5. LINEAR THEORY

5.1. Probability distributions

5.1.1. Surface elevation

Like many other processes in nature, the surface elevation, according to linear theory, is Gaussian distributed. A theoretical explanation of the applicability of this distribution is the *Central Limit Theorem*. Broadly speaking, it says that the sum of large independent random variables (not necessarily Gaussian distributed) is Gaussian distributed. In linear theory the surface elevation can be considered as the sum of a large number of independent harmonic waves and therefore the above mentioned theorem holds.

$$p(\eta) = \frac{1}{\sqrt{2\pi m_0}} \exp\left(-\frac{\eta^2}{2m_0} \right) \qquad \text{for } E\{\eta\} = 0 \tag{5.1}$$

The spectral zero moment m_0 is the variance of the surface elevation:

$$Var(\eta) = E(\eta^2) = \sum_{i=1}^{\infty} a_i^2 E(\cos(\omega_i t + \alpha_i)) = \sum_{i=1}^{\infty} a_i^2 \frac{1}{2} = \sum_{i=1}^{\infty} 2E(f)\Delta f \frac{1}{2} = m_0 \tag{5.2}$$

in which the definition of wave spectrum has been used (Eq. (4.14)).

5.1.2. Wave crest

Once the surface elevation is considered to be Gaussian distributed, the distributions of the crest and the wave height can be found.

Firstly, one can derive the average time between successive up or down crossings through the level η in terms of the spectrum (Rice, 1945):

$$\overline{T}_\eta = \sqrt{\frac{m_0}{m_2}} \Big/ \exp\left(-\frac{\eta^2}{2m_0} \right) \tag{5.3}$$

From Eq. (5.3) the average time between zero up or down crossings directly follows:

$$\overline{T}_0 = \sqrt{\frac{m_0}{m_2}} \tag{5.4}$$

It is now necessary to make a distinction between narrow and wide spectrums. In Figure 5.1, one can appreciate that the wider the spectrum, the more irregular the character of the surface elevation. In the first case (narrow spectrum), associated to more regular waves, a crest height can be directly defined as a maximum surface elevation per wave.

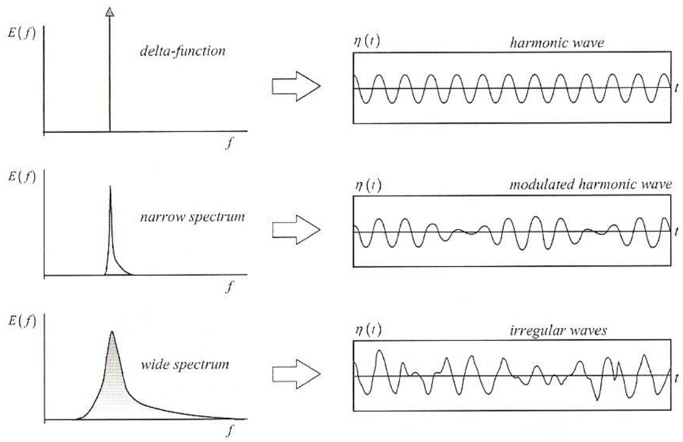

Figure 5.1 Surface elevation for different widths of the spectrum (Holthuijsen,2007)

The relative number of high crests exceeding a certain level η can be derived from Eq. (5.3):

$$\frac{number\ of\ crests\ with\ \underline{\eta}_{crest} > \eta\ in\ duration\ D}{total\ number\ of\ crests\ in\ duration\ D} = \frac{D/\overline{T}_\eta}{D/\overline{T}_0} = \frac{\overline{T}_0}{\overline{T}_\eta} \tag{5.5}$$

This expression is, in fact, the exceedance probability of the crest height:

$$P(\underline{\eta}_{crest} > \eta) = \frac{\overline{T}_0}{\overline{T}_\eta} = \exp\left(-\frac{\eta^2}{2m_0}\right) \tag{5.6}$$

