JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 91, NO. C1, PAGES 995-1006, JANUARY 15, 1986

ON THE JOINT DISTRIBUTION OF SURFACE ELEVATION AND SLOPES FOR A NONLINEAR RANDOM SEA, WITH AN APPLICATION TO RADAR ALTIMETRY

M. A. Srokosz

Institute of Oceanographic Sciences, Wormley, Godalming, Surrey, England

Abstract. The work of Longuet-Higgins (1963), on the effects of wave nonlinearity on the statistics of the sea surface, is extended here to obtain the joint distribution of surface elevation and slopes for a nonlinear random sea. The results are used to examine the effect of wave nonlinearity on the form of a radar altimeter pulse which is reflected from the sea surface. It is shown how wave information may be derived from the return pulse and how this can be used to correct errors in the altimeter mean sea level estimate (a problem known as "sea state bias").

1. Introduction

Until fairly recently most measurements of waves on the sea surface, such as those obtained from a wave rider buoy, gave a time series of the surface elevation at a given point. Exceptions were the studies of Chase et al. [1957] and Holthuijsen [1983], who used stereophotography, and Cox and Munk [1956], who used sun glitter to obtain spatial information about the wave field (at a given instant of time). However, such spatial studies of waves are difficult to carry out (see above quoted papers) and have not been very common.

With the advent of satellite and airborne radar altimetry it has become possible to obtain spatial information about the sea surface on a large scale. In order to obtain such information from the radar return it is necessary to make some assumptions about the statistics of the sea surface elevation and slopes. In particular, knowledge of the form of the joint distribution of the surface elevation and surface slopes is necessary [Barrick, 1972; Lipa and Barrick, 1981]. Early studies [Barrick, 1972] of the form of the radar return from the sea surface assumed Gaussian statistics for the surface elevation and slopes. More recently attempts have been made to allow for the nonlinearity of the surface wave field by using a non-Gaussian theory to obtain the form of the radar return [Jackson, 1979; Lipa and Barrick, 1981]. The non-Gaussian theory used is based on the work of Longuet-Higgins [1963] who, using a weakly nonlinear dynamical model of the waves, obtained the distribution of the surface elevation and also the joint distribution of surface slopes (measured in two mutually orthogonal directions).

Jackson [1979] used this theory to obtain the joint distribution of the surface elevation and slope (in one direction). He related the parameters of the distribution to the wave spectrum under the assumption that the wave field was unidirectional. The theory was then used to examine the form of the radar return from the sea surface and the effect of nonlinearity on the return [Jackson, 1979; Lipa and Barrick, 1981]. Clearly the assumption of unidirectionality is somewhat restrictive in that generally sea waves have marked directional spread. Furthermore, for a complete description of the sea surface the joint distribution of the elevation and slopes in two mutually orthogonal directions is required (rather than the slope in one direction).

In this paper the work of Longuet-Higgins [1963] and Jackson [1979] will be extended to obtain the joint distribution of surface elevation and two slopes. The results obtained will be applied to the problem of determining the radar return from the sea surface and used to illustrate the effects of wave nonlinearity on the return. The "sea state bias" produced by nonlinear wave effects on the estimates of the mean level of the sea surface from the return will also be discussed.

2. Formulation

Consider the joint distribution $p(\zeta, \zeta_x, \zeta_y)$ of surface elevation ζ and slopes ζ_x, ζ_y. This may be expressed as the Fourier transform of the moment generating function and thence related to cumulants of the distribution [see Longuet-Higgins, 1963], so that

$$p(\zeta, \zeta_x, \zeta_y) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp \Big[-i(\zeta t + \zeta_x t' + \zeta_y t'') + \sum_{\substack{i,j,k \\ \neq (o,o,o)}} \frac{\kappa_{ijk}}{i!j!k!} (it)^i (it')^j (it'')^k \Big] \, dt \, dt' \, dt'' \tag{1}$$

The cumulants κ_{ijk} are related to the moments of the distribution μ_{ijk} by

$$\sum_{i,j,k} \frac{\kappa_{ijk}}{i!j!k!} (it)^i (it')^j (it'')^k = \log \Big[\sum_{i,j,k} \frac{\mu_{ijk}}{i!j!k!} (it)^i (it')^j (it'')^k \Big] \tag{2}$$

and may be calculated by equating the coefficients of combinations of powers of t, t', t''. The moments μ_{ijk} are given by

$$\mu_{ijk} = \langle \zeta^i \zeta_x{}^j \zeta_y{}^k \rangle \tag{3}$$

where the angle brackets enclose the ensemble average.

A generalization of the argument given by

Paper number 5C0582.
0148-0227/86/005C-0582$05.00

Longuet-Higgins [1963] allows the neglect of cumulants κ_{ijk} such that i+j+k>3 and so leads to the approximation

$$p(\zeta, \zeta_x, \zeta_y) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty}$$

$$\exp \Big[-i\,(\zeta t + \zeta_x t' + \zeta_y t'')$$

$$+ \sum_{\substack{i,j,k \neq (o,o,o) \\ i+j+k \leq 3}} \frac{\kappa_{ijk}}{i!j!k!} \; (it)^i \; (it')^j$$

$$(it'')^k \Big] \, dt\, dt'\, dt'' \qquad (4)$$

For the Gaussian case, $\kappa_{ijk} = 0$ for $i+j+k > 3$. It is from this approximation that the joint distribution will be obtained. In order to do this it is necessary to know the cumulants κ_{ijk} for $i+j+k \leq 3$ and these may be obtained from the moments μ_{ijk} for $i+j+k \leq 3$ via equation (2). In the next section we will calculate these moments.

3. Calculation of the moments

From Longuet-Higgins [1963] we note that the sea surface elevation ζ may be written, to second order, as

$$\zeta = \zeta_1 + \zeta_2 \qquad (5)$$

where

$$\zeta_1 = \sum_n a_n \cos \phi_n \qquad (6a)$$

$$\zeta_2 = \frac{1}{2} \sum_{i,j} a_i a_j$$

$$(C_{ij} \cos \phi_i \cos \phi_j + S_{ij} \sin \phi_i \sin \phi_j) \qquad (6b)$$

$$C_{ij} = (k_i k_j)^{-1/2}$$

$$\Big[B_{ij}^- - B_{ij}^+ - \underline{k_i} \cdot \underline{k_j} + (k_i + k_j)(k_i k_j)^{1/2} \Big] \qquad (7a)$$

$$S_{ij} = (k_i k_j)^{-1/2} \, [B_{ij}^- - B_{ij}^+ - k_i k_j] \qquad (7b)$$

$$B_{ij}^{\pm} = \frac{(k_i^{1/2} \pm k_j^{1/2})^2 \quad (\underline{k_i} \cdot \underline{k_j} \mp k_i k_j)}{(k_i^{1/2} \pm k_j^{1/2})^2 - \left| \underline{k_i} \pm \underline{k_j} \right|} \qquad (7c)$$

$$\phi_n = \underline{k_n} \cdot \underline{x} - \omega_n t + \theta_n \qquad (8a)$$

$$\omega_n^2 = g \; \underline{k_n} = g k_n \qquad (8b)$$

Here k_n and ω_n are the wave numbers and frequency of the nth wave component. The amplitudes and phases a_n, θ_n are assumed to be independent random variables, with a_n Rayleigh distributed on $[0, \infty]$ and θ_n uniformly distributed on $[0, 2\pi]$. Note therefore that the ϕ_n may be regarded as random variables distributed uniformly on $[0, 2\pi]$. If the second order term ζ_2 is neglected we obtain the Gaussian representation of the sea surface [Longuet-Higgins, 1957]. The above results are basically those given by Longuet-Higgins [1963] and used by Jackson [1979], but have a correction factor of one half in the expression for ζ_2 which was missing from their papers.

