# Wave statistics in nonlinear random sea

Ulla Machado and Igor Rychlik *

Centre for Mathematical Sciences

Mathematical Statistics

Lund University, Sweden

**Abstract**

The sea elevation at a fixed point is modelled as a quadratic form of a vector valued Gaussian process with arbitrary mean. An apparent wave is a part of the sea record observed between two following upcrossings of the still water level. Saddlepoint method is used to approximate the intensity $\mu(u)$, say, the sea level crosses the level $u$. The intensity is then used to estimate more complicated wave characteristics like wave period and crest height. Numerical examples are given.

**Keywords:** Crest distribution, non-Gaussian sea, Rice's formula, wave period and amplitude, Stokes waves.

## 1 Introduction

An accurate description of environmental loads, such as wind and ocean waves, is important in design of offshore structures and evaluation of risks. The safety of a structure may depend on extreme and rare events such as loads which exceed the strength of a component, or on

*Research supported in part by the Knut and Alice Wallenberg Foundation through the Gothenburg Stochastic Centre

everyday load variability that may cause changes in the properties of the material, e.g. cracking (fatigue) or other types of ageing processes.

From experience it is known that an irregular sea surface with high waves is stationary for a relatively short period of time, between 20 minutes and a few hours. Consequently, a measured sea surface at a fixed point contains approximately between 100 and 1000 waves, which can be considered to be equally distributed. This, limits the possibility to estimate frequencies and properties of relatively rare and possibly dangerous waves. Since, in safety analysis, the knowledge of the frequencies of big waves is important one often relays on mathematical models of the sea surface and compute the so called short term distributions of wave characteristics. Finally, one uses the long term statistics, describing the distribution of times that a structure will be exposed to different sea conditions, to mix the short term distribution. The mixed distribution gives the frequencies of waves, having wave characteristics of interest, that can be expected during service time of a structure.

The deep water sea with moderate waves (not too steep) is often well described using Gaussian fields based on the linear wave theory. However, for severe sea states considerable asymmetry is observed; troughs are shallower while crests are higher than predicted by the Gaussian model. Since the crests can be up to 20% higher, these effects can not be neglected. In the literature one is often modelling the observed wave asymmetry by adding a random quadratic correction to the Gaussian model term.

For Gaussian fields there are many tools developed to estimate the wave characteristic distributions. The non-linear sea models are more difficult to analyse. The simulation techniques can still be used but the analytical computations of the wave characteristic distributions are very difficult. For example, even the one-dimensional density of the sea elevation has to be computed using approximate methods; see Machado (2002) for a review of different approaches. Here we shall be concerned with computations of the crossing intensity $\mu(u)$, i.e. expected frequency the sea level passes a fixed level $u$. This quantity is crucial in safety analysis of offshore structures. In Section 2 we shall shortly present different applications of $\mu(u)$. In Section 3 a review of the basic properties of the second order sea model is given, while in Section 4 the saddlepoint method will be employed to approximate the crossing intensity. Finally, Section 5 contains numerical examples.

## 2 Rice's formula and its applications

Assume that $\eta$ is a strictly stationary process. For a fixed $u$, let $\mu(u)$ be the expected number of times the process $\eta$ crosses the level $u$ in the upward direction. The following classical result is called Rice's formula; see Leadbetter et al. (1983) for a proof.

**Theorem 1** *Let $\eta(t)$ be a stationary, zero mean Gaussian process. If the derivative $\dot{\eta}(t)$ exists (in a quadratic mean) then*

$$\mu(u) = \frac{1}{2\pi}\sqrt{\frac{\lambda_2}{\lambda_0}}e^{-\frac{u^2}{2\lambda_0}}, \qquad (1)$$

$$= \int_0^{+\infty} z f_{\eta(0),\dot{\eta}(0)}(u,z)\, dz \qquad (2)$$

*where $f_{\eta(0),\dot{\eta}(0)}(u,z)$ is the joint density of $\eta(0), \dot{\eta}(0)$, and $\lambda_0, \lambda_2$ are spectral moments equal to the variances of $\eta(0), \dot{\eta}(0)$, respectively.*

Rice (1944, 1945) and Kac (1943), obtained (1) for a certain class of Gaussian processes and polynomials with random coefficients, respectively.

In engineering literature formula (2) is used to compute the upcrossing intensity $\mu(u)$ even for non-Gaussian processes as long as the density of $\eta(0), \dot{\eta}(0)$ is available. Since the density is not uniquely defined, equation (2) can not be true without some additional conditions. However, the following general result can be shown:

$$\mu(u) \overset{\text{a.a.}u}{=} \int_0^{+\infty} z f_{\eta(0),\dot{\eta}(0)}(u,z)\, dz,$$

where $\overset{\text{a.a.}u}{=}$ means that equality is valid for almost all $u$. For process $\eta$ for which the right side of the last equation is a continuous function of $u$, this somewhat weaker version of Rice's formula can still be applied in many engineering problems; see Rychlik (2000) and the appendix in Baxevani et al. (2002) for discussion of validity and applications of Rice's formula in oceanography.

Rice's formula extends also to non-stationary processes. Let $E[N_T(u)]$ be the expected number of times the process $\eta(t)$, $t \in [0,T]$, crosses the level $u$; then

$$E[N_T(u)] \overset{\text{a.a.}u}{=} \int_0^T \int_0^{+\infty} z\, f_{\eta(t),\dot\eta(t)}(u,z)\, dz\, dt = \int_0^T \mu_t(u)\, dt.$$

Here $\mu_t(u)$ is the time variable crossing intensity. An important application of the crossing intensity is to bound the distribution of the size of the highest wave crest $M_T = \max_{0\le t\le T} \eta(t)$ the offshore structure is exposed to during a specified period of time $T$. If $\mu_t(t)$ is known then the distribution of $M_T$ can be bound as follows

$$P(M_T > u) \le P(\eta(0) > u) + E[N_T(u)]. \tag{3}$$

(The last inequality is referred as the "Rice's method".) The time variable crossing intensity $\mu_t(u)$ is constant for the duration of the sea state, and can be computed using (2). The last inequality is sharp for a high level $u$, long periods $T$ and when $\int_0^T \mu_t(u)\, dt$ is small.

Another important problem in safety analysis is to model the ageing process in the material, as damage accumulates with a rate depending both on the frequency and the magnitude of waves. Here the knowledge of the crest height distribution is essential. Consider first the stationary case and denote by $m$ the still water level. Then often one assumes that the intensity of waves is equal to $\mu(m)$. (We propose to choose $m$ as the level which is most frequently crossed by $\eta$.) An important characteristic in fatigue analysis is the distribution of the height of wave crest $A_c$, say. As shown in Rychlik (1993) and, Rychlik and Leadbetter (2000), we have

$$P(A_c > h) \le \frac{\mu(h)}{\mu(m)}, \tag{4}$$

with equality as $h$ tends to infinity. By means of (1) one has that $P(A_c > h) \le \exp(-h^2/2m_0)$, for a Gaussian sea.

Next, let us consider the distribution of crest heights that a structure may experience during its service time $T$. For long $T$ periods the stationarity assumption can not be justified. In such a case the probability $P(A_c > h)$ is defined as a normalised expected number of waves with crest heights exceeding $h$ in $[0,T]$. Similarly as in the stationary case, the probability can be bounded as follows

$$P(A_c > h) \leq \frac{\int_0^T \mu_t(h)\,dt}{\int_0^T \mu_t(m)\,dt}.$$

Here we have also assumed that the still water level $m$ is constant during the whole service period $T$.

Finally, we shortly address the problem of computing the joint distributions of wave characteristics, for example the joint distribution of crest period $T_c$ and crest amplitude $A_c$, defined as the distance between upcrossing and the following downcrossing of the still water level $m$ and the maximum height in the interval, respectively. This is a very complicated problem even for the Gaussian sea model; see Podgórski et al. (2000) for a review of different approaches. For the non-Gaussian sea one possible approach is to approximate the sea elevation by some simpler process for which one has methods to compute the distributions of wave characteristics.

