J. Fluid Mech. (2009), *vol.* 620, *pp.* 221–239.
doi:10.1017/S0022112008004424 Printed in the United Kingdom

On nonlinear wave groups and crest statistics

FRANCESCO FEDELE† AND M. AZIZ TAYFUN

School of Civil and Environmental Engineering, Georgia Institute of Technology, Savannah, Georgia, USA

Civil Engineering Department, College of Engineering and Petroleum, Kuwait University, Kuwait

(Received 25 October 2007 and in revised form 17 September 2008)

We present a second-order stochastic model of weakly nonlinear waves and develop theoretical expressions for the expected shape of large surface displacements. The model also leads to an exact theoretical expression for the statistical distribution of large wave crests in a form that generalizes the Tayfun distribution (Tayfun, *J. Geophys. Res.*, vol. 85, 1980, p. 1548). The generalized distribution depends on a steepness parameter given by $\mu = \lambda_3/3$, where λ_3 represents the skewness coefficient of surface displacements. It converges to the Tayfun distribution in narrowband waves, where both distributions describe the crests of all waves well. In broadband waves, the generalized distribution represents the crests of large waves just as well whereas the Tayfun distribution appears as an upper bound and tends to overestimate them. However, the theoretical nature of the generalized distribution presents practical difficulties in oceanic applications. We circumvent these by adopting an appropriate approximation for the steepness parameter. Comparisons with wind-wave measurements from the North Sea suggest that this approximation allows both distributions to assume an identical form with which we can describe the distribution of large wave crests fairly accurately. The same comparisons also show that third-order nonlinear effects do not appear to have any discernable effect on the statistics of large surface displacements or wave crests.

1. Introduction

In linear random seas, relatively large waves appear more regular or 'narrowband' in the sense that they do not display any secondary maxima or minima, but a single dominant crest (Longuet-Higgins 1957; Tayfun in press). This allows the expected shape of large surface displacements and the distribution of large crests to be described reasonably well based on Gaussian statistics and the Rayleigh distribution, respectively (Longuet-Higgins 1957; Lindgren 1970, 1972; Lindgren & Rychlik 1991; Phillips, Gu & Donelan 1993; Boccotti 2000). However, oceanic waves are nonlinear, displaying sharper narrower crests and shallower more rounded troughs. Accordingly, the distribution of surface displacements tends to deviate from the Gaussian form with a positive skewness. For such waves, the exact forms of the theoretical expressions describing the distributions of surface displacements and wave crests are not known under general conditions. Longuet-Higgins (1963) approximated the distribution of surface elevations in a Gram-Charlier type series. Tayfun (1980) explored a Stokes-like model, and devised a simple theoretical model for describing the distribution of nonlinear crests in narrowband waves. Since 1980s, numerous nonlinear crest-height models have also been proposed by Winterstein (1988), Marthinsen &

† Email address for correspondence: ffedele3@gtsav.gatech.edu

Winterstein (1992), Kriebel & Dawson (1993), Forristall (2000), Prevosto, Krogstad & Robin (2000), Arena & Fedele (2002), Prevosto & Forristall (2002), Tromans & Vanderschuren (2004), Milder (2007) and others. These provide valuable insight and work well with varying degrees of success in describing the distribution of crests of large waves. Some models are basically statistical in nature in the sense that they either use nonlinear transformations of Gaussian variables analogous to Gram-Charlier type expansions or fit observational data to a theoretical distribution defined in terms of parameters which reflect second-order nonlinear corrections (Winterstein 1988; Forristall 2000). Some represent extensions of the narrowband approximations to shallow-water waves (Marthinsen & Winterstein 1992; Prevosto *et al.* 2000) or heuristic modifications of the Tayfun model (Kriebel & Dawson 1993), and others require intricate analytics, directional spectra or numerical computation (Arena & Fedele 2002; Prevesto & Forristall 2002; Tromans & Vanderschuren 2004; Milder 2007). As for the expected shape of large waves in second-order seas, Jensen (1996, 2005) developed theoretical approximations based on Gram-Charlier expansions, and Tayfun & Fedele (2007*a*) considered simpler narrowband approximations and also formulated an exact theoretical expression, extending the linear quasi-deterministic theory of Boccotti (2000) to second-order waves.

Simulations of relatively broadband waves, linear or nonlinear, and oceanic observations of wind waves show that wave crests in general follow a complex bimodal distribution characterized by a relatively sharp narrow mode over the range of low waves, and a more familiar and wider second mode over larger waves (Tayfun 2006). This structure cannot be predicted by any theoretical expression devised so far in any of the aforementioned studies. Exact theoretical representations in the Rayleigh form for linear waves and in the Tayfun form for nonlinear waves are valid only for the crests of extremely long-crested and unrealistically narrowband waves. Amplitudes of such waves are also prone to symmetric amplifications induced by third-order modulational instabilities which can modify the distribution of various surface features well beyond the second-order predictions (Socquet-Juglard *et al.* 2005; Mori & Janssen 2006; Onorato *et al.* 2006). Although the Tayfun model coupled with third-order Gram-Charlier expansions seems to work reasonably well in such cases (Tayfun & Fedele 2007*b*), modulational instabilities do not appear likely in short-crested oceanic wind waves. This is confirmed by directional simulations of relatively broadband waves by way of the modified nonlinear Schödinger (NLS) equation (Dysthe 1979; Socquet-Juglard *et al.* 2005; Gramstad & Trulsen 2007) as well as by the statistics derived from oceanic measurements (Forristall 2000, 2007; Tayfun & Fedele 2007*b*, 2008).

Herein, we first briefly review and extend Boccotti's linear quasi-deterministic theory (Boccotti 1989, 2000) to second-order random waves, and derive a new stochastic representation of wave groups for describing the statistical structure of the sea surface around a large crest. This enables us to obtain exact theoretical expressions describing the expected shape of large surface displacements and the statistical distribution of associated wave crests. The first of these results agrees with a previous theoretical expression derived somewhat heuristically in Tayfun & Fedele (2007*b*), where it is also compared with the approximations of Jensen (1996, 2005), Tayfun & Fedele (2007*a*) and oceanic observations. We, therefore, focus our attention in the remainder of the paper on the statistical structure of large wave crests, and derive their statistical distribution in a form that generalizes the Tayfun model.

The relative validity and accuracy of the generalized model in representing the statistics of large wave crests depend on the selection of a dimensionless steepness

parameter μ. In essence, μ characterizes the nonlinearity of large waves, and it is closely associated with the vertical skewness of the nonlinear sea surface. We explore the nature of the generalized model and compare it with the original Tayfun model by way of simple second-order simulations of infinitely long-crested waves. In all comparisons with the original Tayfun model, we stay faithful to the narrowband definition of the steepness parameter in Tayfun (1980). We first show that the generalized distribution describes the distribution of simulated crests over large waves quite well. It also converges to the Tayfun distribution in narrowband waves, where both tend to describe crests of all waves quite accurately. For broadband waves, the generalized distribution describes large wave crests just as well whereas the Tayfun distribution appears as an upper bound and tends to overpredict them. However, the theoretical nature of the generalized distribution, in particular, its dependence on a steepness definition expressed in terms of the first-order integral properties present practical difficulties in oceanic applications. We overcome these by modifying the definition of steepness parameter slightly, rendering the generalized model more practical and the Tayfun model more accurate for applications. Comparisons with wind-wave measurements from the North Sea show that the steepness parameter so modified reduces both models to a simple identical form with which the distribution of large wave crests can be described fairly well. The same comparisons also suggest that third-order nonlinear effects such as NLS type modulational instabilities do not appear to have any discernable effect on the statistics of large surface displacements or wave crests.

