J. Fluid Mech. (1998), *vol.* 355, *pp.* 113–138. Printed in the United Kingdom

On the chance of freak waves at sea

By BENJAMIN S. WHITE[1] AND BENGT FORNBERG[2]

[1] Exxon Research and Engineering Co., Route 22 East, Annandale, NJ 08801, USA

[2] Department of Applied Mathematics, University of Colorado at Boulder, Campus Box 526, Boulder, CO 80309-0526, USA

(Received 28 October 1996 and in revised form 4 July 1997)

When deep-water surface gravity waves traverse an area with a curved or otherwise variable current, the current can act analogously to an optical lens, to focus wave action into a caustic region. In this region, waves of surprisingly large size, alternatively called freak, rogue, or giant waves are produced. We show how this mechanism produces freak waves at random locations when ocean swell traverses an area of random current. When the current has a constant (possibly zero) mean with small random fluctuations, we show that the probability distribution for the formation of a freak wave is universal, that is, it does not depend on the statistics of the current, but only on a single distance scale parameter, provided that this parameter is finite and non-zero. Our numerical simulations show excellent agreement with the theory, even for current standard deviation as large as 1.0 m s^{-1}. Since many of these results are derived for arbitrary dispersion relations with certain general properties, they include as a special case previously published work on caustics in geometrical optics.

1. Introduction

Waves of surprisingly large size, alternatively called freak, rogue, or giant waves, are a well-documented hazard to mariners. Perhaps the most celebrated incident occurred during the world's first solo circumnavigation, when, in 1896, well off the Patagonian coast, the *Spray*'*s* hull was completely submerged by a giant wave as Captain Joshua Slocum (1899) clung to the peak halyards. Captain Mallory (1974) analysed eleven more-recent incidents, off the south-east coast of South Africa, of freak waves which caused damage to large vessels, including one ship which was cleaved in half. All these occurrences were in an area renowned for producing freak waves when a large ocean swell opposes the swift Agulhas current.

Peregrine (1976) suggested that, in areas of strong current such as the Agulhas, abnormally large waves could be produced when wave action is concentrated by refraction into a caustic region. In this scenario, a curved or otherwise variable current acts analogously to an optical lens to focus wave action. Gerber (1996) applied this idea, and the theory of Gerber (1993), to explain the large waves encountered in the Agulhas current, examining in particular the 1986 incident involving the semi-submersible *Actinia*. A related theory was given by Gutshabash & Lavrenov (1986). Irvine & Tilley (1988) analysed Synthetic Aperture Radar (SAR) data of the Agulhas current, and concluded that caustics caused by meanders in the Agulhas could produce giant waves.

Further support for the caustic theory of giant waves is given by Smith (1976), who calculated the shape of a wave near a caustic, and produced an asymmetry, as is reported by mariners. Mallory (1974) describes freak waves as having a steeper forward face preceded by a deep trough, or 'hole in the sea'.

We adopt here the nomenclature of Bacon (1991, in the classic text of Coles), to distinguish between extreme waves, representing the tail of some typical statistical distribution of wave heights (generally a Rayleigh distribution), and freaks, 'defined as waves of a height occurring more often than would be expected from the "background" probability distribution'. Note however that not all authorities accept this distinction, e.g. Van Dorn (1993). Although Sand *et al.* (1990), in their analysis of waves on the Danish Continental Shelf, do not use this nomenclature, they do confirm the existence of large waves that Bacon would call freak, rather than extreme. These waves, in their words, 'do not belong to the traditional short term statistical distributions used for ocean waves. The waves are too high, too asymmetric and too steep.'

The expected structure of extreme waves was calculated theoretically by Phillips, Gu & Donelan (1993*a*), who obtained good agreement with data, as did Phillips, Gu & Walsh (1993*b*). In contrast to Smith's (1976) analysis of giant waves, this expected structure of extreme waves is symmetric. A more definitive mathematical theory applicable to determining the structure of extreme waves is given by Lindgren (1970).

In this paper, we compute the probability distribution for the formation of a freak wave when a regular ocean swell traverses a region of deep water with random current fluctuations. It is assumed that freak waves are produced by caustics resulting from ray focusing when the swell interacts with the random current. Under wide hypotheses, when the mean current is constant (possibly zero) and the random current fluctuations are small, the probability distribution for the formation of a freak is universal, i.e. it does not depend on the details of the current distribution, but is described by a universal mathematical form controlled by a single distance scale parameter, provided that this parameter is finite and non-zero. The theory is verified with Monte Carlo simulations, and good agreement is obtained, even when the standard deviation of the current is as large as 1 m s^{-1}. Since many of the main results are derived for arbitrary ray systems generated by two-dimensional dispersion relations, they include, for example, previously published results on caustics in geometrical optics.

Note that in this model caustics occupy a fixed position in space near which freak waves are generated. Conditions like this were encountered by the ketch *Tzu Hang* when she was pitchpoled in the South Pacific, 900 miles offshore, in the first of her two unsuccessful attempts to round Cape Horn (Smeeton 1959). Henderson (1991) postulated the existence of uncharted seamounts to explain crew member John Guzzwell's description of the sea state as similar to 'what would be encountered with a long swell passing over a shoal area'. A caustic region is an alternative explanation. Peregrine (1976) suggested that a better description of wave behaviour in a caustic region might be given by considering the behaviour of a short wave group, that, he surmised, might 'show just one or two large waves persisting for a limited time'.

To study the focusing of waves, we introduce, in §2, the ray theory for water waves (Longuet-Higgins & Stewart 1961; Whitham 1974; Peregrine 1976), in which wave rays are refracted by a spatially varying current. In this theory, a singularity in the ray field develops when two rays, which are initially (infinitesimally) close, coalesce. A caustic is the locus of such points, where rays are pinched together, and large amplitudes occur. In numerical simulations, the location of a caustic may be detected by monitoring the 'Jacobian', or equivalently the 'raytube area', which is the distance between two initially close rays, normalized by their initial separation. A caustic is then determined as the locus of points in space where the raytube area vanishes.

In §2 we also derive new equations for the propagation of raytube area (or, equivalently, the Jacobian) along a ray. This is a new development in the ray theory for water waves, with many possible applications beyond that of determining the location of caustics, for which these equations are utilized here. In the assessment of Irvine & Tilley (1988), the Jacobian 'is the most important factor governing energy density in situations dominated by refraction'. Furthermore, we derive these equations in great generality, for arbitrary dispersion relations in two dimensions, so that they are also applicable to a wide range of wave propagation problems other than those of surface gravity waves in deep water. In this way, our equations can be seen as a generalization of the raytube area equations that are known for non-dispersive waves (see, e.g. Kulkarny & White 1982).

The principle of conservation of wave action (CWA) (Longuet-Higgins & Stewart 1961; Whitham 1974; Peregrine 1976) may be used, in conjunction with ray theory, to determine wave amplitudes at points in space that are not near a caustic. For completeness, and to demonstrate the amplitude singularity at a caustic, we use CWA in §2 to express the wave amplitude in terms of the raytube area. However, all other results in this paper, including the raytube area equations and the universal rogue probability curve, are independent of the assumption of CWA. This point is important because CWA has, in general, only been demonstrated mathematically for irrotational currents and we intend to apply our results for example to random eddy fields with various statistics. Since our calculations are only concerned with the location of caustics, and not their amplitude, our only assumption is that an amplitude singularity results from the geometric one. That is, we assume, as the other authors cited above, that freak waves are associated with caustics. A further discussion of wave amplitudes and related matters is contained in the Appendix.

In §3 we consider the propagation of a ray and its associated raytube area through a region of random current. It is assumed that the current has a constant (possibly zero) mean and that the random fluctuations have a standard deviation σ which is much smaller than the phase velocity, c_p, of the waves. The random fluctuations are assumed to have an intrinsic length scale $\tilde{l}$, and to satisfy a 'mixing property', roughly, that correlations between any two points decay rapidly to zero when the two points are separated by large distances in space. The equations are scaled for the proper distance scale to see caustics developing, a propagation distance of $O((\sigma/c_p)^{-2/3}\tilde{l})$.

Application of the general limit theorem of Papanicolaou & Kohler (1974) next yields an approximation in terms of a diffusion Markov process of a particularly simple form. Actually, we demonstrate this limit not for surface gravity waves only but, again, for an arbitrary ray system generated by any two-dimensional dispersion relation with certain general properties. Because of this, our results contain as a special case previously published work on geometrical optics, or acoustics (Kulkarny & White 1982; White, Nair & Bayliss 1988; Klyatskin 1993). In addition, this previous work can be used to deduce further properties of the limit process. It is this substantial generality of the limit, which does not depend on the details of the random fluctuations, or, for that matter, on the precise form of the dispersion relation, that provides universality.

In §4 we examine the universal limit, and derive more explicit formulas for the relevant parameters. In particular, we study a model for two-dimensional incompressible flow with a random stream function, which is the model used in the Monte Carlo simulations of §5.

Several effects are not included in our mathematical model. (i) Swell is represented as a regular time-harmonic wave train, so that a more realistic wave variability is neglected. (ii) As discussed above, we do not attempt to predict the actual wave heights, but only the caustic locations. Thus nonlinear effects are neglected as are the effects of wave breaking and other dissipation mechanisms. (iii) We do not consider nonlinear instabilities, or the possible defocusing effects of nonlinearities. (iv) We do not consider the generation of waves by wind forces, a subject with an extensive literature (Komen *et al.* 1994). We consider only the focusing of wave action, using linear monochromatic theory, after the swell has been generated. Some of these other topics are discussed further in the Appendix.

This manuscript, submitted in the centennial year of Slocum's Great Wave, is our contribution to the Joshua Slocum Centennial festivities of 1995–8, commemorating his historic voyage. Whether a current-induced caustic caused the Great Wave is, of course, uncertain. However, it is plausible, since Slocum was near an area of what might be considered random current, the notorious tide-races. He had gone well offshore because 'Hoping that she might go clear of the destructive tide-races, the dread of big craft or little along this coast, I gave all the capes a berth of about fifty miles, for these dangers extend many miles from the land. But where the sloop avoided one danger she encountered another', that is, the rogue wave (Slocum 1899). So it is possible that the danger Slocum encountered was another manifestation of the one he had sought to avoid.

2. Propagation equations for raytube area

The dispersion relation for surface gravity waves in the presence of a constant current $\boldsymbol{U} \in \mathscr{R}^2$ is (Peregrine 1976)

$$\Omega = \bar{\Omega}(|\boldsymbol{k}|) + \boldsymbol{k} \cdot \boldsymbol{U}, \tag{2.1}$$

where

$$\bar{\Omega}(|\boldsymbol{k}|) = \pm (g|\boldsymbol{k}|)^{1/2} \tag{2.2}$$

is the dispersion relation with no current, and g is the acceleration due to gravity. For $\boldsymbol{U} = \boldsymbol{U}(\boldsymbol{x})$ a slowly varying function of $\boldsymbol{x} \in \mathscr{R}^2$, an approximate theory (Whitham 1974; Peregrine 1976) provides a generalization for the phase, ϕ, of a wave, which in the case of a constant current is of the form $\phi = \boldsymbol{k} \cdot \boldsymbol{x} - \omega t$, with $\omega = \Omega(\boldsymbol{k})$. A local phase, $\phi(t, \boldsymbol{x})$, is constructed from the local frequency $\omega = -\phi_t$ and the local wavenumber $\boldsymbol{k} = \boldsymbol{\nabla}\phi \equiv \phi_{\boldsymbol{x}}$, by using the dispersion relation locally, so that

$$\phi_t + \Omega(\boldsymbol{x}, \phi_{\boldsymbol{x}}) = 0. \tag{2.3}$$

The amplitude, a, of the wave is determined by conservation of wave action

$$\frac{\partial}{\partial t}\left(\frac{a^2}{\bar{\Omega}}\right) + \boldsymbol{\nabla} \cdot \left(\frac{a^2}{\bar{\Omega}} \Omega_k\right) = 0. \tag{2.4}$$

Consider the steady problem of ocean swells of constant frequency ω incident on a region of random current. Putting $\phi(t, \boldsymbol{x}) = -\omega t + \hat{\phi}(\boldsymbol{x})$, and dropping hats, (2.3) and (2.4) become

$$\Omega(\boldsymbol{x}, \phi_{\boldsymbol{x}}) = \omega, \tag{2.5}$$

$$\boldsymbol{\nabla} \cdot \left(\frac{a^2}{\bar{\Omega}} \Omega_k\right) = 0. \tag{2.6}$$

Let $\phi = \phi_0(\boldsymbol{x})$ be given along some initial curve $\boldsymbol{x}_0(\alpha)$ parametrized by arclength α. In particular, an initially plane wavefront is analysed below, so that ϕ_0 is constant along a straight line $\boldsymbol{x}_0(\alpha)$. Then (2.5) can be solved by the method of characteristics (Courant & Hilbert 1962). The rays are the characteristic curves $\bar{\boldsymbol{x}}(t, \alpha), \bar{\boldsymbol{k}}(t, \alpha) \in \mathscr{R}^2$, where α denotes the starting point of the ray on $\boldsymbol{x}_0(\alpha)$, and the parameter along the ray, which has the dimension of time, is denoted by t. The rays satisfy the characteristic equations

$$\frac{\partial \bar{\boldsymbol{x}}(t,\alpha)}{\partial t} = \Omega_k(\bar{\boldsymbol{x}},\bar{\boldsymbol{k}}), \quad \frac{\partial \bar{\boldsymbol{k}}(t,\alpha)}{\partial t} = -\Omega_x(\bar{\boldsymbol{x}},\bar{\boldsymbol{k}}), \tag{2.7}$$

with the initial condition

$$\bar{\boldsymbol{x}}(0,\alpha) = \boldsymbol{x}_0(\alpha), \quad \bar{\boldsymbol{k}}(0,\alpha) = \phi_{0,x}(\boldsymbol{x}_0(\alpha)). \tag{2.8}$$

