far field but also an exact solution. For reasons that are obvious from figure 2, the quantum literature often calls it a "dislocated" wave field. In the quantum problem there is no restriction to small α. That is easily verified; the relevant Schrödinger equation can be written in suitable units as

$$\mathrm{i}\,\frac{\partial\phi}{\partial t} \;+\; \frac{c_0}{k_0}\big(\boldsymbol{\nabla} + \mathrm{i}\alpha r^{-1}\hat{\boldsymbol{\theta}}\big)^2\phi \;=\; 0\;, \tag{A 1}$$

where the square denotes a scalar product. When $\phi = \exp(\mathrm{i}\mathit{\Phi})$, with the error term deleted from (1.2), we have $\partial\phi/\partial t = -\mathrm{i}c_0k_0\phi$ and $\nabla\phi = (\mathrm{i}k_0\hat{\boldsymbol{x}} - \mathrm{i}\alpha r^{-1}\hat{\boldsymbol{\theta}})\phi$, satisfying (A 1) exactly. This wavefunction ϕ is part of a solution to (A 1) that describes nonrelativistic electrons going past an infinitely long, thin magnetic solenoid, whose total magnetic flux and magnetic vector potential $\propto r^{-1}\hat{\boldsymbol{\theta}}$ play the roles of $\mathit{\Gamma}$ and $\boldsymbol{u}_0$ in the vortex problem.

In the complete solution, originally derived by Aharonov & Bohm and generalized to

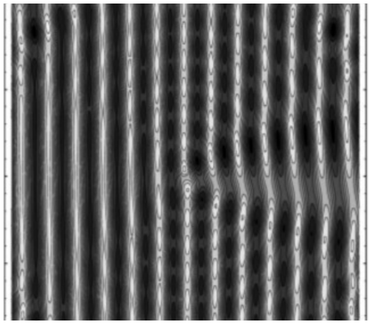

FIGURE 6. Numerical solution of the original Aharonov–Bohm problem (A 1), from Stone (2000*a*). The real part of ϕ is plotted. Here $\alpha = 0.25$, just large enough to make the phase change across the Fresnel wake easily visible. Also visible, very faintly, is a Born-scattered contribution recognizable by its approximately circular wavecrests. Reprinted, with permission, from figure 1 of Stone (2000*a*); copyright 2000 by the American Physical Society.

a solenoid of arbitrary diameter by Berry *et al.* (1980), there is in addition a Fresnel diffractive wake and a smaller, $O(r^{-1/2})$ contribution outside the wake region. The Fresnel wake is exactly centred on the positive x axis, for arbitrary α, and smooths out the discontinuity in Φ. In the thin-solenoid case the smaller, $O(r^{-1/2})$ contribution is describable as Born scattering off the solenoid, whose approximately circular wavecrests are faintly visible in figure 6, from a paper by Stone (2000*a*). The figure shows a numerical calculation of the thin-solenoid solution of (A 1) with $\alpha = 0.25$, large enough to make visible the phase change across the Fresnel wake.

All these features are qualitatively the same as those found by FLS in their analysis for $k_0 r_0 \ll 1$ of the linear wave field in the vortex problem. However, in stark contrast with the Schrödinger problem, the wave field becomes qualitatively different (e.g. Coste *et al.* 1999) as soon as α goes outside the very restricted range of values permitted by (1.3), when $\epsilon \ll 1$ as well as $k_0 r_0 \ll 1$.

Appendix B. Secular changes in problem (iii)

In deriving (2.18) in §8 we ignored a subtlety worth remarking on. The argument for taking $\tilde{q} = q_0$, even though correct within the quasigeostrophic framework, does not by itself exclude secular changes in $\tilde{q}$ over very long times in the *exact* dynamics of problem (iii). However, such changes can be excluded by appealing to the exact conservation of the Kelvin circulation around material loops of all sizes, shapes and orientations as fluid particles travel around the vortex, and in and out of the wavetrain. The $O(a^4)$ terms neglected in going from (8.3) to (8.5) describe only slight, reversible distortions, within the wave layer, of such material loops and of the absolute vortex lines threading them – the inertially-stiff lines of $\boldsymbol{f} + \widetilde{\boldsymbol{\omega}}$. Equation (8.7) is only an approximate expression of the exact statement that the Kelvin circulation is constant for all material loops, including those whose parts outside the wavetrain lie in horizontal planes, at all altitudes z. The circulation of such loops cannot change secularly unless the qualitative geometry of the picture changes, such that the loops and the absolute vortex lines deform irreversibly. Physically, this would correspond to the presence of large-amplitude *breaking* waves, in this case breaking surface gravity waves, or breaking inertia waves, or both. As shown by GLM theory, the irreversible deformation of otherwise wavy material contours can usefully be taken as the defining property of wave breaking (McIntyre & Palmer 1985). Our thought-experiments assume that no such wave breaking occurs.

