J. Fluid Mech. (2006), *vol.* 549, *pp.* 403–428.
doi:10.1017/S0022112005008335 Printed in the United Kingdom

On the instability leading to rip currents due to wave–current interaction

By JIE YU

School of Mechanical, Aerospace and Civil Engineering, The University of Manchester, Manchester M60 1QD, UK

(Received 8 December 2004 and in revised form 23 August 2005)

We examine the instability leading to depth-averaged circulations, which are related to rip currents in the surf zone, due to wave–current interactions. Our intention is to clarify some issues which are critical to the determination of instability properties, as yet unresolved from previous studies. Those issues are also of interest for hydrodynamic instability problems in general. Attention is restricted to normally incident waves and the region inside, and just offshore of, the surf zone where linearized shallow-water waves are applied. The coupling of waves and circulations is modelled using the concept of wave radiation stresses and the classical ray theory. The instability properties, in terms of the neutral modes, the most unstable mode and the corresponding maximum growth rate, are examined in the domain of parameters which represent the effects of offshore wave height and bottom friction dissipation. Comparisons with observations of natural rip currents are made, and qualitative agreements are achieved.

1. Introduction

This paper examines the linear instability in the formation of rip currents in the surf zone, examining some aspects, of the mathematical procedure of determining instability properties, that were not handled properly in previous studies. This instability concerns particularly situations where the incident wave field and the beach are both alongshore uniform, and depends on the interactions of waves and the currents induced by the waves.

When a wave climbs up a beach, the loss of wave momentum due to shoaling and breaking induces nearshore currents. Alongshore and rip currents, the two major systems whose flow structures are predominantly horizontal, are of this origin. The latter are the focus of this study. Rip currents are narrow seaward flows, and often observed when waves are normally incident or nearly so. They are parts of the depth-averaged cellular circulations which are characterized by narrow offshore directed flows (rips), broad onshore flows in between and intense alongshore flows 'feeding' the rips close to the shoreline. Rip currents usually display a relatively regular spacing alongshore, ranging from 50 to 1000 m (Short 1999).

The importance of alongshore variations in the wave field for the generation of rip currents has long been established in the literature (Shepard, Emery & LaFond 1941; Shepard & Inman 1951). With an imposed alongshore variation in the wave height, Bowen (1969) first gave a mathematical description of the cellular circulations using the concept of wave radiation stresses (Longuet-Higgins & Stewart 1960, 1961, 1962). He also confirmed that the nonlinear convective inertia is responsible for narrowing

the offshore flows and broadening the onshore returning flows, as was speculated by Arthur (1962). What remains in question is the cause of such alongshore variations, which probably arise from diverse processes in view of the wide range of the observed rip spacing. This is somewhat less ambiguous on alongshore variable beaches when the topographic effects on the wave field are considered. Indeed, rip currents are often associated with rhythmic beaches. However, the presence of such a topography does not explain whether it is a consequence of the currents or a cause. There are studies that investigate the development of cellular circulations and rhythmic topography by coupling the hydrodynamics of currents and sediment transport (e.g. Hino 1974; Damgaard *et al.* 2002). On the other hand, it has also been argued that the necessary alongshore variation in the wave field may arise as a result of an instability due to the interactions of waves and currents. From the wave hydrodynamic point of view, when a wavetrain of uniform crests encounters a current, refraction due to the current can cause the wave crests to bend, resulting in focusing and defocusing of the wave energy alongshore. In a study of rip currents driven by topography, Yu & Slinn (2003) showed that the alongshore variability of the wave field can be altered significantly even when circulations are weak. Therefore, a question to be put forward is: If an alongshore-uniform wave field is perturbed, can the subsequent development of circulations enhance the alongshore variability in the wave field?†

Early studies to explore this mechanism include LeBlond & Tang (1974), Iwata (1976), Mizuguchi (1976). Realizing the weaknesses in those previous analyses, Dalrymple & Lozano (1978, hereafter called DL78) attempted this approach again, with a focus on the neutrally stable (or steady-state) circulation cells of small amplitude on a planar beach. They found that equilibrium cells do not exist if the effects of currents on the wavenumber field are not taken into account. This conclusion was subsequently confirmed by Falqués, Montoto & Vila (1999) with a nice argument using a Lyapounov functional. Upon including the effects of currents on both the wavenumber and wave energy fields, DL78 did find linearized steady circulations. However, there are uncertainties about their results. First, though we have used the same basic set of equations, we cannot recover the solutions far offshore asserted in DL78. Second, the asymptotic solutions close to the shoreline were not determined properly in DL78. Some comments on these can be found in § 4.2. These solutions are needed to facilitate the shoreward and seaward integrations toward the breaking line where matching the physical variables leads to the determination of the normal modes. Third, DL78 have neglected the perturbation in the location of the breaking line. While this seems to have been recognized in their paper (see the discussion around their equation (A6)), the conclusion drawn there is only appropriate if all the variables of the basic state are zero. As will be seen in § 4.3, the perturbation of the breaking line location plays an important role in determining the proper matching conditions. Since DL78 considered only neutral modes, Falqués *et al.* (1999) made an attempt to investigate if there is indeed an instability. They considered a constant still-water depth in front of a vertical wall, as an attempt to avoid dealing with wave breaking and hence any matching at the breaking line. However, they did not consider any reflection from the vertical shore face, nor any boundary condition at the wall. It is then not clear how the normal modes of the linearized system were

† Murray & Reydellet (2001) studied rip currents on planar beaches based on a hypothesized wave–current interaction process which causes waves to be dissipated more in the presence of rip currents. That hypothesized wave–current process is outside the scope of this study. Some comments on that instability analysis can be found in § 4.1.

found. While they expressed some reservations about their model, they concluded that the instability could occur, and estimated a typical growth time about 0.06 s for a water depth of 1 m. This is unreasonably fast.

In view of the above discussion, and that the depth-averaged circulations and their interactions with waves are still of significant interest in nearshore studies, it seems useful to re-examine this instability problem by conducting a systematic mathematical analysis. For this purpose, we will follow, as closely as possible, the assumptions in DL78. Our particular attention is to carefully address the issues of the boundary condition at infinity, the appropriate unboundedness at the shoreline, and the matching conditions at the breaking line where the basic state is not smooth due to wave breaking. In all of these we find significant differences from DL78. Furthermore we do not restrict our attention to steady-state solutions, but investigate the conditions for existence and the properties of growing perturbations. Nevertheless, the extent of the applicability of this mechanism to the formation of rip currents on natural beaches remains open. There seem to have been no careful laboratory studies relevant to it, and the difficulties of observationally distinguishing this instability mechanism from other possibilities are considerable, to say nothing of the uncertainty in parameters like the breaking index and the bottom friction coefficient. Despite this, the issues addressed here are of interest for hydrodynamic instability problems in general, in particular for systems where differential equations are not consistent (in the number and the form) throughout the domain due to inherent discontinuities.

The outline of the paper is as follows. In order to be self-contained, equations describing the dynamics of the depth-averaged circulations and the average wave field are presented in the next section. The basic-state solution, which is well-known, is given in § 3 without great detail. Linear instability analysis is carried out in § 4 where the mathematical procedure to determine the normal modes is discussed at length. The properties of the instability are discussed in § 5, where comparisons with observations of natural rip currents are also attempted. Some concluding remarks are made in § 6.

2. Formulation

We consider monochromatic waves on alongshore-uniform beaches. At incidence, there is no variation along the wave crests. In the horizontal plane, x is pointing offshore and y is alongshore. In the absence of fluid motions, the shoreline is located at $x = 0$ and the water depth $h(x)$ describes the beach topography. As is customary in nearshore studies, circulations are described using the two-dimensional shallow-water equations, which are derived by averaging the continuity and Navier–Stokes equations in time with respect to the wave period and in depth (Mei 1989). Let $\boldsymbol{u} = (u, v)$ be the across-shore and alongshore velocities of the current and η the mean surface elevation (pressure field). We then write

$$\eta_t + \nabla \cdot (d\boldsymbol{u}) = 0, \tag{2.1}$$

$$\boldsymbol{u}_t + \boldsymbol{u} \cdot \nabla \boldsymbol{u} = -g\nabla\eta + \boldsymbol{\tau}_w - \boldsymbol{\tau}_b, \tag{2.2}$$

where $d = h(x) + \eta$ is the total water depth, g the acceleration due to gravity, $\boldsymbol{\tau}_w$ the forcing due to waves and $\boldsymbol{\tau}_b$ the bottom friction arising from the depth integration. The horizontal mixing terms, representing the effects due to small-scale turbulence and/or vertical fluctuations in the current velocities, are neglected in this linear instability analysis, following LeBlond & Tang (1974), DL78 and Falqués *et al.* (1999). The forcing term $\boldsymbol{\tau}_w$ represents the transfer of wave momentum to the currents. It is

commonly modelled as the convergence of the wave radiation stress tensor after Longuet-Higgins (1970), i.e.

$$\boldsymbol{\tau}_w = -(\rho d)^{-1}\nabla \cdot \boldsymbol{S}, \tag{2.3}$$

where ρ is the water density and

$$S_{xx} = E\left(\tfrac{1}{2} + \frac{k^2}{k^2+l^2}\right), \quad S_{xy} = S_{yx} = E\frac{kl}{k^2+l^2}, \quad S_{yy} = E\left(\tfrac{1}{2} + \frac{l^2}{k^2+l^2}\right) \tag{2.4}$$

are the elements of $\boldsymbol{S}$ for linear shallow-water waves. $E = \frac{1}{8}\rho g H^2$ is the wave energy, where H is the wave height, and $\boldsymbol{k} = (k, l)$ is the wavenumber vector. The parameterization of the bottom friction is a direct adaptation of studies in open channel hydraulics. For weak currents, it can be approximated as

$$\boldsymbol{\tau}_b = d^{-1}\mu\boldsymbol{u}, \tag{2.5}$$

where $\mu = (2/\pi)c_f u_w$ is the friction velocity defined using the amplitude of the wave orbital velocity, and c_f is the empirical friction coefficient. For linear shallow-water waves, $u_w = \frac{1}{2}Hd^{-1}\sqrt{gd}$.

We assume that the topography and current both vary slowly such that the modulations of the averaged wave field can be described using the ray theory. The conservation of wave crests requires

$$\boldsymbol{k}_t + \nabla\omega = 0, \tag{2.6}$$

where the wavenumber $\boldsymbol{k}$ and wave angular frequency ω satisfy the dispersion relationship which, for linear shallow-water waves, is

$$(\omega - \boldsymbol{u}\cdot\boldsymbol{k})^2 = gd(\boldsymbol{k}\cdot\boldsymbol{k}). \tag{2.7}$$

By definition, $\nabla \times \boldsymbol{k} = 0$. Before the wave breaks, the wave energy is described by

$$E_t + \nabla\cdot[(\boldsymbol{C_g} + \boldsymbol{u})E] + S_{xx}u_x + \tfrac{1}{2}S_{xy}(u_y + v_x) + S_{yy}v_y = 0, \tag{2.8}$$

where $\boldsymbol{C_g} = \sqrt{gd}(\boldsymbol{k}\cdot\boldsymbol{k})^{-1/2}\boldsymbol{k}$ is the group velocity vector. After breaking, a saturated surf zone is assumed, i.e.

$$H = \gamma d, \tag{2.9}$$

where γ is the empirical breaking index. This means that the wave height inside the surf zone is controlled by the local water depth, rather than by the offshore wave conditions. The value of γ cited in the literature can vary from 0.3 to 1.2, accounting for both monochromatic and random waves (Lippmann, Brookins & Thornton 1996). The wave considered in this study may be regarded as the ensemble average of a random wave. Note from (2.8) that the wave energy is not conserved, even before the breaking, due to the terms for wave radiation stresses. These terms, unlike the energy flux divergence, represent the work done by radiation stresses acting on the strain of the velocity field of the current. Thus, they act as spatially distributed energy sources/sinks. Equations (2.1)–(2.9) are the same as in DL78.

The depth-averaged circulation is not divergence free, see (2.1). An external forcing on a water column can be balanced by a change of the total water depth of the column without necessarily setting up a net, vertically averaged, transport. The dynamics of the circulation is better revealed by the vorticity balance. Let $q = (v_x - u_y)d^{-1}$ be the potential vorticity in the vertical direction. Then from (2.1) and (2.2)

$$q_t + \boldsymbol{u}\cdot\nabla q = d^{-1}[\nabla\times\boldsymbol{\tau}_w - \nabla\times\boldsymbol{\tau}_b]. \tag{2.10}$$

Clearly, vorticity is generated due to $\nabla\times\boldsymbol{\tau}_w$, the curl of the wave forcing, and dissipated by the bottom friction. For normally incident waves in the absence of alongshore variations, the wave forcing is non-zero and also in the across-shore direction due to wave breaking and refraction by the topography, i.e. $\partial S_{xx}/\partial x \neq 0$. However, it has no curl because of the alongshore uniformity, and hence does not contribute to the vorticity generation. This is the scenario of the basic state where $S_{xx,x}$ is balanced by an across-shore pressure gradient and no flow is induced, see § 3. An inhomogeneity alongshore, for instance due to a disturbance in the wave field, will lead to $\nabla\times\boldsymbol{\tau}_w \neq 0$ and the initiation of circulations. The subsequent evolution of the circulation then depends on the balance between the source and dissipation of the vorticity.