Then it can be shown that

$$\bar{\boldsymbol{k}}(t,\alpha) = \phi_x(\bar{\boldsymbol{x}}(t,\alpha)) \tag{2.9}$$

and $\phi(\bar{\boldsymbol{x}}(t,\alpha)) = \bar{\phi}(t,\alpha)$, where $\bar{\phi}$ satisfies

$$\frac{\partial \bar{\phi}(t,\alpha)}{\partial t} = \boldsymbol{k}\cdot\Omega_k, \quad \bar{\phi}(0,\alpha) = \phi_0(\boldsymbol{x}_0(\alpha)). \tag{2.10}$$

Auxiliary equations, which are convenient for determining the amplitude, a, will now be derived. The unit tangent along a ray is

$$\boldsymbol{e}_1(t,\alpha) = \frac{\Omega_k(\bar{\boldsymbol{x}},\bar{\boldsymbol{k}})}{|\Omega_k(\bar{\boldsymbol{x}},\bar{\boldsymbol{k}})|}. \tag{2.11}$$

Let $\boldsymbol{e}_2 \perp \boldsymbol{e}_1$ be the unit normal, so that $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ are right-handed. Then

$$\frac{\partial \boldsymbol{e}_1}{\partial t} = |\Omega_k|\,\kappa \boldsymbol{e}_2, \quad \frac{\partial \boldsymbol{e}_2}{\partial t} = -|\Omega_k|\,\kappa \boldsymbol{e}_1, \tag{2.12}$$

where

$$\kappa = \frac{1}{|\Omega_k|}\left(\boldsymbol{e}_2^T\cdot\Omega_{kx}\cdot\boldsymbol{e}_1 - \boldsymbol{e}_2^T\cdot\Omega_{kk}\cdot\frac{\Omega_x}{|\Omega_k|}\right) \tag{2.13}$$

is the ray curvature. Here Ω_{kx} is the 2×2 matrix with ijth entry $\partial^2\Omega/(\partial k_i\,\partial x_j)$, and a similar notation is used for Ω_{kk}, Ω_{xx} and $\Omega_{xk} = \Omega_{kx}^T$.

Let

$$\gamma = (t,\alpha)^T \tag{2.14}$$

so that $\bar{\boldsymbol{x}}_\gamma$ is the 2×2 matrix of derivatives of the transformation from ray coordinates to physical space. From (2.7), (2.11) the Jacobian, J, of the transformation can be obtained:

$$J = \det(\bar{\boldsymbol{x}}_\gamma) = |\Omega_k| A, \tag{2.15}$$

where

$$A = \boldsymbol{e}_2^T\cdot\bar{\boldsymbol{x}}_\alpha \tag{2.16}$$

is the raytube area. That is, $A(t,\alpha)\,\mathrm{d}\alpha$ is the distance of the point $\bar{\boldsymbol{x}}(t,\alpha)$ from the infinitesimally close ray $\boldsymbol{x}(\cdot,\alpha+\mathrm{d}\alpha)$. Propagation equations for $\bar{\boldsymbol{x}}_\gamma, \bar{\boldsymbol{k}}_\gamma$ along a ray are obtained by differentiation of (2.7):

$$\frac{\partial}{\partial t}\bar{\boldsymbol{x}}_\gamma = \Omega_{kx}\cdot\bar{\boldsymbol{x}}_\gamma + \Omega_{kk}\cdot\bar{\boldsymbol{k}}_\gamma, \quad \frac{\partial}{\partial t}\bar{\boldsymbol{k}}_\gamma = -\Omega_{xx}\cdot\bar{\boldsymbol{x}}_\gamma - \Omega_{xk}\cdot\bar{\boldsymbol{k}}_\gamma. \tag{2.17}$$

From (2.17) a propagation equation can be obtained for J in terms of

$$\boldsymbol{k}_x(\bar{\boldsymbol{x}}(t,\alpha)) = \phi_{xx}(\bar{\boldsymbol{x}}(t,\alpha)) = \bar{\boldsymbol{k}}_\gamma(t,\alpha))\,\bar{\boldsymbol{x}}_\gamma^{-1}(t,\alpha). \tag{2.18}$$

Thus

$$\frac{1}{J}\frac{\partial}{\partial t}J = \mathrm{Trace}\,\{\Omega_{kx} + \Omega_{kk}\,\boldsymbol{k}_x). \tag{2.19}$$

From (2.6), (2.7)

$$\frac{\partial}{\partial t}\left(\frac{a^2(\bar{\boldsymbol{x}}(t,\alpha))}{\bar{\Omega}(\phi_x(\bar{\boldsymbol{x}}(t,\alpha)))}\right) = -\left(\frac{\alpha^2}{\bar{\Omega}}\right)\left(\boldsymbol{\nabla}_x\cdot\Omega_k\right)\Big|_{x=\bar{x}} = -\left(\frac{a^2}{\bar{\Omega}}\right)\mathrm{Trace}\,\{\Omega_{kx}+\Omega_{kk}\,\boldsymbol{k}_x\}. \quad (2.20)$$

From (2.19), (2.20), (2.15) the amplitude is written in terms of the raytube area:

$$a(t,\alpha) = a^0\left(\frac{\bar{\Omega}}{\bar{\Omega}^0}\frac{|\Omega_k^0|}{|\Omega_k|}\frac{A^0}{A}\right)^{1/2}, \quad (2.21)$$

where superscript 0 denotes values at $t = 0$.

Propagation equations for A along a ray can now be derived. From (2.17), (2.18) a matrix Riccati equation is obtained for $\boldsymbol{k}_x$:

$$\frac{\partial}{\partial t}\boldsymbol{k}_x(\bar{\boldsymbol{x}}(t,\alpha)) = -\Omega_{xx}-\Omega_{xk}\,\boldsymbol{k}_x-\boldsymbol{k}_x\,\Omega_{kx}-\boldsymbol{k}_x\,\Omega_{kk}\,\boldsymbol{k}_x. \quad (2.22)$$

Let $\boldsymbol{Q}$ be the 2×2 matrix with $\boldsymbol{e}_j, j = 1,2$ as its jth column. Changing to the $\boldsymbol{e}_1, \boldsymbol{e}_2$ basis, and using symmetry of $\boldsymbol{k}_x$ gives

$$\hat{\boldsymbol{k}}_x = \boldsymbol{Q}^T\boldsymbol{k}_x\,\boldsymbol{Q} = \begin{bmatrix} -\boldsymbol{e}_1^T\cdot\dfrac{\Omega_x}{|\Omega_k|} & -\boldsymbol{e}_2^T\cdot\dfrac{\Omega_x}{|\Omega_k|} \\ -\boldsymbol{e}_2^T\cdot\dfrac{\Omega_x}{|\Omega_k|} & R \end{bmatrix} \quad (2.23)$$

where

$$R = \boldsymbol{e}_2^T\cdot\boldsymbol{k}_x\cdot\boldsymbol{e}_2 = \frac{B}{A} \quad (2.24)$$

and

$$B = \boldsymbol{e}_2^T\cdot\frac{\Omega_x}{|\Omega_k|}\cdot\boldsymbol{e}_1^T\,\bar{\boldsymbol{x}}_\alpha+\boldsymbol{e}_2^T\cdot\bar{\boldsymbol{k}}_\alpha. \quad (2.25)$$

Change of basis in (2.22) and use of (2.12) and (2.23) yields a scalar Riccati equation for R:

$$\frac{\partial R}{\partial t} = \mu_1-2\mu_2\,R-\mu_3\,R^2, \quad (2.26)$$

where

$$\left.\begin{aligned} \mu_1 = {}& -\boldsymbol{e}_2^T\cdot\Omega_{xx}\cdot\boldsymbol{e}_2+2\frac{\boldsymbol{e}_2^T\,\Omega_x}{|\Omega_k|}[\boldsymbol{e}_1^T\cdot\Omega_{kx}\cdot\boldsymbol{e}_2+\boldsymbol{e}_2^T\cdot\Omega_{kx}\cdot\boldsymbol{e}_1] \\ & -2\left(\frac{\boldsymbol{e}_1^T\cdot\Omega_x}{|\Omega_k|}\right)\left(\frac{\boldsymbol{e}_2^T\cdot\Omega_x}{|\Omega_k|}\right)\boldsymbol{e}_1^T\cdot\Omega_{kk}\cdot\boldsymbol{e}_2-\left(\frac{\boldsymbol{e}_2^T\cdot\Omega_x}{|\Omega_k|}\right)^2[\boldsymbol{e}_1^T\Omega_{kk}\cdot\boldsymbol{e}_1+2\boldsymbol{e}_2^T\cdot\Omega_{kk}\cdot\boldsymbol{e}_2], \\ \mu_2 = {}& -\boldsymbol{e}_1^T\cdot\Omega_{kk}\cdot\boldsymbol{e}_2\left(\frac{\boldsymbol{e}_2^T\cdot\Omega_x}{|\Omega_k|}\right)+\boldsymbol{e}_2^T\cdot\Omega_{kx}\cdot\boldsymbol{e}_2 \quad \mu_3 = \boldsymbol{e}_2^T\,\Omega_{kk}\,\boldsymbol{e}_2. \end{aligned}\right\} \quad (2.27)$$

By differentiating (2.5) with respect to α an expression for $\boldsymbol{e}_1^T\cdot\bar{\boldsymbol{k}}_\alpha$ is obtained, which, when combined with (2.25), gives

$$\bar{\boldsymbol{k}}_\alpha = -A\frac{\boldsymbol{e}_2^T\cdot\Omega_x}{|\Omega_k|}\boldsymbol{e}_1+B\boldsymbol{e}_2-\boldsymbol{e}_1^T\cdot\bar{\boldsymbol{x}}_\alpha\frac{\Omega_x}{|\Omega_k|}. \quad (2.28)$$

Differentiation of (2.16), and use of (2.28) yields

$$\frac{\partial A}{\partial t} = \mu_2\,A+\mu_3\,B. \quad (2.29)$$

Differentiating $B = RA$ and use of (2.26) and (2.29) yields

$$\frac{\partial B}{\partial t} = \mu_1 A - \mu_2 B. \tag{2.30}$$

To summarize, the ray position $\bar{\boldsymbol{x}}$ and local wavenumber $\bar{\boldsymbol{k}}$ are determined by the ray equations (2.7), which are four nonlinear scalar equations. The raytube area A and auxiliary variable B can then be computed along a ray from the two additional scalar linear equations (2.29) and (2.30), with coefficients (2.27). The amplitude a is then determined from (2.21). The derivatives of Ω used in (2.7) and (2.27) are

$$\left.\begin{aligned}
&(\Omega_{\boldsymbol{k}})_i = \pm\frac{g^{1/2}}{2|\boldsymbol{k}|^{3/2}}k_i + U_i, \quad (\Omega_{\boldsymbol{x}})_i = \sum_{j=1}^{2} k_j \frac{\partial U_j}{\partial x_i}, \quad (\Omega_{\boldsymbol{xx}})_{ij} = \sum_{l=1}^{2} k_l \frac{\partial^2 U_l}{\partial x_i \partial x_j},\\
&(\Omega_{\boldsymbol{kx}})_{ij} = (\Omega_{\boldsymbol{xk}})_{ji} = \frac{\partial U_i}{\partial x_j}, \quad (\Omega_{\boldsymbol{kk}})_{ij} = \pm\left(\frac{g^{1/2}}{2|\boldsymbol{k}|^{3/2}}\delta_{ij} - \frac{3g^{1/2}}{4|\boldsymbol{k}|^{7/2}}k_i k_j\right).
\end{aligned}\right\} \tag{2.31}$$

3. A stochastic limit

To determine the distance scale over which caustics may occur, the dispersion relation (2.5) is first non-dimensionalized. Let $\tilde{l}$ be a typical length and $\tilde{k}$ a typical (scalar) wavenumber. For the validity of ray theory $\tilde{k}\tilde{l}$ must be large. Non-dimensional position, wavenumber and phase are defined by

$$\boldsymbol{x}' = \frac{\boldsymbol{x}}{\tilde{l}}, \quad \boldsymbol{k}' = \frac{\boldsymbol{k}}{\tilde{k}}, \quad \phi' = \frac{\phi}{\tilde{k}\tilde{l}}. \tag{3.1}$$

Note that if $\boldsymbol{k} = \phi_{\boldsymbol{x}}$, then $\boldsymbol{k}' = \phi'_{\boldsymbol{x}'}$. Define the non-dimensional dispersion relation by

$$\Omega'(\boldsymbol{x}', \boldsymbol{k}') = \Omega(\tilde{l}\boldsymbol{x}', \tilde{k}\boldsymbol{k}')/\omega. \tag{3.2}$$

Then the dispersion (2.5) holds for the primed variables, with $\omega' = 1$.