In order to calculate the moments we will substitute the above results into (3) and evaluate the resulting ensemble averages. To do this will require the use of results of the following type, which depend on the properties of the random variables a_n, ϕ_n:

$$\langle \cos \phi_n \rangle = 0$$

$$\langle \cos \phi_i \;\; \sin \phi_j \rangle = 0$$

$$\langle \cos \phi_i \;\; \cos \phi_j \rangle = \frac{1}{2} \delta_{ij}$$

$$\langle a_n^4 \rangle = 2 \langle a_n^2 \rangle^2$$

etc. ...

Note that all odd order correlations of this type vanish. Furthermore if α and β are independent random variables and $f(\alpha)$, $g(\beta)$ are functions of α and β then

$$\langle f(\alpha)\, g(\beta) \rangle = \langle f(\alpha) \rangle \; \langle g(\beta) \rangle$$

Finally note that μ_{mnp} may be obtained from μ_{mpn} by interchanging the x and y coordinates.

It is straightforward, though tedious, to show that

$$\mu_{000} = 1 \qquad (9a)$$

$$\mu_{100} = \mu_{010} = \mu_{001} = \mu_{101} = \mu_{110} = 0 \qquad (9b)$$

$$\mu_{200} = \frac{1}{2} \sum_n \;\; \langle a_n^2 \rangle \qquad (9c)$$

$$\mu_{020} = \frac{1}{2} \sum_n \; k_{nx}^2 \quad \langle a_n^2 \rangle \qquad (9d)$$

$$\mu_{002} = \frac{1}{2} \sum_n \; k_{ny}^2 \quad \langle a_n^2 \rangle \qquad (9e)$$

$$\mu_{011} = \frac{1}{2} \sum_n \; k_{nx} \, k_{ny} \; \langle a_n^2 \rangle \qquad (9f)$$

$$\mu_{030} = \mu_{003} = \mu_{210} = \mu_{201} = \mu_{012} = \mu_{021} = 0 \qquad (9g)$$

$$\mu_{300} = \frac{3}{4} \sum_{i,j} \; \langle a_i^2 \rangle \langle a_j^2 \rangle \; C_{ij} \qquad (9h)$$

$$\mu_{120} = \frac{1}{4} \sum_{i,j} \; \langle a_i^2 \rangle \langle a_j^2 \rangle$$

$$[(k_{xi}^2 + k_{xj}^2) C_{ij} - k_{xi} k_{xj} S_{ij}] \qquad (9i)$$

$$\mu_{102} = \frac{1}{4} \sum_{i,j} \; \langle a_i^2 \rangle \langle a_j^2 \rangle$$

$$[(k_{yi}^2 + k_{yk}^2) C_{ij} - k_{yi} k_{yj} S_{ij}] \qquad (9j)$$

$$\mu_{111} = \frac{1}{4} \sum_{i,j} \; \langle a_i^2 \rangle \langle a_j^2 \rangle$$

$$[(k_{xi} \; k_{yi} + k_{xj} \; k_{yj}) \; C_{ij} - k_{xi} \; k_{yj} \; S_{ij}] \qquad (9k)$$

The details of the calculations are omitted but a sample calculation for μ_{120} is given in the appendix. The results are correct to leading order; for example μ_{030} and μ_{003} are not strictly zero but are of higher order in wave slope than μ_{300}, μ_{120}, μ_{102}, μ_{111} and are therefore neglected.

From the above and (3) it can be shown that

$$\kappa_{mnp} = \mu_{mnp} \qquad m+n+p \leq 3 \tag{10}$$

so that the only nonzero cumulants are κ_{200}, κ_{020}, κ_{002}, κ_{011}, κ_{300}, κ_{120}, κ_{102}, κ_{111}. Thus the number of parameters the distribution depends on has been reduced from the original 19 cumulants $\kappa_{ijk}(0 \leq i+j+k \leq 3)$ to 8. It is useful to define normalized cumulants λ_{mnp} as

$$\lambda_{mnp} = \frac{\kappa_{mnp}}{\kappa_{200}{}^{m/2}\,\kappa_{020}{}^{n/2}\,\kappa_{002}{}^{p/2}} \tag{11}$$

With these results we are able to simplify (4) to obtain $p(\zeta, \zeta_x, \zeta_y)$.

4. The Joint Distribution of Surface Elevation and Slopes

By using the results of the previous section, expanding the integrand in (4) and neglecting higher order terms [following Longuet-Higgins, 1963] we obtain

$$p(\zeta, \zeta_x, \zeta_y) = \frac{1}{(2\pi)^3}\,\frac{1}{\sqrt{\kappa_{200}\,\kappa_{020}\,\kappa_{002}}}$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp[-i(\eta u + \eta_x u' + \eta_y u'') - \tfrac{1}{2}(u^2 + u'^2 + u''^2 + 2\lambda_{011}u'u'')]$$

$$\times \left[1 + \frac{i^3}{6}(\lambda_{300}u^3 + \lambda_{120}uu'^2 + \lambda_{102}uu''^2 + \lambda_{111}uu'u'')\right] du\,du'\,du''$$

where

$$\eta = \frac{\zeta}{\kappa_{200}{}^{1/2}} \qquad \eta_x = \frac{\zeta_x}{\kappa_{020}{}^{1/2}} \qquad \eta_y = \frac{\zeta_y}{\kappa_{002}{}^{1/2}} \tag{12}$$

and the following change of variables has been made:

$$u = \frac{t}{\kappa_{200}{}^{1/2}} \qquad u' = \frac{t'}{\kappa_{020}{}^{1/2}} \qquad u'' = \frac{t''}{\kappa_{002}{}^{1/2}}$$

The integral may be evaluated to obtain

$$p(\zeta, \zeta_x, \zeta_y) = \frac{1}{(2\pi)^{3/2}}\,\frac{1}{\mu_{200}{}^{1/2}\sqrt{\mu_{020}\mu_{002}-\mu_{011}{}^2}}$$

$$\exp\left[-\tfrac{1}{2}\eta^2 - \tfrac{1}{2}(\eta_x{}^2 - 2\lambda_{011}\eta_x\eta_y + \eta_y{}^2)/(1-\lambda_{011}{}^2)\right] \left[1 + \tfrac{1}{6}(\lambda_{300}H_{300} + 3(\lambda_{120}H_{120} + \lambda_{102}H_{102} + 2\lambda_{111}H_{111}))\right] \tag{13}$$

where the H_{mnp} are generalized Hermite polynomials given by

$$H_{mnp}(\eta, \eta_x, \eta_y; \rho) = (-1)^{m+n+p}$$

$$\exp\{\tfrac{1}{2}\eta^2 + \tfrac{1}{2}(\eta_x{}^2 - 2\rho\eta_x\eta_y + \eta_y{}^2)/(1-\rho^2)\}$$

$$\frac{\partial^m}{\partial\eta^m}\frac{\partial^n}{\partial\eta_x{}^n}\frac{\partial^p}{\partial\eta_y{}^p}\left[\exp\{-\tfrac{1}{2}\eta^2 - \tfrac{1}{2}(\eta_x{}^2 - 2\rho\,\eta_x\eta_y + \eta_y{}^2)/(1-\rho^2)\}\right]$$

with $\rho = \lambda_{011}$. This result may be obtained by a straightforward generalization of the approach given by Longuet-Higgins [1963]. For the purposes of this paper we only require

$$H_{300} = \eta^3 - 3\eta \tag{14a}$$

$$H_{120} = \eta(1-\rho^2)^{-1}[(\eta_x - \rho\eta_y)^2(1-\rho^2)^{-1} - 1] \tag{14b}$$

$$H_{102} = \eta(1-\rho^2)^{-1}[(\eta_y - \rho\eta_x)^2(1-\rho^2)^{-1} - 1] \tag{14c}$$

$$H_{111} = \eta(1-\rho^2)^{-1}[(\eta_x - \rho\eta_y)(\eta_y - \rho\eta_x)(1-\rho^2)^{-1} + \rho] \tag{14d}$$

Thus (13), together with (14) and the results of section 3, gives the joint distribution of surface elevation and slopes for a weakly nonlinear random wave field.

