# LINEAR WAVE THEORY PART B

**Random waves and wave statistics** 

HARALD E. KROGSTAD

AND

# ØIVIND A. ARNTSEN

# NORWEGIAN UNIVERSITY OF SCIENCE AND TECHNOLOGY TRONDHEIM NORWAY



February 2000



# TABLE OF CONTENTS

# PART B - RANDOM WAVES AND WAVE STATISTICS

1.	INTRODUCTION	1
2.	RANDOM WAVES AND WAVE SPECTRA	3
3.	SEA STATE PARAMETERS AND ENGINEERING WAVE SPECTRA	7
4.	SHORT TERM STATISTICS	15
5.	LONG TERM STATISTICS	19
	5.1 Joint probabilities	19
	<b>5.2</b> Long term distributions of the significant wave height $H_{m0}$	24
6.	EXTREME VALUE ANALYSIS	28
7.	REFERENCES	33
AF	PPENDIX A. CENTRAL CONCEPTS IN PROBABILITY THEORY	35
AF	PPENDIX B WEIBULL SCALED DIAGRAM	37

## **1 INTRODUCTION**

In Part A, the emphasis was on one or a few regular waves specified in terms of the wavelengths, the amplitudes and the phase factors. The dispersion relation defined the angular frequency. However, if we watch the sea on a windy day, the waves are far from being regular, and it is impossible to keep track of, say an individual wave crest for more than a few periods. It thus appears that the theory we have developed will be of little or no use in this situation. Whereas the theory of ocean waves as we know it from Part A was developed in the last century, a usable description of the complex sea surface at the open ocean came much later, basically after the Second World War.

The only waves in the ocean which resemble what we considered in Part A is *swell* generated by a distant storm. Swell can be surprisingly regular, but never pure sine waves. In order to model a sea surface consisting of a wind sea and swell, we must introduce what is commonly referred to as *random waves*. This is a description which, in a way, is not as detailed as the full description of the surface motion we considered in Part A. Nevertheless, linear random wave theory is basically the theory in Part A put in a random or probabilistic setting. Knowledge of elementary probability theory is therefore necessary for the understanding of the theory.

The most important concept for random waves is the *wave spectrum* to be defined below. Within the approximations which are built into linear wave theory itself, the spectrum basically gives us all properties we need about the waves, that is, it defines what is called the sea *state*.

Under normal conditions the wave spectrum and hence the sea state is likely to be constant over, say, half an hour. The properties of the sea for a constant sea state is covered by what is denoted *short term wave statistics*. Short term wave statistics deals with the properties of the individual waves, typically the probability distributions of wavelength, period, height and so on. All these and other concepts will be fully defined and discussed below.

For time periods longer than a few hours, the sea state is likely to vary. Variations in the sea states are covered by a (different) random theory and described by *long term wave statistics*.

For coastal and ocean engineering, it is very important to know how rough conditions the structures are likely to encounter during their lifetime, and this part of the long term statistics is treated by *extreme wave statistics*. Extreme wave statistics provides methods to estimate how rough conditions are likely to happen at a given location over a time span of, say 100 years.

In Part B we shall introduce some of the central topics of random waves starting with the wave spectrum and the main wave parameters. Short and long term statistics are treated briefly, and finally some points of extreme wave statistics are introduced.



# THE RANDOM SURFACE AS A SUPERPOSITION OF PLANE WAVES

Figure 1 . Interpretation of the random sea surface

## **2 RANDOM WAVES AND WAVE SPECTRA**

In the open ocean, it is impossible to keep track on each individual wave. Actually, this does not make much sense even if we had very good instruments! What we need is a manageable way to describe *essential properties* of the wave conditions. Fortunately, it turns out that the linear wave theory we discussed in Part A is a good starting point. In fact, it turns out be a good model of the sea surface to consider it made up from *a lot of regular plane waves*. This model is called *random linear wave theory*. As the name suggests, the theory requires some understanding of probability theory and a very short review of the necessary concepts are introduced in Appendix A.