More specifically, consider (2.1), (2.2), where $\boldsymbol{U} = \boldsymbol{U}(\boldsymbol{x}/\tilde{l})$ is a random function of position, and $\tilde{l}$ is an intrinsic length scale. For given ω, let

$$\tilde{k} = \omega^2/g, \quad c_p^0 = \omega/\tilde{k} \tag{3.3}$$

be the wavenumber and phase velocity, respectively, in the absence of current, and let

$$\boldsymbol{U}' = \boldsymbol{U}/c_p^0. \tag{3.4}$$

Then (2.1), (2.2) hold for the primed variables, with $g' = 1$. For notational convenience, primes will be dropped in what follows.

Consider propagation across a region of current with small random fluctuations

$$\boldsymbol{U}(\boldsymbol{x}) = \boldsymbol{U}_0 + \sigma\hat{\boldsymbol{U}}(\boldsymbol{x}), \tag{3.5}$$

where $\boldsymbol{U}_0$ is a constant, non-random mean current, and $\sigma\hat{\boldsymbol{U}}(\boldsymbol{x})$ is a mean zero homogeneous random field with standard deviation σ, which is assumed small. We will follow a ray, with its associated raytube area, to determine the probability of a caustic developing (i.e. $A = 0$) within a give propagation distance.

The main tool used below is the probabilistic limit theorem of Papanicolaou & Kohler (1974). This theorem applies to random ordinary differential equations of the form

$$\left.\begin{aligned}
\frac{\mathrm{d}\boldsymbol{W}}{\mathrm{d}\tau} &= \frac{1}{\epsilon}\boldsymbol{H}_1\left(\frac{\tau}{\epsilon^2}, \boldsymbol{W}\right) + \boldsymbol{H}_2\left(\frac{\tau}{\epsilon^2}, \boldsymbol{W}\right),\\
\boldsymbol{W}(0) &= \boldsymbol{W}_0,
\end{aligned}\right\} \tag{3.6}$$

where $\boldsymbol{W} \in \mathscr{R}^d$, ε is a small parameter, and for each fixed, non-random value $\bar{\boldsymbol{W}}$ of $\boldsymbol{W}$, $\boldsymbol{H}_1(t, \bar{\boldsymbol{W}})$ and $\boldsymbol{H}_2(t, \bar{\boldsymbol{W}})$ are random functions of t satisfying a 'mixing condition'. Roughly, this means that $(\boldsymbol{H}_1(t, \bar{\boldsymbol{W}}), \boldsymbol{H}_2(t, \bar{\boldsymbol{W}}))$ become asymptotically independent of $(\boldsymbol{H}_1(t+t', \bar{\boldsymbol{W}}), \boldsymbol{H}_2(t+t', \bar{\boldsymbol{W}}))$ as the time difference t' becomes large. Furthermore, for a sensible limit of (3.6), $\boldsymbol{H}_1(t, \bar{\boldsymbol{W}})$ must have mean zero. Then as $\varepsilon \downarrow 0$, the solution $\boldsymbol{W}$ of (3.6) converges (weakly) to a diffusion Markov process, with infinitesimal generator (Kolmogorov backward operator)

$$\mathrm{L} = \sum_{i,j=1}^{d} a^{ij}(\boldsymbol{W}) \frac{\partial^2}{\partial W^i \partial W^j} + \sum_{i=1}^{d} b^i(\boldsymbol{W}) \frac{\partial}{\partial W^i}. \tag{3.7}$$

Formulas for a^{ij}, b^i are as follows:

$$\left.\begin{aligned} a^{ij}(\bar{\boldsymbol{W}}) &= \lim_{T\to\infty} \frac{1}{T} \int_t^{t+T} \int_t^{s_1} \langle H_1^i(s_1, \bar{\boldsymbol{W}}) H_1^j(s_2, \bar{\boldsymbol{W}}) \rangle \, \mathrm{d}s_2 \, \mathrm{d}s_1, \\ b^i(\bar{\boldsymbol{W}}) &= \lim_{T\to\infty} \frac{1}{T} \int_t^{t+T} \langle H_2^i(s_1, \bar{\boldsymbol{W}}) \rangle \, \mathrm{d}s_1 \\ &\quad + \lim_{T\to\infty} \frac{1}{T} \int_t^{t+T} \int_t^{s_1} \left\langle H_1^j(s_1, \bar{\boldsymbol{W}}) \frac{\partial H_1^i(s_2, \bar{\boldsymbol{W}})}{\partial \bar{W}^j} \right\rangle \mathrm{d}s_2 \, \mathrm{d}s_1, \end{aligned}\right\} \tag{3.8}$$

where $\langle \cdot \rangle$ denotes probability average.

The operator L determines the statistics of the limiting Markov process through, for example, partial differential equations for its probability density. In particular, let $\boldsymbol{\Sigma}$ be the square root of the symmetric part of the matrix $\boldsymbol{a}$ whose components are given in (3.8), i.e. $\boldsymbol{\Sigma}$ satisfies

$$\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T = \tfrac{1}{2}(\boldsymbol{a} + \boldsymbol{a}^T). \tag{3.9}$$

Let $\boldsymbol{b}$ be the vector with components given in (3.9), and let $\boldsymbol{\beta}$ be a vector of independent Brownian motions. Then the limit for $\boldsymbol{W}$ can be characterized as the solution of the Ito stochastic differential equations (white noise equations)

$$\mathrm{d}\boldsymbol{W} = \boldsymbol{b}(\boldsymbol{W}) \, \mathrm{d}t + \surd 2 \boldsymbol{\Sigma}(\boldsymbol{W}) \, \mathrm{d}\boldsymbol{\beta}. \tag{3.10}$$

To apply these general formulas to the case of water waves, let

$$\varepsilon = \sigma^{1/3}, \quad F(\boldsymbol{k}) = (g|\boldsymbol{k}|)^{1/2} + \boldsymbol{k} \cdot \boldsymbol{U}_0, \quad G(\boldsymbol{x}, \boldsymbol{k}) = \boldsymbol{k} \cdot \hat{\boldsymbol{U}}(\boldsymbol{x}), \tag{3.11}$$

so that

$$\Omega(\boldsymbol{x}, \boldsymbol{k}) = F(\boldsymbol{k}) + \varepsilon^3 G(\boldsymbol{x}, \boldsymbol{k}), \tag{3.12}$$

where $G(\boldsymbol{x}, \boldsymbol{k})$ is, for fixed $\boldsymbol{k}$, a homogeneous random field, i.e., its statistics are translation invariant. The limit theorem will be applied to general dispersion relations of the form (3.12). Thus the results will be applicable not only to water waves, but, for example, to geometrical optics in a random medium, where $F(\boldsymbol{k}) = |\boldsymbol{k}|$, $G(\boldsymbol{x}, \boldsymbol{k}) = |\boldsymbol{k}| \hat{c}(\boldsymbol{x})$, $\sigma \hat{c}(\boldsymbol{x})$ is the mean zero random perturbation in the local propagation speed (which has mean one), σ is the standard deviation, and $\varepsilon = \sigma^{1/3}$.

The ray equations may be written on a long propagation distance scale of $O(\sigma^{-2/3})$ by defining a scaled time as

$$\tau = \varepsilon^2 t = \sigma^{2/3} t. \tag{3.13}$$

Substitution of (3.12), (3.13) into (2.7) gives

$$\left.\begin{aligned} \frac{\mathrm{d}\bar{\boldsymbol{x}}}{\mathrm{d}\tau} &= \frac{1}{\varepsilon^2} F_{\boldsymbol{k}}(\bar{\boldsymbol{k}}) + \varepsilon G_{\boldsymbol{k}}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{k}}), \quad \bar{\boldsymbol{x}}(0) = \boldsymbol{x}_0, \\ \frac{\mathrm{d}\bar{\boldsymbol{k}}}{\mathrm{d}\tau} &= -\varepsilon G_{\boldsymbol{x}}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{k}}), \quad \bar{\boldsymbol{k}}(0) = \boldsymbol{k}_0, \end{aligned}\right\} \tag{3.14}$$

where the α dependence has been suppressed, to write ordinary differential equations.

From (3.14) $\bar{\boldsymbol{k}}$ deviates at most $O(\varepsilon)$ from $\boldsymbol{k}_0$, and will be expanded about

$$\hat{\boldsymbol{k}} = \lim_{\varepsilon \downarrow 0} \boldsymbol{k}_0. \tag{3.15}$$

$\boldsymbol{k}_0 = \phi_x(\boldsymbol{x}_0(\alpha))$, and hence $\hat{\boldsymbol{k}}$ are parallel to the normal to $\boldsymbol{x}_0(\cdot)$. Let c_g be the group velocity in the absence of random fluctuations, and $\hat{\boldsymbol{e}}_1$ the corresponding direction:

$$c_g = |F_k(\hat{\boldsymbol{k}})| > 0, \quad \hat{\boldsymbol{e}}_1 = \frac{F_k(\hat{\boldsymbol{k}})}{c_g}. \tag{3.16}$$

Let $\hat{\boldsymbol{e}}_2 \perp \hat{\boldsymbol{e}}_1$ so that $\hat{\boldsymbol{e}}_1$, $\hat{\boldsymbol{e}}_2$ are right handed. Let the initial strip for (2.5) be non-characteristic:

$$\cos\psi \equiv \hat{\boldsymbol{e}}_1^T \frac{\hat{\boldsymbol{k}}}{|\hat{\boldsymbol{k}}|} \neq 0. \tag{3.17}$$

Substitution into (3.12) yields

$$\boldsymbol{k}_0 = \hat{\boldsymbol{k}} - \frac{\varepsilon^3}{c_g} \frac{G(\boldsymbol{x}_0, \hat{\boldsymbol{k}})}{\cos\psi} \frac{\hat{\boldsymbol{k}}}{|\hat{\boldsymbol{k}}|} + O(\varepsilon^6). \tag{3.18}$$

Let $\boldsymbol{x}_1$, $\boldsymbol{k}_1$ be defined by

$$\bar{\boldsymbol{x}} = \boldsymbol{x}_0 + c_g \frac{\tau}{\varepsilon^2} \hat{\boldsymbol{e}}_1 + \boldsymbol{x}_1, \quad \bar{\boldsymbol{k}} = \hat{\boldsymbol{k}} + \varepsilon^2 \boldsymbol{k}_1. \tag{3.19}$$

The scaling for $\bar{\boldsymbol{k}}$ is suggested by comparison of (3.14) with (3.6), i.e. $\varepsilon^{-1} G(\bar{\boldsymbol{x}}, \bar{\boldsymbol{k}})$ is expected to make an $O(1)$ contribution to the limit; $\boldsymbol{x}_1$, $\boldsymbol{k}_1$ will be assumed $O(1)$, and determined self-consistently as $\varepsilon \downarrow 0$. Substitution of (3.19) into (3.14) yields

$$\frac{\mathrm{d}\boldsymbol{x}_1}{\mathrm{d}\tau} = F_{kk}(\hat{\boldsymbol{k}})\, \boldsymbol{k}_1 + \varepsilon G_k(\bar{\boldsymbol{x}}, \hat{\boldsymbol{k}}) + O(\varepsilon^2), \tag{3.20}$$

$$\frac{\mathrm{d}\boldsymbol{k}_1}{\mathrm{d}\tau} = -\frac{1}{\varepsilon} G_x(\bar{\boldsymbol{x}}, \hat{\boldsymbol{k}}) - \varepsilon G_{xk}(\bar{\boldsymbol{x}}, \hat{\boldsymbol{k}})\, \boldsymbol{k}_1 + O(\varepsilon^3). \tag{3.21}$$

From the identity

$$\frac{\mathrm{d}}{\mathrm{d}\tau} G(\bar{\boldsymbol{x}}, \hat{\boldsymbol{k}}) = G_x^T(\bar{\boldsymbol{x}}, \hat{\boldsymbol{k}}) \left[\frac{c_g \hat{\boldsymbol{e}}_1}{\varepsilon^2} + F_{kk}(\hat{\boldsymbol{k}})\, \boldsymbol{k}_1 + \varepsilon G_k(\bar{\boldsymbol{x}}, \hat{\boldsymbol{k}}) + O(\varepsilon^2) \right] \tag{3.22}$$

it is apparent that

$$\boldsymbol{k}_1 = \varepsilon \eta_1 \hat{\boldsymbol{e}}_1 + \eta_2 \hat{\boldsymbol{e}}_2 \tag{3.23}$$

with η_1, η_2 of $O(1)$, since substitution of (3.23) into (3.21) and use of (3.22) yields