From this distribution several results can be deduced. Firstly, by integrating out the surface elevation ζ it is possible to obtain the joint distribution of the slopes, which proves to be Gaussian. This is consistent with the results of Longuet-Higgins [1963] who showed that, to this order, the cumulants for the slope distribution vanish and hence the distribution of slopes is Gaussian. More generally if the coefficients of skewness λ_{300}, λ_{120}, λ_{102}, λ_{111} vanish we obtain the Gaussian result for the joint distribution of surface elevation and slopes [Longuet-Higgins, 1957].

Secondly, by integrating out a slope variable ζ_y we obtain the joint distribution of surface elevation and slope, as found previously by Jackson (1979),

$$p(\zeta, \zeta_x) = \frac{1}{2\pi\sqrt{\mu_{200}\,\mu_{020}}} \exp\{-\tfrac{1}{2}(\eta^2 + \eta_x{}^2)\} \left[1 + \tfrac{1}{6}\lambda_{300}H_{300} + \tfrac{1}{2}\lambda_{120}\eta(\eta_x{}^2 - 1)\right] \tag{15}$$

Now by integrating out the second slope variable ζ_x we obtain Longuet-Higgins' [1963] original result for the distribution of the surface elevation

$$p(\zeta) = \frac{1}{\sqrt{2\pi\,\mu_{200}}}\,e^{-\eta^2/2}\left[1 + \tfrac{1}{6}\lambda_{300}H_{300}\right] \tag{16}$$

Returning to the joint distribution of surface

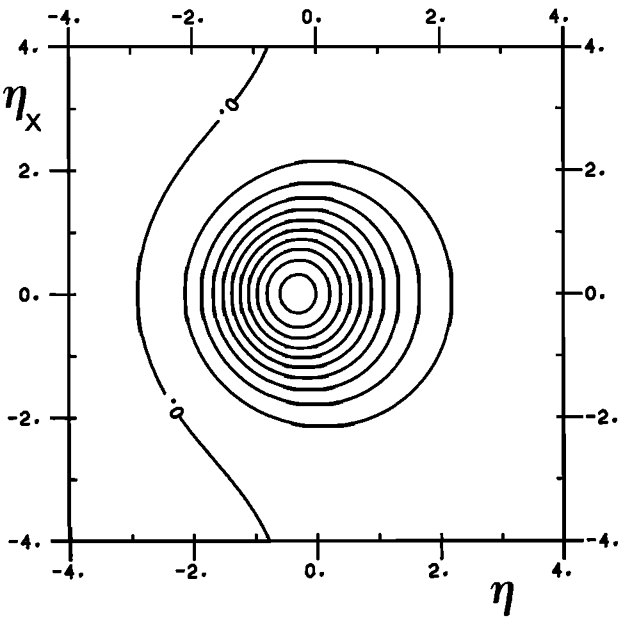


Fig. 1. Contour plot of the distribution $p(\eta, \eta_x)$ for $\lambda_{300} = 0.5$, $\lambda_{120} = 0.2$. Contour interval 0.01592. The region to the left of the zero contour is negative.

elevation and slopes we note that it depends on the following eight parameters:

μ_{200}	variance of the sea surface elevation;
μ_{020}, μ_{002}	variances of the slopes;
λ_{011}	correlation coefficient between the two slopes;
λ_{300}	skewness of the sea surface elevation;
$\lambda_{120}, \lambda_{102}, \lambda_{111}$	other skewness coefficients.

(The physical significance of the coefficients λ_{120}, λ_{102}, λ_{111} is not clear). All these quantities may be related to the wave number spectrum $E(\underline{k})$ by noting that

$$\sum_{\underline{dk}} \langle a_n^2/2 \rangle = E(\underline{k})\underline{dk}$$

and using the results of section 3, but this aspect of the problem will not be pursued here (see Jackson [1979] for details of how this can be done in the unidirectional case).

Before applying these results to the problem of obtaining wave information from an altimeter return we will attempt to illustrate them graphically and also discuss their relationship with the work of Huang et al. [1984]. As $p(\zeta, \zeta_x, \zeta_y)$ is a function of three variables, it cannot be plotted in a simple manner; instead we will contour the marginal distribution of elevation and one slope given in equation (15). Two examples of the (normalized) distribution are shown in Figures 1 and 2, for $\lambda_{300} = 0.5$, $\lambda_{120} = 0.2$ and $\lambda_{300} = 0.2$, $\lambda_{120} = 0.5$ respectively. For a Gaussian distribution the contours would be circles centered on the origin. Here it can be seen that large values of λ_{120} lead to a greater departure from the Gaussian distribution than do large values of λ_{300}. Furthermore the plots show up a problem associated with the use of Gram-Charlier series to represent probability density functions, in that there is a region where the distribution is negative (strictly it should be positive everywhere). However, the bulk of the distribution is positive so the results are still useful.

Huang et al. [1984] have recently derived a joint distribution $p(\zeta, \zeta_x)$ for the case of unidirectional, narrow band, nonlinear waves. Their distribution is positive everywhere and their results are similar to the ones shown in Figures 1 and 2. Unlike the distribution given in (15) for elevation and slope, which has two parameters λ_{300}, λ_{120} to describe nonlinear effects, the nonlinear effects enter their distribution through only one parameter, the significant slope. Unfortunately, their method of deriving the joint distribution of elevation and one slope cannot be extended to obtain the joint distribution of elevation and two slopes as studied in this paper, their approach being based on a unidirectional, narrow band representation of the wave field. In practice the waves on the ocean surface are likely to be both directionally distributed and broadband and allowance needs to be made for these effects.

5. Application to Radar Altimetry

Rather than give a detailed discussion of how the above results might be used in obtaining information on sea waves from an altimeter radar return we will concentrate on a particular problem. This is done for illustrative purposes, but it also has practical applications, particularly in terms of determining sea state bias in mean level measurements due to nonlinear wave effects (see below, section 6).

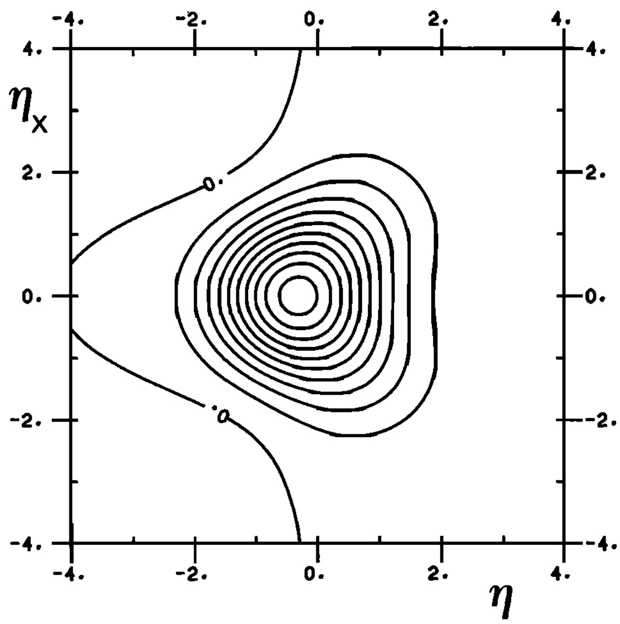


Fig. 2. Contour plot of the distribution $p(\eta, \eta_x)$ for $\lambda_{300} = 0.2$, $\lambda_{120} = 0.5$. Contour interval 0.01592. The region to the left of the zero contour is negative.

Lipa and Barrick [1981] give the following formula for the leading edge of the radar return for a pulse limited radar:

$$\sigma(t) = K \int_{-\infty}^{\infty} \left[\int_0^{\infty} P\left(\xi + \frac{ct}{2} - u\right) du\right] p_s(\xi)\, d\xi \tag{17}$$

where

$$K = 2\pi^2 H \left|R(0)\right|^2 / (1 + H/a) \tag{18}$$

Here P( ) is the spatial compressed altimeter power pulse and $p_s(\xi)$ the joint distribution of surface elevation and slopes at elevation ξ and zero slope. H is the altimeter altitude, a the earth's radius and R(0) the radar Fresnel reflection coefficient at normal incidence for the sea surface. This formula is applicable to the radar return for the altimeters carried by the satellites GEOS 3 and Seasat and for that to be carried by the European satellite ERS 1.