The first step towards the random model is to consider a surface made up by a finite sum of plane waves

$$\eta(\mathbf{x},t) = \sum_{n=1}^{N} a_n \sin(\boldsymbol{\omega}_n t - \mathbf{k}_n \mathbf{x} + \boldsymbol{\phi}_n)$$

where  $a_n$  is the amplitude,  $\mathbf{k}_n$  and  $\omega_n$  are related by the dispersion relation and  $\phi_n$  is the phase. Recall that this *is* a solution of the linearized equations, cf. Part A, Sec. 8. Whereas the amplitudes and the wavenumbers in the sum are constant characteristics independent of where we are, the phase is more arbitrary. Actually it depends on where we choose our origin. Since it is impossible to give a unique definition of where to put an origin in the open ocean, we simply abandon to specify the phase. In random linear wave theory one thus says that the phase is arbitrary, or *random*. Without going into more details, one assumes that  $\phi_1, \phi_{2,\dots}, \phi_N$  are independent stochastic variables, uniformly distributed on  $[0,2\pi]$ . This means that there is no relation whatsoever between the phase angles of the different waves, and that the phase angle can take any value between 0 and  $2\pi$  with equal probability.

For a given **x** and *t*, set  $\omega t - \mathbf{k}\mathbf{x} = \phi_0$ . Let us compute the expectation value and the variance of the function  $\sin(\phi_0 + \phi)$  where  $\phi$  is uniformly distributed on  $[0, 2\pi]$ :

$$E(\sin(\phi_0 + \phi)) = \frac{1}{2\pi} \int_{\phi=0}^{2\pi} \sin(\phi_0 + \phi) d\phi = 0,$$
  

$$Var(\sin(\phi_0 + \phi)) = \frac{1}{2\pi} \int_{\phi=0}^{2\pi} (\sin(\phi_0 + \phi) - 0)^2 d\phi$$
  

$$= \frac{1}{2\pi} \int_{\phi=0}^{2\pi} \frac{1 - \cos(2(\phi_0 + \phi))}{2} d\phi$$
  

$$= \frac{1}{2}$$

Since we have assumed that  $\phi_1, \phi_2, \dots, \phi_N$  are uniformly distributed random variables, we now obtain by the rules for sums of independent stochastic variables that

$$E(\eta(\mathbf{x},t)) = \sum_{n=1}^{N} a_n E(\sin(\phi_0 + \phi)) = 0,$$
  

$$Var(\eta(\mathbf{x},t)) = \sum_{n=1}^{N} a_n^2 E(\sin^2(\phi_0 + \phi))$$
  

$$= \sum_{n=1}^{N} a_n^2 \frac{1}{2}$$

In practice, instead of a finite number of waves, one assumes that the surface is made up of infinitely many waves, all having different amplitudes and wavenumbers. It is then convenient to write  $a_n = a(\mathbf{k}_n)$  replace the summation

$$\sum_{n=1}^{N} a_n^2 \frac{1}{2}$$

by an integral

$$\sum_{n=1}^{N} a(\mathbf{k}_{n})^{2} \frac{1}{2} \rightarrow \int_{k_{x}=-\infty}^{+\infty} \int_{k_{y}=-\infty}^{+\infty} \Psi(k_{x},k_{y}) dk_{x} dk_{y}$$

The function  $\Psi(k_x, k_y) = \Psi(\mathbf{k})$  signifies the "density of waves" around the wavenumber  $\mathbf{k}$  since the part of the sum to the left having wavenumbers in the area  $dk_x dk_y$  around  $\mathbf{k}$  is just  $\Psi(\mathbf{k}) dk_x dk_y$ . The function may also be called the variance density of the wave field for the obvious reason

$$\operatorname{Var}(\eta(\mathbf{x},t)) = \operatorname{E}((\eta(\mathbf{x},t)-0)^2) = \int_{\mathbf{k}} \Psi(\mathbf{k}) dk^2$$

(Note the shorthand notation for the integral). However, the most common name for  $\Psi(\mathbf{k})$  is the *wavenumber spectrum*. The wavenumber spectrum and the other related spectra discussed below are the most important concepts in random linear wave theory.