Let us consider a topography which consists of two sections: a beach with constant slope β and a horizontal flat bed extending to infinity, i.e.

$$h(x) = \beta x \quad \text{for} \quad x \leqslant x_F, \quad h = h_F \quad \text{for} \quad x \geqslant x_F. \tag{2.11}$$

The incoming waves are normally incident from $x > x_F$, with the wave height H_F and angular frequency ω_F in water of depth h_F over the horizontal bed. With the assumption of shallow-water waves, attention is restricted to waves near and inside the surf zone. The choice of the topography implies that the waves over the horizontal bed have undergone considerable shoaling and are nearly breaking. In other words, H_F and h_F should have values close to the breaker height and breaking depth, respectively. This is merely for theoretical expediency, since nearshore circulations are primarily due to wave breaking. Denoting dimensionless variables by primes, we choose the following scaling for waves and currents:

$$\left.\begin{aligned} &(x, y) = (x', y')x_F, \quad d = d'h_F, \quad \eta = \eta' h_F, \quad \boldsymbol{u} = \boldsymbol{u}'\sqrt{gh_F}, \quad t = t'\frac{x_F}{\sqrt{gh_F}}, \\ &H = H'H_F, \quad \omega = \omega'\omega_F, \quad \boldsymbol{k} = \boldsymbol{k}'\frac{\omega_F}{\sqrt{gh_F}}, \quad \boldsymbol{C_g} = \boldsymbol{C_g}'\sqrt{gh_F}. \end{aligned}\right\} \tag{2.12}$$

The dimensionless equations are readily obtained from (2.1)–(2.9), and are omitted here. Two control parameters are identified. The first,

$$R_0 = \frac{H_F}{h_F}, \tag{2.13}$$

represents the wave condition close to the breaking point. It provides a measure of the relative strength of the incident waves. The second,

$$R_1 = \frac{\pi}{16}\frac{\beta}{c_f}R_0, \tag{2.14}$$

indicates the relative importance of the bottom friction. According to the scaling in (2.12),

$$\frac{\nabla\times\boldsymbol{\tau}_w}{\nabla\times\boldsymbol{\tau}_b} \sim \frac{\frac{1}{8}gH_F^2h_F^{-1}x_F^{-1}x_F^{-1}}{(2/\pi)c_fH_Fh_F^{-1}\sqrt{gh_F}\sqrt{gh_F}h_F^{-1}x_F^{-1}} = \frac{\pi}{16}\frac{\beta}{c_f}R_0. \tag{2.15}$$

Thus, R_1 compares the rate of vorticity generation by $\nabla\times\boldsymbol{\tau}_w$ to the rate of its dissipation by bottom friction. The beach slope β appears because the scales for the depth and for the horizontal lengths are different. As will be seen in § 5, β is only a geometric parameter relating the scales rather than a dynamical one controlling the instabilities. Hereafter, we shall work only with dimensionless variables, and drop the primes for simplicity.

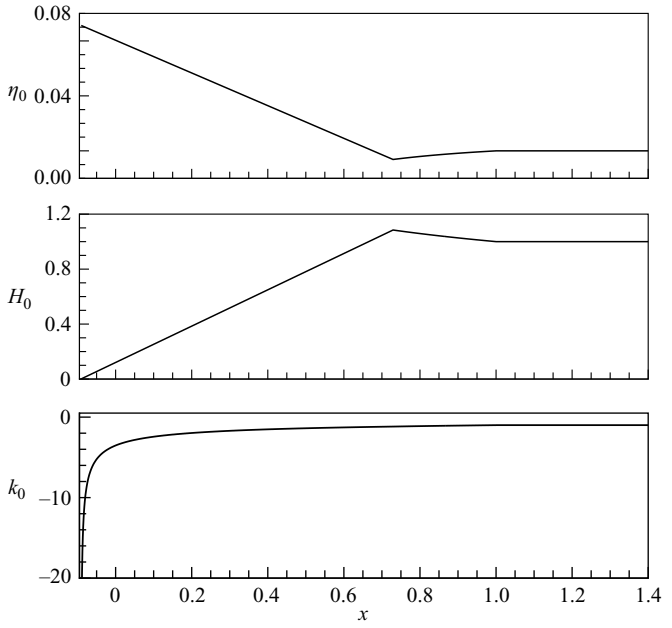

FIGURE 1. The across-shore profile of the basic state. $R_0 = 0.4$ and $\gamma = 0.6$.

3. Basic state

Let us denote the basic state by the subscript 0. For normally incident waves in the absence of any alongshore variations, the basic state is independent of y and t, and hence is motionless, i.e. $u_0 = v_0 = 0$. Thus, $\omega_0 \equiv 1$, $l_0 \equiv 0$, $S_{0xy} \equiv 0$ and $\partial S_{0yy}/\partial x \equiv 0$. The momentum balance reduces to that between the across-shore wave forcing $\partial S_{0xx}/\partial x$ and the gradient of the surface elevation η_0 (set-down/up). The equations for η_0, k_0 and H_0 are readily obtained from (2.2), (2.7), (2.8) and (2.9). To integrate these equations, we require $\eta_0 \to 0$ as $x \to \infty$, and the solution to be continuous at the breaking line. The results are presented below. Details can be found in Mei (1989) and other publications.

For $x > 1$, $\eta_0 = 0$, $H_0 = 1$, $k_0 = -1$, and so $d_0 = 1$. Inside $x = 1$, it is convenient to express the solution in terms of the total water depth. For $x_{b0} \leqslant x \leqslant 1$, $H_0 = d_0^{-1/4}$ and d_0 is implicitly given by

$$d_0 = x + \left(R_0^2/16\right)\left(1 - d_0^{-3/2}\right), \tag{3.1}$$

where

$$x_{b0} = (1 + \gamma^2/16)(R_0/\gamma)^{4/5} - R_0^2/16 \tag{3.2}$$

is the location of the breaking line, measured from $x = 0$. For $x < x_{b0}$, $H_0 = \gamma R_0^{-1} d_0$ and

$$d_0 = \widetilde{\beta}(x - x_{s0}), \quad \text{where} \quad \widetilde{\beta} \equiv (1 + 3\gamma^2/8)^{-1}, \tag{3.3}$$

and

$$x_{s0} = -\tfrac{1}{16}\left[R_0^2 + 5\gamma^2(R_0/\gamma)^{4/5}\right] \tag{3.4}$$

is the location of the shoreline accounting for the set-up η_0 and measured from $x = 0$. For $x < 1$, $k_0 = -d_0^{-1/2}$ and is unbounded at $x = x_{s0}$ because the wavelength vanishes at zero water depth. A sample plot of the basic state is shown in figure 1 for $R_0 = 0.4$

and $\gamma = 0.6$. Taking for example $h_F = 3.0\,\mathrm{m}$, the value of R_0 is for $H_F = 1.2\,\mathrm{m}$. Note that the basic state has discontinuous gradients at $x = x_{b0}$, due to a sudden increase in dissipation of the wave energy. It is also not smooth at $x = 1$ where the topography changes.

It is interesting to note that the basic state has no depth-averaged current, but a pressure gradient due to the presence of waves. It is the instability of the set-up that is of interest here, and this instability depends on the interaction of waves and currents: alongshore perturbations in the wave field produce sources of vorticity; the resulting circulation could enhance the perturbations by their refractive effects on the wave field when certain conditions are met, leading to the growth of the perturbations. This is analogous to Rayleigh–Bénard instability in which the instability of the basic temperature gradient generates motions (the basic state is otherwise motionless).

4. Linear instability

To study the instability of the basic state subject to infinitesimal alongshore perturbations, we assume

$$(u, v, \eta, H, k, l, \omega) = (0, 0, \eta_0, H_0, k_0, 0, 1) + \epsilon(u_1, v_1, \eta_1, H_1, k_1, l_1, \omega_1), \tag{4.1}$$

where ϵ is arbitrarily small. Linearizing the dimensionless versions of (2.1)–(2.9) around the basic state, we obtain the dimensionless equations for the perturbations as follows. For circulations,

$$\eta_{1,t} + d_{0,x}u_1 + d_0u_{1,x} + d_0v_{1,y} = 0, \tag{4.2}$$

$$u_{1,t} = -\eta_{1,x} - R_0^2\left[d_0^{-1}(S_{1xx,x} + S_{1xy,y}) - \eta_1 d_0^{-2}S_{0xx,x}\right] - \tfrac{1}{8}R_0^2R_1^{-1}\mu_0 d_0^{-1}u_1, \tag{4.3}$$

$$v_{1,t} = -\eta_{1,y} - R_0^2\left[d_0^{-1}(S_{1xy,x} + S_{1yy,y})\right] - \tfrac{1}{8}R_0^2R_1^{-1}\mu_0 d_0^{-1}v_1, \tag{4.4}$$

where a subscript proceeded by a comma stands for a partial derivative and

$$S_{1yy} = \tfrac{1}{2}E_1, \quad S_{1xy} = k_0^{-1}l_1E_0, \quad S_{1xx} = \tfrac{3}{2}E_1. \tag{4.5}$$

$E_0 = \frac{1}{8}H_0^2$ and $E_1 = \frac{1}{4}H_0H_1$ are the basic and perturbed wave energies, respectively, and $\mu_0 = \frac{1}{2}H_0d_0^{-1/2}$ is the friction velocity. For waves,

$$l_{1,t} + \omega_{1,y} = 0, \tag{4.6}$$

where

$$\omega_1 = k_0u_1 + \tfrac{1}{2}k_0^2\eta_1 + d_0k_0k_1. \tag{4.7}$$

The irrotationality of the wavenumber field leads to

$$l_{1,x} - k_{1,y} = 0. \tag{4.8}$$

Before the wave breaks,

$$E_{1,t} + \left[\left(u_1 - \tfrac{1}{2}d_0^{-1/2}\eta_1\right)E_0 - d_0^{1/2}E_1\right]_{,x} + \left[v_{1,y} - d_0^{1/2}k_0^{-1}l_{1,y}\right]E_0 + S_{0xx}u_{1,x} + S_{0yy}v_{1,y} = 0, \tag{4.9}$$

where $S_{0yy} = \frac{1}{2}E_0$. After the breaking,

$$H_1 = \gamma R_0^{-1}\eta_1. \tag{4.10}$$

Apart from scaling, equations (4.2)–(4.10) are the same as in DL78.

Once the circulations develop, the shoreline becomes a moving boundary, at which vanishing normal flow must be satisfied, i.e.

$$-x_{s,t} + u = v x_{s,y} \quad \text{at} \quad x = x_s(t, y). \tag{4.11}$$

Expanding the shoreline position as $x_s = x_{s0} + \epsilon x_{s1}(t, y)$, we have $x_{s1} = -\widetilde{\beta}^{-1}\eta_1(x_{s0})$ following from $d(x_s) = 0$. Linearizing (4.11) around $x = x_{s0}$, the kinematic boundary condition at the shoreline is approximated as

$$u_1 = -\widetilde{\beta}^{-1}\eta_{1,t} \quad \text{at} \quad x = x_{s0}. \tag{4.12}$$

As $x \to \infty$, the appropriate boundary condition for a hyperbolic system is the so-called 'radiation condition', which states that disturbances must propagate outward at infinity (Courant & Hilbert 1962). As will be seen in §4.2, simply rejecting the families of solution which increase spatially at large x is not sufficient to ensure the radiation condition.

Hydrodynamic instability problems which can be solved more or less explicitly are commonly attacked by immediately examining wave-like or cellular perturbations, a procedure which goes back at least to Rayleigh in his early studies of inviscid parallel flows and convection. There are good technical reasons for this, and this tradition will be followed here. However, it is still desirable to keep in mind that the underlying problem is the initial value problem for the linearized partial differential equations: we want to know if there are any finite (though small) disturbances which can evolve from arbitrarily small initial perturbations. When the linearized equations constitute a hyperbolic system, as is the case here, the theory of characteristics gives a good overall picture of the initial value problem. Sometimes this broader perspective is very helpful in clarifying various questions, such as the precise meaning of a 'radiation condition' or the number and character of the appropriate boundary conditions. This has been the case in this study. In this section, we shall still follow the traditional approach of wave-like perturbations. When necessary, we make use of the theory of characteristics, but defer more detailed considerations to the Appendix. Let f_1 represent the perturbation of a variable and

$$f_1 = \widehat{f}_1(x)\mathrm{e}^{\mathrm{i}\alpha y + \mathrm{i}\sigma t} + \text{c.c.}, \tag{4.13}$$

where α is the wavenumber of the alongshore perturbations and σ the complex frequency. For generally complex values of σ, the circulation cells described by (4.13) propagate alongshore and grow (or decay) in strength as time passes, depending on the sign of $\mathrm{Im}\,\sigma$. We shall focus on non-propagating cells, i.e. $\mathrm{i}\sigma$ is real or zero. These are directly associated with rip current circulations.