Therefore, the cumulative distribution function (cdf), that is to say, the non exceedance probability is:

$$F(\eta) = P(\underline{\eta}_{crest} < \eta) = 1 - \exp\left(-\frac{\eta^2}{2m_0}\right) \tag{5.7}$$

By deriving Eq. (5.7), the probability density function (pdf) is obtained:

$$p_{\underline{\eta}_{crest}}(\eta) = \frac{\eta}{m_0}\exp\left(-\frac{\eta^2}{2m_0}\right) \tag{5.8}$$

which is the well-known Rayleigh distribution.

Another way of deriving the expression of Eq. (5.8) is using the wave envelope theory. Using the peak frequency (ω_p), one can rewrite the surface elevation as follows:

$$\eta(t) = \mathrm{Re}\left[e^{i\omega_p t}\sum_{i=1}^{\infty} a_n e^{i\left[(\omega_n - \omega_p)t + \alpha_n\right]}\right] \tag{5.9}$$

which can be rewritten as:

$$\eta(t) = \eta_c(t) - \eta_s(t) = A_c(t)\cos(\omega_p t) - A_s(t)\sin(\omega_p t) \tag{5.10}$$

where:

$$\left.\begin{aligned} A_c(t) &= \sum_{i=1}^{\infty} a_i \cos\left[\left(\omega_i - \omega_p\right)t + \alpha_i\right] \\ A_s(t) &= \sum_{i=1}^{\infty} a_i \sin\left[\left(\omega_i - \omega_p\right)t + \alpha_i\right] \end{aligned}\right\} \tag{5.11}$$

Eq. (5.11) can be expressed in terms of the amplitude[3] $A(t)$ and phase $\phi(t)$:

$$\eta(t) = A(t)\cos\left\lfloor\omega_p t + \phi(t)\right\rfloor \tag{5.12}$$

where:

$$A(t) = \sqrt{A_c^2(t) + A_s^2(t)}, \quad \varphi(t) = \arctan\left[A_s(t) / A_c(t)\right]$$

or, equivalent: (5.13)

$$A_c(t) = A(t)\cos\varphi(t), \quad A_s(t) = A(t)\sin\varphi(t)$$

Owing to the narrowness of the process, the amplitudes $A_c(t)$ and $A_s(t)$ vary gradually over time (Massel, 1996) as the spectrum is centred on ω_p. Therefore, $A(t)$, the wave envelope also varies slowly over time and practically equals the wave amplitude as illustrated in Figure 5.2.

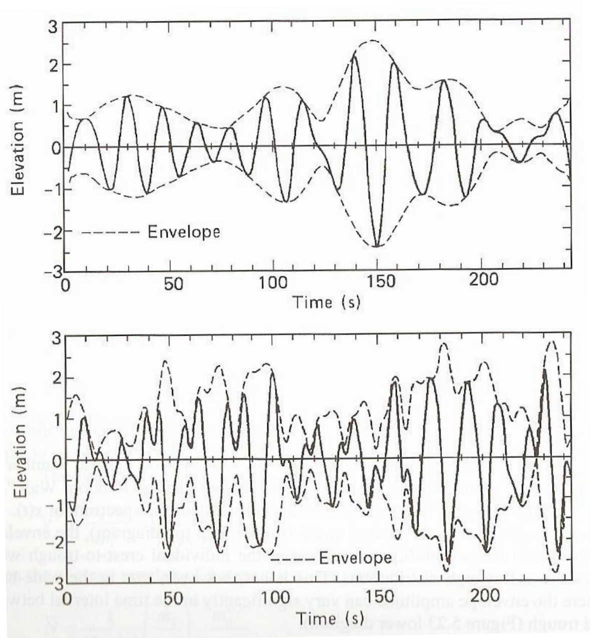

Figure 5.2 Wave envelope for: narrow spectrum (upper), wide spectrum simulated with a Pierson-Moskowitz spectrum (lower); from Carter et al., 1986 (Tucker & Pitt, 2001)

[3] The term *amplitude* can have different meanings. In the spectral analysis, as in more general applications, it is usually associated to the amplitude of the harmonic waves in which the surface elevation is split. However, in most cases of the probability distribution's theories presented in this study, it is related to the amplitude of the theoretical wave envelope which can describe the crest and trough levels of surface elevation and therefore it is a function of time instead of a constant value.