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left[\eta_1 + \frac{1}{c_g} G(\bar{\boldsymbol{x}}, \hat{\boldsymbol{k}}) \right] = \eta_2 \frac{1}{c_g} G_x^T(\bar{\boldsymbol{x}}, \hat{\boldsymbol{k}})\, F_{kk}(\hat{\boldsymbol{k}})\, \hat{\boldsymbol{e}}_2 - \eta_2 \hat{\boldsymbol{e}}_1^T\, G_{xk}(\bar{\boldsymbol{x}}, \hat{\boldsymbol{k}})\, \hat{\boldsymbol{e}}_2 + O(\varepsilon), \tag{3.24}$$

$$\frac{\mathrm{d}\eta_2}{\mathrm{d}\tau} = -\frac{1}{\varepsilon} \hat{\boldsymbol{e}}_2^T\, G_x(\bar{\boldsymbol{x}}, \hat{\boldsymbol{k}}) + O(\varepsilon). \tag{3.25}$$

Also, substitution of (3.23) into (3.20) yields

$$\frac{\mathrm{d}\boldsymbol{x}_1}{\mathrm{d}\tau} = \eta_2 F_{kk}(\hat{\boldsymbol{k}})\, \hat{\boldsymbol{e}}_2 + \varepsilon \eta_1 F_{kk}(\hat{\boldsymbol{k}})\, \hat{\boldsymbol{e}}_1 + \varepsilon G_k(\bar{\boldsymbol{x}}, \hat{\boldsymbol{k}}) + O(\varepsilon^2). \tag{3.26}$$

Let ξ_2 be defined by

$$\frac{d\xi_2}{d\tau} = \eta_2, \quad \xi_2(0) = 0, \tag{3.27}$$

so that

$$\boldsymbol{x}_1 = \xi_2 F_{kk}(\hat{\boldsymbol{k}})\, \boldsymbol{e}_2 + \varepsilon \boldsymbol{x}_2, \tag{3.28}$$

where

$$\frac{d\boldsymbol{x}_2}{d\tau} = \eta_1 F_{kk}(\hat{\boldsymbol{k}})\, \hat{\boldsymbol{e}}_1 + G_k(\bar{\boldsymbol{x}}, \hat{\boldsymbol{k}}) + O(\varepsilon). \tag{3.29}$$

The ray equations will next be expanded about

$$\tilde{\boldsymbol{x}} = \boldsymbol{x}_0 + \frac{c_g \tau}{\varepsilon^2} \hat{\boldsymbol{e}}_1 + \xi_2 F_{kk}(\hat{\boldsymbol{k}})\, \hat{\boldsymbol{e}}_2. \tag{3.30}$$

From (3.19), (3.28), (3.30)

$$\bar{\boldsymbol{x}} = \tilde{\boldsymbol{x}} + \varepsilon \boldsymbol{x}_2. \tag{3.31}$$

Let

$$\eta_3 = \eta_1 + \frac{1}{c_g} G(\bar{\boldsymbol{x}}, \hat{\boldsymbol{k}}) = \eta_1 + \frac{1}{c_g} G(\tilde{\boldsymbol{x}}, \hat{\boldsymbol{k}}) + O(\varepsilon). \tag{3.32}$$

Substitution of (3.31), (3.32) into (3.24), (3.25) yields

$$\frac{d\eta_3}{d\tau} = \eta_2 \frac{1}{c_g} G_x^T(\tilde{\boldsymbol{x}}, \hat{\boldsymbol{k}}) \cdot F_{kk}(\hat{\boldsymbol{k}}) \cdot \hat{\boldsymbol{e}}_2 - \eta_2 \hat{\boldsymbol{e}}_1^T \cdot G_{xk}(\tilde{\boldsymbol{x}}, \hat{\boldsymbol{k}}) \cdot \hat{\boldsymbol{e}}_2 + O(\epsilon), \tag{3.33}$$

$$\frac{d\eta_2}{d\tau} = -\frac{1}{\varepsilon} \hat{\boldsymbol{e}}_2^T \cdot G_x(\tilde{\boldsymbol{x}}, \hat{\boldsymbol{k}}) - \hat{\boldsymbol{e}}_2^T \cdot G_{xx}(\tilde{\boldsymbol{x}}, \hat{\boldsymbol{k}}) \cdot \boldsymbol{x}_2 + O(\varepsilon). \tag{3.34}$$

From (3.29), (3.31), (3.32)

$$\frac{d\boldsymbol{x}_2}{d\tau} = \left[\eta_3 - \frac{1}{c_g} G(\tilde{\boldsymbol{x}}, \hat{\boldsymbol{k}})\right] F_{kk}(\hat{\boldsymbol{k}})\, \hat{\boldsymbol{e}}_1 + G_k(\tilde{\boldsymbol{x}}, \hat{\boldsymbol{k}}) + O(\varepsilon). \tag{3.35}$$

For initial conditions, first note, from (2.25), (3.12), that $B(0) = O(\varepsilon^3)$. A limit will be derived for A and $B' = B/\epsilon^2$. Dropping primes, the initial conditions are, using also (2.26), (3.18), that

$$\left.\begin{aligned} &\eta_1(0) = -\frac{G(\boldsymbol{x}_0, \hat{\boldsymbol{k}})}{c_g} + O(\varepsilon^3), \quad \eta_2(0) = O(\varepsilon), \quad \eta_3(0) = O(\varepsilon^3), \\ &\boldsymbol{x}_2(0) = 0, \quad A(0) = \cos\psi + O(\varepsilon^3), \quad B(0) = O(\varepsilon). \end{aligned}\right\} \tag{3.36}$$

Since only $A/A(0)$ is relevant, and the A, B equations are linear, we may take $A(0) = 1$ in (3.36).

From (2.27), (3.12), and the estimates $\boldsymbol{e}_j = \hat{\boldsymbol{e}}_j + O(\varepsilon^2)$, $j = 1, 2$, the coefficients of the A, B equations may be written as

$$\mu_1 = -\varepsilon^3 \hat{\boldsymbol{e}}_2^T \cdot G_{xx}(\bar{\boldsymbol{x}}, \hat{\boldsymbol{k}}) \cdot \hat{\boldsymbol{e}}_2 + O(\varepsilon^5), \quad \mu_2 = O(\epsilon^3), \quad \mu_3 = \hat{\boldsymbol{e}}_2^T \cdot F_{kk}(\hat{\boldsymbol{k}}) \cdot \hat{\boldsymbol{e}}_2 + O(\varepsilon^2). \tag{3.37}$$

Further expansion about $\tilde{\boldsymbol{x}}$ and substitution into (2.29), (2.30) now yields

$$\frac{dA}{d\tau} = \Delta B + O(\varepsilon), \tag{3.38}$$

$$\frac{dB}{d\tau} = -\frac{1}{\varepsilon} [\hat{\boldsymbol{e}}_2^T \cdot G_{xx}(\tilde{\boldsymbol{x}}, \hat{\boldsymbol{k}}) \cdot \hat{\boldsymbol{e}}_2]\, A - [\hat{\boldsymbol{e}}_2^T (G_{xxx}(\tilde{\boldsymbol{x}}, \hat{\boldsymbol{k}})\, \boldsymbol{x}_2) \cdot \hat{\boldsymbol{e}}_2]\, A + O(\varepsilon), \tag{3.39}$$

where the matrix $G_{xxx}\boldsymbol{x}_2$ has ijth component $\Sigma_{l=1}^{2}(\partial^2 G/\partial x_i\,\partial x_j\,\partial x_l)(\boldsymbol{x}_2)_l$, and

$$\Delta = \hat{\boldsymbol{e}}_2^T\cdot F_{\boldsymbol{kk}}(\hat{\boldsymbol{k}})\cdot\hat{\boldsymbol{e}}_2. \tag{3.40}$$

To apply the limit theorem, consider the seven-dimensional vector $\boldsymbol{W}=(\xi_2,\eta_2,\eta_3,\boldsymbol{x}_2,A,B)^T$ satisfying equations (3.27), (3.34), (3.33), (3.35), (3.38), (3.39), where the $O(\varepsilon)$ terms have been neglected. These equations are of the form (3.6), where $\boldsymbol{H}_i$, $i=1,2$ are function of $\tilde{\boldsymbol{x}}$, and so are of the form

$$\boldsymbol{H}_i(t,\boldsymbol{W}) = \boldsymbol{H}_i(\boldsymbol{x}_0+c_g\,t\hat{\boldsymbol{e}}_1+\xi_2\,F_{\boldsymbol{kk}}(\hat{\boldsymbol{k}})\,\hat{\boldsymbol{e}}_2,\boldsymbol{W}), \tag{3.41}$$

where $t=\tau/\epsilon^2$. Now the mixing condition must be inferred from some corresponding property of the random field $G(\boldsymbol{x},\hat{\boldsymbol{k}})=\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{U}}(\boldsymbol{x})$. We assume that if two sets $S_1,S_2\in\mathscr{R}^2$ are separated by a large distance, then the sigma-algebra generated $\{G(\boldsymbol{x},\hat{\boldsymbol{k}}):\boldsymbol{x}\in S_1\}$ is approximately independent of that generated by $\{G(\boldsymbol{x},\hat{\boldsymbol{k}}):\boldsymbol{x}\in S_2\}$. Roughly, this means that $\hat{\boldsymbol{U}}(\boldsymbol{x})$ cannot be predicted from its values in a region far from $\boldsymbol{x}$.

From (3.41) it is apparent that for fixed non-random $\overline{\boldsymbol{W}},\boldsymbol{H}_i(t,\overline{\boldsymbol{W}})$ is determined by the values of $G(\boldsymbol{x},\hat{\boldsymbol{k}})$ and its first three $\boldsymbol{x}$-derivatives along a line ℓ_t in the direction of $F_{\boldsymbol{kk}}(\hat{\boldsymbol{k}})\,\hat{\boldsymbol{e}}_2$, passing through the point $\boldsymbol{x}_0+c_g\,t\hat{\boldsymbol{e}}_1$. Similarly, $\boldsymbol{H}_i(t+t',\boldsymbol{W})$ is determined by the values of G and its first three derivatives along a parallel line $\ell_{t+t'}$ through the point $\boldsymbol{x}_0+c_g(t+t')\,\hat{\boldsymbol{e}}_1$. First assume that $\Delta\neq 0$, so that these lines are not parallel to $\hat{\boldsymbol{e}}_1$ ($\Delta=0$ will be treated below). Then $\ell_t,\ell_{t+t'}$ are separated by the distance

$$\frac{|\Delta|}{|F_{\boldsymbol{kk}}(\hat{\boldsymbol{k}})\,\hat{\boldsymbol{e}}_2|}(c_g\,t')\rightarrow\infty \quad \text{as} \quad t'\rightarrow\infty. \tag{3.42}$$

So $(\boldsymbol{H}_1(t,\overline{\boldsymbol{W}}),\ \boldsymbol{H}_2(t,\overline{\boldsymbol{W}}))$ becomes asymptotically independent of $(\boldsymbol{H}_1(t+t',\overline{\boldsymbol{W}}),\ \boldsymbol{H}_2(t+t',\overline{\boldsymbol{W}}))$ as the time difference t' becomes large, since they depend on the values of $\boldsymbol{G}(\boldsymbol{x},\hat{\boldsymbol{k}})$ on sets separated by a large distance. (Technically, the sigma-algebra $\mathscr{F}_t^t$ is the smallest sigma-algebra with respect to which $G(\boldsymbol{x},\hat{\boldsymbol{k}})$, $\boldsymbol{x}\in\ell_t$ and its first three $\boldsymbol{x}$- derivatives are measurable.)

Next suppose $\Delta=0$, so that ℓ_t is parallel to $\hat{\boldsymbol{e}}_1$. Let

$$\tau' = \tau+\epsilon^2\frac{|F_{\boldsymbol{kk}}(\hat{\boldsymbol{k}})\,\hat{\boldsymbol{e}}_2|}{c_g}\xi_2.$$

Then if all equations are written in terms of τ', the function evaluations are at $\tilde{\boldsymbol{x}}=\boldsymbol{x}_0+(c_g\,\tau'/\epsilon_2)\,\hat{\boldsymbol{e}}_1$, so that mixing follows. Also, since $\mathrm{d}/\mathrm{d}\tau=(1+O(\epsilon^2))\,\mathrm{d}/\mathrm{d}\tau'$, the limit is not affected by this change of variables.