4.1. *On the neglect of effects on the wavenumber field and the resulting stability*

In their study of neutral modes, DL78 found that (4.2)–(4.10) do not have non-trivial steady-state solutions if one assumes that the effects of currents on the wavenumber field are insignificant and neglects k_1 and l_1. This indicates that alongshore perturbations may not be sustained due to insufficient wave forcing if variations in the wavenumber field are not taken into account. This was confirmed by Falqués *et al.* (1999), who proved that the basic state is stable by judiciously constructing a Lyapounov function. In both DL78 and Falqués *et al.* (1999), the wave dispersion relationship (2.7) was retained with the wave frequency being unperturbed, which is a consequence of conserving the wave crests. From the mathematical point of view, the system in that case is over-determined, since the four unknowns, u_1, v_1, η_1 and H_1, are required to satisfy five equations: the continuity, the momentum balances in the

x- and y-directions, the wave energy balance, and the wave dispersion relationship. From the physical point of view, the stability of the basic state is a result of an inconsistency in the demands made on the pressure field (surface elevation) by the conservation of wave crests and by the alongshore momentum. The explanation may be as follows.

When k_1 and l_1 are neglected, the waves remain normally incident and $S_{xy} \equiv 0$. Furthermore, the perturbations of $S_{xx,x}$ and $S_{yy,y}$ are due to the variations in the wave height solely. Inside the surf zone, the momentum equations for the currents, (4.3) and (4.4), reduce to

$$u_{1,t} = -\left(1 + \tfrac{3}{8}\gamma^2\right)\eta_{1,x} - \tfrac{1}{8}R_0^2 R_1^{-1}\mu_0 d_0^{-1} u_1, \tag{4.14a}$$

$$v_{1,t} = -\left(1 + \tfrac{1}{8}\gamma^2\right)\eta_{1,y} - \tfrac{1}{8}R_0^2 R_1^{-1}\mu_0 d_0^{-1} v_1. \tag{4.14b}$$

Because of the assumption of a saturated surf zone, i.e. (4.10), the perturbed wave forcing due to the variations in the wave height simply enhances the magnitudes of the pressure gradient terms, and the flow inside the surf zone is effectively pressure driven as seen in (4.14a) and (4.14b). On the other hand, the wave frequency is not altered in the absence of k_1 and l_1, and the wave dispersion relationship (4.7) leads to $u_1 = \frac{1}{2}d_0^{-1/2}\eta_1$. This means that to conserve wave crests the water depth must be increased at places where offshore flows are developed. Such a flow field, however, cannot be sustained, because the alongshore flows driven by the pressure gradient diverge from the offshore flows (high-pressure fields), and hence tend to stop the circulations.

The importance of the perturbations in the wavenumber field is to disrupt the alongshore uniformity of the wave crests so that across-shore transport of the alongshore wave momentum, S_{xy}, is produced, and consequently a wave forcing $S_{xy,x}$. Thus, the flow field given by (4.3) and (4.4) is no longer purely pressure driven. Specifically, the refractive effects due to circulations tend to produce an alongshore wave forcing $S_{xy,x}$ directed toward the places where flows are offshore, and so tending to counter-balance the pressure force $-\eta_{1,y}$. With appropriate alongshore variations, the forcing $S_{xy,x}$ can be sufficient that alongshore flows develop and converge toward the offshore flows. This leads to the growth of the circulations and hence the instability of the basic state.

It is worth pointing out that the instability analysis in Murray & Reydellet (2001) also does not include the effects of currents on the wavenumber field. The circulations considered there are therefore also pressure driven, in particular in the alongshore direction. Their finding of instability, however, does not contradict the argument above. That instability is a result of their hypothesized extra dissipation of the wave energy in the presence of offshore directed currents. This process may be regarded as a kind of wave–current interaction, but is quite different from the refractive effects considered in this paper, and affects only the wave height not the wavenumber field. With such a dissipation, an additional wave forcing, directed offshore, is effectively produced in regions where offshore currents are, and therefore tends to accelerate the flows there. As a result, low-pressure fields are formed and associated with the offshore currents. Driven by pressure, alongshore flows then converge toward the offshore flows. One might say that the hypothesized Murray & Reydellet mechanism enhances the circulation by driving its offshore section and sucking it outward, while the refractive wave–current interaction instead drives the alongshore feeding section and pushes out the offshore currents. While the hypothesized dissipation can lead to an instability (perhaps a nonlinear one, as the authors suggest that finite-amplitude

perturbations may be needed to trigger the instability), the existence of the dissipation itself has yet to be established. The refractive wave–current interaction on the other hand is based on well-established physical principles.

4.2. *Asymptotic behaviour*

Since the differential equations are not consistent on crossing the breaking line, we must carry out the integrations of the linearized equations inside and outside the surf zone separately, and construct the complete solution by properly matching the physical variables at the breaking line. To do so, we need to supply the data at the shoreline which satisfy the boundary condition (4.12) and the data at large x which satisfy the radiation condition. These can be provided by the asymptotic solutions as $x \to x_{s0}$ and as $x \to \infty$. Note that the system is singular at the shoreline because $d_0(x_{s0}) = 0$. Therefore, unbounded solutions may exist at the shoreline, and their suitability for constructing a physical solution has to be carefully examined.

4.2.1. *At the shoreline*

Close to the shoreline, we introduce the transformation $\zeta = (x - x_{s0})^{1/2}$, where $x > x_{s0}$. The total water depth is then written as $d_0 = \widetilde{\beta}\zeta^2$, and the bottom friction velocity $\mu_0 = \frac{1}{2}\gamma R_0^{-1}\widetilde{\beta}^{1/2}\zeta$. For sufficiently small ζ, it is seen from (4.4) that v_1 can be at most of $O(\zeta)$ for finite values of R_1. By omitting the terms of $O(\zeta)$ smaller than the leading-order terms, three differential equations for the amplitudes of the perturbations, $\widehat{u}_1(x)$, $\widehat{\eta}_1(x)$ and $\widehat{l}_1(x)$, can be approximated from (4.2), (4.3) and (4.8) as follows:

$$\widehat{u}_{1,\zeta} = -2\zeta^{-1}\widehat{u}_1 - 2\mathrm{i}\sigma\widetilde{\beta}^{-1}\zeta^{-1}\widehat{\eta}_1, \tag{4.15}$$

$$\widehat{\eta}_{1,\zeta} = -\tfrac{1}{8}\gamma R_0 R_1^{-1}\widetilde{\beta}^{1/2}\widehat{u}_1, \tag{4.16}$$

$$\widehat{l}_{1,\zeta} = -2\mathrm{i}\alpha\widetilde{\beta}^{-1}\zeta^{-1}\widehat{u}_1 + \mathrm{i}\alpha\widetilde{\beta}^{-3/2}\zeta^{-2}\widehat{\eta}_1 + 2\mathrm{i}\sigma\widetilde{\beta}^{-1/2}\widehat{l}_1. \tag{4.17}$$

From (4.6) and (4.7), we have

$$\widehat{k}_1 = -d_0^{-1}\widehat{u}_1 + \tfrac{1}{2}d_0^{-3/2}\widehat{\eta}_1 + \alpha^{-1}\sigma d_0^{-1/2}\widehat{l}_1. \tag{4.18}$$

Following (4.10), $\widehat{H}_1 = \gamma R_0^{-1}\widehat{\eta}_1$. We note that as the shoreline is approached, the wave radiation stresses become essentially independent of the wavenumber field, and depend only on the wave height which is determined by the local water depth. This is due to refraction by topography which causes the waves to be increasingly normal to the shoreline as the depth decreases. This means that sufficiently close to the shoreline, the flow field of the circulation is not influenced by the wavenumbers, even though the latter may be large there. This is explicitly evident in the asymptotic flow equations (4.15) and (4.16), where the perturbation of the wavenumber field is absent.

Approximate solutions of (4.15)–(4.17) as $\zeta \to 0$ can be obtained using a Frobenius series in ζ. Of the three families of solutions admitted by the equations, one is bounded at the shoreline. A second family has an unbounded flow field as $x \to x_{s0}$ and thus is not physically appropriate since $\widehat{\eta}_1$ and $\widehat{u}_1$ must remain finite. The third family is also unbounded, but only in the wavenumber components $\widehat{l}_1$ and $\widehat{k}_1$. This singularity is unsurprising since the wave approximation becomes invalid as the water depth approaches zero. However, the wave solution may be disregarded in a small region close to the shoreline, since its influence on the flow properties is negligible there, and the wave information only propagates shoreward. Thus the first and third family are admissible for constructing a solution near the shoreline, satisfying the

shoreline boundary condition on the flow and matching with the incoming wave and flow information. The first family (bounded) is found to be

$$\widehat{u}_1 = 0, \quad \widehat{\eta}_1 = 0, \quad \widehat{l}_1 = \mathrm{i} - 2\sigma\widetilde{\beta}^{-1/2}\zeta. \tag{4.19a,b,c}$$

The third family (admissible unbounded) is

$$\widehat{u}_1 = -\mathrm{i}\sigma\widetilde{\beta}^{-1} + \tfrac{1}{12}\gamma\sigma^2 R_0 R_1^{-1}\widetilde{\beta}^{-3/2}\zeta, \tag{4.20a}$$

$$\widehat{\eta}_1 = 1 + \tfrac{1}{8}\gamma\mathrm{i}\sigma\widetilde{\beta}^{-1/2}R_0 R_1^{-1}\zeta, \tag{4.20b}$$

$$\widehat{l}_1 = -\tfrac{1}{8}\gamma\alpha\sigma R_0 R_1^{-1}\widetilde{\beta}^{-2}\big(1 + 2\mathrm{i}\sigma\widetilde{\beta}^{-1/2}\zeta\big)\ln\zeta - \mathrm{i}\alpha\widetilde{\beta}^{-3/2}\zeta^{-1} + \mathrm{i}. \tag{4.20c}$$

First, the flow variables $\widehat{u}_1$ and $\widehat{\eta}_1$ given by both families satisfy the boundary condition at $d_0 = 0$, i.e. (4.12), as $\zeta \to 0$. Second, in (4.20), $\widehat{l}_1 = O(\zeta^{-1})$. It follows from (4.18) that $\widehat{k}_1 = O(\zeta^{-3})$. This is due to $\nabla \times \boldsymbol{k} = 0$, which is a linear equation. Since $k_0 \sim O(\zeta^{-1})$, strictly speaking, the solution (4.20) should be restricted to $O(\epsilon^{1/2}) < \zeta \ll 1$, i.e. $O(\epsilon) < d_0 \ll 1$, for the perturbation expansions of l and k to be valid. However, as the circulation is not coupled to the wavenumber field the singular behaviour for $d_0 \to 0$ does not affect the flow variables given in (4.20*a*, *b*). Using (4.19) and (4.20) as the initial data, one can start the seaward integrations of (4.2)–(4.8), together with (4.10), to obtain two appropriate independent solutions. A suitable linear combination of these, determined by the matching at the breaking line (see § 4.3), will give the actual perturbation solution in the surf zone. We note that in DL78 only one family of solution was reported close to the shoreline, and hence in the surf zone.

The existence of two families of solution in the vicinity of the shoreline can be justified by analysing the characteristics of the hyperbolic system, (4.2)–(4.8) and (4.10), for the surf zone. As shown in the Appendix § A.1, there are four characteristics across the surf zone; two are incoming (i.e. directed toward the shoreline), one is outgoing (seaward directed) and the other is $x = \text{const}$. The two incoming characteristics propagate the solution from the interior (the surf zone) to the shoreline. One of them carries $\widehat{l}_1$, and the other carries a linear combination of $\widehat{u}_1$ and $\widehat{\eta}_1$. As the solution in the interior is yet to be determined, two degrees of freedom must be allowed in the asymptotic representation near the shoreline in order to avoid any conflict with the incoming solution from the interior. These are in fact the proportionality coefficients of the two families given by (4.19) and (4.20).