A class of processes that can be used in this way is that of the transformed Gaussian processes, i.e., processes $\eta(t)$ such that $\eta(t) = G(\tilde{\eta}(t))$ for some continuous increasing function $G$ and a Gaussian process $\tilde{\eta}(t)$ with variance one. Since the intensity of high crests above $u$ is well approximated by the crossing intensity $\mu(u)$, Rychlik et al. (1997) proposed to choose $G$ so that $G(\tilde{\eta}(t))$ has crossing intensities equal to $\mu(u)$. Actually the inverse function $g(u) = G^{-1}(u)$ is needed in the computations and is given by

$$g(u) = \begin{cases} \sqrt{-2\ln(\mu(u)/\mu(m))} & \text{if } u > m, \\ -\sqrt{-2\ln(\mu(u)/\mu(m))} & \text{if } u \leq m. \end{cases} \tag{5}$$

In Section 5 we shall compare the joint distribution of $T_c, A_c$ computed for $\eta(t) = G(\tilde{\eta}(t))$ with the density estimated from simulations.

# 3 Second order description of the sea surface elevation

We begin with the linear sea model, which postulates that the sea surface is a sum of simple cosine waves. In this paper we consider only long crested sea, i.e. the surface does not depend on the direction $y$. In addition we consider an unidirectional sea, where all waves travel along the $x$ axis with positive velocity. The linear sea $\eta_l$, consisting of $N$ cosine waves, is given by

$$\eta_l(x,t) = \sum_{n=-N}^{N} \frac{A_n}{2}\, e^{i(\omega_n t - \kappa_n x)}, \tag{6}$$

where for each elementary wave $A_n$ denotes its complex valued amplitude, $\omega_n$ angular frequency and $\kappa_n$ wave number. We assume that $A_{-n} = \bar{A}_n$, where $\bar{z}$ denotes complex conjugate of $z$. Since $\eta_l$ should be a real valued field, we need to assume that $\omega_{-j} = -\omega_j$, $\kappa_{-j} = -\kappa_j$. Finally, the linear wave theory postulates that $\kappa$ and $\omega$, are functionally related by the so-called dispersion relation

$$\omega^2 = g\kappa \tanh(h\kappa), \qquad \omega > 0,$$

where $g$ and $h$ are the earth acceleration and water depth, respectively. This equation implies that long waves travel faster than short ones.

Measurements of the real sea show that the linear wave model is often too simplistic and leads to errors in the predicted crest height (in deep water) of about 10-20%. The model can be corrected using "second-order" (or "quadratic") terms allowing for interactions between the elementary cosine waves. Following Hasselmann (1962), where the detailed derivations are given (see also Langley (1987)), the quadratic correction $\eta_q$ is given by

$$\eta_q(x,t) = \sum_{n=-N}^{N} \sum_{m=-N}^{N} \frac{A_n}{2}\frac{A_m}{2}\, E(\omega_n,\omega_m)\, e^{i(\omega_n t - \kappa_n x)} e^{i(\omega_m t - \kappa_m x)}, \tag{7}$$

where the amplitudes $A$, angular frequencies $\omega$ and wave numbers $\kappa$ satisfy the same relations as in the linear model. The quadratic-transfer function (QTF), $E(\omega,\tilde{\omega})$, is given by

$$E(\omega,\tilde{\omega}) = \frac{\frac{g\kappa\tilde{\kappa}}{\omega\tilde{\omega}} - \frac{1}{2g}(\omega^2 + \tilde{\omega}^2 + \omega\tilde{\omega}) + \frac{g}{2}\frac{\omega\tilde{\kappa}^2 + \tilde{\omega}\kappa^2}{\omega\tilde{\omega}(\omega+\tilde{\omega})}}{1 - g\frac{\kappa+\tilde{\kappa}}{(\omega+\tilde{\omega})^2}\tanh(\kappa+\tilde{\kappa})h} - \frac{g\kappa\tilde{\kappa}}{2\omega\tilde{\omega}} + \frac{1}{2g}(\omega^2 + \tilde{\omega}^2 + \omega\tilde{\omega}), \tag{8}$$

where $\kappa, \tilde{\kappa}$ are wave numbers which are computed using the dispersion relation from the angular frequencies $\omega, \tilde{\omega}$, respectively. (We also define $E(0,0) = 0$.) It is important to note that for any positive $\omega$ and $\tilde{\omega}$, the QTF satisfies the symmetry relations $E(\omega,\tilde{\omega}) = E(\tilde{\omega},\omega)$, $E(\omega,\tilde{\omega}) = E(-\omega,-\tilde{\omega})$ and $E(\omega,-\tilde{\omega}) = E(-\omega,\tilde{\omega})$. These properties imply that $\eta_q$ is a real valued field.

For the deep water waves the QTF simplifies, and for positive $\omega, \tilde{\omega}$ we have

$$E(\omega, \tilde{\omega}) = \frac{1}{2g}\,(\omega^2 + \tilde{\omega}^2), \tag{9}$$
$$E(\omega, -\tilde{\omega}) = -\frac{1}{2g}\,|\omega^2 - \tilde{\omega}^2|.$$

Since $\lim_{h\to\infty} \tanh(\kappa\, h) = 1$ the dispersion relation simplifies to $\omega^2 = g\kappa$.

**Definition 2** *The deterministic Second-Order Stokes Wave is defined as*

$$\eta^N(x,t) = \eta_l(x,t) + \eta_q(x,t) \tag{10}$$

*where $\eta_l$ and $\eta_q$ are the linear and quadratic processes given by Eqs. (6) and (7), respectively. The Gaussian second-order sea is obtained by assuming that the complex amplitudes $A_n$, $n > 0$, are independent and normally distributed variables, i.e. $A_n = \sigma_n(U_n - iV_n)$, where $U_n, V_n$ are independent zero mean and variance one Gaussian variables, and $\sigma_n^2$ is the energy of waves with angular frequencies $\omega_n$ and $-\omega_n$.*

Often it is assumed that the linear Gaussian process $\eta_l$ has a spectral density. This case is not covered by Definition 2. The following remark addresses this problem.

**Remark 3** *For a sea model with linear one-sided spectrum $S(\omega)$, $0 \le \omega \le \omega_c$, where $\omega_c$ is the cut off frequency, we define $\eta(x,t)$ as the limit, as $N$ tends to infinity, of the second order Gaussian sea $\eta^N(x,t)$ given by (10). The individual waves have angular frequencies $\omega_j = j\omega_c/N$ and energy*

$$\sigma_j^2 = S(\omega_j)\Delta\omega,$$

*$j = 1,\ldots,N$, while $\Delta\omega = \omega_c/N$. The cut off frequency $\omega_c$ is often used in oceanography to reduce the noise in the models. The small wiggles, caused by the high frequencies, on the top of the main body of an apparent waves are irrelevant in applications. The Gaussian process with band limited spectrum is smooth, i.e. has a.s. derivatives of all orders. The conditions for convergence of distributions of wave characteristics observed in $\eta^N(0,t)$ to the wave characteristics in $\eta(0,t)$ were discussed by several authors, see for example Eplett (1986), Rychlik (1987), Seleznjev (1991) and references therein. We shall not consider this type of problems in this paper.*

In the following section we shall estimate the crossing intensity $\mu(u)$ for the process $\eta^N(0,t)$. In order to simplify the notation we shall write $\eta(t)$ for $\eta^N(0,t)$. By (2), the computation of $\mu(u)$ requires estimate of the joint density of $\eta(0), \dot\eta(0)$. However, an explicit closed form formula for the joint density of $\eta(0), \dot\eta(0)$ is not known at present (except when $N = 1$). Even the marginal densities of $\eta(0)$, $\dot\eta(0)$ have to be derived using approximate methods; see Machado (2002) for a review of the existing approximations. Here we shall use the saddlepoint method to approximate $f_{\eta(0),\dot\eta(0)}(u,z)$. In order to employ the method one needs the explicit formula for the characteristic function of $\eta(0), \dot\eta(0)$, which will be given in the following subsection.

## 3.1 Characteristic function of $\eta(0), \dot\eta(0)$

In order to write the formula for the characteristic function in a transparent way we shall first rewrite the defining equation for the process $\eta(t)$.