Application of the limit theorem can be considerably simplified by noting that, because of the role of $\boldsymbol{H}_2$ in (3.8), all $O(1)$ terms that are rapidly varying on the τ/ϵ^2 scale, and are mean zero, do not contribute to the limit. Equivalently, setting these terms equal to zero before computing the limit, we find that the subsystem $\boldsymbol{W}^1=(\xi_2,\eta_2,A,B)^T$ decouples from the subsystem $\boldsymbol{W}^2=(\eta_3,\boldsymbol{x}_2)^T$, which can be computed trivially. From this calculation and (3.32)

$$\eta_1 = -G(\boldsymbol{x}_0,\hat{\boldsymbol{k}})/c_g, \quad \eta_3=0, \quad \boldsymbol{x}_2=0. \tag{3.43}$$

The other limits may now be calculated from (3.8), and will be written in the form (3.10). Let β_1,β_2 be two independent Brownian motions. Then in the limit

$$\mathrm{d}\xi_2=\eta_2\,\mathrm{d}\tau, \quad \xi_2(0)=0; \qquad \mathrm{d}\eta_2=(\bar{a}^{22})^{1/2}\mathrm{d}\beta_1, \quad \eta_2(0)=0; \tag{3.44}$$

$$\mathrm{d}A=\Delta B\,\mathrm{d}\tau, \quad A(0)=1, \qquad \mathrm{d}B=(\bar{a}^{44})^{1/2}A\,\mathrm{d}\beta_2, \quad B(0)=0. \tag{3.45}$$

The constants $\bar{a}^{22}$, $\bar{a}^{44}$ are expressed in terms of the correlation function

$$\rho_G(\boldsymbol{x}',\boldsymbol{k}') = \langle G(\boldsymbol{x},\boldsymbol{k}')\,G(\boldsymbol{x}+\boldsymbol{x}',\boldsymbol{k}')\rangle. \tag{3.46}$$

Then

$$\bar{a}^{22} = -\frac{1}{c_g}\frac{\partial^2}{\partial s_2^2}\int_{-\infty}^{\infty} \rho_G(s_1\hat{\boldsymbol{e}}_1 + s_2\hat{\boldsymbol{e}}_2)\,\mathrm{d}s_1\bigg|_{s_2=0}, \quad (3.47)$$

$$\bar{a}^{44} = \frac{1}{c_g}\frac{\partial^4}{\partial s_2^4}\int_{-\infty}^{\infty} \rho_G(s_1\hat{\boldsymbol{e}}_1 + s_2\hat{\boldsymbol{e}}_2)\,\mathrm{d}s_1\bigg|_{s_2=0}. \quad (3.48)$$

Let the correlation matrix of $\hat{\boldsymbol{U}}$ be

$$\rho_{\hat{U}}(\boldsymbol{x}') = \langle \hat{\boldsymbol{U}}(\boldsymbol{x})\,\hat{\boldsymbol{U}}^T(\boldsymbol{x}+\boldsymbol{x}')\rangle. \quad (3.49)$$

Then for G given by (3.11)

$$\rho_G(\boldsymbol{x}', \hat{\boldsymbol{k}}) = \hat{\boldsymbol{k}}^T \rho_{\hat{U}}(\boldsymbol{x}')\,\hat{\boldsymbol{k}}. \quad (3.50)$$

4. Universal rogue distribution

From (3.17), (3.40) and (2.31) with $\Omega = F$, we obtain

$$\Delta = \pm\frac{g^{1/2}}{2\hat{k}^{3/2}}\left(\frac{3\cos^2\psi - 1}{2}\right), \quad (4.1)$$

where ψ is the angle between the rays and the waves in the absence of randomness, and $\hat{k} = |\hat{\boldsymbol{k}}|$. From (3.45) $\mathrm{d}A = 0$ if $\Delta = 0$. So $A = 1$ and no caustics are possible (on this $O(\sigma^{-2/3})$ distance scale) if the rays make an angle of $\psi^* = \arccos(1/\surd 3) = 54.7°$ with the waves. This situation would require such swift currents, without significant mean curvature, that the theory is unlikely to find applications where this effect is observed. However, it is interesting to note that the complement of ψ^*, 35.3°, is familiar from the theory of ship waves (Whitham 1974). A ship wake is confined to a wedge of semi-angle $\arctan(1/(2\surd 2)) = 19.5°$. A wave reaching the boundary of this wedge makes an angle of 35.3° to the direction of the ship.

From (3.45) the joint probability density, $P(\tau, A, B)$, of A, B at time τ, is the solution of the forward Kolmogorov (Fokker–Planck) equation

$$\frac{\partial P}{\partial \tau} = \tfrac{1}{2}\bar{a}^{44}A^2\frac{\partial^2 P}{\partial B^2} - \Delta B\frac{\partial P}{\partial A}, \quad P(0, A, B) = \delta(A-1)\,\delta(B). \quad (4.2)$$

From (4.2) it is evident that for $\Delta \neq 0$ and finite $\bar{a}^{44} \neq 0$, the variables B and τ can be re-scaled to effectively set Δ and $\bar{a}^{44}$ arbitrarily. In particular, let

$$\bar{\tau} = \left(\frac{\bar{a}^{44}\Delta^2}{3}\right)^{1/3}\tau, \quad \bar{B} = \left(\frac{3\Delta}{\bar{a}^{44}}\right)^{1/3} B. \quad (4.3)$$

Then when written in terms of $\bar{\tau}$, $\bar{B}$, equation (4.2), and hence (3.45), are of the same form but with Δ replaced by 1, and $\bar{a}^{44}$ replaced by 3.

Thus, for $\Delta \neq 0$ and finite $\bar{a}^{44} \neq 0$ the (A, B) process is universal, except for scale factors. In particular, all the statistics of $A(\tau)$ are determined by a single distance scale factor, irrespective of the details of the random medium. Indeed, the universal statistics of $A(\tau)$ do not even depend on the dispersion relation, except for general properties, such as (3.12). This limit was first discovered for the non-dispersive case by Kulkarny & White (1982), and was verified by the Monte Carlo simulations of Hesselink & Sturtevant (1988). Extensions of these ideas to the non-dispersive case are in Zwillinger & White (1985), Nair & White (1991), and White *et al.* (1988). A three-dimensional theory for geometrical optics is in White (1984).

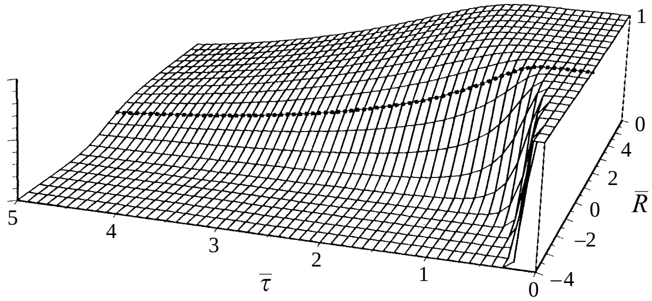


FIGURE 1. Function $u(\bar{\tau}, \bar{R})$, defined through equation (4.6). The curve $u(\bar{\tau}, 0)$ (entering in equation (4.7)) is marked by thick dashes.

For a caustic, we are interested in those values of $\bar{\tau}$ at which $A = 0$. Let

$$\bar{R} = \bar{B}/A. \tag{4.4}$$

A caustic occurs when $\bar{R}$ is infinite. By Ito's lemma (McKean 1969) $\bar{R}$ satisfies the Ito–Riccati equation

$$\mathrm{d}\bar{R} = -\bar{R}^2\,\mathrm{d}\bar{\tau} + \sqrt{3}\,\mathrm{d}\beta, \tag{4.5}$$

where $\beta(\bar{\tau})$ is a standard Brownian motion. From Kulkarny & White (1982) $\bar{R}$ has explosions, i.e. it becomes infinite in finite time, with probability 1. In the language of Feller (1952) $\bar{R} = +\infty$ is an entrance boundary, and $\bar{R} = -\infty$ is an exit boundary. That is, if $\bar{R}$ is large and positive, it will return to $O(1)$ values, but if $\bar{R}$ is large and negative it will soon explode to $-\infty$. So caustics occur when $\bar{R} \to -\infty$, as might be inferred from the minus sign of the quadratic term in (4.5).

Let s be the random distance at which $\bar{R}$ first becomes infinite, i.e. the first place along the ray at which a rogue wave is formed. Let $P_s(\bar{\tau})$ be the probability density of s. Then $P_s(\bar{\tau})$ can be obtained from the solution, $u(\bar{\tau}, \bar{R})$, of the backward Kolmogorov equation

$$\left.\begin{aligned}
&\frac{\partial u}{\partial \bar{\tau}} = \frac{3}{2}\frac{\partial^2 u}{\partial \bar{R}^2} - \bar{R}^2\frac{\partial u}{\partial \bar{R}},\\
&u(0, \bar{R}) = 1,\\
&u \to 0 \quad \text{as} \quad \bar{R} \to -\infty,\\
&u \quad \text{bounded as} \quad \bar{R} \to +\infty.
\end{aligned}\right\} \tag{4.6}$$

Then (Kulkarny & White 1982; White *et al.* 1988)

$$P_s(\bar{\tau}) = -\frac{\partial}{\partial \bar{\tau}} u(\bar{\tau}, 0), \tag{4.7}$$

see figure 1. Setting $\bar{R} = 0$ in (4.7) corresponds to the plane wave initial conditions considered above. However, the probability density of the distance between two rogue waves, formed along the same ray, can be obtained by evaluating u at $\bar{R} = +\infty$ in (4.7).

A closed-form approximation for P_s can be constructed from a composite of the $\bar{\tau} \downarrow 0$ and $\bar{\tau} \uparrow \infty$ asymptotic expansions. From White *et al.* (1988)

$$P_s(\bar{\tau}) \approx \left(\frac{0.7917}{\bar{\tau}^{5/2}} + 0.4528\right) \exp\left\{-\frac{0.6565}{\bar{\tau}^3} - 0.4054\bar{\tau}\right\}. \tag{4.8}$$

This curve is very flat near $\bar{\tau} = 0$, rises to a peak near $\bar{\tau} = 1$, and then decays exponentially to zero.

Let

$$D = c_g \bar{\tau}/\sigma^{2/3} \tag{4.9}$$

be the distance, to leading order, traversed along a ray at time $\bar{\tau}$. Then from (4.1), (4.3)

$$\bar{\tau} = \left(\frac{\sigma^2 \bar{a}^{44} g}{12\hat{k}^3}\left[\frac{3\cos^2\psi - 1}{2}\right]^2\right)^{1/3}\frac{D}{c_g}. \tag{4.10}$$

As a specific model for a random current, let U be two-dimensionally incompressible, so that $\hat{U}$ is determined by a random stream function $\Psi(x)$. Let Ψ have mean zero and correlation function

$$\rho_\Psi(x') = \langle \Psi(x)\,\Psi(x+x')\rangle = \frac{L^2}{4}\exp\left\{-\frac{|x'|^2}{L^2}\right\}. \tag{4.11}$$

Here L is an intrinsic distance scale, and the pre-exponential factor is chosen so that $\langle|\hat{U}|^2\rangle = 1$. For this model,

$$\rho_{\hat{U}}(x') = \exp\left\{\frac{-|x'|^2}{L^2}\right\}\begin{bmatrix} \frac{1}{2} - \frac{x_2'^2}{L^2} & \frac{x_1' x_2'}{L^2} \\ \frac{x_1' x_2'}{L^2} & \frac{1}{2} - \frac{x_1'^2}{L^2}\end{bmatrix}, \tag{4.12}$$

and equation (4.10) becomes

$$\bar{\tau} = (100\pi)^{1/6}\left(\frac{\sigma}{c_g}\right)^{2/3}\left(\frac{c_g^0}{c_g}\right)^{2/3}\left(\cos\psi\left[\frac{3\cos^2\psi - 1}{2}\right]\right)^{2/3}\frac{D}{L}, \tag{4.13}$$

where $c_g^0 = \frac{1}{2}c_p^0$ is the group velocity in the absence of current.

When the mean current is zero, the third and fourth factors in (4.13) are unity. For a non-trivial mean current, the situation may be more or less dangerous depending on the average set and drift, with $\psi = 54.7°$ the least dangerous situation, as discussed at the beginning of this section. The most dangerous situation obtains when the swell is in a direction opposite to the set of the mean current, both according to (4.13) and in accord with common lore. For a fixed mean current strength this situation both maximizes the terms involving $\cos\psi$ and minimizes c_g, and gives a larger value of $\bar{\tau}$ than if the mean drift were zero. Note that this result is non-trivial, since a constant current with no random fluctuations will not by itself produce caustics in an opposing swell.

5. Numerical simulations

5.1. *Description of the numerical ray tracing code*

To illustrate the design of the ray tracing code, we will first refer to the same test case as is shown in figure 16 of Gerber (1993). The annular current he considers is shown in our figure 2. It has a parabolic velocity profile, with a peak velocity of 2 m s^{-1}. Waves with a time period of $T = 10$ s enter from the south-west. To allow our code to work with completely general flow fields (e.g. with currents determined from oceanographic measurements), we represent the u- and v-velocity components on a discrete grid

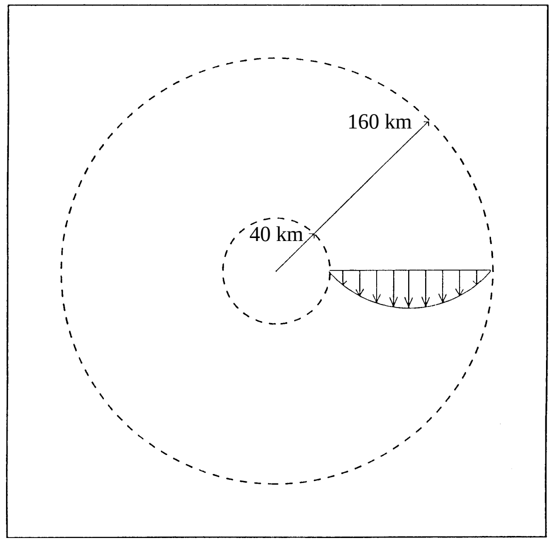


FIGURE 2. Ray tracing test problem: annular current; parabolic velocity profile with maximum velocity 2 m s^{-1}.