Note that  $\Psi(\mathbf{k}) \ge 0$  and that it may be considered to be a kind of "probability density" for the waves. If  $\Psi(\mathbf{k})$  is concentrated around the wavenumber  $\mathbf{k}_0$ , this means that the wavefield is dominated by waves with wavenumber  $\mathbf{k}_0$ , that is, waves going in the direction  $\mathbf{k}_0 / k_0$  with frequency  $\omega_0^2 = gk_0 \tanh(hk_0)$ . Apart from the simple case with a finite number of waves, it is not so obvious how to talk about the amplitude of individual waves in this general setting. Usually  $\Psi$  is considered to be a smooth function of  $\mathbf{k}$ . However, with a finite number of waves,  $\Psi$  will be a sum of "spikes",so-called  $\delta$ -functions. In the following we shall not consider this situation but simply assume that  $\Psi$ is a well-behaved, non-negative function.

An alternative interpretation of the wave spectrum is in terms of the ocean wave surface energy pr. area unit. We recall that the sum of the kinetic and potential energy pr area unit, *E*, was  $\rho g a^2 / 2$  for a regular wave with amplitude *a*. In the present case this generalizes to

$$E = \rho g \operatorname{Var}(\eta) = \rho g \int_{\mathbf{k}} \Psi(\mathbf{k}) d^2 k$$

Thus, the wave spectrum may also be seen as proportional to the wave energy distribution as a function of the wavenumber.

Since  $k(\cos(\theta)\mathbf{i} + \sin(\theta)\mathbf{j})$ , it is possible to write the integral for the variance in (k-) polar coordinates:

$$\operatorname{Var}(\eta) = \int_{\mathbf{k}} \Psi(\mathbf{k}) d^2 k = \int_{k=0}^{\infty} \int_{\theta=0}^{2\pi} \Psi(k,\theta) k dk d\theta$$

Moreover, because of the dispersion relation, we may also change the variable from k to  $\omega$ :

$$\operatorname{Var}(\eta) = \int_{k=0}^{\infty} \int_{\theta=0}^{2\pi} \Psi(k,\theta) k dk d\theta$$
$$= \int_{\omega=0}^{\infty} \int_{\theta=0}^{2\pi} \Psi(k(\omega),\theta) k(\omega) \frac{dk(\omega)}{d\omega} d\omega d\theta$$
$$= \int_{\omega=0}^{\infty} \int_{\theta=0}^{2\pi} E(k,\theta) d\omega d\theta$$

In practice, the function

$$E(k,\theta) = \Psi(k(\omega),\theta)k(\omega)\frac{dk(\omega)}{d\omega}$$

is even more commonly used than  $\Psi$  and is called the *directional wave spectrum*. Again, it that important to note that the integral of E over  $\omega$  and  $\theta$  is equal to the variance of the surface. Note there is no extra  $\omega$  in the integral of E as we had for k in the integral of  $\Psi$ . This is just what has become common. The function E is often written as a product

$$E(k, \theta) = S(\omega)D(\omega, \theta)$$

where the angular dependent part,  $D(\omega, \theta)$ , is normalized such that

$$\int_{\theta=0}^{2\pi} D(\boldsymbol{\omega}, \boldsymbol{\theta}) d\boldsymbol{\theta} = 1 \text{ for all } \boldsymbol{\omega} \ge 0.$$

The function S is called the *wave frequency spectrum* or simply the spectrum. The wave frequency spectrum does not contain any information about the direction of the waves. The directional information is contained in the directional distributions,  $D(\omega, \theta)$ . In general, D will be dependent of  $\omega$ .