Appendix C. Asymptotic validity of equation (2.18)

As well as using $\mathrm{Ro} \ll 1$ in going from (8.3) to (8.4), the derivation of (2.18) used overall hydrostatic balance to determine horizontal pressure gradients on the bottom as $\rho g \boldsymbol{\nabla}_{\mathrm{H}} \tilde{h}$, together with geostrophic balance to give (2.15) with $\tilde{\psi} = g(\tilde{h} - H)/f$ beneath the wavetrain and elsewhere. The Taylor–Proudman effect extends this picture upward into the wavetrain via the stiffness of the vortex lines of $\widetilde{\boldsymbol{\omega}} + \boldsymbol{f}$, which bend away from the vertical only slightly, through small angles $O(\mathrm{Ro})$.

Before proceeding to the asymptotic justification of (2.15) and (2.18), we note that overall hydrostatic balance does actually fail, along with the impulse–pseudomomentum theorem and the pseudomomentum rule, in the somewhat similar problems studied in Thomas *et al.* (2018, & refs.). Those problems assume rotating shallow water dynamics for the wavemotion as well as for the mean flow. The failure is due to the confinement of the wavetrain by the lower boundary. Other cases of confinement by boundaries and consequent pseudomomentum-rule failure include the classic case of one-dimensional sound waves in a rigid tube, with a wavemaker at one end and an absorber at the other (e.g. Brillouin 1936; McIntyre 1981, B14 §12.2.2). In problems like that of Thomas *et al.*, the lower boundary gives the radiation-stress field a foothold – a bottom boundary to react against – allowing the stress divergence to push or pull vertically on the complete fluid column and to disrupt overall hydrostatic balance so as to change the $O(a^2)$ pressure gradients on the bottom. This in turn produces additional terms on the right of equations like (2.18) governing potential-vorticity inversion, breaking the impulse–pseudomomentum theorem by breaking the connection between $\tilde{\psi}$ and $g(\tilde{h} - H)/f$. Remote recoil is still generic, however. Here we are using the term "radiation stress" in the slightly loose sense of any wave-induced momentum flux that arises from averaging the equations of motion in some way, rather than in the stricter sense adopted in Brillouin's writings and for instance in Longuet-Higgins & Stewart (1964), in B14 §10.5, and in Andrews & McIntyre (1978, hereafter AM78, §8.4), to mean the sole effect of the waves on the mean flow – which is definable in some but not all wave–mean interaction problems.

In problem (iii), the impulse–pseudomomentum theorem does hold and with it the pseudomomentum rule – as was independently confirmed in §8 – essentially because the foothold effect is too small to disrupt overall hydrostatic balance, thanks to sufficient separation between the lower boundary and the wavetrain such that $\exp(-kH)$ can be neglected. To verify this in detail and to check for other possible errors it is simplest, again, to work within the GLM framework, thereby avoiding the complications that come from the intersection of the free surface with the horizontal Eulerian coordinate surface $z = 0$, which we take as the undisturbed free surface. We would like to demonstrate asymptotic validity not only for problem (iii), but also for the wider variety of wave–vortex configurations covered by the impulse–pseudomomentum theorem in §3. For scale-analytic purposes we use k and kc to denote a typical wavenumber and frequency, whose orders of magnitude are unaffected by weak refraction.

Clearly a, ϵ, Ro, f/kc and $\exp(-kH)$ must all be treated as small parameters, the last two in order to use deep-water wave dynamics with Coriolis effects neglected and to guarantee negligibility of the foothold effect. We would like to let all five parameters tend toward zero, keeping $a \ll \epsilon$, for a given geometry of the vortex core, or cores, and the incident wave field.

For simplicity's sake we restrict attention to cases in which

$$a \ll \mathrm{Ro} \sim f/kc \sim \epsilon \ll 1 \tag{C 1}$$

in the limit. It will prove expedient, however, to allow $\exp(-kH) \lesssim \epsilon$. As in (8.1) we take

$$\mathrm{Ro} = U/fr_0 \, . \tag{C 2}$$

Given geometry means that horizontal scales such as W and r_0 will be held fixed in the limit. We therefore need not distinguish among those scales, and will take r_0 as their representative. It is convenient also to fix f and c, and to take U toward zero like ϵ. Then (C 1) implies that we must take k toward infinity like ϵ^{-1}. The meaning of "given incident wave field" will therefore have to be relaxed to mean a given amplitude distribution while $k \to \infty$, consistent with ray theory. We must also take gravity g toward infinity like ϵ^{-1}, because $c^2 = g/k$. Restated in a dimensionally consistent way, these conditions can be summarized as

$$U \sim c\epsilon \sim f r_0 \epsilon, \quad k \sim r_0^{-1}\epsilon^{-1}, \quad g \sim c^2 k \sim c^2 r_0^{-1}\epsilon^{-1} \tag{C 3}$$

as $\epsilon \to 0$. The assumption $\exp(-kH) \lesssim \epsilon$ implies that $kH \gtrsim |\ln \epsilon|$ and hence that $H \gtrsim r_0\,\epsilon|\ln\epsilon|$ and $L_{\mathrm{D}} = (gH)^{1/2}/f \gtrsim r_0|\ln\epsilon|^{1/2}$, which allows enough flexibility to accommodate our illustrative results (8.9)–(8.13) alongside the more general wave–vortex configurations considered in §3. It is convenient also to assume that $H \lesssim r_0$, though this is hardly a significant restriction since $H \sim r_0$ would correspond to $L_{\mathrm{D}} \sim r_0\,\epsilon^{-1/2}$, greatly exceeding any other horizontal scale.