Assume that the angular frequencies $\omega_j$ and average energies $\sigma_j$, $j = 1,\ldots,N$, are chosen. Denote by $\boldsymbol\sigma$ the column vector containing $\sigma_j$ while $\boldsymbol\omega$ be the column vector of $\omega_j$. Define

$$\mathbf{Z}(t) = [(U_1 - iV_1)e^{i\omega_1 t} \ldots (U_N - iV_N)e^{i\omega_N t}]^T,$$

and let $\mathbf{X}(t)$, $\mathbf{Y}(t)$ be the real and the imaginary parts of $\mathbf{Z}(t)$, i.e.

$$\mathbf{Z}(t) = \mathbf{X}(t) + i\mathbf{Y}(t).$$

Using the notation, the linear part of the random sea writes

$$\eta_l(t) = \frac{1}{2}\boldsymbol\sigma^T\mathbf{Z}(t) + \frac{1}{2}\bar{\boldsymbol\sigma}^T\bar{\mathbf{Z}}(t) = \boldsymbol\sigma^T\mathbf{X}(t). \tag{11}$$

We shall rewrite the nonlinear part in a similar way. First let us introduce some matrices

$$\begin{aligned}
\mathbf{Q} &= [q_{mn}], \quad q_{mn} = (E(\omega_m,-\omega_n) + E(\omega_m,\omega_n))\sigma_m\sigma_n, & (12)\\
\mathbf{R} &= [r_{mn}], \quad r_{mn} = (E(\omega_m,-\omega_n) - E(\omega_m,\omega_n))\sigma_m\sigma_n, & (13)\\
\mathbf{W} &= [w_{mn}], \quad w_{mm} = -\omega_m, \text{ and } w_{mn} = 0 \text{ if } m \neq n, & (14)\\
\mathbf{S} &= \mathbf{QW} - \mathbf{WR}, & (15)
\end{aligned}$$

where $m, n = 1, \ldots, N$. Some simple algebra shows that

$$\eta_q(t) = \frac{1}{2}\mathbf{X}(t)^T\mathbf{Q}\mathbf{X}(t) + \frac{1}{2}\mathbf{Y}(t)^T\mathbf{R}\mathbf{Y}(t), \tag{16}$$

and hence

$$\eta(t) = \boldsymbol{\sigma}^T\mathbf{X}(t) + \frac{1}{2}\mathbf{X}(t)^T\mathbf{Q}\mathbf{X}(t) + \frac{1}{2}\mathbf{Y}(t)^T\mathbf{R}\mathbf{Y}(t). \tag{17}$$

Furthermore, since

$$\dot{\mathbf{X}}(t) = \mathbf{W}\mathbf{Y}(t), \qquad \dot{\mathbf{Y}}(t) = -\mathbf{W}\mathbf{X}(t), \tag{18}$$

we have the following formula for the derivative $\dot{\eta}(t)$:

$$\begin{aligned}\dot{\eta}(t) &= \boldsymbol{\sigma}^T\mathbf{W}\mathbf{Y}(t) + \mathbf{X}(t)^T\mathbf{Q}\mathbf{W}\mathbf{Y}(t) - \mathbf{X}(t)^T\mathbf{W}\mathbf{R}\mathbf{Y}(t)\\ &= \boldsymbol{\sigma}^T\mathbf{W}\mathbf{Y}(t) + \frac{1}{2}\mathbf{X}(t)^T\mathbf{S}\mathbf{Y}(t) + \frac{1}{2}\mathbf{Y}(t)^T\mathbf{S}^T\mathbf{X}(t).\end{aligned} \tag{19}$$

We turn now to derivations of the formula for the characteristic function of $\eta(0), \dot{\eta}(0)$, denoted by $M(\theta_1, \theta_1) = \mathrm{E}[\exp\{i(\theta_1\eta(0) + \theta_2\dot{\eta}(0))\}]$. In the computation of $M(\theta_1, \theta_1)$ we will employ the following well known result from Cramér (1946):

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \frac{1}{(2\pi)^{n/2}} e^{\mathbf{t}^T\mathbf{z} - \frac{1}{2}\mathbf{z}^T\mathbf{A}\mathbf{z}}\, dz_1 \ldots dz_n = \frac{1}{\sqrt{\det(\mathbf{A})}}\, e^{\frac{1}{2}\mathbf{t}^T\mathbf{A}^{-1}\mathbf{t}}. \tag{20}$$

Here, the real $(n, n)$-matrix $\mathbf{A}$ must be symmetric, non-singular and positive definite. (Obviously, $\mathbf{t}$, $\mathbf{z}$ are $n$-dimensional column vectors.)

**Lemma 4** *The characteristic function of the variables $\eta(0), \dot{\eta}(0)$, defined by (17), (19), respectively, is given by*

$$M(\theta_1, \theta_1) = \frac{1}{\sqrt{\det(\mathbf{I} - \mathbf{A})}}\, e^{\frac{1}{2}\mathbf{t}^T(\mathbf{I}-\mathbf{A})^{-1}\mathbf{t}}, \tag{21}$$

*where $\mathbf{I}$ is a $(2N, 2N)$-dimensional identity matrix. The matrix $\mathbf{A} = \mathbf{A}(i\theta_1, i\theta_2)$ and the vector $\mathbf{t} = \mathbf{t}(i\theta_1, i\theta_2)$ are defined as follows*

$$\mathbf{A}(x, y) = \begin{bmatrix} x\mathbf{Q} & y\mathbf{S} \\ y\mathbf{S}^T & x\mathbf{R} \end{bmatrix}, \qquad \mathbf{t}(x, y) = \begin{bmatrix} x\boldsymbol{\sigma} \\ y\mathbf{W}\boldsymbol{\sigma} \end{bmatrix}. \tag{22}$$

*Proof:* The variables $\mathbf{X}(0), \mathbf{Y}(0)$ are independent standard Gaussian and their joint density $f(\mathbf{z})$ is given by

$$f(\mathbf{z}) = \frac{1}{(2\pi)^N} e^{-\frac{1}{2}\mathbf{z}^T \mathbf{I} \mathbf{z}},$$

hence the characteristic function

$$M(\theta_1, \theta_1) = E[\exp(i(\theta_1 \eta(0) + \theta_2 \dot{\eta}(0)))] = \int \frac{1}{(2\pi)^N} e^{\mathbf{t}^T \mathbf{z} - \frac{1}{2}\mathbf{z}^T (\mathbf{I}-\mathbf{A}) \mathbf{z}} \, dz_1 \ldots \, dz_{2N}.$$

Since $\mathbf{I} - \mathbf{A}$ satisfies the assumptions of formula (20) the lemma follows. □

The characteristic function defines uniquely the density function by means of the inverse Fourier transform

$$f(u, y) = \frac{1}{(2\pi)^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-i(\theta_1 u + \theta_2 y)} M(\theta_1, \theta_2) \, d\theta_1 \, d\theta_2, \tag{23}$$

and hence the crossing intensity can be computed by combining formulas (2) and (23), viz.

$$\mu(u) = \frac{1}{(2\pi)^2} \int_0^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y e^{-i(\theta_1 u + \theta_2 y)} M(\theta_1, \theta_2) \, d\theta_1 \, d\theta_2 \, dy. \tag{24}$$

Obviously the integration in (24) has to be done numerically. In previous works, see Næss and Machado (2000) and Michna and Rychlik (1995), numerical procedures to evaluate (24) have been presented. The advantage of the numerical integration procedure is that it usually gives very accurate results. However, since typically the matrix $\mathbf{A}$ is large, may have dimension $(500, 500)$ and more, and since each evaluation of characteristic function requires evaluation of the inverse $(\mathbf{I} - \mathbf{A})^{-1}$ the numerical integration becomes extremely slow. Here we shall propose two simplifications in the estimation of $\mu(u)$. The first is to use the saddlepoint method to approximate the joint density of $\eta(0), \dot{\eta}(0)$; this can be seen as an approximation of the integral in (23); see Daniels (1954) for discussion of the one dimensional case. The second is to reduce the dimension of the matrix $\mathbf{A}$. In examples the proposed method will be called "SPM". The SPM-approximation will be compared with estimated crossing intensities from simulated second order sea and with $\mu(u)$ computed by means of numerical integration in (24). This procedure will be called "NM" and the resulting values of $\mu(u)$ will be treated as being exact.