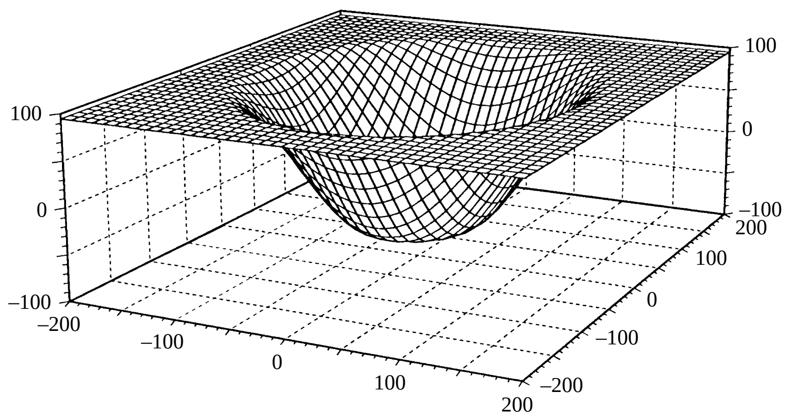


FIGURE 3. Stream function for annular current (with horizontal scales in kilometres, vertical in 1000 m^2 s^{-1}).

(rather than assuming a known functional form, as was the case in the code by Gerber). Figure 3 shows the stream function for this annular current.

In order to solve (2.7) numerically, one needs values for u, v, u_x, u_y, v_x, v_y at arbitrary spatial locations. Since we also solve the ray tube equations (2.29) and (2.30), we additionally need values for all the second derivatives $u_{xx}, u_{xy}, u_{yy}, v_{xx}, v_{xy}, v_{yy}$. All these derivatives can be obtained very rapidly from the grid values of u and v by successive one-dimensional finite difference (FD) approximations as illustrated in figure 4. It is immaterial whether the one-dimensional approximations are first made horizontally on several grid lines, and then vertically at the desired x-location, as shown, or first vertically, etc. – the results are identical.

A fast algorithm for determining weights in general one-dimensional FD formulas was first given by Fornberg (1988). It is explained further, with codes, in Fornberg (1996). This algorithm rapidly provides the optimal weights in FD formulas under the following very general conditions:

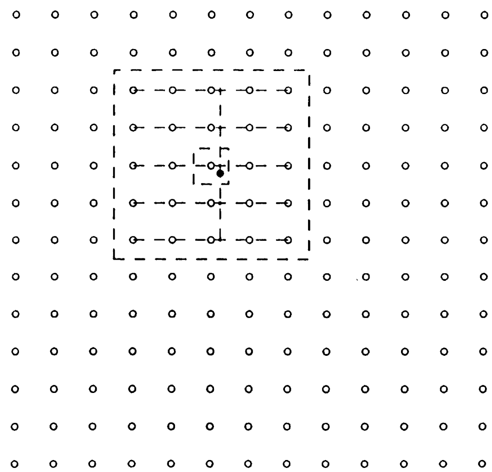


FIGURE 4. Structure of finite difference approximations used to find derivatives at arbitrary locations.

(i) arbitrary location of the point at which the approximations are to be accurate (may, but need not be, at a grid point),
(ii) arbitrary distribution of grid points (need not be equi-distributed, as here),
(iii) any level of accuracy (i.e. any stencil width), and
(iv) any order of derivative to be approximated (in its special case of the zeroth derivative, using this algorithm to find the FD weights and then applying these to the data constitutes the fastest known method for polynomial interpolation).

In the present calculations, all derivatives were obtained using a 5×5 stencil, as shown in figure 4. In the case of $n\times n$ stencils, the total cost for one case (all 12 derivatives) becomes to leading order $10n^2$ arithmetic operations for finding all needed FD weights, and another $12n^2$ operations to apply these weights to the grid values for u and v (if we are only calculating ray paths – six derivatives – the corresponding operation counts become $6n^2$ and $8n^2$ respectively).

A standard two-stage, second-order-accurate Runge–Kutta method was then used to advance the six coupled ODEs (giving ray paths and, along these, the wave vector and the ray tube area).

For the model problem just introduced, computations on a 41×41 grid (i.e. with grid spacing 10 km) give the ray paths shown in figure 5. To within the graphical accuracy, these paths are indistinguishable from those presented by Gerber (1993). The dots along the paths in figure 5 mark where the raytube area variable A first crosses zero. As is to be expected, these locations agree with the locations where closely spaced neighbouring rays first cross.

5.2. *Generation of random eddy fields*

We start the construction of our random eddy fields by first assigning, at the grid points of an $N\times N$ grid,

$$\Psi_{i,j} = \{\text{uniformly distributed random numbers between } -1 \text{ and } +1\}.$$

This field is brought to two-dimensional discrete Fourier space (using one-dimensional FFTs), and each component $\hat{\Psi}_{i,j}$ is damped by multiplying it with the factor $E_{i,j}$:

$$\hat{\Psi}_{i,j} \leftarrow \hat{\Psi}_{i,j}\, E_{i,j} \quad \text{where} \quad E_{i,j} = \mathrm{e}^{-(L/\pi S)^2(i^2+j^2)}.$$

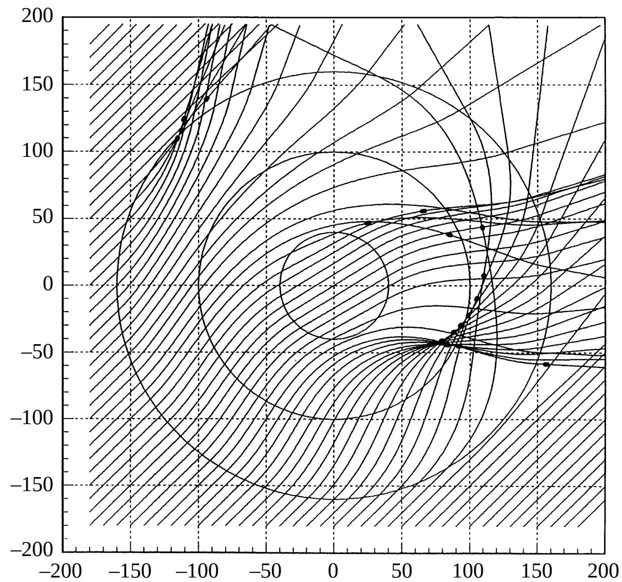


FIGURE 5. Ray paths for annular current test problem (wave period $T = 10$ s). The dots mark locations where raytube areas first cross zero, i.e. the first intersection points for infinitesimally close rays.

Here, S denotes the physical side length of the square domain. After returning $\hat{\Psi}_{i,j}$ to the physical domain, the correlation length will be the specified L. Next we normalize Ψ by the scaling

$$\Psi_{i,j} \leftarrow \Psi_{i,j} \frac{L\sigma\sqrt{3N}}{(2\sum E_{i,j}^2)^{1/2}}.$$

The random stream function Ψ and the velocities $u = \Psi_y$, $v = -\Psi_x$ (calculated by pseudospectral differentiation – taking the analytical derivative in discrete Fourier space) will now have standard deviation as follows:

$$\langle \Psi_{i,j}^2 \rangle^{1/2} = \sigma \frac{L}{2},$$

and

$$\langle u_{i,j}^2 + v_{i,j}^2 \rangle^{1/2} = \sigma.$$

In the ray tracing calculations, a physical 640×640 km domain was discretized with 256×256 grid points. The rays were started with $k_y = 0$, and with k_x chosen to correspond, through equation (2.1), to a wave with $T = 10$ s in the case of $\boldsymbol{U} = \boldsymbol{0}$, i.e. k_x is obtained by solving the quadratic equation

$$gk_x = \left(\frac{2\pi}{T} - k_x u\right)^2.$$

Here $g = 9.8$ m s^{-2}, $T = 10$ s, and u is the velocity component in the x-direction at the ray starting point. Initial conditions for A and B were obtained numerically from equations (2.16) and (2.25).

The domains in figures 6–8 are all of the size 320×320 km; the bottom left of the full periodic random 640×640 km fields, thus avoiding any wrap-around correlations within the ray-traced subdomains. Each of figures 6–8 (a) shows streamlines in an eddy

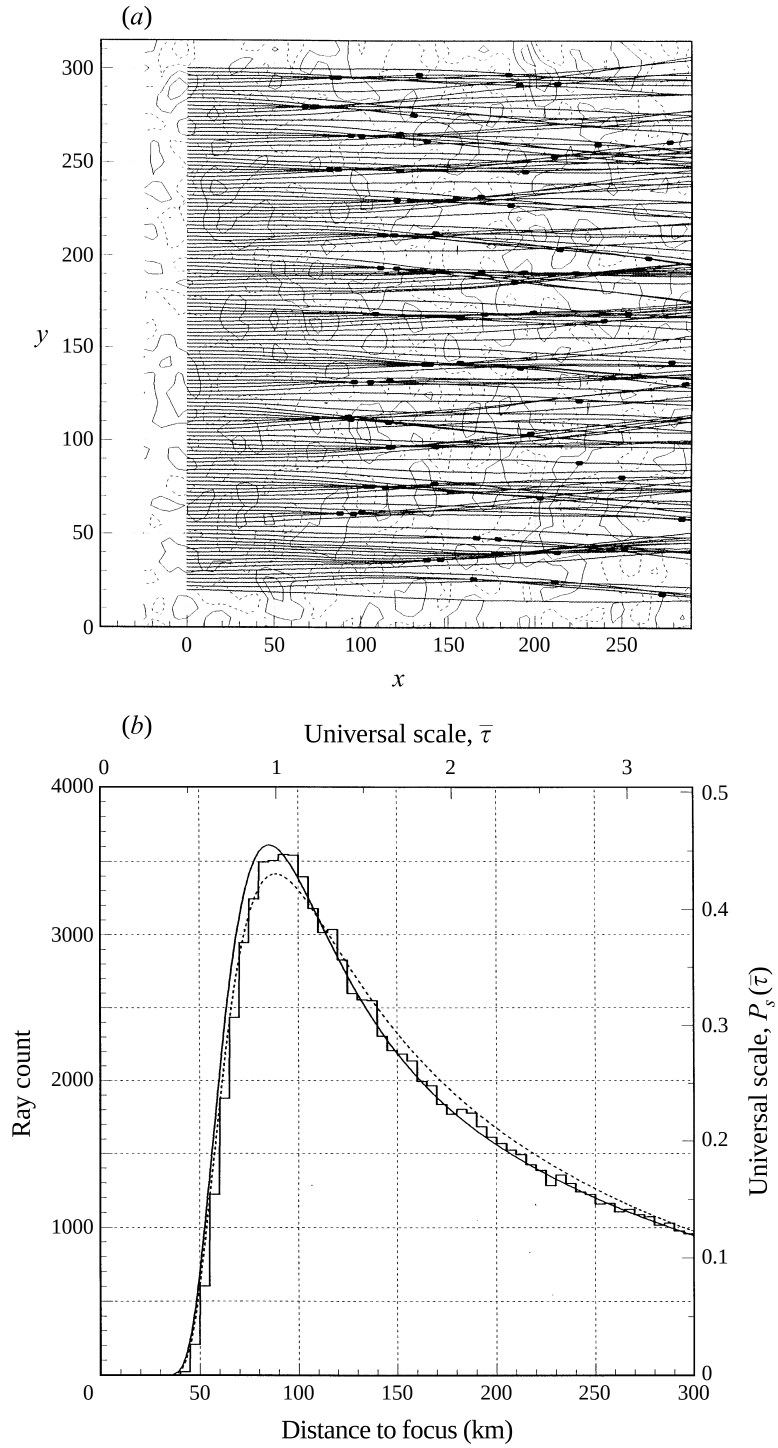


FIGURE 6. (*a*) Traces of 141 rays through a random field with $L = 8$ km, $\sigma = 0.05$ m s^{-1}. (*b*) Histogram of the distance to the first focus along 141 rays in each of 1000 random fields, as in (*a*), Comparison with 'universal curve' prediction (solid) and its approximation (dashed).