$A_c(t)$ and $A_s(t)$ are independent and Gaussian distributed variables (the Central Limit theorem holds because they are the sum of independent cosine and sine waves respectively). Consequently, the two-dimensional probability density function is:

$$f_2(A_c, A_s) = f(A_c)\cdot f(A_s) = \frac{1}{2\pi\sigma_\eta^2}\exp\left(-\frac{A_c^2(t)+A_s^2(t)}{2\sigma_\eta^2}\right) \tag{5.14}$$

In order to transform $A_c(t)$ and $A_s(t)$ into the amplitude and phase variables, the Jacobian of the variable transformation is used:

$$J = \left|\frac{\partial(A_c, A_s)}{\partial(A,\phi)}\right| = A\cos^2\phi + A\sin^2\phi = A \tag{5.15}$$

and therefore:

$$f_2(A,\phi) = f_2\left[A_c(A,\phi), A_s(A,\phi)\right]J = \frac{A}{2\pi\sigma_\eta^2}\exp\left(-\frac{A^2}{2\sigma_\eta^2}\right) \tag{5.16}$$

Finally, by integrating the joint probability and substituting the variance of the surface elevation for the spectral zero moment, the one-dimensional pdf for the amplitude is found:

$$f(A) = \int_{-\pi}^{\pi} f_2(A,\phi)d\phi = \frac{A}{m_0}\exp\left(-\frac{A^2}{2m_0^2}\right) \tag{5.17}$$

Eq. (5.17) is the same as Eq. (5.8) (in which the amplitude is expressed in terms of the wave crest). Although not being of particular relevance, by integrating Eq. (5.18) one can find that the phase is Uniformly distributed between 0 and 2π:

$$f(\phi) = \int_0^{\infty} f_2(A,\phi)dA = \frac{1}{2\pi} \tag{5.18}$$

For a wide spectrum the total number of crests higher than a certain level η differs from the number of down-crossings through this level as one considers only the maximum wave crest per wave, not all the local maxima. From another viewpoint, the envelope does not slowly vary over time and differs from the wave amplitude (see Figure 5.2). Hence, the above mentioned results cannot theoretically be applied.

However, if we consider the crest height as the maximum crest height per wave, observations have shown that this is nearly Rayleigh distributed.

5.1.3. Wave height

From the engineering view point, the wave height is preferred to the crest height. In deep water and under the assumption of a narrow-band spectrum, the following approximation can be made:

$$\underline{H} \cong 2\underline{\eta}_{crest} \tag{5.19}$$

and therefore the cdf and pdf of the wave height are respectively:

$$F(H) = P(\underline{H} < H) = 1 - \exp\left(-\frac{H^2}{8m_0}\right) \tag{5.20}$$

$$p_{\underline{H}}(H) = \frac{H}{4m_0}\exp\left(-\frac{H^2}{8m_0}\right) \tag{5.21}$$

According to these results, the only parameter which one needs to know in order to describe the wave height distribution, is the zeroth-order spectral moment, that is to say, the variance of the surface elevation. In order to compare wave height exceedance probabilities of different records (and therefore with different variance) it is convenient to normalize the wave height as: $\hat{H} = H\big/\sqrt{m_0}$ which results in the following cdf:

$$F(\hat{H}) = 1 - \exp\left(-\frac{\hat{H}^2}{8}\right) \tag{5.22}$$

With the wide-band spectrum, which often occurs in reality, the wave height suffers from the same problem as the wave crest. The reason for this is that often there are different sources of wave energy as for example a distant storm or local winds. The wave crest probability is not theoretically derived and, therefore, the wave height probability. However, once more, observations more or less agree with the Rayleigh distribution, although a certain scaling factor is found: the real waves are slightly smaller than the predicted values (see Figure 5.3).