# 4 Approximation of $\mu(u)$ using the saddlepoint method

The saddlepoint density was first introduced by Daniels (1954) as a formula to approximate the probability density function from its cumulant generating function. We shall apply the method to approximate the density of $\eta(0), \dot{\eta}(0)$. The cumulant generating function $K(s_1, s_2)$ of $\eta(0), \dot{\eta}(0)$ is defined as

$$K(s_1, s_2) = \ln(\mathrm{E}[\exp\{s_1\eta(0) + s_2\dot{\eta}(0)\}]), \quad (s_1, s_2) \in S,$$

where $S$ is the set of arguments for which the last integral converges. (Obviously $(0,0) \in S$). The cumulant generating function can be computed from the characteristic function $M$ by means of the relation $K(s_1, s_2) = \ln M(-is_1, -is_2)$. Using Lemma 4 we have that

$$K(s_1, s_2) = -\frac{1}{2}\ln(\det(\mathbf{I} - \mathbf{A})) + \frac{1}{2}\mathbf{t}^T(\mathbf{I} - \mathbf{A})^{-1}\mathbf{t}, \quad (s_1, s_2) \in S, \tag{25}$$

where $\mathbf{A} = \mathbf{A}(s_1, s_2)$ and $\mathbf{t} = \mathbf{t}(s_1, s_2)$; see (22). Since $A(s_1, s_2)$ is a matrix of continuous functions of $(s_1, s_2)$ it follows that $S$ is an open set. It is also known that $S$ is convex.

*The saddlepoint approximation* $\hat{f}(u, y)$, say, of the density $f_{\eta(0),\dot{\eta}(0)}(u, y)$ is defined by

$$\hat{f}(u, y) = (2\pi)^{-1}\,|\ddot{K}(\hat{s}_1, \hat{s}_2)|^{-\frac{1}{2}}\,\exp\{K(\hat{s}_1, \hat{s}_2) - \hat{s}_1 u - \hat{s}_2 y\}, (u, y) \in \mathbb{R}^2. \tag{26}$$

The so called saddle point $(\hat{s}_1, \hat{s}_2) \in S$ is the unique solution of the following system of equations

$$\begin{cases} K_1(s_1, s_2) = \frac{\partial K(s_1,s_2)}{\partial s_1} = u, \\ K_2(s_1, s_2) = \frac{\partial K(s_1,s_2)}{\partial s_2} = y. \end{cases} \tag{27}$$

Furthermore $\ddot{K}$ is the Hessian matrix (containing second order derivatives) of the cumulant generating function $K$.

The existence and uniqueness of the saddle point for all $(u, y) \in R^2$ is the consequence of the fact that $S$ is open and that the support of the range of the density of $\eta(0), \dot{\eta}(0)$ is $R^2$. Since $K$ is a strictly convex function on $S$ the Hessian $\ddot{K}$ is positive definite. Finally, note that $\hat{f}(u, y)$ usually does not integrate to one and hence it is not a density. Some further properties of cumulant generating function and the saddle point method can be found in Barndoff-Nielsen and Cox (1989).

Although the saddlepoint estimate $\hat{f}(u,y)$ seems to be explicitly defined function this is not the case since it operates on $\hat{s}_1, \hat{s}_2$ which are defined as the solutions of a non-linear equation system (27). Often $\hat{s}_1, \hat{s}_2$ have to be found by means of a numerical method. The use of an efficient algorithm to invert equation (27) is important, since each evaluation of $K(s_1, s_2)$ requires the computation of the inverse of a large matrix. In the examples presented here we have used standard procedures supplied by MATLAB. We now turn to the algorithm used to estimate the crossing intensity $\mu(u)$.

*Algorithm:* Evaluate $K(s_1, s_2)$ on a grid, then, for fixed level $u$, find a contour line

$$\mathcal{C}(u) = \{(s_1, s_2) \in S : K_1(s_1, s_2) = u\}.$$

On the contour a sequence of points $\mathbf{s}^k \in \mathcal{C}(u)$, $k = 1, \ldots, n$, is chosen. For each $\mathbf{s}^k$ compute $y^k = K_2(\mathbf{s}^k)$, and

$$f^k = \hat{f}(u, y^k) = (2\pi)^{-1} \, |\ddot{K}(\mathbf{s}^k)|^{-\frac{1}{2}} \, \exp\{K(\mathbf{s}^k) - s_1^k u - s_2^k K_2(\mathbf{s}^k)\},$$

if $y^k > 0$, $f^k = 0$ otherwise. Finally, use $(y^k, f^k)$, $k = 1, \ldots, n$, to estimate the crossing intensity

$$\mu(u) = \int_0^{+\infty} y f_{\eta(0),\dot{\eta}(0)}(u,y)\, dy \approx \int_0^{+\infty} y \hat{f}(u,y)\, dy \approx \sum y^k \frac{(f^k + f^{k-1})}{2} |y^k - y^{k-1}|.$$

□

In the following remark we present some simple properties of the cumulant generating function $K$ for the Gaussian second order sea which are used in the algorithm.

**Remark 5** *From formulas (17) and (19) it follows that $K(s_1, s_2) = K(s_1, -s_2)$. Consequently one needs only to find the contour $\mathcal{C}(u)$ for $(s_1, s_2) \in S$ and $s_2 \geq 0$. Actually the algorithm first finds the one-dimensional saddle point associated with $(u, 0)$. Since $K_2(s_1, 0) = 0$ then $\hat{s}_2 = 0$ while $\hat{s}_1$ is the solution of the equation $K_1(s_1, 0) = u$. The saddle $(\hat{s}_1, 0)$ can be used to compute $\hat{f}(u)$ (the saddle point approximation of the marginal density of $\eta(0)$) and as the starting point for the algorithm searching for the contour $\mathcal{C}(u)$. Next we give some estimates on the size of the set $S$ in which the contour $\mathcal{C}(u)$ is contained.*

*Let $\alpha_i$, $\beta_i$, $i = 1, \ldots, 2N$, be the eigenvalues of the following two matrices* $\mathbf{A}(1,0), \mathbf{A}(0,1)$, *defined by (22). By symmetry of the matrices all eigenvalues are real. It can be shown, see Langley (1987), that $\sum \alpha_i = 0$ and hence*

$$\alpha_- = \min\{\alpha_i\} < 0 < \max\{\alpha_i\} = \alpha_+.$$

*Further, let $\beta = \max\{|\beta_i|\}$. Now by studying zeros of* $\det(\mathbf{A}(s_1, 0))$ *and* $\det(\mathbf{A}(0, s_2))$ *we obtain that*

$$S \cap \{(s_1, s_2) : s_2 = 0\} = (\alpha_-^{-1}, \alpha_+^{-1}), \quad S \cap \{(s_1, s_2) : s_1 = 0\} = (-\beta^{-2}, \beta^{-2}).$$

*Finally, in examples the derivatives $K_1, K_2$ and the Hessian $\ddot{K}$ are computed numerically. We compared the values of one dimensional saddlepoint density $\hat{f}(u)$ computed using numerical derivatives and based on explicit formulas (exact). The error was negligible.*

Since the number of cosine waves $N$ building the signal $\eta(t)$ can be large, often $N > 250$, the computation of $K(s_1, s_2)$, which involves inverting a $(2N, 2N)$ dimensional matrix $\mathbf{I} - \mathbf{A}$, can be a limiting factor. (The inverse has to be computed for each pair $\mathbf{s} = (s_1, s_2)$.) In the following subsection we shall present some approximations of $K(s_1, s_2)$.

## 4.1 Approximations of the cumulant generating function

As we mentioned before, the computation of $K(s_1, s_2)$, the cumulant generating function of $\eta(0), \dot{\eta}(0)$, requires the evaluation of $(\mathbf{I} - \mathbf{A})^{-1}$. In this section we shall discuss some approximations that will speed up the computations.

The matrix $\mathbf{A}$, see (22), has a block structure. The two blocks on the diagonal, namely the matrices $\mathbf{Q}$ and $\mathbf{R}$, reflect the influence of the non-linear part of the sea model $\eta(t)$; see (17). In the following we shall simplify the quadratic term $\eta_q(t)$, defined by (16), leading to a faster computation of the crossing intensity. This will be accomplished by means of the diagonalization of the quadratic forms $\mathbf{X}(t)^T\mathbf{Q}\mathbf{X}(t)$ and $\mathbf{Y}(t)^T\mathbf{R}\mathbf{Y}(t)$.