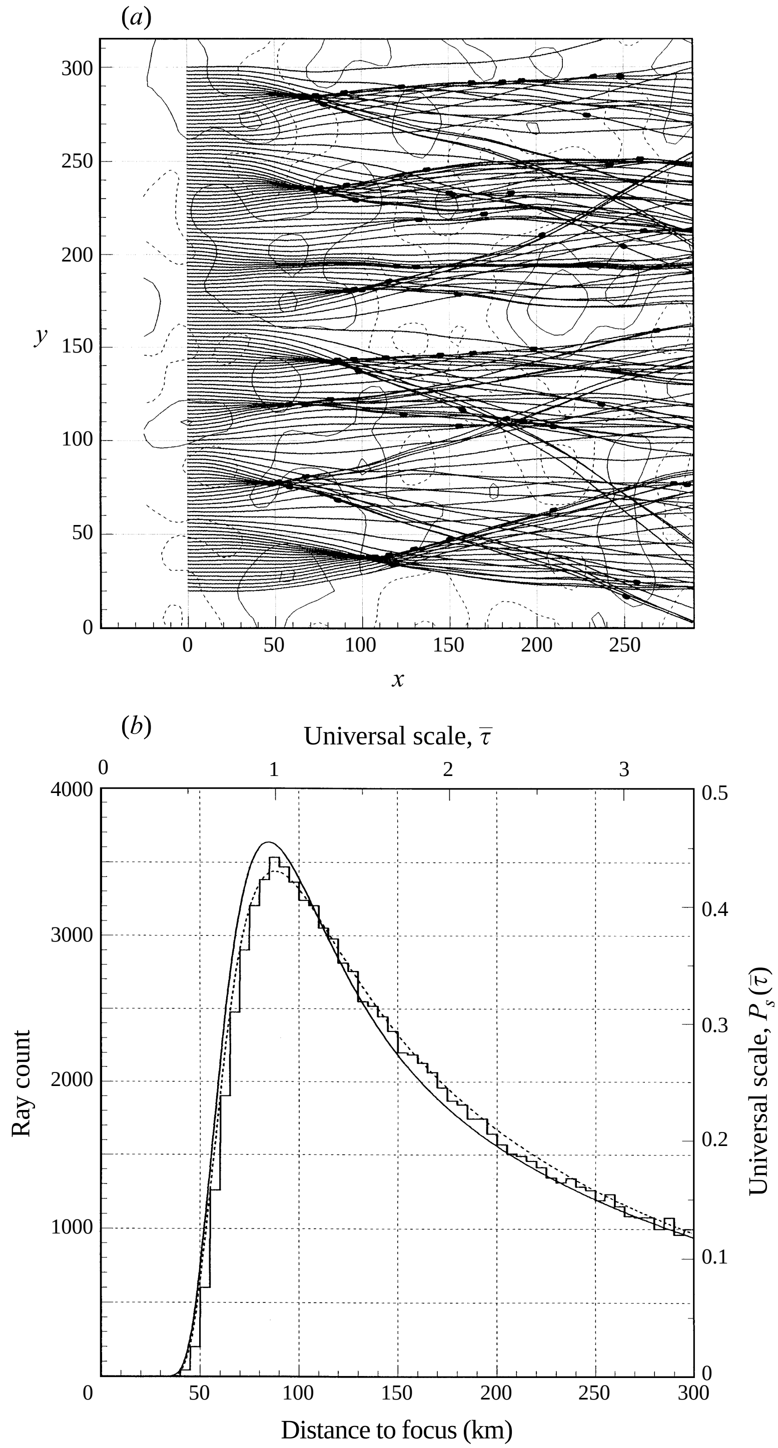
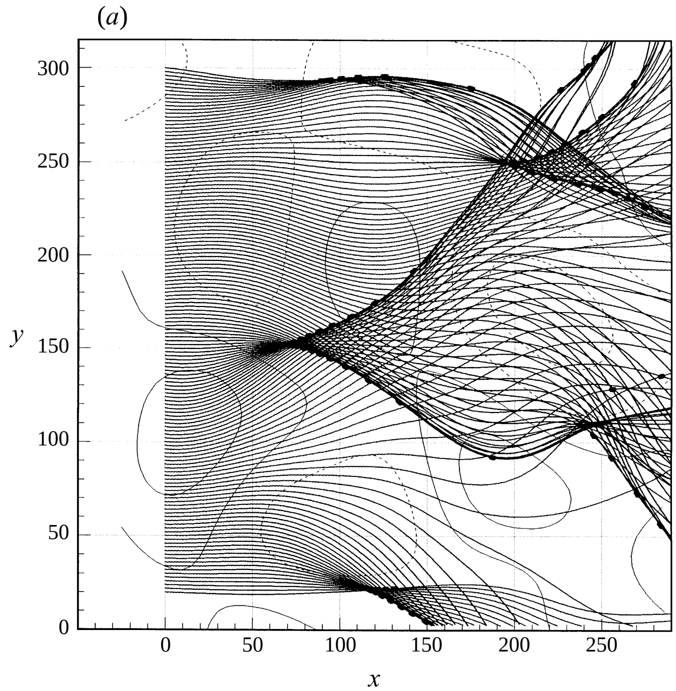


FIGURE 7. (a,b) As figure 6 but with $L = 20$ km, $\sigma = 0.2$ m s^{-1}.



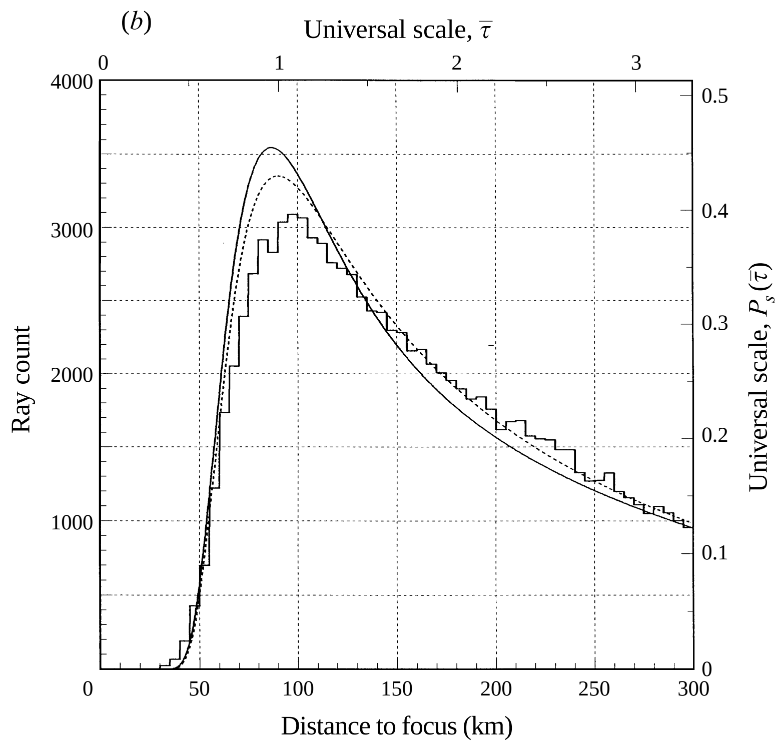


FIGURE 8. (a, b) As figure 6 but with $L = 60$ km, $\sigma = 1.0$ m s^{-1}.

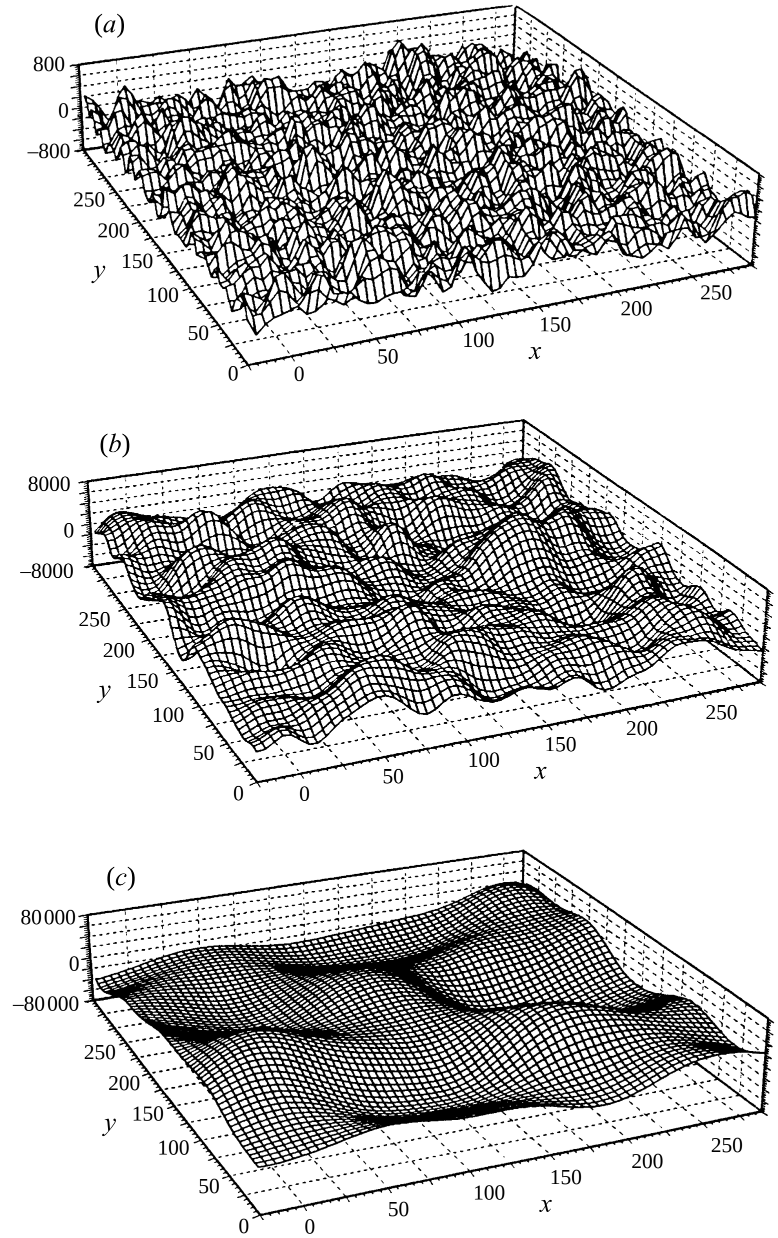


FIGURE 9. Same stream function fields as were shown through contours in figures 6(*b*), 7(*b*) and 8(*b*) respectively. Note that the horizontal axes are unchanged, but that the strengths of the eddies (vertical scale) are increasing.

field with its associated rays and focus points. The histograms in figures 6–8(*b*) show the $\boldsymbol{x}$-coordinates of the first focus point along each ray, summed in each of the three cases over 1000 random fields, i.e. in all, 141000 rays for each histogram. The solid curve in each of these figures is the universal curve $P_s(\bar{\tau})$, given by (4.7), and scaled according to (4.13), with $c_g^0 = c_g$, $\psi = 0$; the dashed curve is its approximation (4.8).

Since $\bar{\tau}$ is a multiple of $\sigma^{2/3}D/L$ (cf. equation (4.13) and the comments following it), the different σ-cases (eddy strengths) could all be studied in the same size domain by suitably adjusting L. Figure 9 (*a–c*) illustrates again these scalings; the three surfaces show the same three stream function fields as were previously seen in figures 6–8 (*a*). Note the differences in the vertical scale in the three cases.

6. Conclusions

We have demonstrated that a single universal curve, when appropriately scaled, will accurately describe the distribution of focus points, that is, areas of particularly intense wave action, which will arise when uniform waves enter a region of random current. Although in the theoretical derivation of this curve it was assumed that the random current fluctuations were small, and therefore that focusing would occur only after the waves have travelled through many random eddies, the numerical calculations showed excellent agreement with theory even when random effects were relatively large – with focusing occurring before even a single eddy has been traversed. Since $\sigma = 1$ m s^{-1}, i.e. a velocity standard deviation of about 2 knots, represents a substantial current, we expect that these results will be generally applicable. If an estimate can be made of the distance scale parameter, and of the distance propagated through the random current field, reference to the universal curve provides a quantitative estimate of the danger of a rogue wave.

Appendix

In this manuscript we are directly concerned only with the chance of a freak wave at sea, and not with the chance factors influencing the height, steepness, asymmetric shape or any other properties of the giant wave itself. That is, we are concerned with the likelihood of encountering a freak wave, rather than an estimation of the expected damage from encountering one. We also neglect nonlinear effects which may cause instability or defocusing of the waves. Since a study of these related issues would make a useful complement to the present work, we will briefly review them here.

As discussed in the Introduction, even for points in space that are not near caustics, principles such as conservation of wave action (CWA), that would enable calculation of amplitudes, lack a firm mathematical basis when the current has non-zero vorticity. For water of finite depth, with wave modulations that are long compared to the depth, CWA can be justified, even with rotational currents, by the work of Stiassnie & Peregrine (1979). However, it might be argued that waves on a deep current do not satisfy CWA, based on the form of the nonlinear Schrödinger equation (NLS) derived for the amplitude by Mei (1989, equation (2.59) on p. 618) (NLS methods are discussed further below). Still, the application of CWA to rotational currents remains controversial.

Many authors have not been inhibited by its deficiencies from using ray theory to compute amplitudes in the presence of rotational currents, e.g. the example of Gerber (1993) which we considered in §5. Also, Peregrine (1976, §II E) used CWA for waves on a shear current for his investigation of freak waves and caustics. In particular, a form of CWA is often used in computing wave spectra, regardless of vorticity. For example, Mathiesen (1987) computed wave refraction by a current whirl, and Mapp, Welch & Munday (1985) got good agreement of their ray theoretic energy densities with those determined by Synthetic Aperture Radar (SAR) data, for observations of waves refracted by warm core rings. Irvine & Tilley (1988) used CWA to interpret their SAR data of the Agulhas current. Their analysis of meanders supports the basic hypothesis of our manuscript: the theory that giant waves are caused by current-induced caustics. In their assessment, CWA 'represents the presently accepted core of belief'.

Near a caustic, even CWA is not sufficient for the determination of amplitudes. Other factors include: finite frequency diffraction effects that mitigate the caustic singularity; nonlinearities that can cause defocusing and instabilities; blurring of the caustic because real ocean waves are a mixture of frequencies rather than the pure monochromatic waves used in the mathematical model; wave energy generation by wind forces; and energy dissipation from breaking waves. An overview of some of these factors is given in the review lecture of Peregrine (1985).