2. Second-order waves

We consider weakly nonlinear random waves propagating in water of uniform depth d. In Cartesian coordinates $\boldsymbol{x}=(x, y)$ coincident with the mean sea level, z-axis pointing upward and t time, the second-order sea surface displacement $\zeta(\boldsymbol{x}, t)$ and the associated velocity potential $\phi(\boldsymbol{x}, z, t)$ are given, respectively, by the real parts of

$$\zeta = \zeta_1 + \zeta_2, \qquad \phi = \phi_1 + \phi_2, \tag{2.1}$$

with

$$\zeta_1(\boldsymbol{x}, t) = \sum_{n=1}^{N} c_n \mathrm{e}^{\mathrm{i}\theta_n}, \tag{2.2}$$

$$\phi_1(\boldsymbol{x}, z, t) = -\mathrm{i}g \sum_{n=1}^{N} \frac{c_n}{\omega_n} \frac{\cosh k_n(z+d)}{\cosh k_n d} \mathrm{e}^{\mathrm{i}\theta_n}, \tag{2.3}$$

$$\zeta_2(\boldsymbol{x}, t) = \frac{1}{4} \sum_{n,m=1}^{N} c_n c_m \left[A_{nm}^{+} \mathrm{e}^{\mathrm{i}(\theta_n+\theta_m)} + A_{nm}^{-} \mathrm{e}^{\mathrm{i}(\theta_n-\theta_m)} \right], \tag{2.4}$$

$$\phi_2(\boldsymbol{x}, z, t) = -\mathrm{i}\frac{g^2}{4} \sum_{n,m=1}^{N} c_n c_m (\omega_n \omega_m)^{-1} \left[B_{nm}^{+} \mathrm{e}^{\mathrm{i}(\theta_n+\theta_m)} + B_{nm}^{-} \mathrm{e}^{\mathrm{i}(\theta_n-\theta_m)} \right], \tag{2.5}$$

where $\theta_n = \boldsymbol{k}_n \cdot \boldsymbol{x} - \omega_n t + \delta_n = k_n x \cos\varphi_n + k_n y \sin\varphi_n - \omega_n t + \delta_n$, and $A_{nm}^{\pm} = A^{\pm}(\boldsymbol{k}_n, \boldsymbol{k}_m)$ and $B_{nm}^{\pm} = B^{\pm}(\boldsymbol{k}_n, \boldsymbol{k}_m, z)$ represent interaction coefficients (cf. Sharma & Dean 1979), $\boldsymbol{k}_n$ is horizontal wavenumber vector, with $k_n = |\boldsymbol{k}_n|$; φ_n is direction measured from the x-axis; and ω_n is angular frequency, related to k_n via $k_n \tanh k_n d = \omega_n^2/g$, with g the gravitational acceleration. Phases δ_n are independent and uniformly distributed

in $[0, 2\pi]$, and amplitudes c_n are related to the spectrum S of ζ_1 in the form

$$S(\boldsymbol{k})\,\mathrm{d}\boldsymbol{k}=S(k,\varphi)k\delta k\delta\varphi=\sum_n \frac{|c_n|^2}{2}, \tag{2.6}$$

where the sum is over $(k_n, \varphi_n) \in ([k, k+\delta k], [\varphi, \varphi+\delta\varphi])$. The jth order spectral moment is

$$m_j = \int_0^\infty \omega^j S(\boldsymbol{k})\,\mathrm{d}\boldsymbol{k}. \tag{2.7}$$

In theory, the validity of the form assumed for ζ requires that the *r.m.s.* surface gradient of ζ_1 be sufficiently small. Specifically,

$$\varepsilon = \sqrt{\langle|\nabla\zeta_1|^2\rangle} = \frac{\sqrt{m_4}}{g} \ll 1, \tag{2.8}$$

where $\langle\cdot\rangle$ denotes expectation. The spectral mean frequency ω_m, and bandwidth ν of S irrespective of direction are defined by

$$\omega_m = \frac{m_1}{m_0}, \qquad \nu = \sqrt{\frac{m_0 m_2}{m_1^2} - 1}. \tag{2.9}$$

To $O(\varepsilon)$, the space–time covariance Ψ of both ζ_1 and ζ is given in continuous form by

$$\Psi(\boldsymbol{X}, T) = \langle\zeta_1(\boldsymbol{x}, t)\,\zeta_1(\boldsymbol{x}+\boldsymbol{X}, t+T)\rangle = \int S(\boldsymbol{k})\cos(\boldsymbol{k}\cdot\boldsymbol{X}-\omega T)\,\mathrm{d}\boldsymbol{k}, \tag{2.10}$$

where $\boldsymbol{X}=(X, Y)$. Setting $\psi(T)=\Psi(\boldsymbol{0}, T)$ for simplicity, $\psi(0)=m_0=\sigma^2$ gives the absolute maximum of ψ. We assume that the first absolute minimum of ψ occurs at $T=T^*$ and that ψ decreases monotonically between $T=0$ and $T=T^*$. Similarly, the cross-covariance of ζ and ϕ is given to $O(\varepsilon)$ by that of ζ_1 and ϕ_1 as

$$\Phi(\boldsymbol{X}, z, T) = \langle\zeta_1(\boldsymbol{x}, t)\,\phi_1(\boldsymbol{x}+\boldsymbol{X}, z, t+T)\rangle = g\int S(\boldsymbol{k})\frac{\cosh k(z+d)}{\omega\cosh kd}\sin(\boldsymbol{k}\cdot\boldsymbol{X}-\omega T)\,\mathrm{d}\boldsymbol{k}. \tag{2.11}$$

3. Skewness coefficient

The skewness coefficient of ζ is defined by $\lambda_3 = \langle\zeta^3\rangle/\langle\zeta^2\rangle^{3/2} = 3\langle\zeta_1^2\zeta_2\rangle/\langle\zeta_1^2\rangle^{3/2}$ correct to $O(\varepsilon)$. Thus, it follows from (2.2) and (2.4) that

$$\lambda_3 = \frac{3\langle\zeta_1^2\zeta_2\rangle}{\langle\zeta_1^2\rangle^{3/2}} = \frac{3}{2\sigma^3}\int[A^+(\boldsymbol{k}_1, \boldsymbol{k}_2) + A^-(\boldsymbol{k}_1, \boldsymbol{k}_2)]S(\boldsymbol{k}_1)S(\boldsymbol{k}_2)\,\mathrm{d}\boldsymbol{k}_1\mathrm{d}\boldsymbol{k}_2 + O(\varepsilon^2). \tag{3.1}$$

The interaction coefficients A^+ and A^- relate to the second-order bound harmonics with frequencies $\omega_1+\omega_2$ and $\omega_1-\omega_2$, respectively. In the most general case, both coefficients assume rather intricate functional forms, not amenable to general analysis. In deep water, however, $A^\pm$ and thus (3.1) reduces to simpler expressions, which have previously been explored at length by Longuet-Higgins (1963), Srokozs & Longuet-Higgins (1986), Tayfun (1986), Vinje & Haver (1994) and others. In general, $\lambda_3 > 0$ always. It also attains peak values when all component waves travel nearly in the same direction. Clearly, (3.1) can be rewritten as $\lambda_3 = \lambda_3^+ + \lambda_3^-$, where λ_3^+ and λ_3^- represent the contributions from A^+ and A^-, respectively. For waves travelling all in the same direction, $A^\pm = \pm|\omega_1^2 \pm \omega_2^2|/g$ and so

$$\lambda_3^+/3\mu_m = 1 + \nu^2, \tag{3.2}$$

where $\mu_m = \sigma\omega_m^2/g$ is an integral measure of wave steepness. Further, while $\lambda_3^+ > 0$ clearly, λ_3^- is always negative. Thus, bound harmonics representing frequency differences $\omega_1 - \omega_2$ tend to reduce λ_3, as first noted some years ago (Tayfun 1986). Unfortunately, λ_3^- does not admit a simple expression in terms of integral spectral properties as λ_3^+ does. To gain some further insight into the nature of λ_3^-, we can make use of $\omega_1^2 - \omega_2^2 \leqslant |\omega_1^2 - \omega_2^2| = (\omega_1 + \omega_2)|\omega_1 - \omega_2|$ and Schwarz's inequality in (3.1) to obtain (Tayfun 2006)

$$\nu^2 \leqslant |\lambda_3^-|/3\mu_m \leqslant \nu(2 + \nu^2). \tag{3.3}$$

We can now combine (3.2) with $\lambda_3^-/3\mu_m \leqslant -\nu^2$ from the preceding expression to show that $\lambda_3 \leqslant 3\mu_m$ in general, with the equality holding as an upper bound in the narrowband limit as $\nu \to 0$ and $\lambda_3^- \to 0$. The theoretical maximum of λ_3 in directional waves is only about 1 % larger than the preceding upper bound (cf. Longuet-Higgins 1963).

To illustrate the preceding results explicitly and, in particular, how λ_3 varies with the spectral bandwidth, consider waves all travelling in the same direction with a frequency spectrum of the 'generalized Phillips' form

$$S(\omega) = \begin{cases} \dfrac{m_0(n-1)}{\omega_p}\left(\dfrac{\omega_p}{\omega}\right)^n, & \omega \geqslant \omega_p, \\ 0, & \omega < \omega_p, \end{cases} \tag{3.4}$$

where ω_p is the frequency at the spectrum peak. We consider $n \geqslant 4$ as a free variable. On this basis,

$$\nu^2 = 1/(n-1)(n-3), \qquad \lambda_3^+/3\mu_m = 1 + \nu^2, \qquad \lambda_3^-/3\mu_m = -\nu\sqrt{1+\nu^2}, \tag{3.5}$$

and

$$\lambda_3/3\mu_m = 1 - \nu\sqrt{1+\nu^2} + \nu^2. \tag{3.6}$$

These are plotted in figure 1. Evidently, $|\lambda_3^-|/3\mu_m \approx O(\nu)$, and it decreases as ν does for large n. The dominant component $\lambda_3^+/3\mu_m$ is $O(1)$. As $n \to \infty$, $\nu \to 0$, $\lambda_3^- \to 0$ and $\lambda_3 \to \lambda_3^+ \to 3\mu_m$.