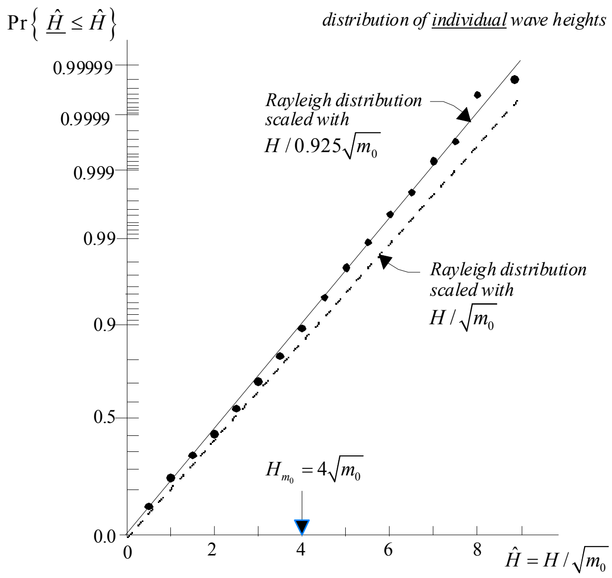

Figure 5.3 Scaling factor between Rayleigh distribution and observations from five hurricanes in the Gulf of Mexico (Forristall, 1978; from Holthuijsen, 2007)

5.1.4. Maximum wave crest/height

The theoretical derivation of extreme values will be done for the wave crest in the case of a narrow banded spectrum. To obtain the associated expression of the wave height only the relationship of Eq. (5.19) should be considered.

In the theory of maximum events, the non-exceedance probability of the maximum value of N, independent and identically distributed random variables, can be written as:

$$F_N(x_{\max}) = [F(x)]^N \tag{5.23}$$

Applying the expression of Eq. (5.23) on Eq. (5.7), the non-exceedance probability (cdf) of the maximum wave crest is found:

$$F_N(\eta_{\max}) = \left[1 - \exp\left(-\frac{\eta^2}{2m_0}\right)\right]^N \tag{5.24}$$

in which N is the number of crests. Then, the pdf is the derivative of Eq. (5.24). If one considers the standardized variable $x = \eta/\sqrt{m_0}$, the cdf becomes:

$$F_N(x_{\max}) = \left[1 - \exp\left(-\frac{x^2}{2}\right)\right]^N \tag{5.25}$$

The *r*th moment, provided it exists, is given by:

$$\begin{aligned} M_r(N) &= \int_{-\infty}^{\infty} x^r dF_N(x_{\max}) = \int_0^{\infty} rx^{r-1}\left(1 - F_N(x_{\max})\right)dx - \int_{-\infty}^{0} rx^{r-1} F_N(x_{\max})dx \\ &= r\int_0^{\infty} x^{r-1}\left(1 - \left(1 - e^{-\frac{1}{2}x^2}\right)^N\right)dx \end{aligned} \tag{5.26}$$