Practical wave analysis uses the *frequency*, f, instead of the angular frequency  $\omega$ . The definition  $\omega = 2\pi f$  must then be used in the dispersion relation. If we have a frequency spectrum,  $S(\omega)$ , then the corresponding frequency spectrum,  $\tilde{S}$ , using f instead of  $\omega$ , will be

$$\widetilde{S}(f) = S(\omega = 2\pi f) \frac{d\omega}{df} = 2\pi \cdot S(2\pi f)$$

Note that

$$\int_{f=0}^{\infty} \widetilde{S}(f) df = \int_{\omega=0}^{\infty} S(\omega) d\omega$$

In the following we omit the ~-symbol and assume that the transformation is carried out properly.

For readers familiar with (stationary) random processes the word *spectrum* will be known as the Fourier transform of the correlation function. It is possible to derive the wave spectrum starting from the space-time correlation function of the ocean surface, and an assumption about linear wave theory. However, even very well known textbooks about waves confuse the wave spectrum and the spectrum of a stochastic surface. The wave spectrum as introduced above requires linear wave theory, or more precisely, that we have a well-defined dispersion relation.

We have now introduced the core of linear wave theory. In the random model we sacrifice some details of the description of the waves in that we do not specify the phase of the waves. The basic new concept we have introduced is the wavenumber spectrum  $\Psi$ . The spectrum is needed because we are considering a continuum of plane regular waves, but one should not forget the obvious link to the simpler case with a finite sum of regular waves. Due to the connection with the surface variance, the spectrum is in effect proportional to the distribution of wave energy as a function of wavenumber. Since we consider the wave phases as random variables, the surface becomes what is commonly denoted a random surface. The wave spectrum gives us a complete knowledge of this random surface under one additional assumption, namely when the surface is suppose to be *Gaussian* (There are some technicalities in the concept of a Gaussian random surface which we do not consider at the moment). Real ocean surfaces have turned out to be very closely Gaussian, but there are slight deviations, the most notable is that the wave crests are somewhat higher than the wave troughs are low. This gives the surface a certain asymmetry which is not reflected in the spectrum under normal conditions.

*Exercise 2.1*: Find the surface variance  $(Var(\eta))$  if

$$\eta(\mathbf{x},t) = (1m)\sin(\omega_1 t - \mathbf{k}_1 \mathbf{x} - \phi_1) + (2m)\sin(\omega_2 t - \mathbf{k}_2 \mathbf{x} - \phi_2)$$

and  $\mathbf{k}_1 \neq \mathbf{k}_2$ .

Exercise 2.2: A very simple form of the wavenumber spectrum for wind generated waves is

$$\Psi(\mathbf{k}) = \begin{cases} 0, & |\mathbf{k}| < k_0 \\ A |\mathbf{k}|^{-4} \cos^2(\theta/2), |\mathbf{k}| > k_0 \end{cases}$$

where  $\theta$  is the direction of **k**, i.e.  $\mathbf{k} = k(\cos\theta \mathbf{i} + \sin\theta \mathbf{j})$ . Determine the corresponding directional spectra as functions of  $\omega$  and  $\theta$  in deep and very shallow water.

#### Review questions:

- 1. What is the *wavenumber spectrum* and how do we obtain the variance of the sea surface from the spectrum?
- 2. What is the *sea state*?
- 3. How can we consider the surface as an infinite sum of plane waves?
- 4. How do we move from the wavenumber spectrum to the *directional* spectrum?
- 5. The directional spectrum is often written as a product of two functions, how?

### **3 SEA STATE PARAMETERS AND ENGINEERING WAVE SPECTRA**

We recall that the *sea state* is the condition of the ocean surface considered as a stochastic field and characterized by the wave spectrum. This is the modern use of the word. Traditionally sea state is a scale for the average *wave height* somewhat similar to the Baufourt scale for wind.

The most complete wave spectra we are considering are the wavenumber spectrum  $\Psi(\mathbf{k})$ , and the directional spectrum  $E(\omega, \theta) = S(\omega)D(\theta, \omega)$ . As discussed in the previous section, practical wave analysis uses the frequency f instead of  $\omega = 2\pi f$ . Moreover, engineers tend to use the directional spectrum instead of the wavenumber spectrum. Below we shall follow this convention and write the directional spectrum as  $S(f)D(\theta, f)$ .