4.2.2. *At large x*

For $x > 1$, $d_0 = 1$ and $\mu_0 = 1$. Four differential equations are obtained from (4.2), (4.3), (4.8) and (4.9) for the amplitudes $\widehat{u}_1$, $\widehat{\eta}_1$, $\widehat{l}_1$ and $\widehat{H}_1$:

$$\widehat{u}_{1,x} = -\mathrm{i}\alpha\widehat{v}_1 - \mathrm{i}\sigma\widehat{\eta}_1, \tag{4.21}$$

$$\widehat{\eta}_{1,x} = -\big(\tfrac{1}{16}R_0^2 R_1^{-1} + \mathrm{i}\sigma\big)\widehat{u}_1 - \tfrac{3}{8}R_0^2\widehat{H}_{1,x} + \tfrac{1}{8}\mathrm{i}\alpha R_0^2\widehat{l}_1, \tag{4.22}$$

$$\widehat{l}_{1,x} = \mathrm{i}\alpha\widehat{k}_1, \tag{4.23}$$

$$\widehat{H}_{1,x} = \tfrac{5}{4}\widehat{u}_{1,x} - \tfrac{1}{4}\widehat{\eta}_{1,x} + \tfrac{3}{4}\mathrm{i}\alpha\widehat{v}_1 + \mathrm{i}\sigma\widehat{H}_1 + \tfrac{1}{2}\mathrm{i}\alpha\widehat{l}_1. \tag{4.24}$$

The alongshore velocity $\widehat{v}_1$ and the wavenumber $\widehat{k}_1$ are given by algebraic equations following (4.4) and (4.6), i.e.

$$\big(\tfrac{1}{16}R_0^2 R_1^{-1} + \mathrm{i}\sigma\big)\widehat{v}_1 = -\mathrm{i}\alpha\widehat{\eta}_1 + \tfrac{1}{8}\mathrm{i}\alpha R_0^2(\widehat{k}_1 - \widehat{H}_1), \tag{4.25}$$

$$\widehat{k}_1 = -\widehat{u}_1 + \tfrac{1}{2}\widehat{\eta}_1 + \sigma\alpha^{-1}\widehat{l}_1. \tag{4.26}$$

The solutions admissible by (4.21)–(4.24) behave exponentially in x for real iσ. For $\sigma = 0$ (neutral modes) and for sufficiently small real iσ, two families of spatially decaying solutions can be found. Suppose both families are used to construct the asymptotic behaviour at large x. By letting the radiation condition be satisfied at a large distance, say x_e, the proportionality of each family can be determined, and is an exponential function of x_e. It is readily seen that when we let $x_e \to \infty$ either the proportionality of the faster decaying family approaches zero exponentially or the proportionality of the slower decaying family approaches infinity exponentially, depending on the normalization. This suggests that a linear combination of the two families cannot uniformly satisfy the radiation condition. Therefore, only one family can be allowed, namely the faster decaying one. For sufficiently large real iσ, (4.21)–(4.24) admit only one family of spatially decaying solutions. This structure of the asymptotic behaviour at large x is justified by the results of the characteristic analysis in the Appendix § A.2. At large x, there is only one outgoing characteristic which propagates the solution from the interior (the region outside the surf zone) to infinity. Connecting consistently to the solution coming from the interior implies that one degree of freedom is allowed in the asymptotic representation at large x.

For the same equations, DL78 reported the existence of three families of spatially decaying solutions at large x. All of them were used to determine the asymptotic behaviour at large x. Their solution of the across-shore profile of the perturbed wave energy exhibits a sharp peak just offshore of the breaking point, see their figures 2 and 3. This seems to suggest an abnormality in the matching.

4.3. *Matching conditions at the breaking line*

The location of the breaking line x_b is defined by

$$H^+(x_b) = \gamma R_0^{-1} d^+(x_b), \tag{4.27}$$

where the superscript + indicates the variables obtained for outside the surf zone. Correspondingly, the superscript − is used below for the inside. Assuming the perturbation expansion $x_b = x_{b0} + \epsilon \delta x_{b1} \mathrm{e}^{\mathrm{i}\alpha y + \mathrm{i}\sigma t} + \text{c.c.}$, where δx_{b1} is a constant, we find from (4.27) that

$$\delta x_{b1} = \frac{4}{5}\left[\frac{(R_0/\gamma)\widehat{H}_1^+ - \widehat{\eta}_1^+}{d_{0,x}^+}\right]_{x=x_{b0}}. \tag{4.28}$$

Clearly, the displacement of the breaking line is a consequence of requiring the wave height (i.e. wave energy) to be continuous. Let $f = f_0 + \epsilon \widehat{f}_1(x)\mathrm{e}^{\mathrm{i}\alpha y + \mathrm{i}\sigma t} + \text{c.c.}$ represent a physical variable to be solved. The continuity of f at $x = x_b$ can be approximated as

$$\widehat{f}_1^- - \widehat{f}_1^+ + \Delta f_{0,x}\delta x_{b1} = 0 \quad \text{at} \quad x = x_{b0}, \tag{4.29}$$

where $\Delta f_{0,x} = f_{0,x}^-(x_{b0}) - f_{0,x}^+(x_{b0})$ is the discontinuity of the slope of the basic state at the breaking line. Notice the significance of δx_{b1} when the basic state is not zero. There are constraints applied to (4.29). First, from (4.6) and (4.7), $\widehat{k}_1$ depends linearly on $\widehat{u}_1$, $\widehat{\eta}_1$ and $\widehat{l}_1$. Second, $\widehat{H}_1$ and $\widehat{\eta}_1$ are linearly related at $x = x_{b0}$ because of (4.28) and (4.10). Third, since the flow is inviscid, a discontinuity in $\widehat{v}_1$ can be allowed. Therefore, of the six unknowns (i.e. $\widehat{u}_1$, $\widehat{v}_1$, $\widehat{\eta}_1$, $\widehat{H}_1$, $\widehat{l}_1$ and $\widehat{k}_1$) to be determined, only three independent relations can be derived from (4.29), regarding $\widehat{u}_1$, $\widehat{\eta}_1$ and $\widehat{l}_1$. The same conclusion is drawn by analysing the characteristics on both sides of the breaking line $x = x_{b0}$, see Appendix A.

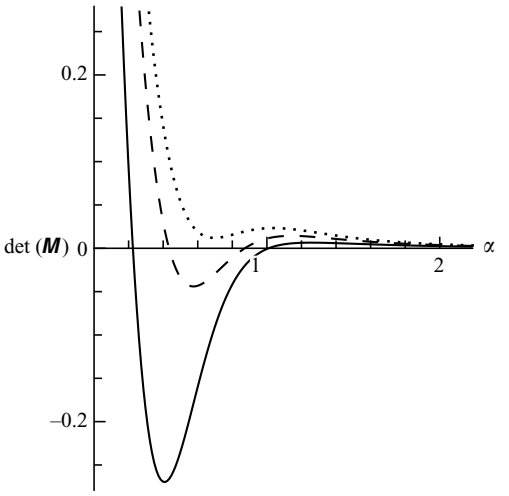

FIGURE 2. det($\boldsymbol{M}$) as a function of α for different values of iσ. $R_0 = 0.4$, $R_1 = 5.4920$ and $\gamma = 0.6$. ——, i$\sigma = 0$; – – –, i$\sigma = 0.002$; · · · · ·, i$\sigma = 0.003$.

In view of the asymptotic behaviour discussed in § 4.2, we then obtain three linear equations from (4.29), determining the proportionality coefficients of the three admissible families of solution. Let $(\widehat{u}_{11}^-, \widehat{\eta}_{11}^-, \widehat{l}_{11}^-)$ and $(\widehat{u}_{12}^-, \widehat{\eta}_{12}^-, \widehat{l}_{12}^-)$ be the two families inside the surf zone, and $(\widehat{u}_1^+, \widehat{\eta}_1^+, \widehat{l}_1^+)$ the one outside the surf zone. The matrix of the homogeneous system for the coefficients is then written as

$$\boldsymbol{M} = \begin{pmatrix} \widehat{u}_{11}^- & \widehat{u}_{12}^- & -\widehat{u}_1^+ \\ \widehat{l}_{11}^- & \widehat{l}_{12}^- & -\widehat{l}_1^+ \\ \widehat{\eta}_{11}^- & \widehat{\eta}_{12}^- & -\widehat{\eta}_1^+ - \frac{4}{5}\left(\Delta\eta_{0,x}/d_{0,x}^+\right)\left[\widehat{\eta}_1^+ - (R_0/\gamma)\widehat{H}_1^+\right] \end{pmatrix}, \tag{4.30}$$

where all the entries are evaluated at $x = x_{b0}$. Note that $\boldsymbol{M} = \boldsymbol{M}(\alpha, \sigma; R_0, R_1, \gamma)$, i.e. a function of α and σ given the parameters R_0, R_1 and γ. For non-trivial coefficients to exist, the determinant of $\boldsymbol{M}$ must vanish. Thus,

$$\det[\boldsymbol{M}(\alpha, \sigma; R_0, R_1, \gamma)] = 0 \tag{4.31}$$

is the dispersion relation function for the normal modes. For purely growing modes, iσ is real and positive, and (4.31) gives the growth rate curve.

5. Instability properties

Typical plots of det($\boldsymbol{M}$) as a function of α, given iσ, are presented in figure 2, showing the existence of growing modes. For real iσ, $\boldsymbol{M}$ can be rescaled such that det($\boldsymbol{M}$) is real. The parameters are $R_0 = 0.4$, $R_1 = 5.4920$, and $\gamma = 0.6$. The growth rate curve is shown in figure 3. The curves for $R_0 = 0.3$ and $R_0 = 0.5$ are also included. For all three growth rate curves, $R_0/R_1 = 0.0728$. From (2.14), R_0/R_1 is effectively the ratio of c_f to β, thus representing the beach conditions in terms of the the roughness and geometry. Rip currents and rhythmic beach morphology are often observed on beaches with moderate slope from 2° to 6° (Short 1999, p. 178). If we take $\beta = \tan 4° \simeq 0.07$, the value of R_0/R_1 in figure 3 is for $c_f = 0.001$. It is seen that the growth rate decreases rapidly as the wave condition weakens (smaller R_0) and becomes negative when R_0 is sufficiently small, indicating that the basic state is stable. Note from (2.12) that the horizontal length and time are both scaled using

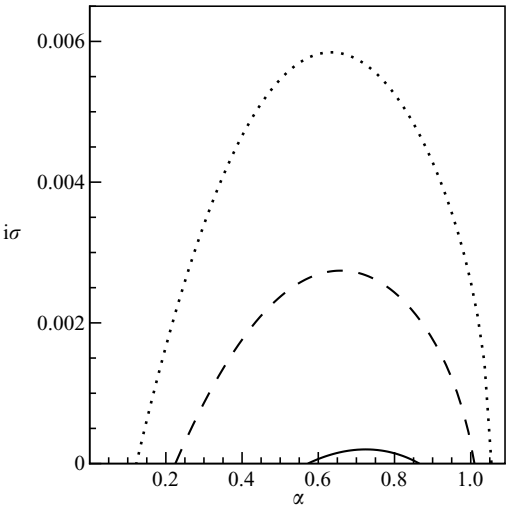

FIGURE 3. Growth rate curves for different values of R_0. $R_0/R_1 = 0.0728$ and $\gamma = 0.6$. ——, $R_0 = 0.3$; – – –, $R_0 = 0.4$; · · · · · ·, $R_0 = 0.5$.

$x_F = h_F/\beta$. Thus, the growth rate curves remain the same when β is changed, except for a different value of R_0/R_1. This indicates that the beach slope β is not a dynamic parameter in this linear analysis, but rather a scaling factor.

Below we will examine the instability properties in the domain of parameters representing the effects of offshore wave height and of bottom friction c_f. The breaking index γ controls not only the onset of breaking but also the transformation of waves across the surf zone. Its effects on the predictions of the instability properties are therefore also examined, and shown to be related to those of wave height and of bottom friction.

5.1. *Neutral modes and threshold wave condition*

For each growth rate curve, there are two neutral modes α_c at $\sigma = 0$, which define the range of unstable wavenumbers for the chosen parameters. In figure 4, α_c is plotted as a function of R_0 for several values of R_B, where $R_B \equiv R_0/R_1 = (16/\pi)c_f/\beta$ is introduced for the convenience of discussions. Note that R_0 cannot exceed γ; otherwise the offshore waves are breaking over the horizontal bed and the breaking line is indefinite. The range of unstable wavenumbers increases with R_0, and as R_B is reduced. So the stronger the offshore waves and the smaller the bottom friction, the more likely it is that the instability occurs. For given R_B, there is a threshold wave condition R_{0cr} above which the neutral modes exist. In other words, when $R_0 < R_{0cr}$ waves are so weak that the instability cannot occur. As the bottom friction is increased (larger R_B), R_{0cr} becomes greater, indicating that stronger waves are required for the instabilities. This threshold satisfies $i\sigma = \partial(i\sigma)/\partial\alpha = 0$. In figure 5, R_{0cr} *vs.* R_B is shown for $\gamma = 0.6$. The curve separates the parameter plane into the stable and unstable regimes. As was discussed in § 2, the dynamics of the circulations is controlled by the competition between vorticity generation due to waves and its dissipation by bottom friction, which is measured by R_1. At the threshold, $R_1 = R_{0cr}/R_B$. The physical significance of the threshold wave condition is then clear: given the beach conditions (roughness and geometry), waves must be so energetic that $R_0 > R_{0cr}$, in order to produce a rate of vorticity generation that is sufficient to sustain the circulations.