From (13), by setting $\zeta_x=\zeta_y=0$, we obtain

$$p_s(\zeta) = \frac{1}{(2\pi)^{3/2}} \frac{1}{\mu_{200}{}^{1/2}\sqrt{\mu_{020}\,\mu_{002} - \mu_{011}{}^2}}\, e^{-\eta^2/2}$$

$$\left[1 + \frac{1}{6}\left(\lambda_{300}(\eta^3 - 3\eta) - 3\eta\,\gamma\right)\right] \tag{19}$$

where η is given by (12) and

$$\gamma = (\lambda_{120} + \lambda_{102} - 2\lambda_{011}\,\lambda_{111})/(1-\lambda_{011}{}^2) \tag{20}$$

The above represents the distribution of the elevation of points on the sea surface with zero slope. Strictly speaking, to make it a probability density function it needs to be normalized by dividing by

$$\int_{-\infty}^{\infty} p_s(\zeta)\, d\zeta = (2\pi)^{-1}\sqrt{(\mu_{020}\,\mu_{002} - \mu_{011}{}^2)} \tag{21}$$

in order to make the area under the curve equal to 1. For future reference we will denote the normalised form by

$$q(\zeta) = p_s(\zeta)/\int_{-\infty}^{\infty} p_s(\zeta) d\zeta \tag{22}$$

Comparison of the result (19) with that used by Lipa and Barrick [1981, equation (20)] shows that both have the same functional form, despite the inclusion of the extra slope information in our formulation. The difference lies in the parameter γ, which in their case is simply λ_{120} (λ_2 in their notation). The physical significance of the parameter γ is unclear although its form suggests that it might give some indication of the long/short-crestedness of the wave field (this is a purely speculative suggestion and needs to be investigated further).

Lipa and Barrick [1981] analyze the radar return from the Seasat altimeter using Jackson's [1979] result for $p_s(\zeta)$ which is a special case of the one derived here. However, they are forced to carry out a numerical evaluation of the integral (17) because the function P( ) representing the compressed altimeter power pulse cannot be described by a simple analytical form for the Seasat altimeter [see Lipa and Barrick, 1981, Figure 8]. For the planned ERS 1 altimeter, P( ) is given by

$$P(x) = P_o \exp\left\{-x^2/\nu^2\right\} \tag{23}$$

which is a Gaussian pulse shape. Here ν is related to the half-power width of the signal pulse. Using this result, together with (19), it is possible to evaluate the integral (17) to obtain

$$\sigma(t) = \frac{1}{2} K P_o \pi\nu \left(\mu_{020}\,\mu_{002} - \mu_{011}{}^2\right)^{-1/2}$$

$$\left[1 + \mathrm{erf}(t/t_p)\right.$$

$$\left. + \frac{e^{-(t/t_p)^2}}{\sqrt{\pi}} \left\{A\left(\frac{t}{t_p}\right)^2 + A - B\right\}\right] \tag{24}$$

where

$$t_p = 2\sqrt{\nu^2 + 2\mu_{200}}\,/c \tag{25a}$$

$$A = \frac{4}{3}\,\lambda_{300}\left[2 + \frac{\nu^2}{\mu_{200}}\right]^{-3/2} \tag{25b}$$

$$B = (\lambda_{300} + \gamma)\left[2 + \frac{\nu^2}{\mu_{200}}\right]^{-1/2}$$

$$-\lambda_{300}\left(\frac{\nu^2}{\mu_{200}}\right)\left[2 + \frac{\nu^2}{\mu_{200}}\right]^{-3/2} \tag{25c}$$

(The details of the integration are given in the appendix.) When $\lambda_{300}=\gamma=0$ (that is, nonlinear wave effects are neglected) the above reduces to the Gaussian result given by Barrick [1972]. It can be seen that the following information about the sea surface may be estimated from the radar return $\sigma(t)$:

The first parameter is μ_{200}, the variance of the sea surface elevation and hence the significant wave height $H_s = 4\sqrt{\mu_{200}}$.

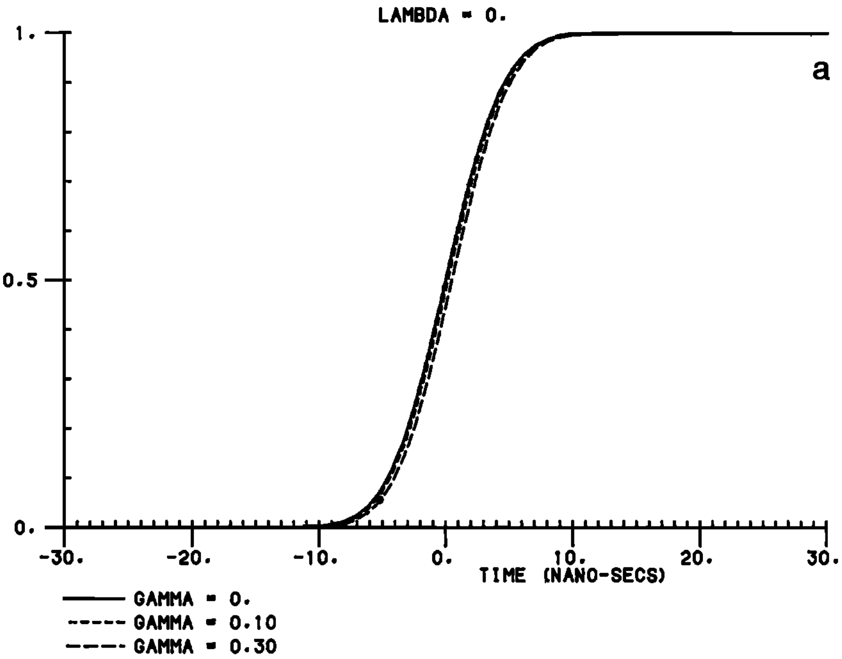
The second parameter is $\sqrt{(\mu_{020}\,\mu_{002} - \mu_{011}{}^2)}$, a measure of the mean square surface slope [Longuet-Higgins, 1957]. In fact, if principal axes are chosen, μ_{011} vanishes and this reduces to $\sqrt{\mu_{020}\,\mu_{002}}$, the geometric mean of the maximum and minimum slope variances. For an isotropic surface this gives the actual mean square slope. This may be used to estimate the wind speed [Brown, 1979] or wave period [Challenor & Srokosz, 1984].

The third parameter is λ_{300}, the skewness of the sea surface elevation distribution. Physically this is related to the peakier crests and flatter troughs of nonlinear waves and is a measure of the nonlinearity.

The fourth parameter is γ (given by (20)), a skewness coefficient with unknown physical significance. This parameter is important for estimating sea state bias (see section 6, below).

In practice these parameters are estimated from the radar return by fitting the theoretical form to the actual return (using least squares, maximum likelihood or some other technique; see for example Lipa and Barrick [1981]).