We assume that the pressure on the free surface is constant, with or without disturbances, and take the constant to be zero without loss of generality. Within the wavetrain there is a three-dimensional $O(a^2\epsilon^0)$ radiation stress or wave-induced momentum flux $\mathit{\Pi}_{ij}$, say, which dies off exponentially with depth like $\exp(2kz)$, for deep-water waves with vertical structure $\exp(kz)$, as well as vanishing at the free surface $z = 0$. The Cartesian-tensor indices i, j now take values $(1, 2, 3)$, corresponding to (x, y, z). The sign convention will be such that the force per unit volume felt by the mean flow is $-\mathit{\Pi}_{ij,j}$. The most convenient formula for $\mathit{\Pi}_{ij}$, which is an $O(a^2)$ wave property, is

$$\mathit{\Pi}_{ij} \;=\; -\,\overline{p}^{\mathrm{L}}\big\{\tfrac{1}{2}\overline{(\xi_l\xi_m)}_{,lm}\delta_{ij} - \overline{(\xi_{l,i}\xi_j)}_{,l}\big\} - \overline{p^\ell\xi_{j,i}} \tag{C 4}$$

where $\overline{p}^{\mathrm{L}}$ is the Lagrangian-mean pressure and p^ℓ the $O(a)$ Lagrangian disturbance pressure, while $\boldsymbol{\xi}$ is the $O(a)$ disturbance particle-displacement field, with Cartesian components $\boldsymbol{\xi} = (\xi_1,\, \xi_2,\, \xi_3)$ and zero divergence $\xi_{l,l} = 0$ correct to $O(a)$. The formula (C 4) can be read off from AM78 (8.6), (8.10) and (9.3), or from B14 (10.43), (10.57), (10.73), (10.77) and (10.84).†

We use a ray-theoretic description of the waves, relative to suitably-oriented horizontal axes. The x or x_1 axis is chosen parallel to the local wavenumber, whose magnitude is asymptotically large like ϵ^{-1} according to (C 3). Zooming in to the local plane-wave structure, we have

$$(\xi_1,\, \xi_2,\, \xi_3) \;=\; b\exp(kz)\{\cos\mathit{\Phi} + O(\epsilon),\; O(\epsilon),\; \sin\mathit{\Phi} + O(\epsilon)\} \tag{C 5}$$

where $\mathit{\Phi} = k(x - ct) + \text{const.}$, with k, c and the displacement amplitude b all locally constant. We take $a = bk$, so that $a \ll \epsilon$ is the dimensionless wave slope. The relative errors $O(\epsilon)$ include weak-refractive effects as well as a small transverse displacement $\xi_2 = O(\epsilon)$ whose magnitude arises from our assumption in (C 1) that $f/kc \sim \epsilon$, in agreement with Pollard's exact solution, which incidentally has p^ℓ exactly zero. However,

† In AM78, $\mathit{\Pi}_{ij}$ is denoted by $-R_{ij}$, and in B14 by $\tilde{\mathit{\Pi}}_{ij} - \overline{p}^{\mathrm{L}}\delta_{ij}$. When using AM78 (8.10) we can neglect the divergence of $\boldsymbol{\xi}$ as well as an $O(a^2)$ term k_{ij}, before substituting into (8.6) and discarding terms $\propto a^3$ or higher. In B14, (10.73) is rewritten as $K_{km} = J\delta_{km} - \xi_{i,k}K_{im}$ before substituting it into (10.84) in the same way. Then use is made of (10.43), (10.57), and (10.77). The equation numbers in B14 correspond to (10.43), (10.57), (10.71), (10.75) and (10.82) in the original, 2009 edition. Though not needed here, it may be of interest to note that substitution of the leading order deep-water plane wave structure (which has $p^\ell = 0$) into the horizontal components of (C 4) leads to the standard $O(a^2)$ result $\int \mathsf{p}\boldsymbol{C}\,dz$ for the depth-integrated horizontal momentum flux (e.g. (24), (33) of Longuet-Higgins & Stewart 1964).

to allow for weak refraction we will use a more conservative estimate $p^\ell \lesssim \epsilon\rho g b \exp(kz)$, which is $O(\epsilon)$ times the Eulerian disturbance pressure.