Let $\mathbf{P}_1$ be the $(N, N)$ matrix whose rows are the eigenvectors of $\mathbf{Q}$ ordered according to increasing absolute value of their eigenvalues, which we denote by $\lambda_j$, $j = 1, \ldots, N$, and let

$\mathbf{\Lambda}$ be a diagonal matrix with the $\lambda_j$ as elements. Similarly, let denote by $\gamma_j$, $j = 1, \ldots, N$ the eigenvalues of $\mathbf{R}$, let $\mathbf{\Gamma}$ be a diagonal matrix with the $\gamma_j$ as elements, and let $\mathbf{P}_2$ be the $(N, N)$ matrix whose rows are the eigenvectors of $\mathbf{R}$ ordered by increasing values of the $|\gamma_j|$. Since $\mathbf{Q}, \mathbf{R}$ are symmetric its eigenvalues are real and

$$\mathbf{Q} = \mathbf{P}_1^T \mathbf{\Lambda} \mathbf{P}_1, \quad \mathbf{R} = \mathbf{P}_2^T \mathbf{\Gamma} \mathbf{P}_2.$$

Similarly as in Langley (1987), let us introduce

$$\mathbf{Z}_1(t) = \mathbf{P}_1 \mathbf{X}(t), \quad \mathbf{Z}_2(t) = \mathbf{P}_2 \mathbf{Y}(t).$$

Using (18) and the fact that $\mathbf{P}_1$, $\mathbf{P}_2$ are orthonormal matrices, we obtain

$$\dot{\mathbf{Z}}_1(t) = \mathbf{P}_1 \mathbf{W} \mathbf{P}_2^T \mathbf{Z}_2(t), \quad \dot{\mathbf{Z}}_2(t) = -\mathbf{P}_2 \mathbf{W} \mathbf{P}_1^T \mathbf{Z}_1(t).$$

The sea surface $\eta(t)$ and its derivative can now be written as

$$\begin{aligned} \eta(t) &= (\mathbf{P_1}\boldsymbol{\sigma})^T \mathbf{Z}_1(t) + \frac{1}{2}\mathbf{Z}_1(t)^T \mathbf{\Lambda} \mathbf{Z}_1(t) + \frac{1}{2}\mathbf{Z}_2(t)^T \mathbf{\Gamma} \mathbf{Z}_2(t), && (28) \\ \dot{\eta}(t) &= \boldsymbol{\sigma}^T \mathbf{W} \mathbf{P}_2^T \mathbf{Z}_2(t) + \mathbf{Z}_1(t)^T \left( \mathbf{\Lambda} \mathbf{P}_1 \mathbf{W} \mathbf{P}_2^T - \mathbf{P}_1 \mathbf{W} \mathbf{P}_2^T \mathbf{\Gamma} \right) \mathbf{Z}_2(t) \\ &= \boldsymbol{\sigma}^T \mathbf{W} \mathbf{P}_2^T \mathbf{Z}_2(t) + \frac{1}{2}\mathbf{Z}_1(t)^T \mathbf{S} \mathbf{Z}_2(t) + \frac{1}{2}\mathbf{Z}_2(t)^T \mathbf{S}^T \mathbf{Z}_1(t). && (29) \end{aligned}$$

Observe that $\mathbf{Z}_1^T(0), \mathbf{Z}_2^T(0)$ are vectors of independent standard Gaussian variables and hence have the same distributions as $\mathbf{X}^T(0), \mathbf{Y}^T(0)$, respectively. Furthermore, equations (17), (19) and (28), (29), defining processes $\eta(t), \dot{\eta}(t)$, have the same structure. Consequently, $K(s_1, s_2)$ can be written in the alternative form

$$K(s_1, s_2) = -\frac{1}{2} \ln(\det(\mathbf{I} - \mathbf{A})) + \frac{1}{2} \mathbf{t}^T (\mathbf{I} - \mathbf{A})^{-1} \mathbf{t}, \tag{30}$$

where now

$$\begin{aligned} \mathbf{A}(s_1, s_2) &= \begin{bmatrix} s_1 \mathbf{\Lambda} & s_2 \mathbf{S} \\ s_2 \mathbf{S}^T & s_1 \mathbf{\Gamma} \end{bmatrix}, \qquad \mathbf{t}(s_1, s_2) = \begin{bmatrix} s_1 \mathbf{x} \\ s_2 \mathbf{y} \end{bmatrix}, && (31) \\ \mathbf{S} &= \mathbf{\Lambda} \mathbf{P}_1 \mathbf{W} \mathbf{P}_2^T - \mathbf{P}_1 \mathbf{W} \mathbf{P}_2^T \mathbf{\Gamma}, \quad \mathbf{x} = \mathbf{P}_1 \boldsymbol{\sigma}, \quad \mathbf{y} = \mathbf{P}_2 \mathbf{W} \boldsymbol{\sigma}. && (32) \end{aligned}$$

Obviously, (30) can be derived by means of matrix algebra, i.e. there is no need to introduce the processes $\mathbf{Z}_1(t), \mathbf{Z}_2(t)$. However, we found the representation (28) useful to propose an approximation $\eta_{ap}(t)$, say, of $\eta(t)$. The processes $\mathbf{Z}_1(t)$ and $\mathbf{Z}_2(t)$ will also be used to investigate the error caused by the approximation. We turn now to the definition of $\eta_{ap}(t)$.

For most sea spectra, a considerable number of the eigenvalues $\lambda_j$ and $\gamma_j$ are very close to zero. We propose to replace them by zeros and use (28), (29) to define $\eta_{ap}(t)$. More precisely, from this point on assume that the $m$ smallest eigenvalues $\lambda_j$ and $\gamma_j$ are replaced by zeros and let

$$\eta_{ap}(t) = (\mathbf{P_1}\boldsymbol{\sigma})^T\mathbf{Z}_1(t) + \frac{1}{2}\mathbf{Z}_1(t)^T\tilde{\mathbf{\Lambda}}\mathbf{Z}_1(t) + \frac{1}{2}\mathbf{Z}_2(t)^T\tilde{\mathbf{\Gamma}}\mathbf{Z}_2(t), \tag{33}$$

$$\dot{\eta}_{ap}(t) = \boldsymbol{\sigma}^T\mathbf{W}\mathbf{P}_2^T\mathbf{Z}_2(t) + \mathbf{Z}_1(t)^T\tilde{\mathbf{S}}\mathbf{Z}_2(t), \tag{34}$$

where $\tilde{\mathbf{\Lambda}}$, $\tilde{\mathbf{\Gamma}}$ are the matrices $\mathbf{\Lambda}$, $\mathbf{\Gamma}$ with the first $m$ rows replaced by zeros, and $\tilde{\mathbf{S}} = \tilde{\mathbf{\Lambda}}\mathbf{P}_1\mathbf{W}\mathbf{P}_2^T - \mathbf{P}_1\mathbf{W}\mathbf{P}_2^T\tilde{\mathbf{\Gamma}}$. Since

$$V(\eta(0)) - V(\eta_{ap}(0)) = 2(\lambda_1^2 + \ldots + \lambda_m^2 + \gamma_1^2 + \ldots + \gamma_m^2), \tag{35}$$

$m$ can be chosen so that the standard deviation of $\eta_{ap}(t)$ is only a fixed percentage, for example 0.1%, smaller than the standard deviation of $\eta(t)$.

We turn now to the estimation of crossing intensity $\mu(u)$ of $\eta_{ap}(t)$ using the saddle-point method. Denote by $K_{ap}(s_1, s_2)$ the cumulant generating function of $\eta_{ap}(0), \dot{\eta}_{ap}(0)$. Obviously, $K_{ap}(s_1, s_2)$ is given by formula (30) with $\mathbf{A}$ matrix replaced by

$$\tilde{\mathbf{A}}(s_1, s_2) = \begin{bmatrix} s_1\tilde{\mathbf{\Lambda}} & s_2\tilde{\mathbf{S}} \\ s_2\tilde{\mathbf{S}}^T & s_1\tilde{\mathbf{\Gamma}} \end{bmatrix}.$$

The matrix $\tilde{\mathbf{A}}$ is $(2N, 2N)$-dimensional but it contains blocks of zeros. Using this block structure of $\tilde{\mathbf{A}}$, we derive the following, numerically more convenient, expression for $K_{ap}$:

$$K_{ap}(s_1, s_2) = -\frac{1}{2}\ln(\det(\mathbf{B})) + s_1^2\frac{\sum_{j=1}^m x_j^2}{2} + s_2^2\frac{\sum_{j=1}^m y_j^2}{2} + \frac{1}{2}\mathbf{r}^T\mathbf{B}^{-1}\mathbf{r}, \tag{36}$$

where $\mathbf{B}$ is a $(2(N-m), 2(N-m))$ dimensional matrix and $\mathbf{r}$ is a $2(N-m)$ column, both dependent on $s_1, s_2$. The expressions for $\mathbf{B}$ and $\mathbf{r}$ are given in the appendix, see (37), while $\mathbf{x}$, $\mathbf{y}$ are defined by (32).