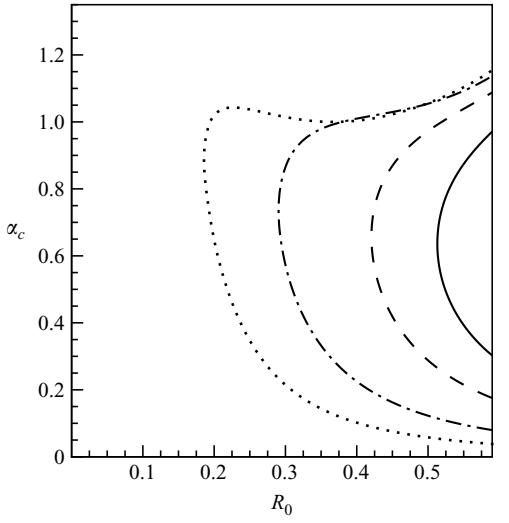

FIGURE 4. Wavenumber of the neutral modes as a function of R_0. $\gamma = 0.6$. $R_B \equiv R_0/R_1$. For a given R_B, the growth rate is zero on the curve, and positive inside the curve. ——, $R_B = 0.2185$; – – –, $R_0 = 0.1457$; – · – · –, $R_B = 0.0728$; · · · · · ·, $R_B = 0.0334$.

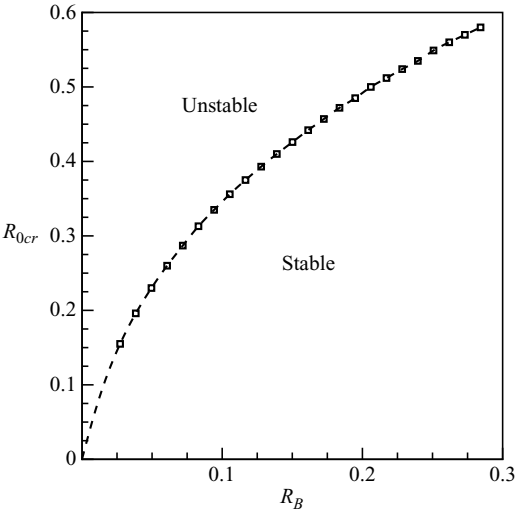

FIGURE 5. Threshold wave condition R_{0cr} as a function of R_B. $\gamma = 0.6$. Given R_B, the basic state is unstable when $R_0 > R_{0cr}$.

Otherwise, the dissipation due to bottom friction is dominant and circulations die away eventually.

We remark that when $R_B = 0$ ($c_f = 0$ for finite β) and $\sigma = 0$, the system is singular because the alongshore momentum equation (4.4) fails to provide information about v_1. This singular behaviour is beyond the scope of this study, as such small values of c_f are not of practical concern. The data points computed in figure 5 cover a range of $0.0004 < c_f < 0.0039$ if $\beta = \tan 4^\circ$ is taken.

When γ is varied, while keeping the conditions of the wave and the beach unchanged (i.e. fixing R_0 and R_B), the dependence of the neutral wavenumbers on the breaking

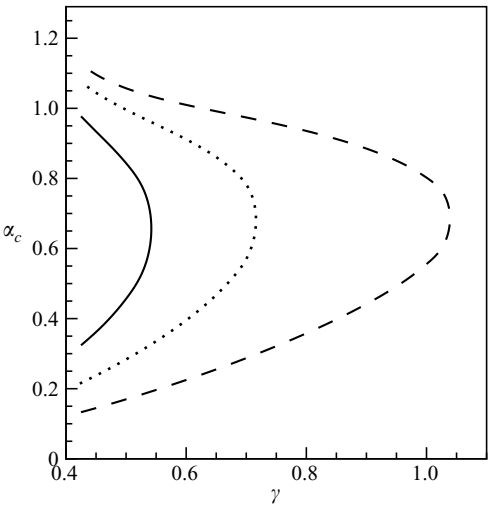

FIGURE 6. Wavenumber of the neutral modes as a function of γ. $R_0 = 0.4$, and $\gamma > R_0$. ——, $R_B = 0.1457$; · · · · · ·, $R_B = 0.1092$; – – –, $R_B = 0.0728$.

index is indicated in figure 6. $R_0 = 0.4$ is chosen so that a reasonable range of γ can be examined, since $\gamma > R_0$ is required. As γ increases, the range of unstable wavenumbers reduces, implying that the instability is less likely to occur. When γ is sufficiently large, the instability does not occur. Given the wave and the beach, varying the breaking index has two effects. First, it affects the width of the surf zone; a smaller γ gives a wider surf zone. This effect can be equivalent to that of varying the offshore wave height with γ being fixed. Second, γ determines the transformation of waves after breaking, and so affects the estimation of the bottom friction force inside the surf zone. Recall that the friction velocity depends on the local wave height, see § 4. A reduction of γ leads to a smaller wave height estimated at a given water depth, and hence a smaller friction velocity. This is similar to the effects due to varying the bottom friction c_f with γ being fixed. In view of these, the results in figure 6 are consistent with the findings in figure 4.

5.2. *Most unstable modes and alongshore spacing of circulations*

The maximum value $(\mathrm{i}\sigma)_m$ of a growth rate curve defines the most unstable mode α_m. The wavelength, $\lambda_m = (2\pi/\alpha_m)x_F$, is referred to as the preferred alongshore spacing of the circulations. It can be used as a measure of the alongshore spacing of rip currents. To see the effects of offshore wave height, we plot α_m and $(\mathrm{i}\sigma)_m$ as functions of R_0 in figures 7(a) and 7(b), respectively, for different values of R_B. Their variations with R_B, given R_0, are shown in figures 8(a) and 8(b), emphasizing the effects of bottom friction. The following observations can be made.

(i) Except for its existence, α_m is independent of R_B, and hence of the bottom friction c_f, see figures 7(a) and 8(a). So once the threshold of instability is exceeded, the preferred alongshore spacing of the circulations is solely a property of the offshore wave condition R_0, though its physical value is proportional to the horizontal extent of the nearshore region. This may be attributed to the fact that the bottom friction appears as a body force, rather than a viscous force which dissipates momentum according to the length scale of the motions. This finding is of particular interest in view of the uncertainties in the friction coefficient c_f.

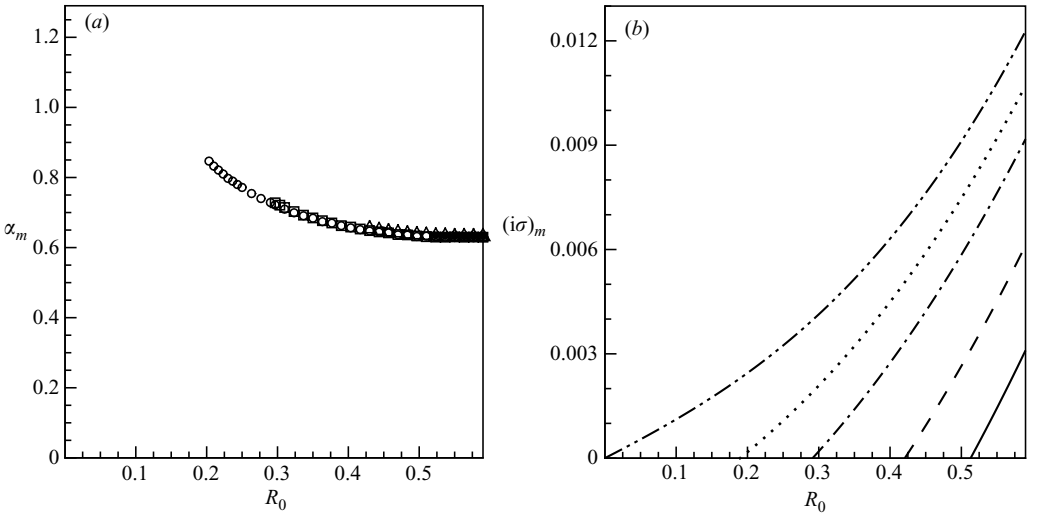

FIGURE 7. (*a*) Wavenumber of the most unstable mode as a function of R_0: ●, $R_B = 0.2185$; △, $R_B = 0.1457$; □, $R_B = 0.0728$; ○, $R_B = 0.0364$. (*b*) Maximum growth rate as a function of R_0: ——, $R_B = 0.2185$; – – –, $R_B = 0.1457$; –·–·–, $R_B = 0.0728$; ·····, $R_B = 0.0364$; –··–··–, $R_B = 0$, $\gamma = 0.6$.

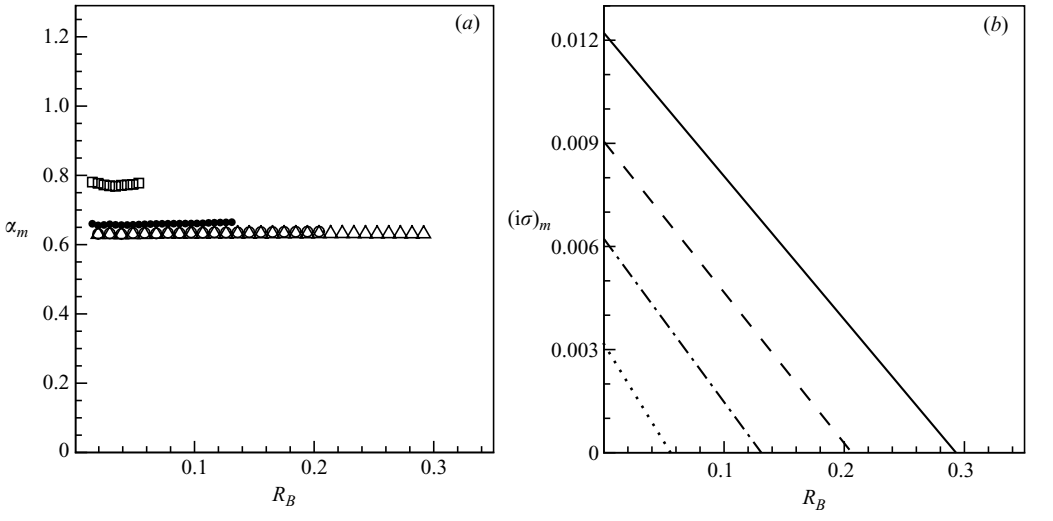

FIGURE 8. (*a*) Wavenumber of the most unstable mode as a function of R_B: △, $R_0 = 0.59$; ○, $R_0 = 0.50$; ●, $R_0 = 0.40$; □, $R_0 = 0.25$. (*b*) Maximum growth rate as a function of R_B: ——, $R_0 = 0.59$; – – –, $R_0 = 0.50$; –·–·–, $R_0 = 0.40$; ·····, $R_0 = 0.25$. $\gamma = 0.6$.

(ii) As R_0 increases, α_m decreases and so λ_m increases. Although the trend is weak, it is consistent with the observations that stronger waves are correlated with greater alongshore spacing of rip currents (Short 1985; Huntley & Short 1992). For sufficiently strong waves, α_m is essentially independent of R_0, and is close to 0.63.

(iii) From figure 7(*b*), $(i\sigma)_m$ increases monotonically with R_0. So the stronger the waves, the faster the circulations develop. The increase is approximately linear when the bottom friction is high (large R_B), and becomes increasingly nonlinear as the friction reduces. At the theoretical limit $c_f = 0$ ($R_B = 0$), the maximum growth

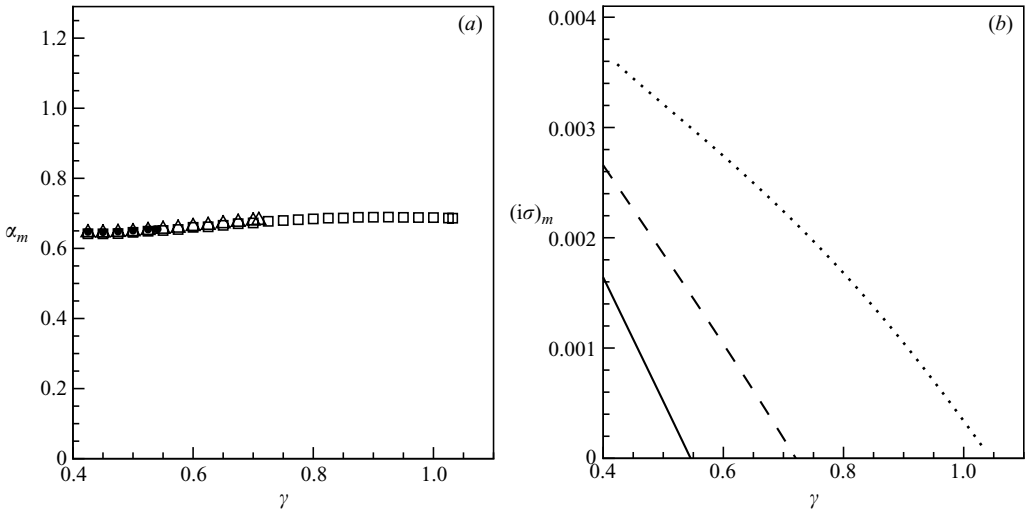

FIGURE 9. (*a*) Wavenumber of the most unstable mode as a function of γ: ●, $R_B = 0.1457$; △, $R_B = 0.1092$; □, $R_B = 0.0728$. (*b*) Maximum growth rate as a function of γ: ——, $R_B = 0.1457$; – – –, $R_B = 0.1092$; · · · · · ·, $R_B = 0.0728$. $R_0 = 0.4$, and $\gamma > R_0$.