The skewness coefficient represents the principal parameter with which we describe the effects of second-order nonlinearities on the geometry and statistics of the sea surface with higher sharper crests and shallower more rounded troughs. For instance, the distribution of the normalized surface displacements is described as (Longuet-Higgins 1963)

$$p_\zeta(z) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{z^2}{2}\right)\left[1 + \frac{\lambda_3}{6}z(z^2-3)\right]. \tag{3.7}$$

In linear waves where $\lambda_3 = 0$, the preceding expression yields the Gaussian density. The latter describes a symmetric surface structure with respect to the mean sea level in that $P^- = Pr\{\zeta \leqslant 0\} = 1/2$ and $P^+ = Pr\{\zeta > 0\} = 1/2$. In other words, surface displacements are equally likely both above and below the mean sea level. In nonlinear waves where $\lambda_3 > 0$, (3.7) leads to

$$P^\pm = \frac{1}{2}\left(1 \mp \frac{\lambda_3}{6\sqrt{2\pi}}\right). \tag{3.8}$$

So, $P^+ < 1/2$ and thus $P^- > 1/2$, more generally, implying that surface displacements are more likely below the mean level than above it. Physically, the surface stays somewhat longer below the mean sea level. If, however, we focus attention only on the surface displacements which in absolute value exceed relatively large thresholds symmetrically located above and below the mean level, we would see that just the

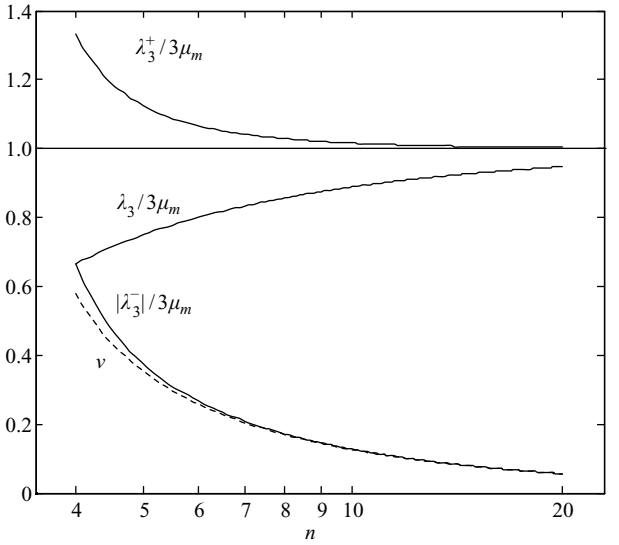

FIGURE 1. Generalized Phillips spectrum: the variations of bandwidth ν, skewness coefficient $\lambda_3 = \lambda_3^+ + \lambda_3^-$, λ_3^+ and λ_3^- with exponent n ($\mu_m = \sigma\omega_m^2/g$).

opposite occurs. These and various other results exploring the physical significance of λ_3 and its effects on the statistics of surface displacements, wave envelopes and phases are elaborated elsewhere (Longuet-Higgins 1963; Srokozs & Longuet-Higgins 1986; Tayfun 1994, in press).

4. Stochastic wave groups and large crests in Gaussian seas

Consider first linear waves and assume that a large crest of amplitude h is observed at $\boldsymbol{x} = \boldsymbol{x}_0$ and $t = t_0$. Boccotti (cf. Boccotti 2000) showed that as $h/\sigma \to \infty$ the large crest occurs when a well-defined wave group $\zeta_{1,c}$ passes through $\boldsymbol{x}_0$. The surface displacement $\zeta_{1,c}$ around $\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{X}$ and $t = t_0 + T$ is asymptotically described by the conditional process

$$\zeta_{1,c} = \{\zeta_1(\boldsymbol{X}, T) \mid \zeta_1(\boldsymbol{0}, 0) = h\} = \zeta_{1,det} + \mathscr{R}_{\zeta_1}, \tag{4.1}$$

where $\zeta_{1,det}$ is of $O(h)$ and given by

$$\zeta_{1,det}(\boldsymbol{X}, T) = \langle \zeta_1(\boldsymbol{X}, T) \mid \zeta_1(\boldsymbol{0}, 0) = h \rangle = h\Psi_c(\boldsymbol{X}, T), \tag{4.2}$$

with $\Psi_c = \Psi/\sigma^2$ (Lindgren 1970, 1972; Lindgren & Rychlik 1991), and $\mathscr{R}_{\zeta_1}(\boldsymbol{X}, T)$ represents a random residual of $O(h^0)$. Correspondingly, the velocity potential $\phi_{1,c}$ of $\zeta_{1,c}$, conditional on $\zeta_1(\boldsymbol{0}, 0) = h$, is given by

$$\phi_{1,c} = \{\phi_1(\boldsymbol{X}, z, T) \mid \zeta_1(\boldsymbol{0}, 0) = h\} = \phi_{1,det} + \mathscr{R}_{\phi_1}, \tag{4.3}$$

where

$$\phi_{1,det}(\boldsymbol{X}, z, T) = \langle \phi_1(\boldsymbol{X}, z, T) \mid \zeta_1(\boldsymbol{0}, 0) = h \rangle = h\Phi_c(\boldsymbol{X}, z, T), \tag{4.4}$$

with $\Phi_c = \Phi/\sigma^2$ and $\mathscr{R}_{\phi_1}(\boldsymbol{X}, z, T)$ is the random residual of $O(h^0)$.

Physically, $\zeta_{1,c}$ represents a large wave group which evolves linearly through a background wave of $O(h^0)$ represented by the residual $\mathscr{R}_{\zeta_1}$. The largest crest occurs at the apex of the group. As $h/\sigma \to \infty$, both $\mathscr{R}_{\zeta_1}$ and $\mathscr{R}_{\phi_1}$ can be expressed explictly

in terms of Ψ and Φ, respectively (see the Appendix for details). Thus, for $h/\Delta \gg 1$, $\zeta_{1,c}$ and $\phi_{1,c}$ can be written as

$$\zeta_{1,c} = h\Psi_c + \Delta Q_{\zeta_1} + O(h^{-1}), \qquad \phi_{1,c} = h\Phi_c + \Delta Q_{\phi_1} + O(h^{-1}), \tag{4.5}$$

where Δ is of $O(h^0)$ and the deterministic functions Q_{ζ_1} and Q_{ϕ_1} are given, respectively, by

$$Q_{\zeta_1}(\boldsymbol{X}, T) = \frac{\Psi_c(\boldsymbol{X}, T - T^*) - \psi^* \Psi_c(\boldsymbol{X}, T)}{1 - \psi^{*2}},$$

$$Q_{\phi_1}(\boldsymbol{X}, z, T) = \frac{\Phi_c(\boldsymbol{X}, z, T - T^*) - \psi^* \Phi_c(\boldsymbol{X}, z, T)}{1 - \psi^{*2}},$$

with $\psi^* = \psi(T^*)/\sigma^2$. If we interpret h and Δ as random variables, then $\zeta_{1,c}$ identifies a stochastic wave group, which together with $\phi_{1,c}$ describes the kinematics and dynamics locally around a randomly chosen crest. Note that, with h given in (4.5), averaging over Δ yields the conditional means $\zeta_{1,det}$ and $\phi_{1,det}$ as expected.

The joint p.d.f. of $\xi = h/\sigma$ and $\tilde{\Delta} = \Delta/\sigma$ is given, as $\xi \to \infty$, by (Boccotti 1989)

$$p_{\xi,\tilde{\Delta}}(\xi, \tilde{\Delta}) = p_{\xi}(\xi) p_{\tilde{\Delta}}(\tilde{\Delta}), \tag{4.6}$$

where

$$p_{\xi}(\xi) = \xi \exp\left(-\frac{\xi^2}{2}\right), \qquad p_{\tilde{\Delta}}(\tilde{\Delta}) = \frac{\exp(-\tilde{\Delta}^2/(2(1-\psi^{*2})))}{\sqrt{2\pi(1-\psi^{*2})}}. \tag{4.7}$$

Thus, ξ and $\tilde{\Delta}$ are independent. Here, we shall show that large crests in second-order nonlinear seas arise from focusing in stochastic wave groups.

5. Nonlinear waves and crests

The preceding results can be generalized to second-order waves, simply noting that the nonlinear surface ζ is described by a nonlinear mapping of the form (cf. Fedele & Arena 2005; Fedele 2006)

$$\zeta = \zeta_1 + \zeta_2 = f(\zeta_1) \tag{5.1}$$

between ζ_1 and ζ at $(\boldsymbol{X}, T)$ based on (2.1), (2.2) and (2.4). Assume for the moment that we observe a large crest of amplitude h_{nl} at $\boldsymbol{X} = \boldsymbol{0}$ and $T = 0$. The nonlinear surface ζ surrounding that crest locally is given by the conditional process

$$\zeta_c = \{\zeta \mid \zeta(\boldsymbol{0}, 0) = h_{nl}\}. \tag{5.2}$$

We can simplify this expression further by exploiting the weakly nonlinear nature of (5.1) and recalling that ζ_2 is phase-locked to the extremes of the first-order surface ζ_1. So, a large crest of ζ with amplitude h_{nl} occurs simultaneously when ζ_1 itself has a large crest with an amplitude, say, h. Thus, (5.2) is equivalent to

$$\zeta_c = \{\zeta \mid \zeta_1(\boldsymbol{0}, 0) = h\} = \{\zeta \mid \zeta_1 = \zeta_{1,c}\}. \tag{5.3}$$

This can be simplified further, using (5.1), to

$$\zeta_c = f(\zeta_{1,c}), \tag{5.4}$$

where $\zeta_{1,c}$ is the Gaussian group of (4.5).