The square root of the variance of the surface, is the *standard deviation* of the surface. The standard deviation is a common measure for the variations about the mean and is thus a reasonable scale for the surface height variations. For historical reasons it has become a standard to denote *four times the standard deviation* the *significant wave height*. Significant wave height is written *Hs*, *SWH*, *Hm0* or a number of other possibilities. We shall write *Hm0* for a special reason. Related to the spectrum is a series of characteristic numbers called the *spectral moments*. These numbers, denoted  $m_k$ ,  $k = 0,1,\cdots$  are defined as

$$m_k = \int_{f=0}^{\infty} f^k S(f) df$$

The spectral moment  $m_0$  is just the variance of the surface and hence

$$Hm0 = 4\sqrt{m_0}$$

which is the reason for the notation. The "old" definition of the significant wave height was *the mean of the one third largest waves in the sea*, - a definition which is not very easy to apply! The significant wave height is without doubt the most important sea state parameter. Note that when computed from actual wave measurements, the new definition is about 5% higher than the old definition.

The frequency for which S(f) attains its maximum is called the *peak wave frequency*.

$$S(f_p) = \max_{f} S(f).$$

The inverse of  $f_p$  is also used, and is called the peak period,

$$Tp = \frac{1}{f_p} \; .$$

An equally common period parameter is the mean wave period, Tz, also denoted Tm02:

$$Tm02 = \sqrt{\frac{\int_{f=0}^{\infty} S(f)df}{\int_{f=0}^{\infty} f^2 S(f)df}}$$

We shall return to the most important directional wave parameters after we have discussed some engineering forms for the wave spectra. The wave spectrum for a given sea state may be measured by various wave recording devices and the functions we are going to present below have been found to fit the measurements.

The most well known functional form of S is the so-called Pierson-Moskowitz spectrum

$$S(f) = Ae^{-B/f^4} / f^5.$$

In the original work by Pierson and Moskowitz both A and B were related to the wind speed 19.5 m (!) above the mean sea surface, but today A and B are in some way related to the main sea state parameters. One possibility is to set

$$A = \frac{5}{16} Hm0^2 f_p^4$$
$$B = 5f_p^4 / 4$$

Exercise 3.1: Check that for the Pierson-Moskowitz spectrum with A and B defined as above

$$4\left(\int_{f=0}^{\infty} S(f)df\right) = H$$

Another popular form has been the *JONSWAP spectrum*, which is defined by the somewhat curious expression

$$S(f) = S_{PM}(f) \gamma^{\exp(-(f-f_p)^2/(2\sigma^2 f_p^2))}$$

The JONSWAP spectrum is thus a Pierson-Moskowitz like spectrum multiplied by an extra "peak enhancement factor"

$$\gamma^{\exp\left(-(f-f_p)^2/(2\sigma^2 f_p^2)\right)}$$

Verify that this factor is equal to  $\gamma$  when  $f = f_p$  and approaches 1 for large and small frequencies away from  $f_p$ . JONSWAP was the acronym for a large field experiment in the North Sea in 1973. The measured spectra turned out not look like the Pierson-Moskowitz form and the extra, somewhat artificial factor was introduced in order better to fit the measurements.



Fig. 3.1: The Pierson-Moskowitz spectrum and the standard JONSWAP spectrum



Fig.3.2 Example of a measured frequency spectrum for a heavy sea state. Note the use of a logarithmic vertical axis



Fig. 3.3 Non-dimensional measured spectrum compared to the JONSWAP spectrum.

In many practical situations, the directions of the waves are also of great interest and the directional part of the spectrum is now just at the point of being included in engineering analyses. We recall the directonal spectrum as

$$E(f,\theta) = S(f)D(\theta, f).$$

The directional distribution,  $D(\theta, f)$ , may be understood as the distribution of wave energy for a given frequency f over the directions  $\theta$ . Note that D is generally also a function of f. This reflects the fact that there most of the time exists several systems of waves simultaneously. High frequency wind sea and low frequency swell from different direction is a common case.