$R_0 = H_F/h_F$	0.59	0.55	0.50	0.45	0.40	0.35
Wave height H_F[m]	1.880	1.760	1.600	1.440	1.280	1.120
Breaker height H_b[m]	1.894	1.791	1.659	1.525	1.388	1.247
Alongshore spacing λ_m[m]	456.40	456.06	454.02	446.96	436.40	419.98
Surf zone width b_s[m]	51.24	48.45	44.89	41.26	37.55	33.75
λ_m/b_s	8.91	9.41	10.11	10.83	11.62	12.44
Growth time t_e[min] (for $c_f = 0.001$)	14.83	17.84	23.28	32.34	48.73	97.20

TABLE 1. Predictions for a beach with $\beta = \tan 4^\circ$ and $h_F = 3.2$ m. Horizontal and time scales: $x_F = h_F/\beta = 45.76$ m, $t^* = x_F/(gh_F)^{1/2} = 8.17$ s; $H_b = H_0(x_{b0})H_F$, $\lambda_m = (2\pi/\alpha_m)x_F$, $b_s = (x_{b0} - x_{s0})x_F$, $t_e = t^*/(\mathrm{i}\sigma)_m$.

rate remains finite, and so is completely determined by the wave condition. From figure 8(*b*), $(\mathrm{i}\sigma)_m$ decreases linearly as the bottom friction is increased for a given wave condition.

In figures 9(*a*) and 9(*b*), the wavenumber of the most unstable modes and the corresponding maximum growth rate are plotted against γ, showing the effects of breaking index on the predictions. The preferred alongshore spacing of the circulations is seen to be insensitive to the value of γ. The maximum growth rate decreases as γ is increased. The decrease is approximately linear for large R_B, and becomes weaker as R_B gets smaller, in particular for $0.4 < \gamma < 0.7$. As pointed out in § 5.1, the two effects of varying γ can be equivalent to those of varying the offshore wave height and of varying the bottom friction c_f. Therefore, the findings in figure 9 are consistent with those in figures 7 and 8.

To give a quantitative idea, we present in table 1 the physical values of the predicted alongshore spacing and growth time (e-folding) of the circulations, based on the most unstable modes. The beach geometry is $\beta = \tan 4^\circ$ and $h_F = 3.2$ m. The breaking index is $\gamma = 0.6$. The breaker height is defined as $H_b = H_0(x_{b0})H_F$. For $R_0 = 0.59$ to 0.35,

$\Omega = H_b/(w_s T)$	2.7 (0.8)	3.0 (0.9)	4.0 (0.9)	>4.2 (1.2)
Sediment fall velocity w_s[m s^{-1}]	0.055	0.055	0.045	0.045
Breaker height H_b[m]	1.49	1.65	1.80	1.89
Rip spacing [m]	252 (105)	298 (109)	367 (104)	463 (123)

TABLE 2. Observed spacing of erosion rips on Narrabeen Beach (Short 1985). For Narrabeen Beach, the wave period $T \simeq 10$ s. Numbers inside the parentheses are standard deviations.

the breaker height is about 1.894 to 1.247 m, and we estimate an alongshore spacing of 456.40 to 419.98 m for the circulations. Correspondingly, the estimated growth time, for $c_f = 0.001$, is 14.83 to 97.20 min. These values will be somewhat greater if a smaller slope β, or a deeper water depth h_F, is chosen since the horizontal scale will be larger.

A thorough test of this instability theory should perhaps be done by laboratory experiments, in which the topography and waves could both be controlled to match the conditions considered. To our knowledge, such experiments have not yet been attempted. On the other hand, comparisons with observations of rip currents on natural beaches can only be qualitative, if not impossible, in view of the idealizations of the theory and the difficulties of observationally distinguishing this mechanism from other possibilities. Short (1985) reported on an extensive series of rip current observations on Narrabeen Beach, Australia, over a 19 month period. The beach changed significantly during the period of observations, but the slope seemed to be moderate and around 4° (see figure 12 in Short's paper). Short emphasizes that the spacing and apparently also the persistence of rips depend not only on the current wave conditions, but also on the recent history of the wave climate and the state of the beach. Erosion rips, as they were called by Short, are initiated during the period of increasing wave height and accompanied by general beach erosion. According to Short, erosion rips are not prominently controlled by the topography, and are spaced at 300 ~ 500 m (with a standard deviation 100 ~ 200 m). They are highly variable both temporally and spatially, and have no preferred locations, suggesting that erosion rips are controlled by the surf zone hydrodynamics rather than the antecedent morphology. Erosion rips eventually rework the morphology to suit the rip spacing. It therefore seems that the data of erosion rips are more appropriate for comparison with the instability theory. In table 2, four data points are extracted from figure 6b in Short's (1985) paper (and the text below that figure). Each point represents an average of a number of individual rip observations, and the standard deviations are also given. These points are also plotted in Short's figure 7b, where they are labelled 6, 7, 8 and 9. They all correspond to erosion rips. The two of them with large breaker height are of particular interest, because the associated beach states are much less variable alongshore. Comparing these two with the two largest wave cases in table 1, which have nearly the same breaker heights, we find a fairly good agreement on the alongshore spacing, especially in view of the large deviations in table 2. Over the whole range, however, the variation of rip spacing with breaker height is greater in table 2 than in table 1. In attempting to make this comparison we have assumed that the instability leads to the dominant growth of the most unstable mode, and that its wavelength does not change significantly in the subsequent nonlinear evolution into a developed rip current. It has also been assumed that the beach parameters β and h_F used in constructing table 1 are appropriate for Narrabeen Beach in all the states involved in the data for table 2 – perhaps an even more questionable assumption.

Almost two-thirds of the rip data in Short (1985) were collected during periods of decreasing wave height when the beach was being rebuilt. These rips, called 'accretion rips' by Short, are more closely spaced and topographically controlled (therefore may persist at one location for days). By grouping the data according to beach state, and adding separately four individual days of especially high waves, Short effectively reduced the weight of accretion rips in the total data set and obtained a regression line of the mean rip spacing versus $\bar{\Omega}$, which is the average over a group of the dimensionless parameter $\Omega = H_b/(w_s T)$ used in table 2. It is evident from his figure 7b that the regression line would be very little changed if these accretion rip data (labelled 1 to 5) were completely omitted. For the relevant beach types, the averaged sediment fall velocity may be taken as $w_s = 0.05\,\mathrm{m\,s^{-1}}$, and the wave period was approximately $T = 10\,\mathrm{s}$ for Narrabeen Beach. Thus we may take $\bar{\Omega} = 2H_b$, with H_b being expressed in m. The regression line obtained by Short is

$$\bar{y}_s[\mathrm{m}] = 124\bar{\Omega} - 53 \quad \text{or} \quad \bar{y}_s[\mathrm{m}] = 248H_b - 53. \tag{5.1}$$

There is considerable variance in the fit, and as Short remarks, parameters other than just breaker height are probably required to determine the spacing. From the point of view of the instability theory, $H_b = (R_0/\gamma)^{4/5}\gamma\beta x_F$, and $y_s = \lambda_m = (2\pi/\alpha_m)x_F$. Upon eliminating x_F, it is interesting to see that

$$y_s = \lambda_m = (2\pi/\alpha_m)H_b\left[(R_0/\gamma)^{4/5}\gamma\beta\right]^{-1}. \tag{5.2}$$

When the offshore waves are so strong that $R_0 \simeq \gamma$, $\alpha_m = 0.63$; for $\gamma = 0.6$ and $\beta = \tan 4^\circ$, (5.2) reduces to $y_s[\mathrm{m}] = 238H_b$, which is very close to Short's regression line (5.1) over the whole range covered in his figure 7b, also within one standard deviation of the means used in his fitting. This seems to suggest that the whole sequence listed in table 2 is actually nearly at $R_0 \simeq \gamma$, with the lower breaker height corresponding to a smaller depth h_F which then increases through the sequence because of the general beach erosion. Thus, for the sequence of breaker heights in table 2, our formula just obtained would predict $y_s = 355$, 392, 428 and 450 m. This obviously gives a better agreement with the observed rip spacing than the predictions in table 1, where the water depth is assumed to be unchanged. In fact the four data points of table 2 are still better fitted by the line $y_s[\mathrm{m}] = 200H_b$ which is obtained from (5.2) if $R_0 \simeq \gamma$, and $\beta = \tan 4^\circ$, $\gamma = 0.7$ (or $\beta = \tan 4.75^\circ$, $\gamma = 0.6$). On the other hand, if $\beta = \tan 3^\circ$ and $\gamma = 0.7$ (or $\beta = \tan 3.5^\circ$, $\gamma = 0.6$), we could obtain from (5.2) that $y_s[\mathrm{m}] = 272H_b$ for $R_0 \simeq \gamma$, which over-estimates the spacing slightly. All these values are plausible, in view of the uncertainty of the topography when the measurements were taken, and the idealizations of the theory. As remarked before, these attempts at comparing the instability theory with observations on natural beaches must be regarded as very speculative.

Another comparison with Short's data (see also Huntley & Short 1992) can be attempted with his figure 5(*b*), in which a number of points are given for erosion rips in a plot for rip spacing against surf zone width. These are quite spread out, but indicate a ratio of spacing to surf zone width ranging from 1.5 to 8. Bowen & Inman (1969) also found this value. From table 1, the present instability theory gives this ratio of 9 when waves are strong, and still larger when waves are weaker. Clearly, the theory over-estimates this ratio. A possible reason of this discrepancy is that the observational definition of the surf zone may differ from that of the theory either because of a different topography, notably a longshore bar (which is often present), or because the waves are random both in direction and in amplitude, leading to an ill-defined point of initiation of breaking.

5.3. *Perturbation fields*

For any particular variable, the total solution is

$$f(x, y, t, \epsilon) = f_0(x) + \epsilon \widehat{f}_1(x) e^{i\alpha y + i\sigma t} + \text{c.c.} + \cdots,$$

where f_0 and $\widehat{f}_1$ are both obtained on the unperturbed domain, i.e. $x_{s0} \leqslant x \leqslant x_{b0}$ for the inside of the surf zone and $x \geqslant x_{b0}$ for the outside. The matching condition (4.29) ensures the continuity of the total solution at the actual breaking line $x = x_b$, except for the alongshore velocity v. This means that in the region bounded by $x = x_b$ and $x = x_{b0}$, an extrapolation of f_0 toward the actual breaking line is required for the continuity of f. Thus, in this region $\widehat{f}_1$ does not in general represent all of the deviation of the total solution from its basic state. The extrapolation of f_0, if not zero, makes an additional contribution to the perturbation field of f, as a result of the deformation of the domain. This is in fact the manifestation of the matching condition (4.29).