For long-crested narrowband waves in deep water, ζ_c becomes (Tayfun 1980, 1986)

$$\zeta_c = \zeta_{1,c} + \frac{\omega_m^2}{2g}\left(\zeta_{1,c}^2 - \hat{\zeta}_{1,c}^2\right) + O(\nu), \tag{5.5}$$

where $\hat{\zeta}_{1,c}$ denotes the Hilbert transform of $\zeta_{1,c}$ with respect to T. To describe ζ_c, or equivalently $f(\zeta_{1,c})$, explicitly under more general conditions, we consider the discrete form of the surface $\zeta_{1,c}$ given by the real part of

$$\zeta_{1,c} = \sum_{n=1}^{N} C_n \mathrm{e}^{\mathrm{i}(\boldsymbol{k}_n \cdot \boldsymbol{X} - \omega_n T)}, \qquad C_n \simeq \left(h + \Delta \frac{\mathrm{e}^{\mathrm{i}\omega_n T^*} - \psi^*}{1 - \psi^{*2}} \right) \frac{S(\boldsymbol{k}_n)}{\sigma^2} d\boldsymbol{k}_n.$$

We note that (2.1) along with (2.2) and (2.4) *not only* defines weakly nonlinear random waves *but also* the general solution for the second-order surface, if c_n and $\boldsymbol{\theta}_n$ are regarded as deterministic variables. Thus, we replace in (2.1) the linear component ζ_1 of the surface ζ with $\zeta_{1,c}$ by setting

$$c_n = C_n, \qquad \boldsymbol{\theta}_n = \boldsymbol{k}_n \cdot \boldsymbol{X} - \omega_n T,$$

and take the real part of the resulting ζ to express ζ_c in a continuous form as

$$\zeta_c = f(\zeta_{1,c}) = h\Psi_c + \frac{h^2}{4\sigma}\mathscr{F} + \frac{h\Delta}{2\sigma}\frac{\mathscr{G} - \psi^*\mathscr{F}}{1 - \psi^{*2}} + \frac{\Delta^2}{4\sigma}\frac{\mathscr{H} - 2\psi^*\mathscr{G} + \psi^{*2}\mathscr{F}}{(1 - \psi^{*2})^2}, \tag{5.6}$$

where

$$\mathscr{F}(\boldsymbol{X}, T) = \int \frac{S_1 S_2}{\sigma^3}(A_{12}^{+} \cos\beta_{12}^{+} + A_{12}^{-} \cos\beta_{12}^{-})\mathrm{d}\boldsymbol{k}_1 \mathrm{d}\boldsymbol{k}_2, \tag{5.7}$$

$$\mathscr{G}(\boldsymbol{X}, T) = \int \frac{S_1 S_2}{\sigma^3}[A_{12}^{+} \cos(\beta_{12}^{+} + \omega_1 T^*) + A_{12}^{-} \cos(\beta_{12}^{-} + \omega_1 T^*)]\,\mathrm{d}\boldsymbol{k}_1 \mathrm{d}\boldsymbol{k}_2,$$

$$\mathscr{H}(\boldsymbol{X}, T) = \int \frac{S_1 S_2}{\sigma^3}\{A_{12}^{+} \cos[\beta_{12}^{+} + (\omega_1 + \omega_2)T^*] + A_{12}^{-} \cos[\beta_{12}^{-} + (\omega_1 + \omega_2)T^*]\} \times \mathrm{d}\boldsymbol{k}_1 \mathrm{d}\boldsymbol{k}_2,$$

with $S_j = S(\boldsymbol{k}_j)$, $j = 1, 2$, and $\beta_{12}^{\pm} = (\boldsymbol{k}_1 \pm \boldsymbol{k}_2) \cdot \boldsymbol{X} - (\omega_1 \pm \omega_2)T$. Evidently, the nonlinear effects of $O(h\Delta)$ and $O(\Delta^2)$ are clearly identified in (5.6). As we emphasize wave crests, we shall not dwell on the derivation of the nonlinear velocity potential ϕ_c. The latter can easily be obtained by applying the same approach to the nonlinear mapping between ϕ_c and $\phi_{1,c}$ given by (2.1), (2.3) and (2.5).

5.1. *Expected shape of large waves*

In general, (5.4) can be rewritten by way of (4.5) as

$$\zeta_c = f(\zeta_{1,c}) = f(h\Psi_c + \Delta Q_{\zeta_1}). \tag{5.8}$$

We recall that $h = \xi\sigma$ and $\Delta = \tilde{\Delta}\sigma$ are random variables with the joint p.d.f. (4.6). Thus, given h, we average the preceding expression with respect to Δ to obtain

$$\langle \zeta \mid \zeta_1(\boldsymbol{0}, 0) = h \rangle = \langle \zeta \mid \zeta_1 = \zeta_{1,c} \rangle_\Delta = \langle f(h\Psi_c + \Delta Q_{\zeta_1}) \rangle_\Delta = f(h\Psi_c) + O\left(\sigma_c^2\right), \tag{5.9}$$

where $\sigma_c^2(\boldsymbol{X}, T)$ is the linear conditional variance of $\zeta_{1,c}$, and $\sigma_c^2(\boldsymbol{0}, 0) = 0$. As $\zeta \to \infty$, the second and higher order terms on the right-hand side of (5.9) become insignificant relative to $f(h\Psi_c) = f(\zeta_{1,det})$. Thus, the leading asymptotic term of $\langle \zeta | \zeta_1 = \zeta_{1,c} \rangle$ is

$f(\zeta_{1,det})$, which follows from (5.6) as

$$\langle \zeta \mid \zeta_1(\mathbf{0},0)=h\rangle = h\Psi_c + \frac{h^2}{4\sigma}\mathscr{F} + O(h^0). \tag{5.10}$$

This provides the analytical form of the expected shape of large waves valid under general conditions in deep or transitional water depths. For narrowband waves all travelling in the same direction in deep water, $A_{12}^{+}=2\omega_m^2/g+O(\nu)$, $A_{12}^{-}=O(\nu)$, and so (5.10) reduces to the simpler expression

$$\langle \zeta \mid \zeta_1(\mathbf{0},0)=h\rangle = h\Psi_c + \frac{h^2\omega_m^2}{2g}\left(\Psi_c^2-\hat{\Psi}_c^2\right)+O(\nu), \tag{5.11}$$

where $\hat{\Psi}_c$ denotes the Hilbert transform of Ψ_c with respect to T. The preceding result also follows from the nonlinear mapping (5.5) in (5.9) (cf. Tayfun & Fedele 2007*a*). In general, (5.11) and a more general expression derived by Jensen (1996, 2005) based on Gram–Charlier expansions both represent approximations in contrast with (5.10), which is exact because it satisfies the Stokes equations to second order. The leading asymptotic term $f(h\Psi_c)$ of (5.9) can also be derived by solving the Stokes equations to second order (Arena 2005). However, this approach requires rather cumbersome algebra which can clearly be avoided if one realizes that the second-order solution of Sharma & Dean (1979) has a general form valid for waves characterized with either random or deterministic amplitudes and phases (Tayfun & Fedele 2007*a*).