The overbars in (C 4) are to be read as Eulerian phase averages over the local wave structure, in the standard way. Notice that if $\overline{p}^{\mathrm{L}}$ were constant and p^ℓ zero then the divergence $\varPi_{ij,j}$ would vanish. Therefore an additive constant in the pressure has no effect on the dynamics, confirming that, without loss of generality, we may take $\varPi_{ij} = 0$ at the free surface. Neglecting $O(a^2)$ contributions to $\overline{p}^{\mathrm{L}}$, we can replace it by $-\rho g z$ so that correct to $O(a^2)$

$$\varPi_{ij} \;=\; \rho g z\left\{\tfrac{1}{2}\overline{(\xi_l\xi_m)}_{,\,lm}\delta_{ij} - \overline{(\xi_{l,i}\xi_j)}_{,\,l}\right\} - \overline{p^\ell\xi_{j,i}}\;. \tag{C 6}$$

The resultant vertical force on a complete fluid column per unit horizontal area is

$$\begin{aligned}\int_{-H}^{0}\varPi_{3j,j}\,dz \;&=\; \int_{-H}^{0}\left\{[\tfrac{1}{2}\rho g z\overline{(\xi_l\xi_m)}_{,\,lm}]_{,\,3} - [\rho g z\overline{(\xi_{l,3}\xi_j)}_{,\,l}]_{,\,j} - [\overline{p^\ell\xi_{j,3}}]_{,\,j}\right\}dz\\ &=\; \left\{\tfrac{1}{2}\rho g H\overline{(\xi_l\xi_m)}_{,\,lm} - \rho g H\overline{(\xi_{l,3}\xi_3)}_{,\,l} - \overline{p^\ell\xi_{3,3}}\right\}\Big|_{z=-H}\\ &\qquad -\int_{-H}^{0}\left\{\rho g z\overline{(\xi_{l,3}\xi_\gamma)}_{,\,l\gamma} + [\overline{p^\ell\xi_{\gamma,3}}]_{,\,\gamma}\right\}dz\end{aligned} \tag{C 7}$$

where the greek index γ runs from 1 to 2, but j, l and m still from 1 to 3. This expression is more convenient than the alternative expression obtainable by applying derivatives to the factor $\rho g z$ only, in the first line, giving a result that looks simpler but obscures the foothold effect, the expression in the second line.

Vertical derivatives $\partial/\partial z = \partial/\partial x_3$ have order of magnitude $\sim k \sim \epsilon^{-1}$, and horizontal derivatives order unity or less, $\lesssim \epsilon^0$, as $\epsilon \to 0$, because horizontal scales such as r_0 and W are being held fixed, while $L_{\mathrm{D}} \gtrsim r_0|\ln\epsilon|^{1/2}$. In the last term of the foothold contribution on the second line of (C 7) we use our conservative estimate $p^\ell \lesssim \epsilon\rho g b\exp(kz)$, and the assumptions $k \sim r_0^{-1}\epsilon^{-1}$ and $\exp(-kH) \lesssim \epsilon$ made in (C 1)–(C 3), to show that the term in question has magnitude $\lesssim \epsilon\rho g b^2 k\exp(-2kH) \lesssim \epsilon^2\rho g b^2/r_0$. The first two terms combine to give a larger estimated magnitude $\lesssim \epsilon\rho g b^2/r_0$, as shown next.

In the first two terms we note that the largest, vertical-derivative contributions $\frac{1}{2}\rho g H\overline{(\xi_3\xi_3)}_{,\,33}$ and $-\rho g H\overline{(\xi_{3,3}\xi_3)}_{,\,3}$ cancel each other to leading order. This is because of the special structure of deep-water waves and would not be the case in, for instance, the problems studied by Thomas *et al.* For the local plane-wave structure we have $\overline{\sin^2\varPhi} = \overline{\cos^2\varPhi} = \frac{1}{2}$, hence $\frac{1}{2}\overline{(\xi_3\xi_3)} = \frac{1}{4}b^2\exp(2kz)$, with relative error $O(\epsilon)$. The vertical second derivative $\frac{1}{2}\overline{(\xi_3\xi_3)}_{,\,33} = \frac{1}{4}b^2 4k^2\exp(2kz) = b^2k^2\exp(2kz) = \overline{(\xi_{3,3}\xi_3)}_{,\,3}$. Therefore the sum of the first two terms has its order of magnitude reduced by a factor ϵ or less, and can be estimated as $\lesssim \epsilon\rho g b^2k^2H\exp(-2kH)$. Using our assumptions $k \sim r_0^{-1}\epsilon^{-1}$, $\exp(-kH) \lesssim \epsilon$, and $H \lesssim r_0$, we have $\epsilon\rho g b^2k^2H\exp(-2kH) \lesssim \epsilon\rho g b^2/r_0$ as asserted. Thus the entire foothold contribution, the second line of (C 7), can be estimated as $\sim \epsilon\rho g b^2/r_0$ at most.

In the vertically integrated, non-foothold contribution in the third line of (C 7), each term has magnitude $\lesssim \epsilon\rho g b^2/r_0$ also. To check this for the term in p^ℓ, we may take $\int ...\,dz \sim k^{-1}$, and as before use $p^\ell \lesssim \epsilon\rho g b\exp(kz) \sim \epsilon\rho g b$ in the integrand, so that $[\overline{p^\ell\xi_{\gamma,3}}]_{,\,\gamma} \lesssim \epsilon\rho g b^2k/r_0$, the factors k and r_0^{-1} coming from the vertical and horizontal derivatives respectively. Integration removes the factor k, leaving a contribution $\lesssim \epsilon\rho g b^2/r_0$.