For the convenience of graphing the perturbation fields, we introduce a boundary-conformal, non-orthogonal coordinate system (ξ, y), in which the domain does not appear to be perturbed, i.e.

$$\xi = x_{s0} + \frac{x - x_s}{x_b - x_s}(x_{b0} - x_{s0}) \quad \text{for} \quad x_s \leqslant x \leqslant x_b, \tag{5.3a}$$

$$\xi = x_{b0} + \frac{x - x_b}{1 - x_b}(1 - x_{b0}) \quad \text{for} \quad x_b \leqslant x \leqslant 1, \tag{5.3b}$$

$$\xi = x \quad \text{for} \quad x > 1. \tag{5.3c}$$

Clearly, the actual surf zone $x_s \leqslant x \leqslant x_b$ is mapped onto $x_{s0} \leqslant \xi \leqslant x_{b0}$, the shoaling zone $x_b \leqslant x \leqslant 1$ onto $x_{b0} \leqslant \xi \leqslant 1$ and the section above the flat bed, i.e. $x > 1$, onto $\xi > 1$. In the absence of perturbations, i.e. $\epsilon = 0$, $x(\xi, y, t, 0) = \xi$ or simply $x = \xi$. The mapping of the total solution is

$$f(x, y, t, \epsilon) = F(\xi(x, y, t, \epsilon), y, t, \epsilon) \quad \text{or} \quad F(\xi, y, t, \epsilon) = f(x(\xi, y, t, \epsilon), y, t, \epsilon). \tag{5.4}$$

The basic state must then satisfy

$$F_0(\xi) \equiv F(\xi, y, t, 0) = f(x(\xi, y, t, 0), y, t, 0) \equiv f_0(\xi). \tag{5.5}$$

For clarity of discussion, we shall henceforth suppress explicit mention of y and t. For small ϵ, $x(\xi, \epsilon) = \xi + \epsilon \widehat{\partial x/\partial \epsilon}|_{\epsilon=0} + O(\epsilon^2)$. Thus,

$$\begin{aligned} f(x(\xi, \epsilon), \epsilon) &= f_0(x(\xi, 0)) + \epsilon f_1(x(\xi, \epsilon)) + \text{c.c.} + O(\epsilon^2) \\ &= F_0(\xi) + \epsilon F_1(\xi) + \text{c.c.} + O(\epsilon^2), \end{aligned} \tag{5.6}$$

where

$$F_1(\xi) = f_{0,\xi}(\xi) \left.\frac{\partial x}{\partial \epsilon}\right|_{\epsilon=0} + f_1(\xi) \tag{5.7}$$

is the perturbation field as seen in the ξ coordinate. Utilizing the expansions of x_s and x_b, see § 4, we then find from (5.3) that

$$\left.\frac{\partial x}{\partial \epsilon}\right|_{\epsilon=0} = \left.\widehat{\frac{\partial x}{\partial \epsilon}}\right|_{\epsilon=0} e^{i\alpha y + i\sigma t} + \text{c.c.},$$

where

$$\left.\widehat{\frac{\partial x}{\partial \epsilon}}\right|_{\epsilon=0} = \frac{\delta x_{s1}(x_{b0} - \xi) + \delta x_{b1}(\xi - x_{s0})}{x_{b0} - x_{s0}} \tag{5.8}$$

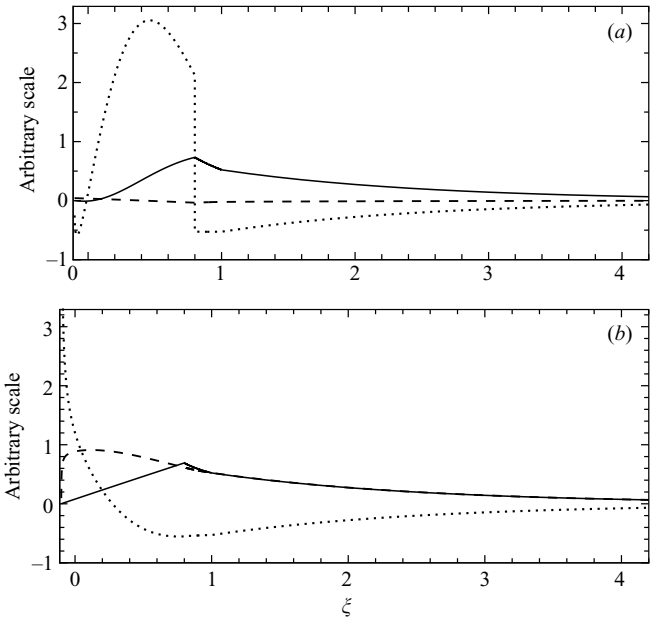

FIGURE 10. Across-shore variations of the amplitudes of the perturbations. $\alpha = 0.64$ and $\mathrm{i}\sigma = 0.00422$. This is the most unstable mode for $R_0 = 0.45$, $R_B = 0.0728$ and $\gamma = 0.6$. The breaking line at the basic state is at $\xi = x_{b0} \simeq 0.80$. (*a*) ——, $\widehat{F}_1^u$; · · · · · ·, $-\mathrm{i}\widehat{F}_1^v$; – – –, $\widehat{F}_1^\eta$. (*b*) ——, $\widehat{F}_1^H$; · · · · · ·, $\widehat{F}_1^k$; – – –, $-\mathrm{i}\widehat{F}_1^l$.

for $x_{s0} \leqslant \xi \leqslant x_{b0}$ and

$$\left.\widehat{\frac{\partial x}{\partial \epsilon}}\right|_{\epsilon=0} = \frac{\delta x_{b1}(1-\xi)}{1-x_{b0}} \tag{5.9}$$

for $x_{b0} \leqslant \xi \leqslant 1$. $\delta x_{s1} = -\widehat{\eta}_1(x_{s0})/\widetilde{\beta}$ and δx_{b1} is given in (4.28). For $\xi > 1$, $(\partial x/\partial \epsilon)_{\epsilon=0} = 0$. Note that $\widehat{\partial x/\partial \epsilon}|_{\epsilon=0} = \delta x_{b1}$ at $\xi = x_{b0}$. Thus, the matching condition (4.29) is in fact the continuity of F_1 at the breaking line. As it explicitly accounts for the deformation of the domain, $F_1(\xi)$ represents more clearly the character of the perturbation than $f_1(x)$, particularly in the neighbourhood of $x = x_{b0}$ where $f_{0,x}$ is discontinuous. For any fixed x away from the breaking line (or shoreline), $\epsilon f_1(x)$ still is the deviation from the basic state.

Upon removing the factor $\mathrm{e}^{\mathrm{i}\alpha y + \mathrm{i}\sigma t}$, we find from (5.7) the amplitude of the perturbation, i.e. $\widehat{F}_1(\xi) = f_{0,\xi}(\xi)\widehat{\partial x/\partial \epsilon}|_{\epsilon=0} + \widehat{f}_1(\xi)$. To refer to a specific variable, we use the appropriate superscript, e.g. $\widehat{F}_1^\eta = \eta_{0,\xi}(\xi)\widehat{\partial x/\partial \epsilon}|_{\epsilon=0} + \widehat{\eta}_1(\xi)$. Figure 10 shows the cross-shore variations of the amplitudes of the perturbations for the variables of the circulation and the wave. The alongshore wavenumber and growth rate are taken as $\alpha = 0.64$ and $\mathrm{i}\sigma = 0.00422$, which correspond to the most unstable mode for $R_0 = 0.45$, $R_B = 0.0728$ and $\gamma = 0.6$. Note that for a real $\mathrm{i}\sigma$, the perturbations can be so normalized that all the variables are real, except for $\widehat{v}_1$ and $\widehat{l}_1$ which are pure imaginary. A few remarks should be made. (i) The cross-shore velocity, i.e. $\widehat{F}_1^u = \widehat{u}_1$, reaches a maximum at the breaking line $\xi = x_{b0} \simeq 0.80$. It is reduced by a factor of e at $\xi \simeq 2.0$, and is less than 20% of the maximum at $\xi \simeq 4x_{b0}$. (ii) The alongshore velocity, i.e. $\widehat{F}_1^v = \widehat{v}_1$, reaches its maximum inside the surf zone. The maximum of $\widehat{v}_1$ is about 4 times of that of $\widehat{u}_1$. This is a consequence of (approximate) flux balance,

as the alongshore flows feed into rip currents over a width (approximately equal to x_{b0}) that is much smaller than the width over which the offshore flux occurs (half of the alongshore wavelength of the normal mode). The discontinuity of $\widehat{v}_1$ at $\xi = x_{b0}$ is expected, as the alongshore velocity is not subject to the matching. (iii) The perturbation of the surface elevation, $\widehat{F}_1^{\eta}$, is small, indicating that the circulation is not primarily pressure driven at this stage. (iv) $\widehat{F}_1^{H}$ has the same sign as $\widehat{u}_1$, and so has $\widehat{H}_1$. Thus, the wave energy is focused when encountering the opposing offshore flows.

6. Concluding remarks

On examining the linear instability in the formation of depth-averaged nearshore circulations due to wave–current interactions, we have given mathematical and physical solutions to those issues which were not handled properly in previous studies. With physically plausible parameters and a moderate beach slope, we have found that circulations with alongshore spacing of a few hundred metres can be initiated by the instability on beaches of typical water depth. The estimated growth time is a few tens of minutes. The predicted alongshore spacing of the circulations is in qualitative agreement with the observed rip spacing. Several conclusions have emerged. (i) The instability occurs when the wave height exceeds a threshold, given a bottom friction condition. (ii) The alongshore spacing of the circulations is independent of the empirical values of the bottom friction coefficient and the breaking index once the threshold of instability is exceeded. For stronger waves, the spacing tends to be greater. (iii) The growth rate increases with the wave height, and as the bottom friction decreases, being approximately a linear function of c_f. It is our intention to extend the theory to deal with more realistic topography, such as beaches with alongshore bars, and to investigate the subsequent nonlinear evolution into developed rip currents.

This work was initiated at Duke University, Durham, NC 27708, when J. Y. was supported by the Andrew W. Mellon Foundation through a grant to A. Brad Murray at the Division of Earth and Ocean Sciences. The preliminary results of this work have been reported at the Theodore Y.-T. Wu Symposium on Engineering Mechanics, joint with the 23rd International Conference on Offshore Mechanics and Arctic Engineering in July 2004. Helpful discussions with Roberto Camassa and Louis N. Howard are gratefully acknowledged.

Appendix. Characteristics for the hyperbolic system

For a hyperbolic system in a finite domain, the number and nature of the boundary conditions are determined by the characteristics of the system. In this Appendix, we show that the characteristic analysis gives rigorous justifications of the asymptotic behaviour at the shoreline and at large x, and the matching conditions, determined in § 4.2 and § 4.3.

Define a column vector $\boldsymbol{\phi}$. For $x < x_{b0}$, $\boldsymbol{\phi} = [u_1, \eta_1, l_1, v_1]^T$, and for $x > x_{b0}$, $\boldsymbol{\phi} = [u_1, \eta_1, H_1, l_1, v_1]^T$, where T stands for the transpose. The variables which do not involve the time derivative are not included, such as H_1 inside the surf zone and k_1. Assuming the periodicity in y such that $\partial/\partial y \to \mathrm{i}\alpha$, but retaining the t and x derivatives, we write (4.2)–(4.10) in the form

$$\boldsymbol{\phi}_t + \boldsymbol{A}\boldsymbol{\phi}_x = \boldsymbol{B}\boldsymbol{\phi}, \tag{A 1}$$

corresponding to the appropriate region. $\boldsymbol{A}$ and $\boldsymbol{B}$ are 4×4 square matrices for the inside of the surf zone, and 5×5 for the outside. Let τ be the eigenvalues of $\boldsymbol{A}$, and $\boldsymbol{e}$ be the left eigenvectors such that $\boldsymbol{e}A=\tau\boldsymbol{e}$. If all the eigenvalues $\tau_1,\ldots,\tau_n$ are real, where $n=4$ for $x<x_{b0}$ and $n=5$ for $x>x_{b0}$, and the corresponding eigenvectors $\boldsymbol{e}_1,\ldots,\boldsymbol{e}_n$ are linearly independent, the system (A 1) is hyperbolic (Courant & Hilbert 1962). The characteristic curves $\mathscr{C}_i$ are then defined as $\mathrm{d}x/\mathrm{d}t=\tau_i$, $i=1,\ldots,n$. Letting $\mathrm{D}_i=\partial/\partial t+\tau_i\partial/\partial x$ be the directional derivative for the curve $\mathscr{C}_i$, we rewrite (A 1) as $\mathrm{D}_i z_i=g_i$, where $z_i=\boldsymbol{e}_i\cdot\boldsymbol{\phi}$ is the quantity that propagates along $\mathscr{C}_i$, and $g_i=\boldsymbol{e}_i\cdot\boldsymbol{B}\cdot\boldsymbol{\phi}$ is the forcing of z_i. We now examine the systems inside and outside the surf zone separately.