By expanding the above expression binomially, and by integrating term by term, Eq. (5.26) becomes (Cartwright, 1958):

$$M_r(N) = 2^{\frac{1}{2}r}\left(\frac{r}{2}\right)!\left[N - \frac{N(N-1)}{2^{\frac{1}{2}r}\cdot 2!} + \frac{N(N-1)(N-2)}{3^{\frac{1}{2}r}\cdot 3!} \ldots + \frac{(-1)^N}{(N-1)^{\frac{1}{2}r}} - \frac{(-1)^N}{N^{\frac{1}{2}r}}\right] \tag{5.27}$$

which is unsuitable for the calculation for large N. However, if N is large and the exceedance probability ($1 - F(x)$) is small, the following approximation can be made:

$$F(x_{\max}) = [F(x)]^N \approx \exp\left(-N\cdot\left(1 - F(x)\right)\right) \tag{5.28}$$

which is called the asymptotic approximation. Cartwright (1958) quantified the difference between exact and asymptotic formulae for the non exceedance probability $F_N(\eta_{\max})$. Figure 5.4 illustrates such a difference, concluding that for $N > 128$, the difference is practically null.

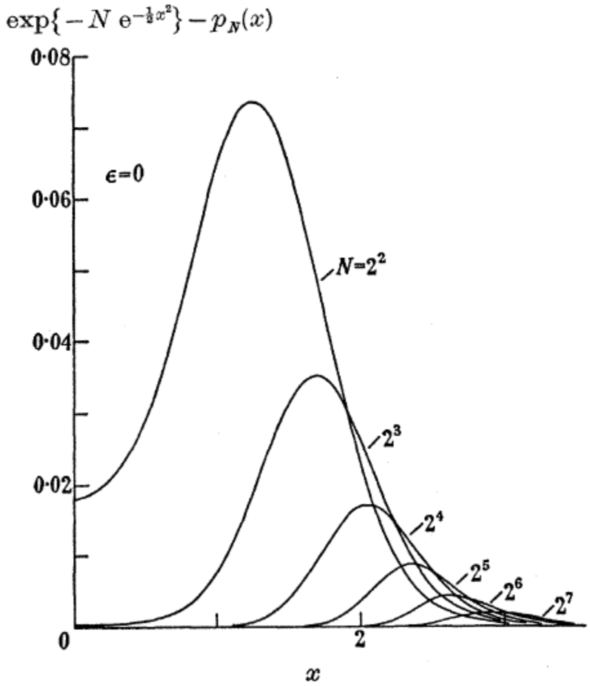

Figure 5.4 Difference between exact and asymptotic formulae for the non-exceedance probability density function (shown as $p_N(x)$) (Cartwright, 1958)

Using this principle the asymptotic formulae for all moments is (Cartwright, 1958):

$$M_r(N) \approx (2\ln N)^{\frac{1}{2}r}\left[1+\frac{1}{2}rA_1(\ln N)^{-1}+\frac{\left(\frac{1}{2}r\right)\left(\frac{1}{2}r-1\right)}{2!}rA_2(\ln N)^{-2}+\ldots\right] \tag{5.29}$$

where:

$$\begin{aligned} A_1 &= 0.5772\ldots, \\ A_2 &= 1.9781\ldots, \end{aligned} \tag{5.30}$$

Eq. (5.29) is correct to an order $(\ln N)^{\frac{1}{2}r} N^{-1}$ and the accuracy therefore increases with increasing N and decreasing *r*. For the expected value of the normalized maximum crest height, $M_1(N)$, the first two terms of Eq. (5.30) are enough: the error is of only 2% for N =20 (Longuet-Higgins, 1952) and rapdly decreases with increasing N:[4]

$$E_N\{x_{\max}\} \approx \left(1+\frac{0.29}{\ln N}\right)\sqrt{2\ln N} \tag{5.31}$$

Another parameter that is sometimes used is the most probable value:

$$\operatorname{mod}_N\{x_{\max}\} \approx \sqrt{2\ln N} \tag{5.32}$$

which represents the value at the maximum in the pdf. For a sample of 200 waves the expected maximum is about 5% higher than the most probable maximum.

[4] E here means expected value rather than E of spectral analysis

5.2. Parameters

In this Section the principal parameters which have been calculated in the study are cited. Most of them relate to the period and the crest/height, which are both calculated directly from the signal and the wave spectrum according to the relationships given in Table 5.1. In order not to be repetitive, the parameters related to the wave crest are not explicitly shown as they are half those of the wave height.