In the other term, the first term on the third line, we have $\overline{(\xi_{l,3}\xi_\gamma)}_{,\,l\gamma} \lesssim \epsilon b^2k^2/r_0$, with a factor k^2 since among the three derivatives at most two are vertical, as happens in the contribution with $l = 3$. The factor ϵ comes from the $O(\epsilon)$ relative magnitude of ξ_γ when $\gamma = 2$ or, when $\gamma = 1$, from the phase difference between $\xi_{3,3}$ and ξ_1, which is $\pi/2 + O(\epsilon)$.

So averaging their product produces a factor ϵ. With $\int ...\, dz \sim k^{-1}$, and $\rho g z \sim \rho g k^{-1}$, this term and therefore the whole third line $\lesssim \epsilon \rho g b^2/r_0$ as asserted.

In summary, then, the resultant vertical force (C 7) per unit horizontal area $\lesssim \epsilon \rho g b^2/r_0$. This is the greatest amount by which the pressure on the bottom boundary can depart from its hydrostatic value $\rho g \tilde{h}$. Let $\delta\tilde{\psi}$ be the corresponding error in $\tilde{\psi} = g(\tilde{h} - H)/f$; then $\delta\tilde{\psi} \lesssim \epsilon g b^2/(f r_0)$. In the operator $(\nabla_{\mathrm{H}}^2 - L_{\mathrm{D}}^{-2})$ on the left-hand side of (2.18) the relevant horizontal scales are either fixed $\sim r_0$, in the limit $\epsilon \to 0$, or expand slightly because $L_{\mathrm{D}} \gtrsim r_0 |\ln \epsilon|^{1/2}$. So the error on the left-hand side of (2.18) is no greater than $\delta\tilde{\psi}$ divided by r_0^2 as $\epsilon \to 0$; so the error $\lesssim \epsilon g b^2/(f r_0^3)$. To neglect this error, we need to show that it is small in comparison with $\hat{\boldsymbol{z}}\cdot\boldsymbol{\nabla}\times\langle\mathbf{p}\rangle$ on the right-hand side of (2.18). Estimating $\hat{\boldsymbol{z}}\cdot\boldsymbol{\nabla}\times\langle\mathbf{p}\rangle$ as $\sim \langle\mathbf{p}\rangle_{\mathrm{typ}}/W \sim \langle\mathbf{p}\rangle_{\mathrm{typ}}/r_0$, where $\langle\mathbf{p}\rangle_{\mathrm{typ}}$ is a typical magnitude of $\langle\mathbf{p}\rangle$, we therefore need to show that

$$\epsilon g b^2/(f r_0^3) \ll \langle\mathbf{p}\rangle_{\mathrm{typ}}/r_0 \tag{C 8}$$

as $\epsilon \to 0$. Now $\langle\mathbf{p}\rangle_{\mathrm{typ}} \sim g b^2/(cH)$ since in ray theory $\langle\mathbf{p}\rangle$ is c^{-1} times the wave-energy per unit horizontal area, $\sim \rho g b^2$, divided by ρH. So in (C 8) the ratio of the left-hand side to the right-hand side is $\epsilon c H/(f r_0^2)$; and, recalling that $\epsilon c \sim U$ and that $\mathrm{Ro} \sim U/(f r_0)$, we see that $\epsilon c H/(f r_0^2) \sim \mathrm{Ro}\, H/r_0 \lesssim \mathrm{Ro} \sim \epsilon$. This estimate is sufficient for our purposes, but is very conservative because it relies again on the assumption $H \lesssim r_0$. If we restrict H more tightly, to its marginal order of magnitude $H \sim r_0\, \epsilon |\ln \epsilon|$, then (C 8) is satisfied more strongly, with ratio $\epsilon^2 |\ln \epsilon|$ instead of ϵ. Either way, (2.15) and (2.18) have now been validated, as required, as leading-order approximations on the basis of which Bretherton flows can be computed correct to $O(a^2\epsilon^0)$ and thence recoil forces correct to $O(a^2\epsilon^1)$.

Although the foregoing is sufficient for our purposes, the results can of course be checked directly from the vertical component of the GLM momentum equation, AM78 (8.7a) or B14 (10.82). In carrying out that check it needs to be remembered that the GLM divergence effect raises the Lagrangian-mean altitudes of the free surface and other isobaric material surfaces. To leading order, in the local plane wave, the surfaces are raised by $O(a^2)$ amounts $\frac{1}{2}\overline{(\xi_3^2)}_{,3}$, as is also necessary to account for the waves' potential energy $\frac{1}{2}\rho g \overline{\xi_3^2}|_{z=0}$ per unit area (McIntyre 1988). The raising of the free surface is accompanied by a compensating $O(a^2)$ reduction, $\frac{1}{2}\rho\overline{(\xi_3^2)}_{,33}$, in the mean density $\tilde{\rho}$ defined such that $\tilde{\rho}\, dx\, dy\, dz$ is the volumetric mass element, consistent with a negligible change in the total mass overlying a horizontal area element of the bottom boundary.