A.1. *Inside the surf zone*

For $x<x_{b0}$, according to the entries in $\boldsymbol{\phi}$, we find from (4.3), (4.2), (4.6) and (4.4) that

$$\boldsymbol{A}=\begin{pmatrix}0 & \left(1+\frac{3}{8}\gamma^2\right) & 0 & 0\\ d_0 & 0 & 0 & 0\\ 0 & 0 & -d_0^{1/2} & 0\\ 0 & 0 & -\frac{1}{8}\gamma^2 d_0^{3/2} & 0\end{pmatrix}, \tag{A 2}$$

$$\boldsymbol{B}=\begin{pmatrix}-\frac{1}{8}R_0^2R_1^{-1}\mu_0 d_0^{-1} & 0 & \frac{1}{8}\mathrm{i}\alpha\gamma^2 d_0^{3/2} & 0\\ -d_{0,x} & 0 & 0 & -\mathrm{i}\alpha d_0\\ \mathrm{i}\alpha d_0^{-1/2} & -\frac{1}{2}\mathrm{i}\alpha d_0^{-1} & 0 & 0\\ 0 & -\mathrm{i}\alpha\left(1+\frac{1}{8}\gamma^2\right) & \frac{5}{16}\gamma^2 d_0^{1/2}d_{0,x} & -\frac{1}{8}R_0^2R_1^{-1}\mu_0 d_0^{-1}\end{pmatrix}. \tag{A 3}$$

In deriving these, we have replaced $k_{1,y}$ in (4.6) by $l_{1,x}$ using (4.8), since k_1 is not an entry of the column vector $\boldsymbol{\phi}$ (as we use (4.8) to maintain the irrotationality of $\boldsymbol{k}$, we do not have an equation of $k_{1,t}$). The column of zeros in the matrix $\boldsymbol{A}$ in (A 2) reflects the fact that $v_{1,x}$ is not involved in the linearized equations. The eigenvalues of $\boldsymbol{A}$ are

$$\tau_1=0,\quad \tau_2=-d_0^{1/2},\quad \tau_3=d_0^{1/2}\left(1+\tfrac{3}{8}\gamma^2\right)^{1/2},\quad \tau_4=-d_0^{1/2}\left(1+\tfrac{3}{8}\gamma^2\right)^{1/2}. \tag{A 4}$$

All the eigenvalues are real and distinct, so a set of linearly independent left eigenvectors exists. Since the eigenvalues do not depend on R_1, the characteristic curves are not affected by the bottom friction coefficient c_f. The curve $\mathscr{C}_1^-$, defined by τ_1, is simply $x=\text{const}$. The curves $\mathscr{C}_2^-$ and $\mathscr{C}_4^-$, defined by τ_2 and τ_4 respectively, are shoreward directed (incoming). $\mathscr{C}_3^-$ is directed seaward (outgoing). Correspondingly,

$$\left.\begin{aligned} z_1&=-\tfrac{1}{8}\gamma^2 d_0 l_1+v_1, & z_2&=l_1,\\ z_3&=\left[d_0/\left(1+\tfrac{3}{8}\gamma^2\right)\right]^{1/2}u_1+\eta_1, & z_4&=-\left[d_0/\left(1+\tfrac{3}{8}\gamma^2\right)\right]^{1/2}u_1+\eta_1.\end{aligned}\right\} \tag{A 5}$$

The shoreline $x=x_{s0}$ is a characteristic curve $\mathscr{C}_1^-$, therefore z_1 should not be prescribed there (Courant & Hilbert 1962). It is physically justified that no boundary conditions are needed for l_1 and v_1 at the shoreline, since the waves propagate onshore and the flow is inviscid. The curve $\mathscr{C}_3^-$ enters the surf zone at the shoreline, so a condition must be imposed on z_3 at $x=x_{s0}$. This is the kinematic boundary condition (4.12) at the shoreline, which concerns u_1 and η_1. The incoming characteristic curves $\mathscr{C}_2^-$ (carrying l_1) and $\mathscr{C}_4^-$ (carrying a linear combination of u_1 and η_1) propagate the solution from the interior (the surf zone) to the shoreline. To connect the asymptotic solution, determined close to the shoreline, to these incoming solutions (to be

determined), one must allow two degrees of freedom in the asymptotic representation. Otherwise, inconsistency may occur.

At the breaking line $x = x_{b0}^-$, two conditions, on z_2 and on z_4, are needed as the curves $\mathscr{C}_2^-$ and $\mathscr{C}_4^-$ are entering the surf zone. These conditions concern l_1, u_1 and η_1, and are to be provided by the solutions at the offshore side of the breaking line (i.e. outside the surf zone). On the other hand, the outgoing curve $\mathscr{C}_3^-$ is leaving the surf zone at $x = x_{b0}^-$, providing the boundary conditions for the region outside the surf zone. Since $x = x_{b0}^-$ is a characteristic curve $\mathscr{C}_1^-$, no prescribed data are applied to z_1. This implies that v_1 is not to be matched at the breaking line.

A.2. *Outside the surf zone*

For $x > x_{b0}$, we find from (4.3), (4.2), (4.9), (4.6) and (4.4) that

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 & \frac{3}{8}R_0^2 d_0^{-5/4} & 0 & 0 \\ d_0 & 0 & 0 & 0 & 0 \\ \frac{5}{4}d_0^{-1/4} & -\frac{1}{4}d_0^{-3/4} & -d_0^{1/2} & 0 & 0 \\ 0 & 0 & 0 & -d_0^{1/2} & 0 \\ 0 & 0 & 0 & -\frac{1}{8}R_0^2 d_0^{-1} & 0 \end{pmatrix}, \tag{A 6}$$

$$\boldsymbol{B} = \begin{pmatrix} -\frac{1}{8}\frac{R_0^2}{R_1}\mu_0 d_0^{-1} & -\frac{3}{32}R_0 d_0^{-7/2} d_{0,x} & \frac{3}{32}R_0^2 d_0^{-9/4} d_{0,x} & \frac{1}{8}\mathrm{i}\alpha R_0^2 d_0^{-1} & 0 \\ -d_{0,x} & 0 & 0 & 0 & -\mathrm{i}\alpha d_0 \\ \frac{1}{4}d_0^{-5/4} d_{0,x} & -\frac{1}{4}d_0^{-7/4} d_{0,x} & -\frac{1}{4}d_0^{-1/2} d_{0,x} & -\frac{1}{2}\mathrm{i}\alpha d_0^{3/4} & -\frac{3}{4}\mathrm{i}\alpha d_0^{-1/4} \\ \mathrm{i}\alpha d_0^{-1/2} & -\frac{1}{2}\mathrm{i}\alpha d_0^{-1} & 0 & 0 & 0 \\ 0 & -\mathrm{i}\alpha & -\frac{1}{8}\mathrm{i}\alpha R_0^2 d_0^{-5/4} & 0 & -\frac{1}{8}\frac{R_0^2}{R_1}\mu_0 d_0^{-1} \end{pmatrix}. \tag{A 7}$$

The eigenvalues of $\boldsymbol{A}$ are $\tau_1 = 0$, $\tau_2 = -d_0^{1/2}$, and τ_3, τ_4 and τ_5 are the roots of

$$\tau^3 + d_0^{1/2}\tau^2 - \left(d_0 + \tfrac{15}{32}R_0^2 d_0^{-3/2}\right)\tau - \left(d_0^{3/2} - \tfrac{3}{32}R_0^2 d_0^{-1}\right) = 0. \tag{A 8}$$

It can be shown that $\tau_1, \ldots, \tau_5$ are real and distinct, and a set of linearly independent left eigenvectors exists. Furthermore, $\tau_3 + \tau_4 + \tau_5 = -d_0^{1/2} < 0$ and $\tau_3\tau_4\tau_5 = (d_0^{3/2} - \frac{3}{32}R_0^2 d_0^{-1}) > 0$; thus (A 8) has two negative roots, say τ_3 and τ_4, and one positive root τ_5. In view of these, the characteristic curves $\mathscr{C}_2^+$, $\mathscr{C}_3^+$ and $\mathscr{C}_4^+$ are directed shoreward; $\mathscr{C}_5^+$ is seaward, and $\mathscr{C}_1^+$ is again $x = \text{const}$. The quantities carried by each curve are

$$z_1 = -\tfrac{1}{8}R_0^2 d_0^{-3/4} l_1 + v_1, \quad z_2 = l_1, \quad z_i = e_{i1}u_1 + e_{i2}\eta_1 + e_{i3}H_1, \ i = 3, 4, 5, \tag{A 9}$$

where e_{i1}, e_{i2} and e_{i3} are the first three entries of the left eigenvector $\boldsymbol{e}_i$.

Since $\mathscr{C}_2^+$, $\mathscr{C}_3^+$ and $\mathscr{C}_4^+$ are incoming, z_2, z_3 and z_4 must vanish as $x \to \infty$ as the radiation condition allows only the outward propagation of the disturbances. $\mathscr{C}_5^+$ is the only outgoing characteristic at infinity, suggesting that one degree of freedom can be allowed in the asymptotic behaviour at large x. This justifies the finding in § 4.2 that only one family of decaying solutions can be used to construct an asymptotic representation which uniformly satisfies the radiation condition.

At $x = x_{b0}^+$, $\mathscr{C}_2^+$, $\mathscr{C}_3^+$ and $\mathscr{C}_4^+$ propagate the solutions (u_1, η_1, l_1 and H_1) in the shoaling zone to the breaking line. These are to be matched with the solutions at $x = x_{b0}^-$. For the outgoing curve $\mathscr{C}_5^+$, a condition on z_5 at $x = x_{b0}^+$ is required, which involves u_1, η_1 and H_1. Such information is provided by the outgoing curve $\mathscr{C}_3^-$ at $x = x_{b0}^-$. Note

that H_1 is linearly related to η_1 at the onset of the breaking, and continues to be so afterward. Since $x = x_{b0}^+$ is a characteristic curve, no condition is applied to z_1. In view of the characteristics at $x = x_{b0}^-$ and at $x = x_{b0}^+$, we conclude that three linearly independent matching conditions can be formulated at the breaking line, concerning u_1, η_1 and l_1.

REFERENCES

Arthur, R. S. 1962 A note on the dynamics of rip currents. *J. Geophys. Res.* **67**, 2777–2779.

Bowen, A. J. 1969 Rip currents, 1. The theoretical investigations. *J. Geophys. Res.* **74**, 5467–5478.

Bowen, A. J. & Inman, D. L. 1969 Rip currents, 2. Laboratory and field observations. *J. Geophys. Res.* **74**, 5479–5490.

Courant, R. & Hilbert, D. 1962 *Methods of Mathematical Physics*, vol. II. Interscience.

Dalrymple, R. A. & Lozano, C. T. 1978 Wave-current interaction models for rip currents. *J. Geophys. Res.* **83**, 6063–6071.

Damgaard, J., Dodd, N., Hall, L. & Chesher, T. 2002 Morphodynamic modelling of rip channel growth. *Coastal Engng* **45**, 199–221.

Falqués, A., Montoto, A. & Vila, D. 1999 A note on hydrodynamic instabilities and horizontal circulations in the surf zone. *J. Geophys. Res.* **104**(C9), 20605–20615.

Hino, M. 1974 Theory on formation of rip-current and cuspidal coast. Copenhagen, Den. Paper presented at *14th Conference on Coastal Engineering,* ASCE.

Huntley, D. A. & Short, A. D. 1992 On the spacing between observed rip currents. *Coastal Engng* **17**, 211–225.

Iwata, N. 1976 Rip current spacing. *J. Oceanogr. Soc. Japan* **32**, 1–10.

LeBlond, P. H. & Tang, C. L. 1974 On energy coupling between waves and rip currents. *J. Geophys. Res.* **79**, 811–816.

Lippmann, T. C., Brookins, A. H. & Thornton, E. B. 1996 Wave energy transformation on natural profiles. *Coastal Engng* **27**, 1–20.

Longuet-Higgins, M. S. 1970 Longshore currents generated by obliquely incident sea waves, 2. *J. Geophys. Res.* **75**, 6790–6801.

Longuet-Higgins, M. S. & Stewart, R. W. 1960 The changes in the form of short gravity waves on long waves and tidal currents. *J. Fluid Mech.* **8**, 565–583.

Longuet-Higgins, M. S. & Stewart, R. W. 1961 The changes in amplitude of short gravity waves on steady non-uniform currents. *J. Fluid Mech.* **10**, 529–549.

Longuet-Higgins, M. S. & Stewart, R. W. 1962 Radiation stress and mass transport in gravity waves, with application to 'surf-beats'. *J. Fluid Mech.* **13**, 481–504.

Mei, C. C. 1989 *The Applied Dynamics of Ocean Surface Waves*. World Scientific.

Mizuguchi, M. 1976 Eigenvalue problems for rip current spacing (in Japanese). *Trans. Japan Soc. Civil Eng.* **248**, 83–88.

Murray, A. B. & Reydellet, G. 2001 A rip-current model based on a hypothesized wave/current interaction. *J. Coastal Res.* **17**, 517–530.

Shepard, F. P., Emery, K. O. & LaFond, E. C. 1941 Rip currents: A process of geological importance. *J. Geol.* **49**(4), 337–369.

Shepard, F. P. & Inman, D. L. 1951 Nearshore circulation. *Proc. 1st Conf. Coastal Engng*, pp. 50–59, University of California, Council on Wave Research.

Short, A. D. 1985 Rip-current type, spacing and persistence, Narrabeen beach, Australia. *Mar. Geo.* **65**, 47–71.

Short, A. (Ed.) 1999 *Handbook of Beach and Shoreface Morphodynamics*. Wiley.

Yu, J. & Slinn, D. N. 2003 Effects of wave-current interaction on rip currents. *J. Geophys. Res.* **108**(C3), 3088, doi:10.1029/2001JC001105.