# 5 Example

In this example we shall compare the crossing intensity and other wave characteristics, estimated for the linear and for the second order Gaussian sea. We shall also investigate the accuracy of the saddlepoint method. In all examples the linear Gaussian sea, $\eta_l$, has continuous one sided spectrum of JONSWAP type. The spectrum is a parametric formula that was derived in the JOint North Sea Wave Project carried out, during 1968 and 1969, in the North Sea; see Hasselmann et al. (1973). This spectrum is often used for deep water conditions. The unknown constants in the spectrum can be computed from the sea state parameters (significant wave height $H_s$ and peak period $T_P$), and the peak-shape parameter $\gamma$. For this example we have considered $H_S = 7$ [m], $T_P = 11$ [sec] and $\gamma = 2.385$. The $\eta_l$ process is approximated by $\eta_l^N$, where the cut off frequency $\omega_c = 3$ [rad/sec] and $N = 257$; see Remark 3.

In Figure 1 (left) we compare the crossing intensity estimated for the second order Gaussian sea $\eta$ computed using the saddlepoint method (SPM), with $m = 231$ eigenvalues replaced by zeros ($N - m = 26$), and the numerical integration (NM). The irregular line is the observed $\mu(u)$ in a simulated sample of $\eta(t)$, $t \in [0, T]$, with $T = 3600$ [sec] and sampling frequency 5 [Hz]. Except the region around the level zero the agreement is good. Observe that the sadlepoint density $\hat{f}(u, y)$ usually does not integrate to one (here the integral is equal 0.92) consequently the crossing intensity estimated using SPM can not be accurate in the whole support. Next we present the accuracy of the approximation in the upper tail, see Figure 1 (right). The agreement is good even there. The crossing intensity computed for the Gaussian part, see (1), is presented in the figure using the dashed-dotted line. As we mentioned before, the use of linear approximation for the sea underestimates frequencies of crossings of high levels. For example the 10-hours wave is underestimated (by means of Gaussian model) by approximately one meter.

To speed up computations, as referred in Section 4.1, we suggest to replace the $m$ smallest $\lambda_j$ and $\gamma_j$ eigenvalues by zeros; introducing $\eta_{ap}(t)$. In the example we have replaced $m = 231$ eigenvalues $\lambda_j$ and equally many $\gamma_j$ by zeros. (The 13 highest (in absolute value) eigenvalues

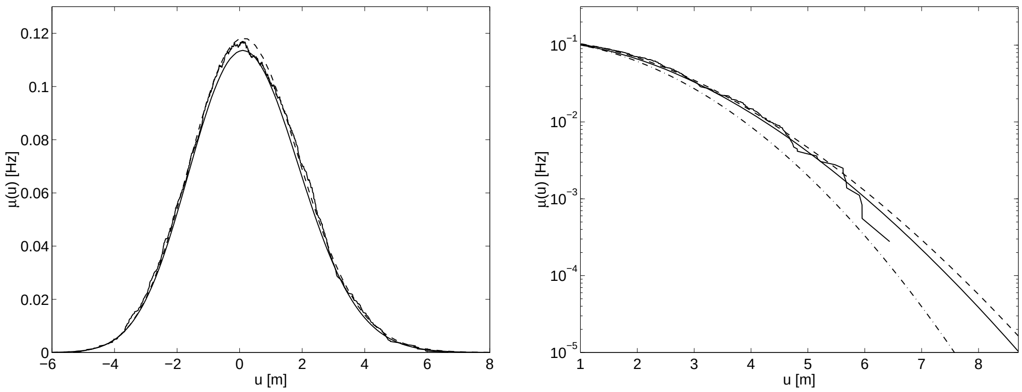


Figure 1: Left: Comparison of crossing intensity $\mu(u)$ estimated from simulated sample path of the second order Gaussian sea $\eta$ (irregular line) and computed using SPM (solid line), NM (dashed line). Right: The tails of the densities in logarithmic scale together with $\mu(u)$ computed for the linear Gaussian sea $\eta_l$ (dash-dotted line).

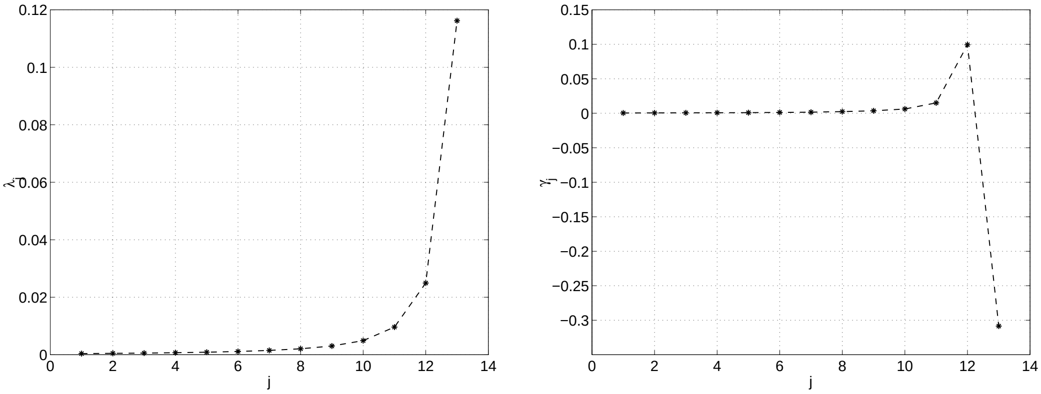


Figure 2: Highest (in absolute value) 13 values of $\lambda_j$ (left) and $\gamma_j$ (right).

are given in Figure 2.) This reduced the dimension of matrices, that have to be inverted, from $(514, 514)$ to $(52, 52)$ without noticeably affecting the accuracy of the SPM method, see Table 1 where we compare $\mu(u)$ computed for $N - m = 100$, 25, 10, 3, 2 and 1. In the table the computed intensities are denoted by $\mu_{N-m}(u)$. It seems that, in this example, only few eigenvalues contributes most to the nonlinear correction of the Gaussian sea. For the Gaussian sea $\eta_l$ the crossing intensities $\mu(u)$, $u = 0, 2, 4, 6, 8, 12$ are 0.1178, 0.0613, 0.0083, $3.301\,10^{-4}$, $3.314\,10^{-6}$, $7.259\,10^{-12}$, respectively.

Table 1: $\mu(u)$ estimates for different $m$ values

| $u$ | $\mu_{100}(u)$ | $\mu_{25}(u)$ | $\mu_{10}(u)$ | $\mu_3(u)$ | $\mu_2(u)$ | $\mu_1(u)$ | $\mu_{NM}(u)$ |
|---|---|---|---|---|---|---|---|
| 0 | 0.1142 | 0.1142 | 0.1142 | 0.1143 | 0.1143 | 0.1139 | 0.1181 |
| 2 | 0.0662 | 0.0661 | 0.0659 | 0.0657 | 0.0636 | 0.0651 | 0.0688 |
| 4 | 0.0131 | 0.0130 | 0.0130 | 0.0129 | 0.0122 | 0.0125 | 0.0141 |
| 6 | 1.043e-03 | 1.039e-03 | 1.033e-03 | 1.018e-03 | 9.544e-04 | 9.611e-04 | 1.270e-03 |
| 8 | 3.860e-05 | 3.845e-05 | 3.816e-05 | 3.746e-05 | 3.484e-05 | 3.432e-05 | 5.696e-05 |
| 12 | 7.885e-09 | 7.844e-09 | 7.765e-09 | 7.550e-09 | 6.940e-09 | 6.477e-09 | 1.708e-08 |

Another measure of the approximation error introduced by replacing some eigenvalues by zeros is $V(\eta(0)) - V(\eta_{ap}(0))$ see (35). In Table 2 the difference between the variances are given for different $m$-values. Even here we can see that almost all eigenvalues could be replaced by zeros without significantly affecting the variance of the process $\eta_{ap}$.