In order to illustrate the effects of wave nonlinearity on the radar return a normalized form of the result given in (24) is plotted, for



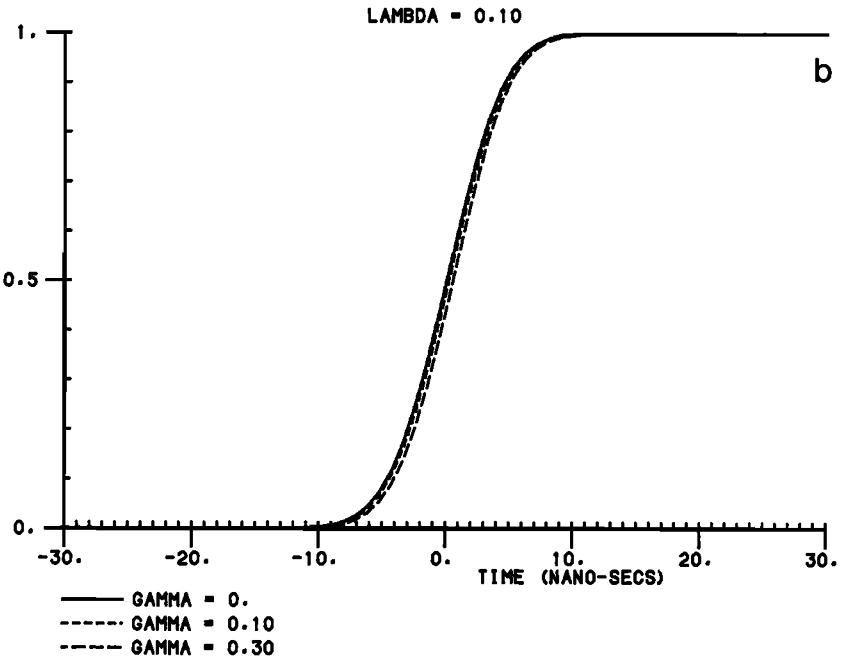


Fig. 3. Return pulse $\sigma(t)$ for $H_s = 2$ m and (a) $\lambda_{300} = 0$, (b) $\lambda_{300} = 0.1$, (c) $\lambda_{300} = 0.3$; with (solid curve) $\gamma = 0$, (short-dashed curve) $\gamma = 0.1$, and (long-dashed curve) $\gamma = 0.3$.

various values of the significant wave height H_s and skewness coefficients λ_{300} and γ, in Figures 3 to 5. Most radar altimeters normalize the return pulse via an automatic gain control; here the normalization used is to make the maximum value of the return equal to 1. The value of the other parameter in the equation has been chosen to correspond to that of the ERS 1 altimeter and is $\nu=0.425$ $cT/\sqrt{2}$ where the altimeter pulse width T is 3 ns. The unit of time used is the nanosecond. The speed of light c is taken as 3×10^8m s^{-1}.

Figures 3 and 4 show the form of the radar return for values of H_s equal to 2 and 8 m respectively and for a range of values of λ_{300} and γ. They show the characteristic stretching of the return pulse with increasing sea state (that is, increasing H_s). Varying λ_{300} and γ produces small changes in the return pulse shape, but these increase with increasing H_s. Figure 5, for H_s=8 m, shows the effect of holding γ fixed and varying λ_{300} and may be compared with Figure 4 where λ_{300} is held fixed and γ varied. Again the changes in pulse shape are small.

The range of values for λ_{300} and γ used in Figures 3 to 5 were chosen to lie in the range of

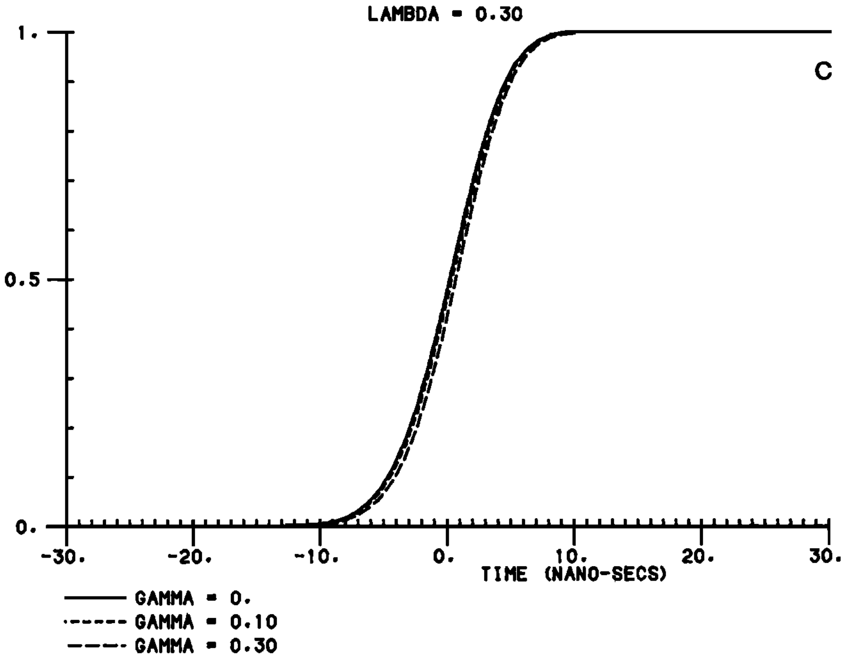


Fig. 3. (continued)

values suggested by Lipa and Barrick's [1981] results from Seasat data. However, independent measurements of the sea surface would appear to be necessary to confirm the accuracy of altimeter-estimated values of λ_{300} and γ, and also to confirm the validity of the non-Gaussian model used here.

6. Sea State Bias

Sea state bias is the error induced in the estimate of the mean sea level from the altimeter return due to nonlinear wave effects. This problem has been discussed by previous authors (see, for example, Hayne and Hancock [1982]; Douglas and Agreen [1983]; and Lipa and Barrick [1981]) but some confusion still appears to exist as to how nonlinear wave effects cause an error in the mean level estimate. As accurate estimates for the mean sea level are important, if the information is to be used to estimate geostrophic surface currents, it seems worthwhile to examine this problem in detail.

The first question to be asked is what do radar altimeters "see" in terms of surface waves? As the radar works by specular reflection [Barrick, 1972] it senses the elevations of points with zero slope (at least, this is a good approximation for the radar return leading edge response considered in the previous section [Barrick, 1972; Lipa and Barrick, 1981]). It does not sense the actual surface elevation directly. In the Gaussian case the distributions of the surface elevation and of points with zero slope are identical statistically, but this is not so in the non-Gaussian case. Some confusion appears to have arisen in the literature because the identity of the two distributions in the Gaussian case has been assumed to hold for the non-Gaussian case as well. A comparison of the surface elevation distribution $p(\zeta)$ (equation (16)) and the distribution of points of zero slope (equation (22)) shows that they are equivalent only if $\gamma=0$ (in the Gaussian case $\lambda_{300}=0$ also), which in general will not be the case. Because of this difference there is a difference between the mean level sensed by the altimeter, given by

$$\int_{-\infty}^{\infty} \zeta q(\zeta) d\zeta = -\frac{\gamma}{2} \mu_{200}^{1/2} = -\frac{\gamma}{8} H_s \qquad (26)$$

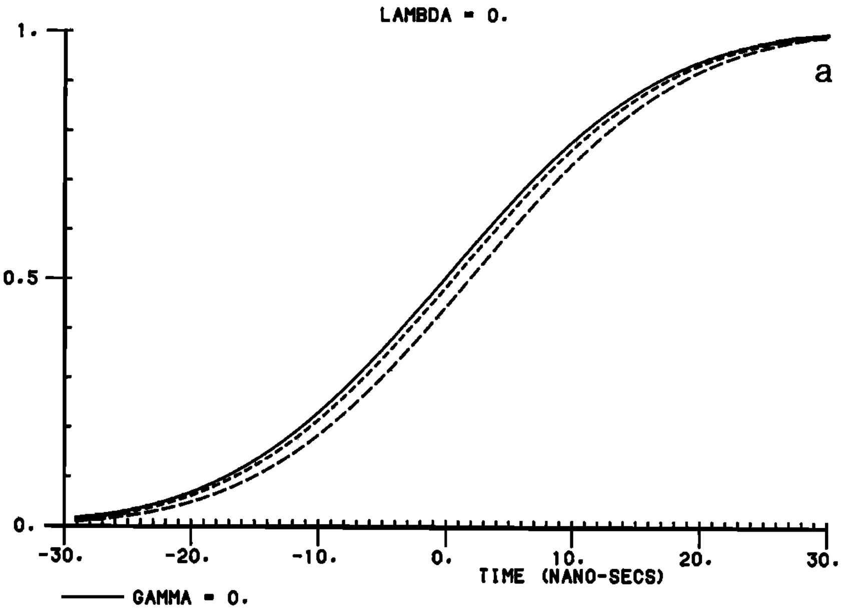
and the actual mean level, given by

$$\int_{-\infty}^{\infty} \zeta \; p(\zeta) \; d\zeta = 0$$

From (26) it can be seen that the mean level estimate is biased downward (if γ positive) and that this bias increases with increasing H_s.