Table 5.1 Principal calculated parameters

Parameter	Statistical	Spectral
T_{mean}	$\overline{T}$	$T_m = m_0/m_1$ $T_0 = \sqrt{m_0/m_2}$
H_{mean}	$\overline{H}$	$\sqrt{2\pi m_0}$
H_{rms}	$\sqrt{\overline{H_i^{\,2}}}$	$\sqrt{8m_0}$
H_s	H1/3	$4\sqrt{m_0}$

5.2.1. The significant wave height

Special attention is paid to the significant wave height (H_s) because it is frequently used in engineering. H_s is defined as the mean of the highest one third wave heights from the record. The reason for choosing one third lies in the fact that with this definition, H_s approximates the visual estimated wave height. Typically, this parameter can be calculated directly from the record or from the spectrum. In order to distinguish them, they are named respectively $H_{1/3}$ and H_{m_0} . The first one is merely the average of the highest one third wave heights. In linear theory, considering the wave height Rayleigh distributed, the second one can be calculated with the following expression:

$$H_{m_0} = 4.004...\sqrt{m_o} \approx 4\sqrt{m_o} \tag{5.33}$$

The derivation of Eq. (5.33) can be summarized in the following steps:

a. Calculation of the threshold wave height $H*$ which has an exceedance probability of 1/3

$$P(H*) = \exp\left(-\frac{H*^2}{8m_0}\right) = \frac{1}{3} \tag{5.34}$$

The threshold is approximately:

$$H* \approx 2.96\sqrt{m_0} \approx 1.048 H_{rms} \tag{5.35}$$

b. Calculation of the mean of all wave heights larger than $H*$

$$H_{m_0} = \frac{\int_{H^*}^{\infty} Hp(H)dH}{\int_{H^*}^{\infty} p(H)dH} = 3\int_{H^*}^{\infty} \frac{H^2}{4m_0}\exp\left(-\frac{H^2}{8m_0}\right)dH \tag{5.36}$$

In fact, considering $H_{1/p,R}$ as the mean of the highest $1/p$ waves and $H^{*}_{p,R}$ the corresponding threshold, generalisations of expressions of Eq. (5.33) and (5.35) are:

$$H_{1/p,R} = \left(2\sqrt{2\ln(p)} + p\sqrt{2\pi}erfc(\sqrt{\ln(p)})\right)\sqrt{m_o} \tag{5.37}$$

$$H_{p,R} = \left(2\sqrt{2\ln(p)}\right)\sqrt{m_o} \tag{5.38}$$

where *erfc* is the complementary error function:

$$erfc(x) = \frac{2}{\sqrt{\pi}}\int_x^{\infty} e^{-t^2}dt \tag{5.39}$$

In the frame of linear theory, $H_{1/3}$ and H_{m_0} are, in mean terms, the same parameter. However, a perfect agreement between them is not found, which manifest the limitations of linear theory. In general, $H_{1/3}$ is 5%-10% lower than H_{m_0} (see Figure 5.5).

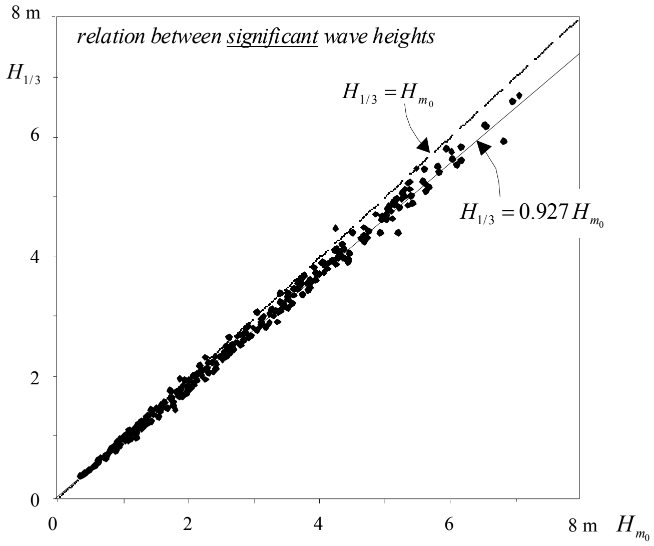

Figure 5.5 The observed and theoretically estimated significant wave height (Holthuijsen, 2007)