Since  $D(\theta, f)$  is a periodic function of  $\theta$  it is convenient to introduce the Fourier series

$$D(\theta, f) = \frac{1}{2\pi} \left( 1 + 2\sum_{n=1}^{\infty} \left\{ a_n(f) \cos n\theta + b_n(f) \sin n\theta \right\} \right).$$

Note that this expression already fulfils already

$$\int_{\theta=0}^{2\pi} D(\theta, f) d\theta = 1$$

It turns out that many of the common instruments that measure directional wave spectra actually measure Fourier coefficients (as a function of the frquency) rather than the directional distribution itself.

The facts that  $D(\theta, f) \ge 0$  and has integral equal to 1 suggests that we may think of D as a kind of probability distribution over *direction*. We recall that the most important parameters for a probability distribution are the mean and the standard deviation. Usually, the stochastic variables we meet are taking integer or real numbers as values. but in the present case, where  $\theta = \theta + n2\pi$ ,  $n = \dots, -1, 0, 1, \dots$  the definitions of the mean and the standard deviations have to be modified. It turns out that the most common definitions, which in the present case are denoted the mean direction and the directional spread for D are defined directly in terms of the Forier coefficients. The mean direction is defined as

$$\theta_1(f) = \operatorname{atan2}(b_1(f), a_1(f)),$$

that is, the direction of the vector with components  $[a_1(f) b_1(f)]$ , or the direction (argument) of the complex number

$$\int_{\theta=0}^{2\pi} e^{i\theta} D(\theta, f) d\theta = \int_{\theta=0}^{2\pi} \{\cos\theta + i\sin\theta\} D(\theta, f) d\theta = a_1(f) + ib_1(f)$$

Note that this also implies that  $\cos \theta_1 = a_1 / (a_1^2 + b_1^2)^{1/2}$  and  $\sin \theta_1 = b_1 / (a_1^2 + b_1^2)^{1/2}$ .

For the directional spread (or rather the directional variance) we would like to have an expression

$$\int_{\theta=0}^{2\pi} (\theta-\theta_1)^2 D(\theta,f) d\theta$$

which, however, is not periodic. One possible choice is instead to use  $2(1 - \cos(\theta - \theta_1))$  which is approximately equal to  $(\theta - \theta_1)^2$  when  $\theta - \theta_1$  small. This gives us

$$\sigma_1^2(f) = 2 \int_{\theta=0}^{2\pi} \{1 - \cos(\theta - \theta_1(f))\} (D(\theta, f) d\theta$$
  
= 2 - 2a\_1(f) \cos \theta\_1(f) - 2b\_1(f) \sin \theta\_1(f) = 2(1 - (a\_1^2(f) + b\_1^2(f))^{1/2}))

which is usually written as  $\sigma_1 = (2(1-r_1))^{1/2}$ ,  $r_1 = (a_1^2 + b_1^2)^{1/2}$  (suppressing the dependence of frequency). In summary, both the mean direction and the directional spread may thus be derived from *the first pair* of Fourier coefficients.

When it comes to actually describing the directional distributions, several functional forms are in use. An often used form is the *cos-2s-distribution* 

$$D(\theta, f) = N(s)\cos^{2s}\left(\frac{\theta - \theta_1}{2}\right)$$

where s(f) and N(s) is a normalization factor such that the integral becomes equal to 1 (s is an integer). It is possible to show that  $N(s) = \frac{1}{\pi} 2^{2s-1} \frac{\Gamma^2(s+1)}{\Gamma(2s+1)}$  where  $\Gamma$  denotes the Gamma function.

 $\Gamma(s+1) = 1 \cdot 2 \cdot 3 \dots (s-1) \cdot s = s!$ 



Fig. 3.4. The cos-2s-distribution for several values of s. s = 1, 3, 12, 20, 40, 80. s = 12 is commonly used as an estimate for practical purposes.