Appendix D. The $O(a^2\epsilon^1)$ pseudomomentum law

The two-dimensional pseudomomentum law (3.4) holds to the order of accuracy required in §3, namely correct to $O(a^2\epsilon^1)$ – the order of magnitude of the refraction term on the right-hand side – as a and ϵ tend toward zero with $a \ll \epsilon$ for a given geometry of the vortices and incident wave field, whose horizontal scales are held fixed in the limit as in Appendix C.

The most secure route to (3.4) is to start with its exact GLM counterparts, in all three problems, so that we can see precisely what is neglected. To save space we refer directly to B14's exact GLM equations (10.123)–(10.126), which are (10.122)–(10.125) in the original, 2009 edition. The second of these equations defines the exact nonadvective flux of pseudomomentum, the exact counterpart of $\mathsf{B}_{ij} - \mathsf{p}_i \overline{u}_j^{\mathrm{L}}$ in (3.4)–(3.5) above. We recall that the nonadvective flux can be rewritten exactly, within the GLM framework, as an isotropic term $\propto \delta_{ij}$ plus the wave-induced flux of *momentum*. This will be useful when considering problem (iii), in which the anisotropic part of the expression on the right of (C 4) will be made use of.

In the gas dynamical version of problems (i) and (ii) the motion is strictly two-dimensional. Equation (3.4) can be read off straightforwardly from its exact counterpart, $\tilde{\rho}$ times B14 (10.126) (see also (10.47)), with indices i, j etc. running from 1 to 2. The two-dimensional mean density $\tilde{\rho}$ is the same as our $\tilde{h}$ and can be approximated as a constant, in its product with the refraction term, the third term on the right. The fractional error involved is small, $O(a^2\epsilon^0) + O(a^0\epsilon^2)$, corresponding to absolute error $O(a^4\epsilon^1) + O(a^2\epsilon^3)$, the first term coming from the hard-spring contribution to the Brillouin radiation stress noted at the end of §2, and the second from the Bernoulli pressure drop surrounding a vortex core. In the last term on the right of B14 (10.126), the gradient $\tilde{\rho}_{,i} = \tilde{h}_{,i}$ is similarly small, $O(a^2\epsilon^0) + O(a^0\epsilon^2)$, but is multiplied by the expression in large curly brackets, which is $O(a^2\epsilon^0)$ rather than $O(a^2\epsilon^1)$. The product $O(a^4\epsilon^0) + O(a^2\epsilon^2)$ is, however, still negligible against $a^2\epsilon^1$, the magnitude of the refraction term. The first term on the right corresponds to $\boldsymbol{\mathcal{F}}$ in (3.4), with the irrotational forcing potential ϕ corresponding to $-\chi'$ in (2.1). The second term on the right is zero, there being no rotational forcing or dissipation. The flux tensor B_{ij} in our (3.4)–(3.5) is given by B14's (10.125) plus the $O(a^2\epsilon^1)$ advective flux $\tilde{\rho}\,\mathsf{p}_i\overline{u}_j^{\mathrm{L}}$, in which $\tilde{\rho}$ can again be approximated as a constant, with the same relative and absolute errors as in the refraction term. Thus (3.4) is established correct to $O(a^2\epsilon^1)$.

In the shallow water version of problems (i) and (ii), the governing equations are the same as in the gas dynamical version with the ratio of specific heats set to 2, and no more need be said.

For problem (iii), we need the vertical average of $\tilde{\rho}$ times the horizontal projection of B14 (10.126) or, more conveniently, of (10.123), with zero right-hand side because there is no rotational forcing or dissipation. In the horizontal projection, the free suffix i takes values $i = 1, 2$, while the dummy suffixes j, k and m run from 1 to 3. The last term on the left of (10.123) corresponds to $\boldsymbol{\mathcal{F}}$, while the second-last term, an elastic-energy term, is zero because the flow is three-dimensionally incompressible.

In the third-last term on the left, the density ρ is constant and the factor $\overline{(p/\rho)}^{\mathrm{L}}$ can be taken as $-gz$ with error $O(a^2\epsilon^0)$, while the factor $\tilde{\rho}_{,i}/\tilde{\rho} = O(a^2\epsilon^0)$, with $\tilde{\rho}$ now the three-dimensional GLM mean density, which contains an $O(a^2\epsilon^0)$ contribution from the GLM divergence effect recalled at the end of Appendix C. This contribution is significant in the third-last term only. Everywhere else it represents a negligible fractional error in $\tilde{\rho}$. The third-last term can be simplified to ρ^{-1} times $-gz\tilde{\rho}_{,i} + O(a^4\epsilon^0)$, whose vertical average is the horizontal gradient of $-\langle gz(\tilde{\rho} - \rho)\rangle$, with error $O(a^4\epsilon^0)$. This gradient can be incorporated without further error into the flux divergence $\boldsymbol{\nabla}_{\mathrm{H}} \cdot \mathbf{B}$ of our (3.4) (in which vertical averaging is understood), via a horizontally isotropic contribution $-\langle gz(\tilde{\rho} - \rho)\rangle\delta_{ij}$ to the averaged flux itself, B_{ij}.