Longuet-Higgins (1980) suggested $H_{1/3} = 0.925 H_{m_0}$ for the data of Forristall (1978) data and proposed that a rescaled Rayleigh distribution should be considered, with a significant wave height slightly smaller than the spectral one (see Figure 5.3). Another similar relation has been found by best-fit linear approximation with data from the North Sea: $H_{1/3} = 0.927 H_{m_0}$ (Holthuijsen, 2007) and also by Forristall (1978): $H_{1/3} = 0.942 H_{m_0}$

5.3. Limitations of Rayleigh theory

The Rayleigh distribution for the wave crest/trough assumes: linearity and a narrow-band process. These two basic assumptions are at the same time its limitations, as the reality is more complex than this.

Although considering that nonlinear effects did not occur and therefore the surface elevation had a Gaussian character, if the spectrum is wide, the Rayleigh theory could not estrictly be applied. That may result in Rayleigh's overprediction. If one comes back to Figure 5.2, it is clear that the wave envelope is equal or higher than the wave amplitude and the wider the spectrum, the more difference there will be between them.

Moreover, in the derivation for the wave height distribution, another important assumption is made: the height is double the crest height. In the linear case, although this is true in the mean, this relationship applyed to random variables is not correct. Making this assumption, the mean wave height is not affected but the statistics of higher wave heights are overpredicted. Actually, the wave height distribution should be derived from the sum of the crest and the trough (in absolute value). Taking double crest height, one assums that crests and troughs are always paired as if they were symmetric. In other words, the possibilty of having a low trough preceding a high crest, or viceversa, is not considered and therefore the probability of high wave heights is artificially enhanced.

In conclusion, the Rayleigh distribution may overpredict the crests/troughs due to the wide character of the spectrum.

Moreover, in reality, nonlinear effects can be also present which further complicates the problem. It is difficult to describe nonlinear effects or to attempt to find a distribution which includes most of them. In deep water, the most important are quadruplet wave- wave interactions and wave breaking due to white capping.

In Chapter 6, the Rayleigh-Edgeworth distribution is presented. It accounts for nonlinear effects related to the interaction between waves which involves the kurtosis of the surface elevation. Such nonlinearities enhance wave heights, especialy higher wave heights. However, it does not consider the asymmetry of the surface profile which can take place in the nonlinear case: the crest becomes higher and more peaked and the trough more rounded and less deep. Such asymmetry is included in the theory of Tayfun (1994).

Wave breaking is a very complicated process which is not dealt with in the present study. This type of nonlinear effect produces the diminution of wave heights, possibly counteracting the effect of wave-wave interactions. Tayfun (1981a) suggested that the overprediction could be due to wave heights being steepness limited by the physical process of breaking. However, others (Chen & Borgman, 1979) have showed that measured wave heights and periods are rather far away from the breaking criteria.

In addition, for the calculation of the expected maxima, the asymptotic assumption is made. In Section 5.1.4, it was commented that for more than 128 waves, the error made is negligible. Moreover, independency between waves is assumed. If each wave is highly correlated to the adjacent waves, the effective number of independent waves is reduced. However, analysing 10.000 waves taken in the Bay of Biscay on 21 and 22 May, Cartwirght (1958) found that this correlation was practically negligible.

To sum up, in the real sea, when all possible effects are combined, it is difficult to determine the reason for the discrepancy between observations and linear theory.