Table 2: Approximation error (35)

| $N - m$ | 100 | 25 | 10 | 3 | 2 | 1 |
|---|---|---|---|---|---|---|
| $V(\eta(0)) - V(\eta_{ap}(0))$ | 2.175e-09 | 4.090e-07 | 5.569e-06 | 2.152e-04 | 8.555e-04 | 0.022 |

In the following subsections we shall illustrate the different applications of the crossing intensity discussed in Section 2. In all examples $m = 231$.

## 5.1 Exceedance probability, "Rice's method"

The most common (and important) application of the crossing intensity $\mu(u)$ is to bound the distribution of the size of the highest wave crest $M_T$, during a specified period of time $T$; see (3). The bound is accurate for high levels $u$ and long time periods $T$. Here we consider the levels above 7 meters and the maximum is taken over the interval $T = 3600$ [sec]. (We assume that the process is stationary for the period of one hour.) Since the probability $P(\eta(0) > u)$ for $u > 7$ is negligible in comparison to $T\mu(u)$ it is omitted in (3). Note that an analytical expression for the distribution of $\eta(0)$ is not known and hence one has to compute it using the saddlepoint- or numerical-method. In Figure 3 the probability to have $M_T$ higher than the level $u$ (in the logarithmic scale) is shown and we can see that the bound is relatively close to the empirical distribution of $M_T$ based on 100 simulated values. Finally we consider also the bound for $M_T$ ($T = 3600$ [sec]) computed for the linear Gaussian sea $\eta_l(t)$ is also presented in Figure 3; the dash-dotted line. We can see that the design values based on the linear model would be ca. 20% lower than the one computed using the nonlinear quadratic model of the sea.

## 5.2 Bound for the distribution of $A_c$

The distribution of heights of crests of individual waves is an important quantity in estimation of fatigue life time (time for creation and growth of cracks in components of offshore structures). The computation of the distribution is a very difficult probabilistic problem related to solving the first passage problem. Here we shall investigate the accuracy of the bound (4) for the crests heights in the quadratic Gaussian sea $\eta$.

From simulated sample path of $\eta$ the probability $P(A_c > h)$ is estimated and presented in Figure 4 as the irregular line. This empirical survival function is compared with the bound (4) with crossing intensity computed using SPM (solid line) and NM (dashed line). We can see that except small crests the bound is close to the estimated survival function.

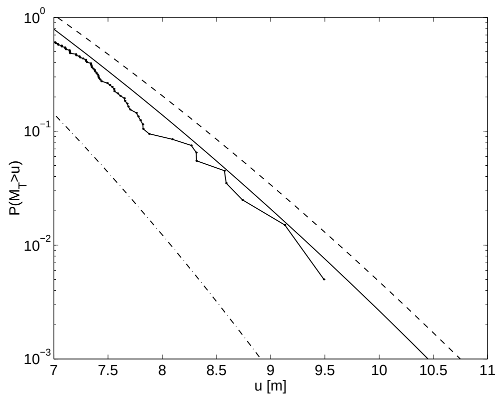


Figure 3: Estimated probability that during one hour the highest crest in $\eta(t)$ exceeds $u$, $P(M_T > u)$, based on 100 simulated values; the irregular solid line, compared with the bound (3) with $\mu(u)$ computed using: SPM (solid line), NM (dashed line). The corresponding bound computed for linear sea $\eta_l$ (the Gaussian sea) dash-dotted line.

## 5.3 Joint distribution of crest period and height

We turn now to the problem of estimating joint distributions of wave characteristics for the non-Gaussian sea. As mentioned in Section 2, one possible approach is to approximate the sea elevation by some simpler process for which one has methods to compute the wave characteristic distributions. Following this idea, we evaluate $g(u)$ defined by (5), with input $\mu(h)$ computed by: SPM (solid line), and NM (dashed line). As we can see from Figure 5 (left), the transformations are almost identical.

We define crest period, $T_c$, as the time between an up-crossing and the next down-crossing while crest $A_c$ is the highest value during the period. In Figure 5 (right) (thin line) the kernel estimate of the $(T_c, A_c)$-density (based on 100000 simulated waves) is presented. Each of the level curves encloses the percentage of waves mentioned in the legend of the plot. For example, if we look to the outer curve; it means that 99.9 % of the simulated waves are inside this curve, and 0.1 % are outside. The estimated density is compared with the computed density of $T_c, A_c$ for transformed Gaussian process, see Podgórski et al (2000)

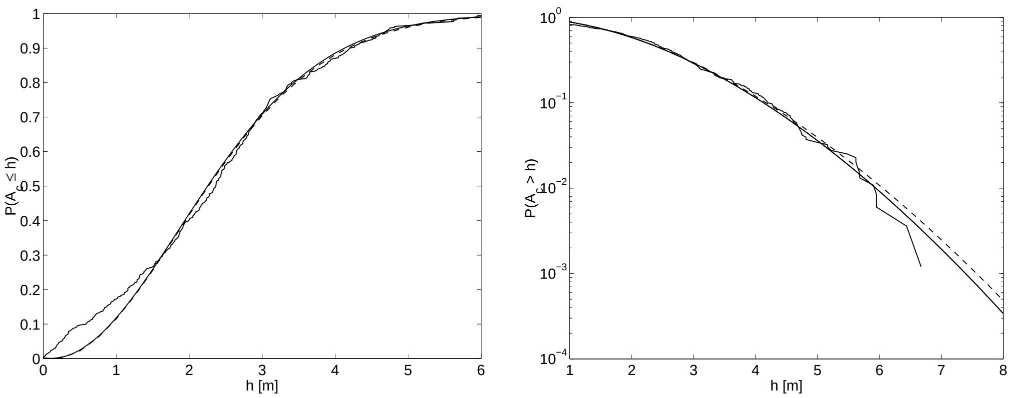


Figure 4: Empirical distribution (left) and exceedance probability (right), of crest height $A_c$ (irregular line) compared to the bound (4) ($P(A_c > h) \leq \mu(h)/\mu(m)$); where $\mu$ is estimated using: SPM (solid line), and NM (dashed line).

for description of the algorithm. The transformation $g$ was derived by means of (5) with $\mu$ computed using the saddlepoint method. The isolines of the density are presented using the bold lines in Figure 5 (right). We can see that the density of $T_c, A_c$ for transformed Gaussian model misses the steepest waves. However, the densities are very similar in all other areas. Since the steep waves are usually the most dangerous our results indicates that further development of tools is needed for prediction of frequencies of steep waves or having other characteristics of interest. All results presented in Figure 5 are computed using WAFO-toolbox which is available free of charge at `http:/www.maths.lth.se/matstat/wafo/`; see Brodtkorb et al. (2000)).


## Acknowledgements

The authors are thankful to Prof. R. Butler for discussion.

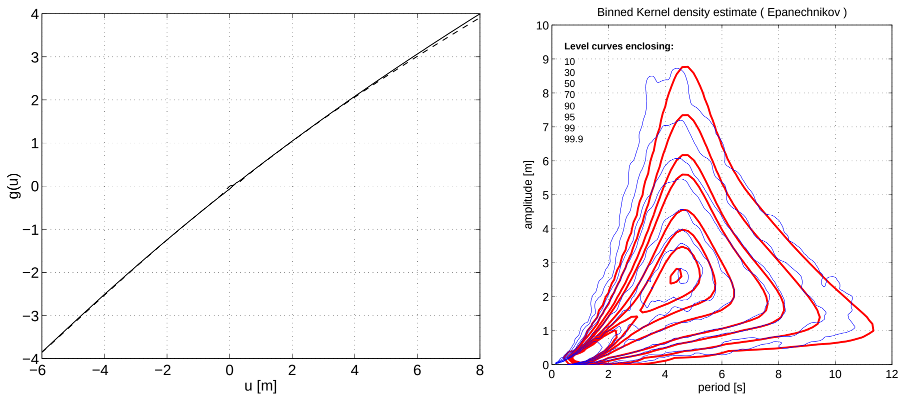


Figure 5: Left: Transformation function $g(u)$ define by (5), where $\mu(u)$ is computed using: SPM (solid line), and NM (dashed line). Right: Isolines of the computed joint density of $T_c$, $A_c$ for the transformed Gaussian process (defined using the transformation from the left figure) - bold line, compared to the kernel estimate of the density based on $10^5$ simulated waves - thin line.

# References


1. Barndoff-Nielsen, O., and Cox, D.R., "Edgeworth and saddle-point approximation with statistical applications," *J. Royal Statist. Soc.* series B 41, 279-312, (1979).