The above effect has been termed "electromagnetic bias" and has been explained as being due to the focusing and scattering of the radar by the troughs and crests of the waves, respectively [Douglas and Agreen, 1983]. As pointed out by Brown and Miller [1977] this is not a correct explanation of this effect, which depends on a redistribution of specular points due to the wave nonlinearity as compared to their distribution in the Gaussian case. Hayne and Hancock [1982] note the difference between what the radar "sees" and the actual sea surface elevation distribution but appear to use the latter distribution in their analysis.

A second source of error in the estimation of the mean level is the determination of the time taken for the radar pulse to travel from the altimeter to the sea surface and back again. For the choice of time origin used here, t=0 corresponds to the position of the mean level of the points of zero slope on the surface. How is this point to be estimated from the radar return? Consider a normalized version of the



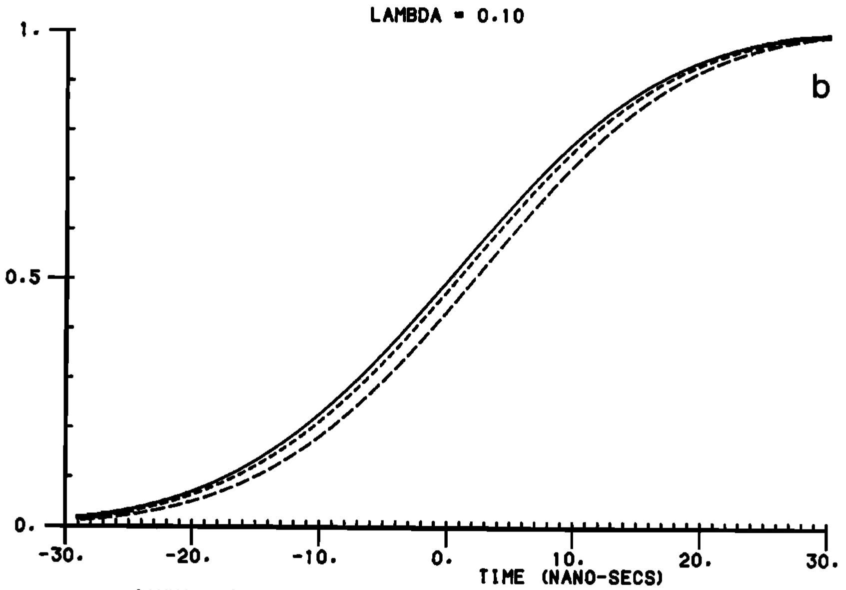


Fig. 4. Return pulse $\sigma(t)$ for H_s = 8 m and (a) λ_{300} = 0, (b) λ_{300} = 0.1, (c) λ_{300} = 0.3; with (solid curve) γ = 0, (short-dashed curve) γ = 0.1, and (long-dashed curve) γ = 0.3.

result given in (24) for the radar return, assume that $\nu^2/2\mu_{200} \ll 1$ and set $\tau=(t/t_p)$, so that

$$\sigma(\tau) = \frac{1}{2}\ \left[1+ \operatorname{erf}(\tau) + \frac{e^{-\tau^2}}{\sqrt{\pi}}\ \left\{\frac{\sqrt{2}\ \lambda_{300}}{3}\ (\tau^2+1) - (\lambda_{300}+\gamma)\ /\ \sqrt{2}\ \right\}\right] \tag{27}$$

approximately. For a Gaussian sea surface, $\lambda_{300}=\gamma=0$, and $\tau=0$ corresponds to the half-power point of the return $\sigma=1/2$. Thus by finding the half-power point of the return the mean sea level may be estimated. (This is the basis for many "on-board" altimeter algorithms.) However, for a nonlinear wave field, $\tau=0$ corresponds to

$$\sigma = \frac{1}{2}\ \left[1 - \sqrt{\frac{2}{\pi}}\ \left(\frac{\lambda_{300}}{6} + \frac{\gamma}{2}\right)\right]$$

and using $\sigma=1/2$ as an estimate of the position of $\tau=0$ will lead to an error. The magnitude of this effect can be judged from the displacement of the curves in Figures 3 to 4 for nonzero λ_{300} and γ from that for the case $\lambda_{300}=\gamma=0$. It can be seen that the displacement in time is of the order of one nanosecond, which would correspond to a 15-cm error in the mean level estimate. As

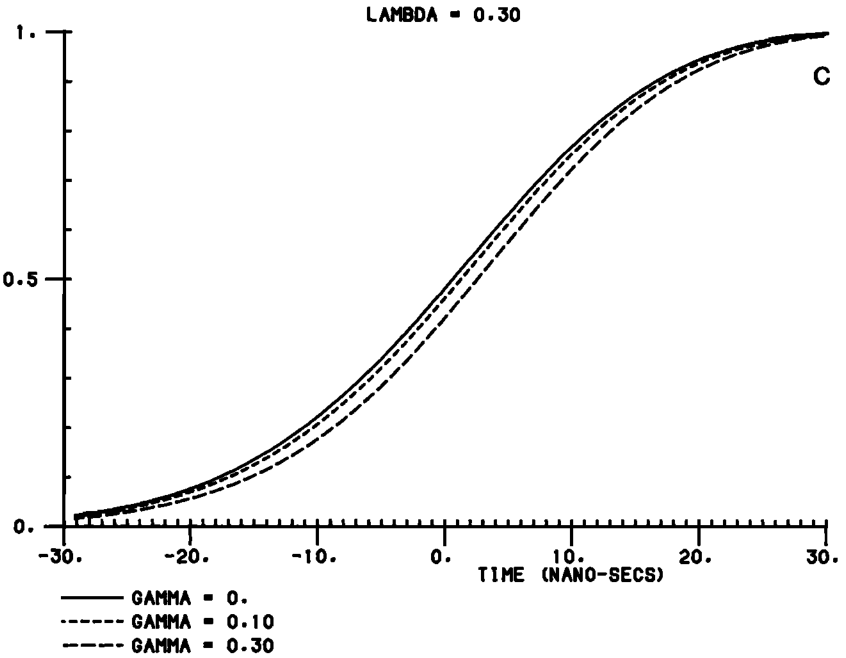


Fig. 4. (continued)

the accuracy required for the calculation of geostrophic surface currents is a few centimeters, this is a significant error, as is the other error discussed above.

The correct method of dealing with both errors is to fit the theoretical form of the radar return (24) to the altimeter data and to estimate the parameters from it (including the time origin t_o). The wave parameters can be used to correct for the sea state bias in the mean level given by (26). The inclusion of the time origin t_o (done by replacing t in (24) by $(t-t_o)$) which is estimated from the return overcomes the problem of using the half-power point to determine the mean level and the associated error in the mean level estimate. It should be noted, from the results shown in Figures 3 and 4, that it might in practice be difficult to distinguish between changes in t_o and γ, variations in γ leading primarily to a shift in the position of the radar return in time. Discussion of the practical implementation of this approach for the analysis of altimeter data is deferred to a future paper.

Before concluding this section we note that as well as the two errors in the mean level estimate discussed above there is the possibility of a third error, which is not due to nonlinear wave effects. This is simply due to the incorrect estimation of the half-power point of the return (see Lipa and Barrick [1981] for a discussion in the case of Seasat). We will not consider this further here, except to note that its presence has led to some confusion in discussions of sea state bias, as it has no connection with nonlinear wave effects.

7. Discussion and Conclusions

It has been shown how the non-Gaussian theory of sea waves, due to Longuet-Higgins [1963], can be extended to obtain the joint distribution of surface elevation and slopes $p(\zeta, \zeta_x, \zeta_y)$ for a weakly nonlinear wave field. The parameters of the distribution can be related to the wave number spectrum $E(\underline{k})$ and thus, in principle, can be derived from measurements. However, the wave number spectrum $E(\underline{k})$ has not been studied much in comparison to the frequency spectrum $S(\omega)$. This is primarily due to the difficulty of obtaining spatial as opposed to temporal measurements of wave elevation [see Chase et al., 1957; Holthuijsen, 1983].