5.2. *Distribution of nonlinear crests*

The highest crest of the nonlinear stochastic wave group ζ_c occurs at $\boldsymbol{X}=\mathbf{0}$ and $T=0$ correct to $O(\varepsilon)$. Further, the dimensionless amplitude $\xi_{max}=h_{nl}/\sigma$ can be expressed in the Tayfun form as

$$\xi_{max} = \xi + \frac{\mu^*}{2}\xi^2, \tag{5.12}$$

where

$$\mu^* = \mu\left(1+K\frac{\tilde{\Delta}}{\xi}\right)+O(\xi^{-2}), \tag{5.13}$$

and

$$\mu = \frac{\mathscr{F}(\mathbf{0},0)}{2}=\frac{\lambda_3}{3}, \qquad K=2\frac{-\psi^*+\kappa_1}{1-\psi^{*2}}, \tag{5.14}$$

with

$$\kappa_1 = \frac{\mathscr{G}(\mathbf{0},0)}{\mathscr{F}(\mathbf{0},0)} = \frac{\langle \zeta_1(\mathbf{0},t)\zeta_1(\mathbf{0},t+T^*)\zeta_2(\mathbf{0},t)\rangle}{\mu}. \tag{5.15}$$

The distribution of μ^*, conditional on ξ, is Gaussian with the mean and standard deviation given, respectively, by

$$\langle \mu^* \mid \xi\rangle = \mu = \frac{\lambda_3}{3}, \qquad \sigma_{\mu^*|\xi} = \frac{\mu K}{\xi}\sqrt{1-\psi^{*2}}. \tag{5.16}$$

Figure 2 illustrates the nature of the parameters in the preceding expressions for the spectrum (3.4). Evidently, $\mu K\sqrt{1-\psi^{*2}}$ is $O(10^{-2})$ as a typical result. Thus, $\sigma_{\mu^*|\xi}\simeq 0$ for $\xi\gg 1$, in which case $\mu^*\to\mu$ and (5.12) converges to

$$\xi_{max} = \xi + \frac{\mu}{2}\xi^2. \tag{5.17}$$

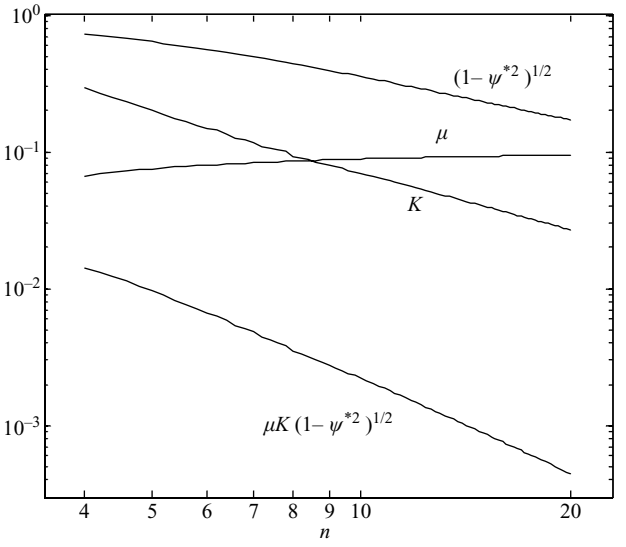

FIGURE 2. Generalized Phillips spectrum: the variations of parameters in (5.16) with n ($m_0 = 1\,\mathrm{m}^2$, $\omega_m = 1\,\mathrm{rad\,s}^{-1}$ and $\mu_m \approx 0.1$).

A parallel derivation for the minimum, say, ξ_{min} of the second-order sea surface around a relatively deep trough gives

$$\xi_{min} = -\xi + \frac{\mu}{2}\xi^2. \tag{5.18}$$

Thus, the most obvious manifestations of the physical effect of λ_3 on the nonlinear sea surface are higher crests and shallower wave troughs, as described quite explicitly by (5.17) and (5.18), respectively.

The exceedance probability distribution and probability density of ξ_{max} are given, respectively, by

$$P(w) = Pr\{\xi_{max} > w\} = \exp\left(-\frac{\xi^2}{2}\right), \tag{5.19}$$

$$p(w) = \frac{\xi}{1+\mu\xi}\exp\left(-\frac{\xi^2}{2}\right), \tag{5.20}$$

where $w = \xi + \mu\xi^2/2$. The corresponding expressions for the trough amplitude $|\xi_{min}|$ are

$$P(w) = Pr\{|\xi_{min}| > w\} = \exp\left(-\frac{\xi^2}{2}\right), \tag{5.21}$$

$$p(w) = \frac{\xi}{|1-\mu\xi|}\exp\left(-\frac{\xi^2}{2}\right), \tag{5.22}$$

where $w = \xi - \mu\xi^2/2$.

The nature of troughs amplitudes, associated statistics and their extensions to waves characterized with third-order nonlinear corrections are considered further in Tayfun & Fedele (2007*b*). Thus, we shall hereafter focus only on wave crests and refer to (5.17), (5.19) and (5.20) as the generalized Tayfun (T) model simply because they have the same form as the original Tayfun model and converge to it as $\mu \to \mu_m$ in

	m_0 (m^2)	ω_m (rad s^{-1})	ν	λ_3	μ	μ_m
Case A	1	1	0.100	0.274	0.091	0.102
Case B	1	1.065	0.428	0.232	0.077	0.116

TABLE 1. Parameters for the first-order spectra (6.1) and (6.2).

deep water. In the more general case, T is an exact theoretical representation of large wave crests in second-order random seas characterized by a well-defined steepness measure $\mu = \lambda_3/3$ in deep or shallow water.

6. Comparisons with simulations

Here, we consider two simple simulation cases, A and B, to illustrate and compare T and the original Tayfun (T_m) model. In both cases, we assume infinitely long-crested deep-water waves. Case A describes narrowband waves with the first-order spectrum

$$S(\omega) = \frac{m_0}{\nu\omega_m\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\omega-\omega_m}{\omega_m\nu}\right)^2\right], \qquad |\omega-\omega_m|/\omega_m \leqslant 6\nu, \tag{6.1}$$

where $m_0 = 1\,\mathrm{m}^2$ and $\omega_m = \omega_p = 1\,\mathrm{rad\,s}^{-1}$. In contrast, case B represents fairly wideband waves characterized by

$$S(\omega) = \frac{\alpha}{\omega_p} u^{-4}\exp(-u^{-4})W(u), \qquad 0.1 \leqslant u = \omega/\omega_p \leqslant 30, \tag{6.2}$$

where

$$W(u) = \begin{cases} 1 & 0 < u < 3.5, \\ (3.5/u)^4 & u \geqslant 3.5, \end{cases} \tag{6.3}$$

$\omega_p = 0.773\,\mathrm{rad\,s}^{-1}$, and α is such that $m_0 = 1.0\,\mathrm{m}^2$ for this case also.

The simulations in both cases follow from a two-step procedure whereby we first generate linear Gaussian displacements at a uniform sampling rate of 10 Hz from (2.2), using a Fast Fourier Transform with 2^{20} random phases (δ_n) and Rayleigh-distributed amplitudes (c_n). These are then modified with second-order corrections, using Tick's (Tick 1959) equivalent form of (2.4) valid for infinitely long-crested waves to generate the eventual nonlinear series efficiently. Each series so simulated is about 29 h long. This process is repeated four times in each case, thus providing us with an ensemble of four linear and nonlinear series for analysing wave crests and related statistics.

Table 1 summarizes the spectral parameters of relevance in the statistics of wave crests. These follow numerically from the linear spectra (6.1) and (6.2). The parameters of the nonlinear series in general differ from these as the second-order bound harmonics with frequencies $\omega_1 + \omega_2$ and $\omega_1 - \omega_2$ render the first-order spectrum slightly larger but noticeably wider over both low and high frequencies, mostly away from the spectrum peak. As a result, second-order spectral moments are generally larger than the first-order spectral moments, and so are some properties such as ω_m and ν. In comparison, λ_3 and so μ are smaller due to the bound harmonics associated with frequency differences. For example, the nonlinear waves simulated in case B are characterized with $m_0 \simeq 1.044\,\mathrm{m}^2$, $\omega_m \simeq 1.142\,\mathrm{rad\,s}^{-1}$, $\nu \simeq 0.604$, $\lambda_3 \simeq 0.201$, $\mu \simeq 0.067$

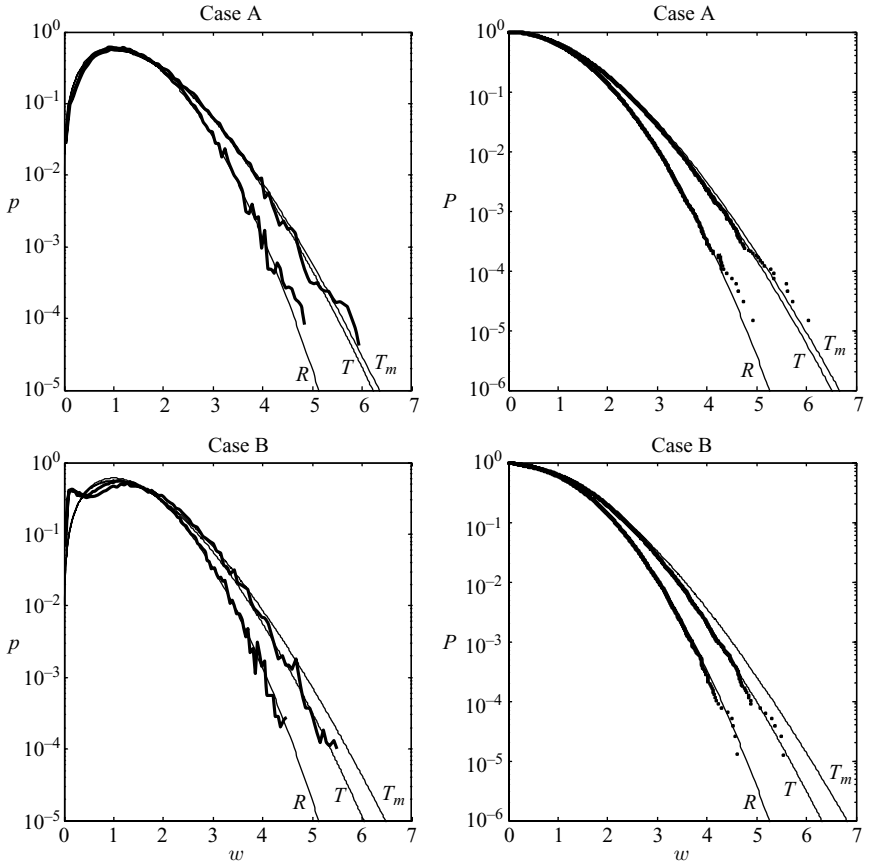

FIGURE 3. Probability densities (thick continuous) and exceedances (points) of simulated linear and nonlinear crests compared to theoretical models. Case A: narrowband spectrum, 67 040 linear and 67 041 nonlinear waves. T_m, Tayfun ($\mu_m = 0.102$) and T, generalized Tayfun ($\mu = 0.091$). Case B: wideband spectrum, 76 996 linear and 78 931 nonlinear waves. T_m ($\mu_m = 0.116$) and T ($\mu = 0.077$). In both cases, R, Rayleigh ($\mu = 0$) appropriate to linear crests.

and $\mu_m \simeq 0.136$. These differ noticeably in the manner just described from the values derived from the first-order spectrum (6.2) of case B in table 1.