The second term on the left of B14 (10.123), a wave kinetic energy term, can be treated in the same way, giving another isotropic contribution to the vertically-averaged flux B_{ij} in (3.4). In the advection term $\overline{\boldsymbol{u}}^{\mathrm{L}}\cdot\boldsymbol{\nabla}\mathsf{p}_i = \boldsymbol{\nabla}\cdot(\boldsymbol{u}_0\mathsf{p}_i) + O(a^4\epsilon^0)$, the z-independent factor $\boldsymbol{u}_0$ can be taken outside the vertical average, as can also be done in the refraction term $\overline{u}^{\mathrm{L}}_{k,i}\mathsf{p}_k = u_{0k,\,i}\mathsf{p}_k + O(a^4\epsilon^0)$. Finally, we note that the $\partial/\partial z$ contribution to the three-dimensional flux divergence, $\tilde{\rho}$ times the fourth term on the left of (10.123), has vertical average zero because of our assumptions, spelt out in Appendix C, that $\exp(-2kH)$ is negligible and that the pressure vanishes or is constant at the free surface, so that the nonadvective 13 and 23 components of the three-dimensional pseudomomentum flux defined in BM (10.124) vanish there. As already mentioned, these anisotropic components are equal to the corresponding components of the wave-induced flux of momentum, which correct to $O(a^2)$ are given by the anisotropic terms in (C 4) or (C 6).

The foregoing is enough to establish for problem (iii) that our (3.4), with vertical averaging understood, holds to the order of accuracy required in §3. However, it may be of interest to note that, to leading order under the scaling assumptions of Appendix C, the quantity $\langle gz(\tilde{\rho}-\rho)\rangle$ is H^{-1} times the potential energy of the deep-water waves per unit area, replacing the elastic energy in the gas dynamical system and, in the isotropic part of B_{ij}, cancelling the wave kinetic energy to leading order as expected from averaged-Lagrangian considerations. Using $\tilde{\rho} = \rho(1 - \frac{1}{2}\overline{(\xi_3^2)}_{,33})$, following on from the end of Appendix C, and continuing to neglect $\exp(-2kH)$ we have, using integration by parts,

$$H\langle gz(\tilde{\rho}-\rho)\rangle = -\tfrac{1}{2}\int_{-H}^{0} \rho g z\,\overline{(\xi_3^2)}_{,33}\,dz = \tfrac{1}{2}\int_{-H}^{0} \rho g\,\overline{(\xi_3^2)}_{,3}\,dz = \tfrac{1}{2}\rho g\,\overline{(\xi_3^2)}\big|_{z=0}\,, \qquad \text{(D 1)}$$

which is the standard formula for the surface-wave potential energy per unit area.

REFERENCES

Aharonov, Y. & Bohm, D. 1959 Significance of electromagnetic potentials in the quantum theory. *Phys. Rev.* **115**, 485–491.

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

Baldwin, M. P., Gray, L. J., Dunkerton, T. J., Hamilton, K., Haynes, P. H., Randel, W. J., Holton, J. R., Alexander, M. J., Hirota, I., Horinouchi, T., Jones, D. B. A., Kinnersley, J. S., Marquardt, C., Sato, K. & Takahashi, M. 2001 The quasi-biennial oscillation. *Revs. Geophys.* **39**, 179–229.

Batchelor, G. K. 1967 *An Introduction to Fluid Dynamics*. Cambridge University Press.

Belyaev, I. V. & Kopiev, V. F. 2008 On the statement of the problem of sound scattering by a cylindrical vortex. *Acoust. Phys.* **54**, 603–614.

Berloff, N. G. 2004 Padé approximations of solitary wave solutions of the Gross-Pitaevskii equation. *J. Phys.* **A 37**, 1617–1632.

Berry, M. V., Chambers, R. G., Large, M. D., Upstill, C. & Walmsley, J. C. 1980 Wavefront dislocations in the Aharonov–Bohm effect and its water wave analogue. *Eur. J. Phys.* **1**, 154–162.

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

Bretherton, F. P. & Garrett, C. J. R. 1968 Wavetrains in inhomogeneous moving media. *Proc. Roy. Soc. Lond.* **A302**, 529–554.

Brillouin, L. 1936 On radiation pressures and stresses (in French). *Revue d'Acoustique* **5**, 99–111.

Bühler, O. 2014 *Waves and Mean Flows,* 2nd edn. (paperback). Cambridge: University Press.

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

Bühler, O. & McIntyre, M. E. 2005 Wave capture and wave–vortex duality. *J. Fluid Mech.* **534**, 67–95.

Coste, C., Lund, F. & Umeki, M. 1999 Scattering of dislocated wave fronts by vertical vorticity and the Aharonov–Bohm effect. I. shallow water. *Phys. Rev. E* **60**, 4908–4916.

Craik, A. D. D. & Leibovich, S. 1976 A rational model for Langmuir circulations. *J. Fluid Mech.* **73**, 401–426.