Linear theory predicts that the amplitude singularity near a caustic of simple ray theory will be mitigated by finite frequency diffraction effects. The maximum amplitude, and the size of the caustic boundary layer within which diffraction corrections are needed, are both determined as powers of spatial frequency k, with the exact powers dependent on the geometry of the caustic. Of course, in the present work, the size of the caustic boundary layer, as well as the wavelength, is assumed to be much smaller than the correlation length of the random medium. For non-dispersive waves a uniform expansion for a smooth caustic was given by Ludwig (1966), and an expansion near the cusp of a caustic was given by Pearcey (1946).

Alternatively, the Gaussian beam summation method can capture caustic corrections for caustics of arbitrary geometric complexity (White *et al.* 1987). An adaptation of this method, combined with stochastic methods that parallel those of the present paper, has been used to compute wave statistics for wave propagation problems where caustics occur at random locations (Nair & White 1991).

Another approach to calculating diffraction at a caustic is the parabolic equation approximation for the amplitude, and a variant of this approach, the nonlinear Schrödinger equation (NLS), can be adapted to include weak nonlinear effects. Smith (1976) used this method in his analysis of giant waves, and estimated that for freak waves in the Agulhas current, there would be 'about a three-fold amplification of the wave height near the caustic'. He also noted that even a doubling of height would be 'quite traumatic'.

Peregrine (1986) gave two simple approximations for wave amplitude at the cusp of a caustic, for water waves propagating over underwater shoals and spurs. One of his approximations is based on the linear theory of Pearcey (1946), and the other, which includes weak nonlinear effects, is based on the NLS. The NLS method for this problem was compared with tank tests by Peregrine *et al.* (1988), and they obtained good quantitative agreement. While they observed the predicted defocusing effects of nonlinearity, they noted that linear diffraction is also important, and can be dominant in practical cases.

A nonlinear mechanism quite different from that of caustic formation has been proposed as an alternative explanation of freak waves. Dean (1990) suggested that the nonlinear superposition of waves might produce waves larger than those suggested by linear superposition. The resulting statistical distribution of wave heights would then have large waves occurring more frequently than predicted by the Rayleigh distribution. These large waves would then be, by definition, freaks.

Nonlinearities can also cause instability of a regular wave train. This was first demonstrated theoretically by Benjamin & Feir (1967), who showed that for weakly nonlinear surface gravity waves in deep water, regular wave trains were unstable to perturbations, a result confirmed in the experimental work of Feir (1967) and others (see Su *et al.* 1982 and references therein). Some of this experimental work showed that the initial instability does not necessarily lead to disintegration of the wave train.

Instead, as in Su *et al.* (1982), the waves can evolve first into a series of crescent-shaped, spilling breakers, and then, finally, to a series of wave groups with a peak frequency that is lower than the frequency of the initial wave train.

Gerber (1987) investigated a variation of the Benjamin–Feir instability, relevant to wave packets propagating in deep water with non-uniform currents. He found an enhancement of the instability for waves travelling against an adverse current, while the current had a stabilizing effect for waves travelling with it. Most relevant for the present work, he found enhanced instability of the wave packet in the neighbourhood of a caustic caused by a shear current.

The Benjamin–Feir instability is one reason why a perfectly regular time-harmonic wave train is an idealized model. In practice, even a relatively regular ocean swell contains a mixture of frequencies. Since surface gravity waves are dispersive, each frequency will give a caustic in a different location, as will, for instance, variations in the initial direction of the waves. Even if all these variations are small, the sharply defined caustic for monochromatic unidirectional waves will be blurred, and this may have a significant effect on amplitudes. However, even though both nonlinearity and a broad spectrum of frequencies are thought to have defocusing effects, the combination of nonlinearity and broad spectrum has been proposed as the cause of freak waves by Trulsen & Dysthe (1996). Consequently, these authors have derived a form of the NLS equation applicable to a broad spectrum of frequencies, but have not yet demonstrated that this equation will generate freak waves.

To return to the question of amplitude, we note that amplitude, or wave height, is not necessarily the best measure of danger for a vessel at sea. Wave steepness is also important, since even a small vessel can ride over large, but long seas provided they are not too steep, a situation that often prevails in the Southern Ocean. Wave breaking is another important feature. In their review of model tests, actual capsizes, and mathematical, statistical and engineering analyses, Kirkman & McCurdy (1987) concluded that 'As a rule, we believe, no non-breaking wave is dangerous' for offshore yachts. Kjeldsen (1990) has derived a theory for calculating wave loads and slamming caused by freak waves which are breaking, and presented some comparisons with oceanographic measurements. Kjeldsen's work is in response to what he perceived as the cause of capsize for the 26 Norwegian ships lost in the period 1970–78, 13 of them with no survivors.

Some of the factors that cause wave breaking were investigated in the numerical computations of Dold & Peregrine (1986). These authors observed the growth of small modulations on regular wave trains, as predicted by the Benjamin–Feir instability, and determined the ranges of wave steepness and modulation length necessary for deep water waves to break.

Details of the wave shape can also be important in assessing danger. For example the deep trough, or 'hole in the sea', often reported preceding a freak wave, presents a special hazard, as explained by Mallory (1974). When a fast ship meets a freak wave head-on, she first steams downward into the trough, burying her bow. Now the forward part of a ship usually has great buoyancy. So the bow forcefully attempts to rise just as the giant wave breaks on deck aft of the ship's forward buoyant area. The resulting shear forces on the hull can cause significant structural damage, usually near the bulkhead between Nos. 1 and 2 hatches.

REFERENCES

BACON, S. 1991 Wind waves. In *Heavy Weather Sailing* (ed. K. Adlard Coles), 4th edn. International Marine.

BENJAMIN, T. B. & FEIR, J. E. 1967 The disintegration of wave trains on deep water. Part 1. Theory. *J. Fluid Mech.* **27**, 417–430.

COURANT, R. & HILBERT, D. 1962 *Methods of Mathematical Physics*, Vol. II. Interscience.

DEAN, R. G. 1990 Freak waves: A possible explanation. In *Water Wave Kinematics* (ed. A. Torum & O. T. Gudmestad), pp. 609–612. Kluwer.

DOLD, J. W. & PEREGRINE, D. H. 1986 Water-wave modulation. In *Proc. 20th Intl Conf. on Coastal Engng* vol. 1, Chap 13, pp. 163–175.

FEIR, J. E. 1967 Discussion: some results from wave pulse experiments. *Proc. R. Soc. Lond.* A **299**, 54–58.

FELLER, W. 1952 The parabolic differential equations and associated semi-groups of transformations. *Ann. Maths* **55**, 468–519.

FORNBERG, B. 1988 Generation of finite difference formulas on arbitrarily spaced grids. *Math. Comput.* **51**, 699–706.

FORNBERG, B. 1996 *A Practical Guide to Pseudospectral Methods*. Cambridge University Press.

GERBER, M. 1987 The Benjamin–Feir instability of a deep-water stokes wavepacket in the presence of a non-uniform medium. *J. Fluid Mech.* **176**, 311–332.

GERBER, M. 1993 The interaction of deep-water gravity waves and an annular current: linear theory. *J. Fluid Mech.* **248**, 153–172.

GERBER, M. 1996 Giant waves and the Agulhas Current. *Deep-Sea Res.* (submitted).

GUTSHABASH, YE. SH & LAVRENOV, I. V. 1986 Swell transformation in the Cape Agulhas current. *Izv. Atmos. Ocean. Phys.* **22**, No. 6, 494–497.

HENDERSON, R. 1991 *Sea Sense*, p. 299. International Marine.

HESSELINK, L. & STURTEVANT, B. 1988 Propagation of weak shocks through a random medium. *J. Fluid Mech.* **196**, 513–553.

IRVINE, D. E. & TILLEY, D. G. 1988 Ocean wave directional spectra and wave-current interaction in the Agulhas from the shuttle imaging radar-B synthetic aperture radar. *J. Geophys. Res.* **93**, 15,389–15,401.

KIRKMAN, K. L. & MCCURDY, R. C. 1987 Avoiding capsize: research work. In *Desirable and Undesirable Characteristics of Offshore Yachts* (ed. J. Rousmaniere), pp. 57–74. Technical Committee of the Cruising Club of America, Nautical Quarterly Books, W. W. Norton.

KJELDSEN, S. P. 1990 Breaking waves. In *Water Wave Kinematics* (ed. A. Torum & O. T. Gudmestad), pp. 453–473. Kluwer.

KLYATSKIN, V. I. 1993 Caustics in random media. *Waves in Random Media* **3**, 93–100.

KOMEN, G. J. *et al.* (Eds.) 1994 *Dynamics and Modelling of Ocean Waves*. (subtitle: Final report of the WAM group) Cambridge University Press.

KULKARNY, V. A. & WHITE, B. S. 1982 Focusing of waves in turbulent inhomogeneous media. *Phys. Fluids* **25**, 1770–1784.

LINDGREN, G. 1970 Some properties of a normal process near a local maximum. *Ann. Math. Statist.* **41**, 1870–1883.

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.

LUDWIG, D. 1966 Uniform asymptotic expansions at a caustic. *Commun. Pure Appl. Maths* **19**, 215–250.

MALLORY, J. K. 1974 Abnormal waves on the south east coast of South Africa. *Intl Hydrog. Rev.* **51**, 99–129.

MAPP, G. R., WELCH, C. S. & MUNDAY, J. C. 1985 Wave refraction by warm core rings. *J. Geophys. Res.* **90**, 7153–7162.

MATHIESEN, M. 1987 Wave refraction by a current whirl. *J. Geophys. Res.* **92**, 3905.

MCKEAN, H. P. 1969 *Stochastic Integrals*. Academic.

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

NAIR, B. & WHITE, B. S. 1991 High-frequency wave propagation in random media – a unified approach. *SIAM J. Appl. Maths* **51**, 374–411.

PAPANICOLAOU, G. & KOHLER, W. 1974 Asymptotic theory of mixing stochastic ordinary differential equations. *Commun. Pure Appl. Maths* **27**, 641–668.

PEARCEY, T. 1946 The structure of an electromagnetic field in the neighborhood of a cusp of a caustic. *Lond. Edinb. Dublin Phil. Mag.* (7) **37**, 311–317.

PEREGRINE, D. H. 1976 Interaction of water waves and currents. *Adv. Appl. Mech.* **16**, 9–117.

PEREGRINE, D. H. 1985 Water waves and their development in space and time. *Proc. R. Soc. Lond.* A **400**, 1–18.

PEREGRINE, D. H. 1986 Approximate descriptions of the focussing of water waves. *Proc. 20th Intl Conf. Coastal Engng*, vol. 1, Chap. 51, pp. 675–685.

PEREGRINE, D. H., SKYNER, D., STIASSNIE, M. & DODD, N. 1988 Nonlinear effects on focussed water waves. *Proc. 21st Intl Conf. Coastal Engng*, vol. 1, Chap. 54, pp. 732–742.

PHILLIPS, O. M., GU, D. & DONELAN, M. 1993*a* Expected structure of extreme waves in a Gaussian sea. Part I: Theory and swade buoy measurements. *J. Phys. Oceanogr.* **23**, 992–1000.

PHILLIPS, O. M., GU, D. & WALSH, E. 1993*b* On the expected structure of extreme waves in a Gaussian sea. Part II: SWADE Scanning Radar Altimeter Measurements. *J. Phys. Oceanogr.* **23**, 2297–2309.

SAND, S. E. *et al.* 1990 Freak wave kinematics. In *Water Wave Kinematics* (ed. A. Torum & O. T. Gudmestad), pp. 535–549. Kluwer.

SLOCUM, J. 1899 *Sailing Alone Around the World*, Chap. VIII. The Century Co. (Reprinted 1954 Sheridan House.)

SMEETON, M. 1959 *Once is Enough*. Rupert Hart-Davis Ltd. (Reprinted 1991, Billings and Sons Ltd.)

SMITH, R. 1976 Giant waves. *J. Fluid Mech.* **77**, 417–431.

STIASSNE, M. & PEREGRINE, D. H. 1979 On averaged equations for finite amplitude water waves. *J. Fluid Mech.* **94**, 401–407.

SU, M.-Y., BERGIN, M., MARLER, P. & MYRICK, R. 1982 Experiments on nonlinear instabilities and evolution of steep gravity-wave trains. *J. Fluid Mech.* **124**, 45–72.

TRULSEN, K. & DYSTHE, K. B. 1996 Freak waves: a 3-D wave simulation. In *Proc. 21st Intl Symp. on Naval Hydrodynamics*. Trondheim, Norway.

VAN DORN, W. G. 1993 *Oceanography and Seamanship*. 2nd Edn. Cornell Maritime Press.

WHITE, B. S. 1984 The stochastic caustic. *SIAM J. Appl. Maths* **44**, 127–149.

WHITE, B., NAIR, B. & BAYLISS, A. 1988 Random rays and seismic amplitude anomalies. *Geophysics* **53**, 903–907.

WHITE, B. S., NORRIS, A., BAYLISS, A. & BURRIDGE, R. 1987 Some remarks on the Gaussian beam summation method. *Geophys. J. R. Astron. Soc.* **89**, 579–636.

WHITHAM, G. B. 1974 *Linear and Non-Linear Waves*. Wiley-Interscience.

ZWILLINGER, D. I. & WHITE, B. S. 1985 Propagation of initially plane waves in the region of random caustics. *Wave Motion* **7**, 202–227.