Radar altimeters are known to sense the spatial, rather than the temporal, statistics of the sea surface, so it has been possible to use the theory derived herein to study the form of the radar return from the sea surface. It has been shown that the form of the radar return depends not only on the sea state, through μ_{200}, but also on the skewness of the sea surface elevation λ_{300}, a skewness parameter γ (given by (20)) and a measure of surface slope $\sqrt{(\mu_{020}\,\mu_{002} - \mu_{011}^2)}$. Thus, as well as the significant wave height H_s, it is possible to derive other information about the sea surface which might be combined to give insight into large scale features of the wave field. For example, by considering simultaneously H_s and λ_{300} it might be possible to distinguish areas of the ocean that are swell dominated (λ_{300} small) from those where active wave generation is occurring (λ_{300} and H_s large).

A by-product of the extra wave information obtainable by using non-Gaussian statistics is the possibility of correcting for sea state bias in altimeter mean level estimates. Other authors [Hayne and Hancock, 1982; Douglas and Agreen, 1983] have suggested simple corrections to remove this bias but the analysis given here shows that this is inadequate. Corrections need to be made on the basis of wave parameters estimated in conjunction with the mean level estimate from the altimeter data. This is necessary if the

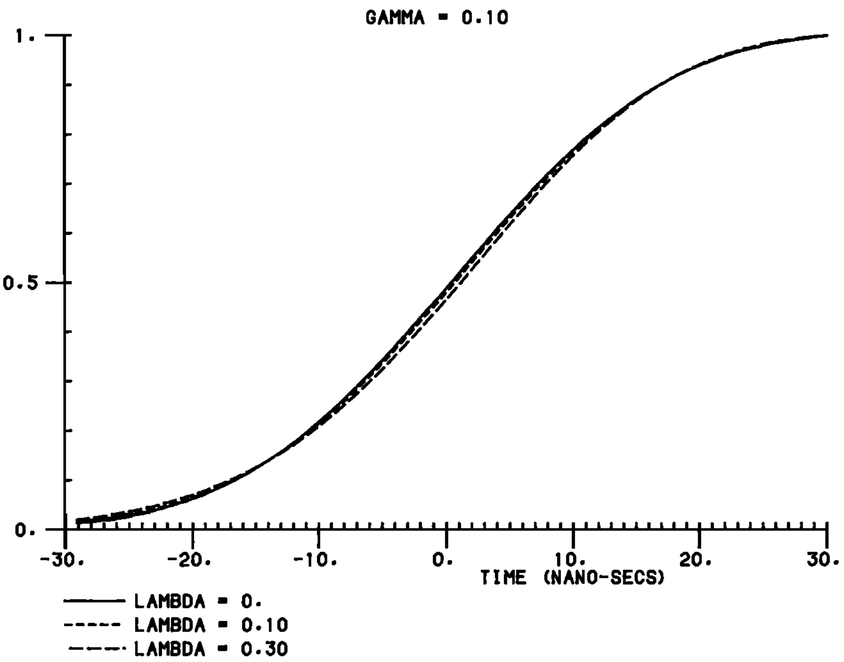


Fig. 5. Return pulse $\sigma(t)$ for H_s = 8 m and γ = 0.1; with (solid curve) λ_{300} = 0, (short-dashed curve) λ_{300} = 0.1, and (long-dashed curve) λ_{300} = 0.3.

accuracy of the mean level estimate is to be of the order of 10 cm or better (necessary for some applications).

It should be noted that, while theoretically all this is possible, comparisons between altimeter-derived skewness parameters and those measured at the sea surface need to be made to verify the results given here (comparisons have been carried out for the significant wave height H_s). This would involve making spatial measurements of the waves in conjunction with altimeter measurements, as some of the parameters (for example, γ) cannot be derived from standard temporal measurements of waves. Two recent studies of electromagnetic bias [Walsh et al. 1984; Choy et al. 1984] have made some progress in this direction. Unfortunately neither paper contains measurements of the parameter γ so a direct check on the theory presented here is not possible from these studies.

During the course of the review of this paper it was brought to the author's attention that Barrick and Lipa [1985] have independently presented some of the results given in this paper. In particular, the result for the distribution of points with zero slope given in (19) is identical to their equation (19). Barrick and Lipa [1985] do not give the more general results for $p(\zeta, \zeta_x, \zeta_y)$ given here, but concentrate instead on evaluating the parameters of the distribution, such as γ, by assuming a JONSWAP form for the wave spectrum. Thus to some extent their results and the ones given here are complementary.

Appendix

Sample Calculation for the Moment μ_{120}

From equations (3) and (5),

$$\mu_{120} = \langle \zeta \, \zeta_x^{\,2} \rangle$$

$$= \langle \zeta_1 \zeta_{1x}^{\,2} \rangle + 2 \langle \zeta_1 \zeta_{1x} \zeta_{2x} \rangle + \langle \zeta_1 \zeta_{2x}^{\,2} \rangle + \langle \zeta_2 \zeta_{1x}^{\,2} \rangle + 2 \langle \zeta_2 \zeta_{1x} \zeta_{2x} \rangle + \langle \zeta_2 \zeta_{2x}^{\,2} \rangle$$

It can easily be seen that,

$$\langle \zeta_1 \zeta_{1x}^{\,2} \rangle = \langle \zeta_1 \zeta_{2x}^{\,2} \rangle = \langle \zeta_2 \zeta_{1x} \zeta_{2x} \rangle = 0$$

as all are odd order correlations. Furthermore $\langle \zeta_2 \zeta_{2x}^{\,2} \rangle$ is of higher order than the remaining terms so it can be neglected and μ_{120} approximated by

$$\mu_{120} = 2 \langle \zeta_1 \zeta_{1x} \zeta_{2x} \rangle + \langle \zeta_2 \zeta_{1x}^{\,2} \rangle$$

Each term will now be calculated separately.

From (6),

$$\langle \zeta_1 \zeta_{1x} \zeta_{2x} \rangle = \frac{1}{2} \sum_{m,n,i,j} \langle a_i a_j a_m a_n \rangle \langle \cos \phi_n (-k_{xm} \sin \phi_m)$$
$$\times [- C_{ij} (k_{xj} \cos \phi_i \sin \phi_j + k_{xi} \sin \phi_i \cos \phi_j)$$
$$+ S_{ij}(k_{xj} \sin \phi_i \cos \phi_j + k_{xi} \cos \phi_i \sin \phi_j)] \rangle$$

There are three cases to consider for which the ensemble average will be nonzero: (case 1) i=n, j=m, i≠j; (case 2) i=m, j=n, i≠j; and (case 3) i=m=j=n. For all other combinations of the subscripts i, j, m, n the ensemble averages with respect to ϕ_n are zero.

Case 1

$$\frac{1}{2} \sum_{\substack{i,j \\ i \neq j}} \langle a_i^{\,2} \rangle \langle a_j^{\,2} \rangle \langle - k_{xj} \cos \phi_i \sin \phi_j$$
$$[- C_{ij} (k_{xj} \cos \phi_i \sin \phi_j + k_{xi} \sin \phi_i \cos \phi_j)$$

$$+ S_{ij} (k_{kj} \sin\phi_i \cos\phi_i + k_{xi} \cos\phi_i \sin\phi_j)] \rangle$$

$$= \frac{1}{8} \sum_{\substack{i,j \\ i \neq j}} \langle a_i^2 \rangle \langle a_j^2 \rangle [k_{xi}^2 C_{ij} - k_{xi} k_{xj} S_{ij}]$$

Case 2: similar to case 1, it gives

$$\frac{1}{8} \sum_{\substack{i,j \\ i \neq j}} \langle a_i^2 \rangle \langle a_j^2 \rangle [k_{xj}^2 C_{ij} - k_{xj} k_{xi} S_{ij}]$$

Case 3

$$\frac{1}{2} \sum_i \langle a_i^4 \rangle \langle -k_{xi} \cos\phi_i \sin\phi_i$$
$$[- C_{ii} k_{xi}\, 2 \sin\phi_i \cos\phi_i$$
$$+ S_{ii} k_{xi}\, 2 \sin\phi_i \cos\phi_i] \rangle$$
$$= \frac{1}{2} \sum_i \langle a_i^4 \rangle \frac{1}{2} k_{xi}^2 (C_{ii} - S_{ii}) \langle \sin^2 2\phi_i \rangle$$
$$= \frac{1}{4} \sum_i \langle a_i^2 \rangle^2 k_{xi}^2 (C_{ii} - S_{ii})$$

Thus

$$\langle \zeta_1 \zeta_{1x} \zeta_{2x} \rangle = \frac{1}{8} \sum_{i,j} \langle a_i^2 \rangle \langle a_j^2 \rangle [(k_{xi}^2 + k_{xj}^2) C_{ij} - 2 S_{ij} k_{xi} k_{xj}]$$

In the above we have used various properties of the random variables a_n, ϕ_n as given in section 3.