2. Baxevani, A. Podgórski, K., and Rychlik, I., "Velocities for random surfaces," *Preprints in Mathematical Sciences, Lund University* 2002:1, 2002.

3. Brodtkorb, P.A., Johannesson, P., Lingren, G., Rychlik, I., Rydén, J., and Sjö, E., "WAFO - a Matlab toolbox for analysis of random waves and loads," *Proceedings of the 10th International Offshore and Polar Engineering Conference* 3, 343-350, (2000).

4. Cramér, H., *Mathematical Methods of Statistics*, Princeton University Press, 1946.

5. Daniels, H.E., "Saddlepoint approximations in statistics," *Ann. Math. Statist.* 25, 631-50, (1954).

6. Eplett, W.J.R., "Approximation theory for simulations of continuous Gaussian processes," *Prob. Theory Rel. Fields* 73, 159-181, (1986).

7. Hasselmann, K., "On the non-linear energy transfer in a gravity-wave spectrum, Part 1, General Theory," *J. Fluid Mech.* 12, 481-500, (1962).

8. Hasselmann, K., Barnett, T.P., Bouws, E., Carlson, H., Cartwright, D.E., Enke, K., Ewing, J.A., Gienapp, H., Hasselmann, D.E., Kruseman, P., Meerburg, A., Müller, P., Olbers, D.J., Richter, K., Sell, W. and Walden, H., "Measurements of Wind Wave Growth and Swell Decay During the Joint North Sea Wave Project (JONSWAP)," *Deutsches Hydro. Zeit.* Reich A12, 1-95, 1973.

9. Langley, R.S., "A statistical analysis of non-linear random waves," *Ocean Engng.* 14(5), 389-407, (1987).

10. Leadbetter, M.R., Lindgren, G. and Rootzén, H., *Extremes and Related Properties of Random Sequences and Processes*, Springer-Verlag, 1983.

11. Machado, U., "Probability density functions for nonlinear random waves and responses," *Preprints in Mathematical Sciences, Lund University* 2002:2, 2002.

12. Michna, Z. and Rychlik, I., "The expected number of level crossings for certain, symmetric $\alpha$-stable processes," *Stochastic Models* 11.1, 1-19, (1995).

13. Næss, A. and Machado, U., "Response statistics of large compliant offshore structures," *Proceedings of the 8th ASCE Specialty Conference on Probabilistic Mechanics and Structure Reliability*, New York, 2000.

14. Podgórski, K., Rychlik, I. and Machado, U.E.B., "Exact distributions for apparent waves in irregular seas," *Ocean Engng.* 27, 979-1016, (2000).

15. Rice, S.O., "The mathematical analysis of random noise," *Bell Syst. Techn. J.* 23, 282-332, (1944).

16. Rice, S.O., "The mathematical analysis of random noise, II," *Bell Syst. Techn. J.* 24, 46-156, (1945).

17. Rychlik, I., "Regression approximations of wavelength and amplitude distributions," *Adv. Appl. Prob.* 19, 396-430, (1987).

18. Rychlik, I., "On the "narrow-band" approximation for expected fatigue damage," *Probabilistic Engineering Mechanics* 8, 1-4, (1993).

19. Rychlik, I., Johannesson, P. and Leadbetter, M.R., "Modelling and statistical analysis of ocean-wave data using transformed Gaussian processes," *Marine Structures* 10, 13-47, (1997).

20. Rychlik, I. and Leadbetter, M.R., "Analysis of ocean waves by crossing and oscillation intensities," *International Journal of Offshore and Polar Engineering* 10, 282-289, (2000).

21. Rychlik, I., "On some reliability applications of Rice formula for intensity of level crossings," *Extremes* 3:4, 331-348, (2000).

22. Seleznjew, O., "Limit theorems for maxima and crossings of a sequence of Gaussian processes and approximation of random processes," *J. Appl. Prob.* 28, 17-32, (1991).

# Appendix

Consider the formula (36). For each pair $(s_1, s_2)$ one needs to evaluate $\det(\mathbf{I} - \tilde{\mathbf{A}})$ and find $(\mathbf{I} - \tilde{\mathbf{A}})^{-1}$. We shall use the block structure of $\mathbf{I} - \tilde{\mathbf{A}}$ to speed up the computations. Let us define the blocks

$$\tilde{\mathbf{A}} = \begin{bmatrix} s_1\mathbf{A}_{11} & s_1\mathbf{A}_{12} & s_2\mathbf{A}_{13} & s_2\mathbf{A}_{14} \\ s_1\mathbf{A}_{12}^T & s_1\mathbf{A}_{22} & s_2\mathbf{A}_{23} & s_2\mathbf{A}_{24} \\ s_2\mathbf{A}_{13}^T & s_2\mathbf{A}_{23}^T & s_1\mathbf{A}_{33} & s_1\mathbf{A}_{34} \\ s_2\mathbf{A}_{14}^T & s_2\mathbf{A}_{24}^T & s_1\mathbf{A}_{34}^T & s_1\mathbf{A}_{44} \end{bmatrix}, \qquad \mathbf{t}(s_1, s_2) = \begin{bmatrix} s_1\mathbf{x}_1 \\ s_1\mathbf{x}_2 \\ s_2\mathbf{x}_3 \\ s_2\mathbf{x}_4 \end{bmatrix},$$

where $\mathbf{A}_{11}$, $\mathbf{A}_{13}$ have dimensions $(m, m)$, and $\mathbf{A}_{12}$, $\mathbf{A}_{14}$ have dimensions $(m, N-m)$. This defines uniquely the dimensions of the remaining blocks. Next, let $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$, $\mathbf{x}_4$, have dimensions $m, N-m, m, N-m$, respectively. Note that $\mathbf{A}_{12}$ and $\mathbf{A}_{34}$ contains only zeros then introduce

$$\mathbf{C} = \begin{bmatrix} \mathbf{0} & s_2\mathbf{A}_{14} \\ s_2\mathbf{A}_{23}^T & \mathbf{0} \end{bmatrix}, \qquad \mathbf{D} = \begin{bmatrix} s_1\mathbf{A}_{22} & s_2\mathbf{A}_{24} \\ s_2\mathbf{A}_{24}^T & s_1\mathbf{A}_{44} \end{bmatrix},$$

then

$$\mathbf{B}(s_1, s_2) = \mathbf{I} - \mathbf{D} - \mathbf{C}^T\mathbf{C}, \qquad \mathbf{r}(s_1, s_2) = \begin{bmatrix} s_1\mathbf{x}_2 \\ s_2\mathbf{x}_4 \end{bmatrix} - \mathbf{C}^T \begin{bmatrix} s_1\mathbf{x}_1 \\ s_2\mathbf{x}_3 \end{bmatrix}. \tag{37}$$

(Note that matrices $\mathbf{A}_{22}$ and $\mathbf{A}_{44}$ are zero outside main diagonals.) The derivations are based on the following lemma, which was used with $\mathbf{E} = \mathbf{I}$.

**Lemma 6** *Assume that*

$$\mathbf{A} = \begin{bmatrix} \mathbf{E} & \mathbf{C} \\ \mathbf{C}^T & \mathbf{D} \end{bmatrix}$$

*is symmetric and non-degenerate matrix, then*

$$\begin{aligned} \mathbf{t}^T\mathbf{A}^{-1}\mathbf{t} &= [s_1\mathbf{x}^T s_2\mathbf{y}^T] \begin{bmatrix} \mathbf{E} & \mathbf{C} \\ \mathbf{C}^T & \mathbf{D} \end{bmatrix}^{-1} \begin{bmatrix} s_1\mathbf{x} \\ s_2\mathbf{y} \end{bmatrix} = s_1^2\mathbf{x}^T\mathbf{E}^{-1}\mathbf{x} + \\ & [s_2\mathbf{y}^T + s_1(\mathbf{C}^T\mathbf{E}^{-1}\mathbf{x})^T] \left\{\mathbf{D} - \mathbf{C}^T\mathbf{E}^{-1}\mathbf{C}\right\}^{-1} [s_2\mathbf{y} + s_1\mathbf{C}^T\mathbf{E}^{-1}\mathbf{x}]. \end{aligned}$$

*Furthermore* $\det(\mathbf{A}) = \det(\mathbf{E}) \det(\mathbf{D} - \mathbf{C}^T\mathbf{E}^{-1}\mathbf{C})$.