Figure 3 shows comparisons of the probability densities and exceedances of linear and nonlinear wave crests simulated in both cases with the theoretical predictions from T, T_m and the Rayleigh (R) distribution appropriate to linear narrowband waves. We recall that T_m is the same as T of (5.19) and (5.20) for $\mu = \mu_m$. The comparisons in the upper part of figure 3 for the narrowband case A show that R describes the probability density and exceedances of linear crests quite well over all waves, if one allows for the relative scarcity of the simulated data and the higher variability of estimates towards the very high-wave tail. The exceedance probability estimate for the jth largest crest in a population of N independent samples is $P_j = j/(N+1)$. For $N \gg j$, the standard deviation of P_j is of the form $\sigma_j \simeq \sqrt{j}/(N+1)$ (cf. Tayfun & Fedele 2007*b*). So, the variability of exceedance estimates for the largest five or so samples is in general relatively large. In particular, σ_1 of the largest value is as large as the estimate P_1 itself, irrespective of how large N is.

As for the probability density and exceedances of the simulated nonlinear crests, they are seen to follow the trends predicted by T faithfully. T_m appears close to T but lies slightly above it. This is all entirely consistent with the expectation that T converges to T_m as an upper bound as $\nu \to 0$.

The lower part of figure 3 shows similar comparisons for case B. Not surprisingly, the probability densities of simulated crests display the bimodal structure typical of relatively wideband waves, both linear and nonlinear. In contrast with the narrowband case A, none of the theoretical predictions describes the distributions of either linear or nonlinear crests well over relatively low waves. However, both the linear and nonlinear scaled crests larger than about 1.5 are predicted quite well by R and T, respectively. T_m appears as a conservative upper bound and overestimates the simulated crests. It is worthwhile to mention that if the theoretical predictions are based on the parameters estimated from the simulated nonlinear waves, the accuracy of predictions from T and T_m deteriorates noticeably. Specifically, T tends to underestimate the simulated crests since $\mu \simeq 0.067$ from the nonlinear waves versus the larger value 0.077 of table 1. And, T_m overestimates the simulated crest rather significantly for $\mu_m \simeq 0.136$ from the nonlinear waves compared to 0.116 of table 1.

7. Comparisons with oceanic measurements

For oceanic comparisons, we consider 9 h of measurements gathered at 5.12 Hz with a Marex radar from the Tern platform in 167 m water depth in the northern North Sea during a severe storm in January, 1993. This data set, hereafter simply referred to as Tern, represents rather energetic storm seas with spectral-peak frequency $\omega_p = 0.433$ rad s^{-1}. The analysis of Tern as a whole gives $\sigma = 3.024$ m, $\omega_m = 0.569$ rad s^{-1}, $\nu = 0.629$, $\lambda_3 = 0.174$, $\mu = \lambda_3/3 = 0.058$ and $\mu_m = 0.1$ as overall averages. However, a similar analysis based on half-hourly series indicates all the preceding parameters vary in time. In particular, $2.723 \text{ m} \leqslant \sigma \leqslant 3.365$ m. So, the half-hourly σ estimates vary relative to the overall average $\sigma = 3.024$ m by as much as $\pm 10\,\%$ approximately. The larger σ values are typically associated with segments where larger waves occur and the smaller values with relatively smaller waves. As a result, we find that scaling elevations with the overall average σ tends to distort their distributions, in particular, causing the crests of the largest few waves to appear noticeably larger than they really are. So, to reduce such distortions, we analysed Tern in half-hourly segments scaled by the corresponding half-hourly σ estimates. When analysed as a whole, Tern has 3173 zero-up-crossing waves. Analysing it in 18 half-hourly segments reduces the wave count by about 16 waves to 3157. This does not lead to any discernable consequences in the statistics of wave crests discussed as follows.

We compare the probability density and exceedances of wave crests observed in figure 4 first with the theoretical predictions from $R(\mu = 0)$, $T(\mu = 0.058)$ and $T_m(\mu_m = 0.1)$. As expected, the observed probability density, shown in figure 4(*a*), displays the typical bimodal structure. None of the theoretical models describes the wave crests smaller than about 1.5 correctly, as in the simulation case B. T_m appears to describe larger wave crests well for the most part whereas T slightly underpredicts them. The differences between T, T_m and the observed data trends are displayed more clearly in the exceedance probability comparisons of figure 4(*b*) and even more so in figure 5. In the latter case, we plot the ratio ξ_{max}/ξ_R versus P, where $\xi_R = \sqrt{-2 \ln P}$ and represents the crest height that would be predicted by R at the same exceedance level as P of T or T_m. The relative stability of the exceedance estimates in figure 4(*b*) are indicated by vertical lines representing $\pm\sigma_j$ bands associated with the exceedance

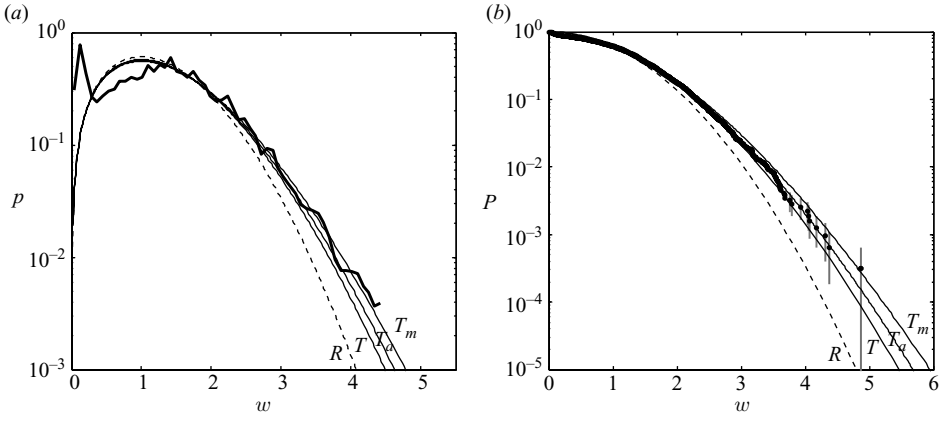

FIGURE 4. Tern: 3157 waves. (*a*) Probability density of crests (thick continuous) and (*b*) crest exceedances (points) compared to theoretical models. T_m, Tayfun ($\mu_m = 0.1$), T_a, generalized Tayfun ($\mu_a = 0.077$), T, generalized Tayfun ($\mu = 0.058$) and R, Rayleigh ($\mu = 0$). Vertical lines indicate $\pm\sigma$ bands for the exceedance probability estimates of 10 largest crests.

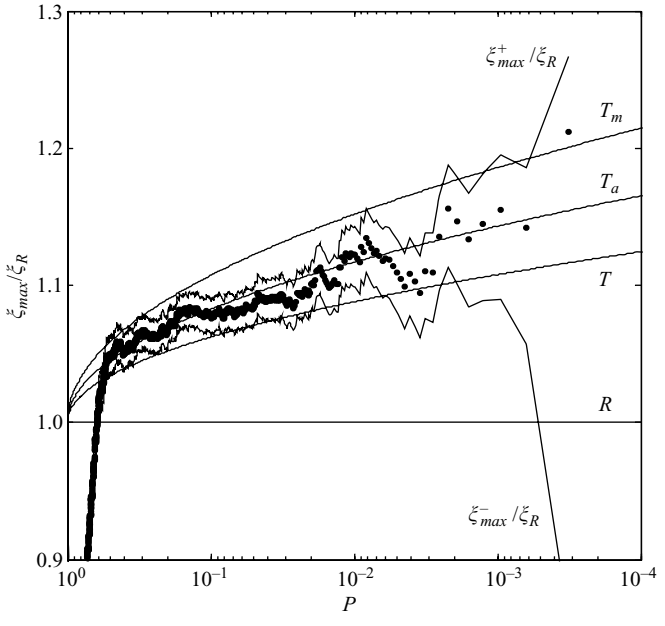

FIGURE 5. Tern: the observed values of the ratio ξ_{max}/ξ_R (points) compared to the theoretical ratios from T_m, T_a and T as for figure 4(*b*). $\xi^{\pm}_{max}/\xi_R$ indicate the upper and lower stability bands for the observed ratios derived from exceedance estimates of the form $P_j \pm \sigma_j$.

estimate P_j for the jth largest crest, but only for $1 \leqslant j \leqslant 10$ for clarity of presentation. The stability of the ratio estimates ξ_{max}/ξ_R in figure 5 is indicated by $\xi^{\pm}_{max}/\xi_R$, where $\xi^{\pm}_{max} = \sqrt{-2\ln(P_j \pm \sigma_j)}$. Both figures 4(*b*) and, particularly, 5 indicate that T tends to underestimate larger wave crests noticeably whereas T_m overestimates them.