A similar calculation to that above shows that

$$\langle \zeta_2 \zeta_{2x}^2 \rangle = \frac{1}{4} \sum_{i,j} \langle a_i^2 \rangle \langle a_j^2 \rangle k_{xi} k_{xj} S_{ij}$$

and so

$$\mu_{120} = \frac{1}{4} \sum_{i,j} \langle a_i^2 \rangle \langle a_j^2 \rangle [(k_{xi}^2 + k_{xj}^2) C_{ij} - k_{xi} k_{xj} S_{ij}]$$

The other results for μ_{mnp} $(m + n + p < 3)$ given in section 3 may be derived similarly.

Evaluation of the integral (17)

From (17) - (19),

$$\sigma(t) = K' \int_{-\infty}^{\infty} [\int_0^{\infty} e^{-(\xi + \frac{ct}{2} - u)^2/\nu^2} du]$$
$$e^{-\xi^2/2\mu_{200}}$$
$$\times [1 + \frac{1}{6} \lambda_{300} (\frac{\xi^3}{\mu_{200}^{3/2}} - \frac{3\xi}{\mu_{200}^{1/2}})$$
$$- \frac{\gamma}{2} \frac{\xi}{\mu_{200}^{1/2}}] \quad d\xi$$

where

$$K' = K P_o (2\pi)^{-1/2} \mu_{200}^{-1/2} (\mu_{020} \mu_{002} - \mu_{011}^2)^{-1/2}$$

Let $x = \xi/\mu_{200}^{1/2}$, and then $\sigma(t)$ may be written as

$$\sigma(t) = K' \mu_{200}^{1/2} \int_0^{\infty} \int_{-\infty}^{\infty}$$
$$[1 + \frac{1}{6} \lambda_{300} x^3 - \frac{1}{2} (\lambda_{300} + \gamma) x]$$
$$\exp \{ - \alpha (x + x_o)^2 - \beta y^2 \} \, dx \, du$$

where

$$\alpha = \frac{1}{2} (1 + 2 \mu_{200}/\nu^2) \quad \beta = (\nu^2 + 2 \mu_{200})^{-1}$$

$$y = \frac{ct}{2} - u \text{ and } x_o = \frac{\mu_{200}^{1/2}}{\alpha \nu^2} y$$

Hence

$$\sigma(t) = K' \mu_{200}^{1/2} \int_0^{\infty} e^{-\beta y^2} \int_{-\infty}^{\infty}$$
$$[1 + \frac{1}{6} \lambda_{300} (x - x_o)^3 - \frac{1}{2} (\lambda_{300} + \gamma)(x - x_o)]$$
$$e^{-\alpha x^2} dx \, du$$
$$= K' \mu_{200}^{1/2} \sqrt{\frac{\pi}{\alpha}} \int_0^{\infty} e^{-\beta y^2} [1 - \frac{1}{6} \lambda_{300} x_o^3$$
$$+ \frac{1}{2} (\lambda_{300} + \gamma) x_o - \frac{1}{4\alpha} \lambda_{300} x_o] du$$
$$= K' \mu_{200}^{1/2} \sqrt{\frac{\pi}{\alpha\beta}} \int_{(-t/t_p)}^{\infty} e^{-v^2} [1 + A v^3 - B v] dv$$

where $v = -\beta^{1/2} y$ and t_p, A and B are given by (25). Evaluating this final integral leads to the result given in (24) for $\sigma(t)$.

Acknowledgments. I am grateful to my colleagues David Carter, Peter Challenor and Trevor Guymer for many helpful discussions and to Andy Margetts for programming the results to produce the diagrams in the paper. I am also grateful to D. E. Barrick for sending me a preprint of his paper (Barrick and Lipa, 1985) and for drawing my attention to the papers by Walsh et al. (1984) and Choy et al. (1984). This research was funded by the European Space Agency and the Department of Trade and Industry.

References

Barrick, D. E., Remote sensing of sea-state by radar, in Remote Sensing of the Troposphere, edited by V. E. Derr, chap. 12, U.S. Government Printing Office, Washington, D. C., 1972.

Barrick, D. E., and B. J. Lipa, Analysis and interpretation of altimeter sea echo, Adv. Geophys., 27, 60-99, 1985.

Brown, G. S., Estimation of surface wind speeds using satellite-borne radar measurements at normal incidence, J. Geophys. Res., 84, 3974-3978, 1979.

Brown, G. S., and L. S. Miller, Microwave backscattering theory and active remote sensing of the ocean surface, NASA Contract. Rep. CR-141423, 1977.

Challenor, P. G., and M. A. Srokosz, Extraction of wave period from altimeter data, Eur. Space Agency Spec. Publ., ESA SP-221, 121-124, 1984.
Chase, J., L. J. Cote, W. Marks, E. Mehr, W. J. Pierson, F. C. Ronne, G. Stephenson, R. C. Vetter, and R. G. Walden, The directional spectrum of a wind generated sea as determined from data obtained by the Stereo Wave Observation Project, Dep. of Meteorol. and Oceanogr. and Eng. Statist. Group, New York Univ., New York, 1957.
Choy, L. W., D. L. Hammond, and E. A. Uliana, Electromagnetic bias of 10-GHz radar altimeter measurements of MSL, Mar. Geod., 8, 297-312, 1984.
Cox, C., and W. H. Munk, Slopes of the sea surface deduced from photographs of sun glitter, Bull. Scripps Inst. Oceanogr., 6, 410-487, 1956.
Douglas, B. C., and R. W., Agreen, The sea state correction for GEOS 3 and SEASAT satellite altimeter data, J. Geophys. Res., 88, 1655-1661, 1983.
Hayne, G. S., and D. W. Hancock, Sea-state-related altitude error in the SEASAT radar altimeter, J. Geophys. Res., 87, 3227-3231, 1982.
Holthuijsen, L. H., Stereophotography of ocean waves, Appl. Ocean Res., 5, 204-209, 1983.
Huang, N. E., S. R. Long, L. F., Bliven, and C. C. Tung, The non-Gaussian joint probability density function of slope and elevation for a nonlinear gravity wave field, J. Geophys. Res., 89, 1961-1972, 1984.
Jackson, F. C., The reflection of impulses from a nonlinear random sea, J. Geophys. Res., 84, 4939-4943, 1979.
Lipa, B. J., and D. E. Barrick, Ocean surface height-slope probability density function from SEASAT altimeter echo, J. Geophys. Res., 86, 10,921-10,930, 1981.
Longuet-Higgins, M. S., The statistical analysis of a random, moving surface, Philos. Trans. R. Soc. London, Ser. A, 249, 321-387, 1957.
Longuet-Higgins, M. S., The effect of nonlinearities on statistical distributions in the theory of sea waves, J. Fluid Mech., 17, 459-480, 1963.
Walsh, E. J., D. W. Hancock, and D. E. Hines, Electromagnetic bias of 36-GHz radar altimeter measurements of MSL, Mar. Geod., 8, 265-296, 1984.

M. A. Srokosz, Institute of Oceanographic Sciences, Brook Road, Wormley, Godalming, Surrey, GU8 5UB, England.

(Received February 7, 1985;
accepted July 9, 1985.)