Dritschel, D. G. & McIntyre, M. E. 2008 Multiple jets as PV staircases: the Phillips effect and the resilience of eddy-transport barriers. *J. Atmos. Sci.* **65**, 855–874.

Ford, R. & Llewellyn Smith, S. G. 1999 Scattering of acoustic waves by a vortex. *J. Fluid Mech.* **386**, 305–328.

Fritts, D. C. 1984 Gravity wave saturation in the middle atmosphere: a review of theory and observations. *Revs. Geophys. Space Phys.* **22**, 275–308.

Garcia, R. R., Smith, A. K., Kinnison, D. E., de la Cámara, Á. & Murphy, D. J. 2017 Modification of the gravity wave parameterization in the Whole Atmosphere Community Climate Model: motivation and results. *J. Atmos. Sci.* **74**, 275–291.

Guo, Y. & Bühler, O. 2014 Wave–vortex interactions in the nonlinear Schrödinger equation. *Phys. Fluids* **26**, 027105.

Haney, S. & Young, W. R. 2017 Radiation of internal waves from groups of surface gravity waves. *J. Fluid Mech.* **829**, 280–303.

Hasselmann, K. 1970 Wave driven inertial oscillations. *Geophys. Fluid Dyn.* **1**, 463–502.

Holton, J. R., Haynes, P. H., McIntyre, M. E., Douglass, A. R., Rood, R. B. & Pfister, L. 1995 Stratosphere–troposphere exchange. *Revs. Geophys.* **33**, 403–439.

Humbert, T., Aumaître, S. & Gallet, B. 2017 Wave-induced vortex recoil and nonlinear refraction. *Phys. Rev. Fluids* **2**, 094701, 1–14.

Kida, S. 1981 Motion of an elliptic vortex in a uniform shear flow. *J. Phys. Soc. Japan* **50**, 3517–3520.

Lane, E. M., Restrepo, J. M. & McWilliams, J. C. 2007 Wave-current interaction: a comparison of radiation-stress and vortex-force representations. *J. Phys. Oceanogr.* **37**, 1122–1141.

Leibovich, S. 1980 On wave–current interaction theories of Langmuir circulations. *J. Fluid Mech.* **99**, 715–724.

Lelong, M.-P. & Riley, J. J. 1991 Internal wave-vortical mode interactions in strongly stratified flows. *J. Fluid Mech.* **232**, 1–19.

Longuet-Higgins, M. S. & Stewart, R. W. 1964 Radiation stress in water waves; a physical discussion, with applications. *Deep-Sea Res.* **11**, 529–562.

McComas, C. H. & Bretherton, F. P. 1977 Resonant interaction of oceanic internal waves. *J. Geophys. Res.* **82**, 1397–1412.

McIntyre, M. E. 1973 Mean motions and impulse of a guided internal gravity wave packet. *J. Fluid Mech.* **60**, 801–811.

McIntyre, M. E. 1981 On the "wave momentum" myth. *J. Fluid Mech.* **106**, 331–347.

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

McIntyre, M. E. 2017 On multi-level thinking and scientific understanding. *Adv. Atmos. Sci.* **34**, 1150–1158.

McIntyre, M. E. & Palmer, T. N. 1985 A note on the general concept of wave breaking for Rossby and gravity waves. *Pure Appl. Geophys.* **123**, 964–975.

Pollard, R. T. 1970 Surface waves with rotation: an exact solution. *J. Geophys. Res.* **75**, 5895–5898.

Sakov, P. V. 1993 Sound scattering by a vortex filament. *Acoust. Phys.* **39**, 280–282.

Sonin, E. 1997 Magnus force in superfluids and superconductors. *Physical Review B* **55**, 485–501.

Stone, M. 2000*a* Iordanskii force and the gravitational Aharonov–Bohm effect for a moving vortex. *Phys. Rev. B* **61**, 11780–11786.

Stone, M. 2000*b* Acoustic energy and momentum in a moving medium. *Phys. Rev. E* **62**, 1341–1350.

Thomas, J. 2017 New model for acoustic waves propagating through a vortical flow. *J. Fluid Mech.* **823**, 658–674.

Thomas, J., Bühler, O. & Smith, K. S. 2018 Wave-induced mean flows in rotating shallow water with uniform potential vorticity. *J. Fluid Mech.* **839**, 408–429.

Thomas, J. & Yamada, R. 2019 Geophysical turbulence dominated by inertia–gravity waves. *J. Fluid Mech.* **875**, 71–100.

Ursell, F. 1950 On the theoretical form of ocean swell on a rotating earth. *Mon. Not. Roy. Astron. Soc. Geophys. Suppl.* **6**, 1–8.

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

Wallace, J. M. & Holton, J. R. 1968 A diagnostic numerical model of the quasi-biennial oscillation. *J. Atmos. Sci.* **25**, 280–292.

Ward, M.L. & Dewar, W. K. 2010 Scattering of gravity waves by potential vorticity in a shallow-water fluid. *J. Fluid Mech.* **663**, 478–506.

Wexler, C. & Thouless, D. J. 1998 Scattering of phonons by a vortex in a superfluid. *Physical Review B* **58**, R8897(R).