In Tern, the prediction inaccuracies do occur unavoidably because the parameters μ and μ_m are derived from the nonlinear surface series. In theory, both parameters are

based on an underlying 'linear' surface structure, plausibly comprised of waves with a narrower spectrum and directional spread, for which we expect λ_3 and thus $\mu = \lambda_3/3$ to be larger than 0.058. For the linear surface, μ_m is smaller than 0.1 also, but T_m would still overestimate wave crests as an upper bound distribution. Consequently, the steepness parameter needs to be modified to render T more practical and accurate for typical oceanic applications. One can consider several possible alternatives for this purpose. A particularly promising one which uses parameters derived from Forristall's (Forristall 2000) directional simulations and leads to consistently accurate predictions for the statistics of large wave crests in deep and shallow water applications is elaborated elsewhere (Tayfun 2006), and thus not repeated here. A simpler alternative that works fairly well and consistently for deep-water waves follows from the general nature of the skewness coefficient in long-crested seas, as we described in § 3. We recall that if large waves behave as long-crested narrowband waves locally, then $\lambda_3^+ \simeq 3\mu_m(1+\nu^2)$ and $\lambda_3^- \simeq -3\mu_m\nu$. On this basis $\lambda_3 \simeq 3\mu_m(1-\nu+\nu^2)$. Thus, the modified steepness parameter sought is approximated by

$$\mu_a \simeq \mu_m(1-\nu+\nu^2). \tag{7.1}$$

Setting $\mu = \mu_a$ in T and T_m model reduces both to an identical form to which we shall hereafter simply refer as T_a. For Tern, (7.1) gives $\mu_a \simeq 0.077$, and the resulting T_a predictions compare noticeably better with Tern, as seen in figure 4(*b*) and more clearly in figure 5.

The expected maximum crest, say $\langle Z \rangle$, in a population of N independent second-order waves, each of which follows a parent distribution of the form T_a, is given by (Tayfun & Fedele 2007*b*)

$$\langle Z \rangle = \sqrt{2\ln N} + \frac{\gamma}{\sqrt{2\ln N}} + \mu_a(\gamma + \ln N), \tag{7.2}$$

where $\gamma = 0.5772\ldots$ is the Euler constant. For $N = 3,173$ and $\mu_a \simeq 0.077$ in Tern, the preceding expression yields $\langle Z \rangle \simeq 4.824$ as compared to the largest crest height 4.863 actually measured. The two differ less than 0.8 %.

In weakly nonlinear waves, third-order corrections to surface displacements, wave envelopes and phases are all of $O(\lambda_3^2)$ and thus negligible relative to second-order effects (Longuet-Higgins 1963, Tayfun and Lo 1990, Tayfun 1994). This has been so assumed in developing the present theoretical results, and the favourable nature of the Tern comparisons validates this assumption for the most part. However, whenever third-order quasi-resonant interactions between free waves are at least as significant as second-order corrections due to bound waves, surface statistics tend to deviate from the second-order predictions. Such interactions and associated modulational instabilities can be generated in wave flumes or numerically simulated under special conditions, generally requiring that waves be rather long-crested and narrowband (Socquet-Juglard *et al.* 2005; Onorato *et al.* 2006; Gramstad & Trulsen 2007). It is uncertain if these represent oceanic wind waves correctly. The numerical simulations of relatively wideband waves based on the Dysthe equation (Socquet-Juglard *et al.* 2005) and the fairly recent analyses of oceanic measurements gathered during a number of hurricanes in the Gulf of Mexico and North Sea storms (Forristall 2000, 2007; Tayfun & Fedele 2007*b*; Tayfun in press) suggest otherwise. Also, quasi-resonant modulational type distortions observed in wave statistics, numerically simulated or mechanically generated in wave flumes, follow a fairly predictable and systematic pattern in the form of a progressive excess of large waves, starting at wave height

and crest levels near their significant values. This is predicted reasonably well by third-order Gram-Charlier type expansions describing the statistics of long-crested narrowband waves (Tayfun and Lo 1990; Mori & Janssen 2006; Tayfun and Fedele 2007*b*). Neither the present Tern results nor any of the aforementioned comparisons of storm-generated extreme waves display any data trend even remotely similar to this systematic pattern. Thus, the theoretical and oceanic evidence so far available seems to suggest that the statistics representative of large oceanic wind waves are not affected in any discernible way by third-order nonlinearities, including NLS-type modulational instabilities.

8. Summary and conclusions

We have presented a new model for weakly nonlinear second-order random waves and their statistics based on the concept of stochastic wave groups. The new model provides a rigorous theoretical framework, enabling us to obtain closed-form expressions describing the expected shape of large waves and the statistical structure of associated crests. We emphasize that these expressions are valid in general for large directional waves characterized by second-order nonlinearities in deep and shallower water depths. There is no difficulty in developing similar expressions for various other kinematic and dynamic properties of large waves. The concept of stochastic wave groups also allows for further generalizations, in particular, to third-order nonlinear waves (cf. Fedele 2006, 2008).

The generalized Tayfun distribution represents an exact result for large second-order waves. It has a simple quadratic form dependent on a steepness parameter defined as $\mu = \lambda_3/3$ exactly. In long-crested narrowband waves in deep water where λ_3 tends to $3\mu_m$ as an upper bound, $\mu \to \mu_m$ and the generalized distribution converges to the original Tayfun model. In this case, either model describes the statistics of crests of all waves quite accurately. In the more general case, the statistical distribution of wave crests presents a more complex bimodal structure. Neither the generalized model nor any other theoretical model so far devised can predict this structure for relatively small wave crests. However, over the range of relatively large waves, the generalized model represents wave crests just as accurately whereas the original Tayfun model appears as a conservative upper bound and tends to overpredict them.

In theory, λ_3 and thus μ are defined in terms of the spectrum of the first-order linear surface. Oceanic measurements often provide fixed-point time series of nonlinear surface displacements. An estimate of λ_3 derived from such measurements is in general smaller than the theoretical value implied by the underlying first-order linear spectrum. So is the steepness parameter, which causes the generalized distribution to underpredict the wave crests observed, as we have shown in the Tern comparisons. This source of inaccuracy is practically remedied if the steepness parameter is approximated in the form $\mu_a = \mu_m(1 - \nu + \nu^2)$, where μ_m and ν represent estimates derived from the frequency spectrum observed. Our experience indicates that this approximation works fairly well for oceanic applications in deep water, but not always so well in shallower water depths. A more effective alternative for shallow-water applications is described in Tayfun (2006).

The authors are indebted to G. Z. Forristall for the Tern data utilized in this paper.

Appendix

To describe $\mathscr{R}_{\zeta_1}$ in terms of Ψ, we first express the wave profile at $\boldsymbol{X}=0$, say $\eta_c(T)$, in terms of an $O(h)$ contribution $\eta_{det}(T)$ and a random residual $r(T)=\mathscr{R}_{\zeta_1}(\mathbf{0}, T)$ of $O(h^0)$ as

$$\eta_c(T) = \eta_{det}(T) + r(T), \tag{A 1}$$

where

$$\eta_{det}(T) = \frac{h}{\sigma^2}\psi(T). \tag{A 2}$$

We can now determine the effects of r on η_c. As $h/\sigma \to \infty$, the surface profile around a large crest tends to assume the shape given by η_{det} (Lindgren 1972; Boccotti 2000). The latter is characterized by a crest of amplitude h at time $T=0$ and a local minimum of amplitude $\eta_{det}(T^*)$ at $T=T^*$, with T^* being the abscissa of the first local minimum of $\psi(T)$. However, the actual local minimum or simply the trough η_c following the crest of amplitude h at $T=0$ really occurs not at $T=T^*$ exactly but at $T=T^*+u$, where u represents a random variable. Further, η_c and $\dot{\eta}_c$ at $T=T^*$ attain values given, correct to $O(h^0)$, by

$$\begin{aligned} \eta_c(T^*) &= \eta_{det}(T^*) + \Delta + O(h^{-1}), \\ \dot{\eta}_c(T^*) &= -\ddot{\eta}_{det}(T^*)u + O(h^{-1}). \end{aligned} \tag{A 3}$$

Conversely, if the conditions in (A 3) hold, then a crest of amplitude h at $T=0$ is followed by a trough at $T=T^*+u$.

Drawing upon Rychlik (1987) and Lindgren & Rychlik (1991), a regression satisfying both conditions of (A 3) exactly is given by

$$\eta_c(T) = B_1\psi(T) + B_2\psi(T-T^*-u), \tag{A 4}$$

where

$$B_1 = \frac{\sigma^2 h - \psi(T^*+u)[h\psi(T^*)/\sigma^2+\Delta]}{\sigma^4 - \psi(T^*+u)^2}, \quad B_2 = \frac{\sigma^2[h\psi(T^*)/\sigma^2+\Delta] - \psi(T^*+u)h}{\sigma^4 - \psi(T^*+u)^2}.$$

To $O(h^0)$, u drops out, and η_c becomes

$$\eta_c(T) = \eta_{det}(T) + \frac{\Delta}{\sigma^2}\frac{\psi(T-T^*) - \psi(T^*)\,\psi(T)/\sigma^2}{1-\psi(T^*)^2/\sigma^4} + O(h^{-1}). \tag{A 5}$$

It is straightforward to extend the above time-domain formulation to the space–time domain obtaining new approximations for the stochastic wave group $\zeta_{1,c}$ of (4.1) and the associated potential $\phi_{1,c}$ of (4.5).

REFERENCES

Arena, F. 2005 On non-linear very large sea wave groups. *Ocean Engng* **32** (11–12), 1311–1331.

Arena, F. & Fedele, F. 2002 A family of narrow-band nonlinear stochastic processes for the mechanics of sea waves. *Eur. J. Mech.* B/*Fluids* **21**, 125–137.

Boccotti, P. 1989 On mechanics of irregular gravity waves. *Atti Acc. Naz. Lincei* Memorie VIII **19**, 111–170.

Boccotti, P. 2000 *Wave Mechanics for Ocean Engineering*. Elsevier Science.

Dysthe, K. B. 1979 Note on a modification to the nonlinear Schrödinger equation for application to deep water waves. *Proc. R. Soc. Lond.* A **369**, 105–114.

Fedele, F. 2006 Extreme events in nonlinear random seas. *ASME J. Offshore Mech. Arctic Engng* **128** (1), 11–16.

Fedele, F. 2008 Rogue waves in oceanic turbulence. *Phys.* D **237** (14–17), 2127–2131.

Fedele, F. & Arena, F. 2005 Weakly nonlinear statistics of high non-linear random waves. *Phys. Fluids* **17** (1), 026601.

Forristall, G. Z. 2000 Wave crest distributions: observations and second-order theory. *J. Phys. Oceanogr.* **30** (8), 1931–1943.

Forristall, G. Z. 2007 Comparing hindcasts with measurements from hurricanes Lili, Ivan, and Katrina. In *Proc. 10th Inter. Workshop Wave Hindcast. & Forecast*, Oahu, Hawaii, pp. 11–16.

Gramstad, O. & Trulsen, K. 2007 Influence of crest and group length on the occurrence of freak waves. *J. Fluid Mech.* **582** , 463–472.

Jensen, J. J. 1996 Second-order wave kinematics conditional on a given crest. *Appl. Ocean Res.* **18**, 119–128.

Jensen, J. J. 2005 Conditional second-order short-crested water waves applied to extreme wave episodes. *J. Fluid Mech.* **545**, 29–40.

Kriebel, D. L. & Dawson, T. H. 1993 Nonlinearity in crest height statistics. In *Proc. 2nd Intl. Conf. Wave Measurement & Analysis, ASCE*, New Orleans, pp. 61–75.

Lindgren, G. 1970 Some properties of a normal process near a local maximum. *Ann. Math. Statist.* **4** (6), 1870–1883.

Lindgren, G. 1972 Local maxima of Gaussian fields. *Arkiv för Matematik* **10**, 195–218.

Lindgren, G. & Rychlik, I. 1991. Slepian models and regression approximations in crossing and extreme value theory. *Intl Statist. Rev.* **59** (2), 195–225.

Longuet-Higgins, M. S. 1957 The statistical analysis of a random moving surface. *Philos. Trans. R. Soc. Lond.* A **966**, 321–387.

Longuet-Higgins, M. S. 1963 The effects of non-linearities on statistical distributions in the theory of sea waves. *J. Fluid Mech.* **17**, 459–480.

Marthinsen, T. & Winterstein, S. R. 1992 On the skewness of random surface waves. In *Proc. 2nd International Offshore and Polar Engineering Conference, ISOPE* San Francisco, vol. 3, pp. 472–478.

Milder, D. M. 2007 On the leading nonlinear correction to gravity-wave dynamics. *J. Fluid Mech.* **579**, 163–172.

Mori, N. & Janssen, P. A. E. M. 2006 On kurtosis and occurrence probability of freak waves. *J. Phys. Oceanogr.* **36** , 1471–1483.

Onorato, M., Osborne, A. R., Serio, L., Cavaleri, L., Brandini, C. & Stansberg, C. T. 2006 Extreme waves, modulational instability and second order theory: wave flume experiments on irregular waves. *Eur. J. Mech.* B/*Fluids* **25** (5), 586–601.

Phillips, O. M., Gu, D. & Donelan, M. 1993 On the expected structure of extreme waves in a Gaussian sea. I. Theory and SWADE buoy measurements. *J. Phys. Oceanogr.* **23**, 992–1000.

Prevosto, M. & Forristall, G. Z. 2002 Statistics of wave crests from models vs. measurements. In *Proc. 21st OMAE Conf.* Oslo, Paper No. OMAE2002-28443, ASME.

Prevosto, M., Krogstad, H. E. & Robin, A. 2000 Probability distributions for maximum wave and crest heights. *Coastal Engng* **40**, 329–360.

Rychlik, I. 1987 Joint distribution of successive zero crossing distances for stationary Gaussian processes. *J. Appl. Prob.* **24**, 378–385.

Sharma, J. N. & Dean, R. G. 1979 Development and evaluation of a procedure for simulating a random directional second order sea surface and associated wave forces. *Ocean Engng Rep.* No. 20, Civil Engineering Department, University of Delaware, Newark.

Socquet-Juglard, H., Dysthe, K., Trulsen, K., Krogstad, H. E. & Liu, J. 2005 Probability distributions of surface gravity waves during spectral changes. *J. Fluid Mech.* **542**, 195–216.

Srokozs, M. A. & Longuet-Higgins, M. S. 1986 On the skewness of sea-surface elevation. *J. Fluid Mech.* **164**, 487–497.

Tayfun, M. A. 1980 Narrow-band nonlinear sea waves. *J. Geophys. Res.* **85** (C3), 1548–1552.

Tayfun, M. A. 1986 On narrow-band representation of ocean waves. Part I. Theory. *J. Geophys. Res.* **91** (C6), 7743–7752.

Tayfun, M. A. 1994 Distributions of envelope and phase in weakly nonlinear random waves. *J. Engng Mech.* **120**, 1009–1025.

Tayfun, M. A. 2006 Statistics of nonlinear wave crests and groups. *Ocean Engng* **33** (11–12), 1589–1622.

Tayfun, M. A. 2008 Distributions of envelope and phase in wind waves. *J. Phys. Oceanogr.* **38**, 2754–2800.

Tayfun, M. A. & Lo, J.-M. 1990 Non-linear effects on wave envelope and phase. *J. Waterway, Port, Coastal, and Ocean Engng* **116**, 79–100.

Tayfun, M. A. & Fedele, F. 2007*a* Expected shape of extreme waves in storm seas. In *Proc. 26th OMAE Conf.,* San Diego, Paper No. OMAE2007-29073, ASME.

Tayfun, M. A. & Fedele, F. 2007*b* Wave-height distributions and nonlinear effects. *Ocean Engng* **34**, 1631–1649. *Proc. 25th OMAE Conf.,* Hamburg, Paper No. OMAE2007-29073, ASME.

Tayfun, M. A. & Fedele, F. 2008 Envelope and phase statistics of large waves. In *Proc. 27th OMAE Conf.,* Estoril, Paper No. OMAE2008-57219, ASME.

Tick, L. J. 1959 A nonlinear random wave model of gravity waves. *J. Math. Mech.* **8**, 643–652.

Tromans, P. S. & Vanderschuren, L. 2004 A spectral response surface method for calculating crest elevation statistics. *J. Offshore Mech. Arctic Engng* **126**, 51–53.

Vinje, T. & Haver, S. 1994 On the non-Gaussian structure of ocean waves. In *Proc. BOSS'94,* Boston, vol. 2, 453–479.

Winterstein, S. R. 1988 Nonlinear vibration models for extremes and fatigue. *J. Engng Mech.* **114**, 1772–1790.