*Exercise 3.2:* Show that the JONSWAP spectrum reduces to the Pierson-Moskowitz spectral form when  $\gamma = 1$ . Also show that the Pierson-Moskowitz spectrum is approximately proportional to  $f^{-5}$  for high values of f.

Exercise 3.3: Consider the directional spectrum

$$S(f)D(\theta, f) = \begin{cases} 0, f < f_0 \\ Af^{-5}\sin^2\theta, f > f_0 \end{cases}$$

Find expressions for *Hm0*, *Tm02*,  $\theta_1$  and  $\sigma_1$  (Hint: Determine S(f) and  $D(\theta, f)$ ).

Exercise 3.3: Consider the following simple directional distribution

$$D(\theta, f) = \begin{cases} 1/(2a) & |\theta| < a \\ 0 & a \le |\theta| < \pi \end{cases}$$

Determine  $\theta_1$  and  $\sigma_1$ .

Exercise 3.4: Consider the directional spectrum

$$S(f)D(\theta, f) = \begin{cases} 0, f < f_0 \\ Af^{-5}\sin^2\theta, f > f_0 \end{cases}$$

Find expressions for *Hm0*, *Tm02*,  $\theta_1$  and  $\sigma_1$  (Hint: Determine S(f) and  $D(\theta, f)$ ).

Review questions:

1. How do we pass from  $\omega$  to f as the frequency variable?

2. How is significant wave height, peak period, mean period, mean direction and directional spread defined?

3. What is the Pierson-Moskowitz spectrum?

# 4. SHORT TERM STATISTICS

The figure shows a typical wave record obtained by a wave staff, a buoy or a similar device. Note how the wave height and wave period are defined!



Fig. 4.1 Definition of wave parameters in a wave record

A recording of the sea surface at a single point gives a series of pairs of H and T, say  $(H_1, T_1), \dots$ 

 $(H_{\rm N}, T_{\rm N})$ . In short term wave statistics, the wave height and wave period are considered as stochastic variables.

To a reasonable accuracy, the cumulative probability distribution for the wave height has been found to follow the Rayleigh probability curve:

$$P(H \le h) = 1 - e^{-2(h/H_{m0})^2}$$
.

Thus, the probability that H is larger than *h* is exp(-  $2(h/H_{m0})2$ ). E.g., the probability that an arbitrary chosen wave has a larger height than  $H_{m0}$  is

$$\exp\left(-2(H_{m0} / H_{m0})^2\right) = e^{-2} \approx 0.14$$
.

Similarly, a wave is larger than 2 · Hm0 with a probability equal to

$$\exp(-2 \cdot 2^2) = e^{-2} \approx 3.3 \cdot 10^{-4}$$

# **Exercise 4.1** (difficult):

Show that the mean value of the 1/3 largest waves,  $H_{1/3}$  is equal to  $1.0011 \cdot \text{Hm0}$  for Rayleigh distributed waves.

The distribution for wave period has been difficult to describe in simple terms. For the larger

waves  $(H > H_{m0})$ , the period is approximately Gaussian with a mean equal to  $1.3 \cdot T_{m02}$  and a standard deviation  $(1.8 / (H/H_{m0})) \cdot T_{m02}$ .





Fig. 4.2 Time series of a heavy sea state recorded in Norway Trænabanken 14 January 1982, 2140 GMT. Significant wave height = 14.8 m.



Fig. 4.3 Nondimensional joint wave height/wave period distribution for heavy sea states  $(H_{m0} > 8.0 \text{ m})$ 

# Review questions: SHORT TERM STATISTICS

- 1. How do we define the wave height and the wave period for a record of the sea elevation taken at a fixed horizontal location?
- 2. How is the Rayleigh distribution for wave heights defined?
- 3. Which distribution can we use for the wave period?

# 5. LONG TERM WAVE STATISTICS

# 5.1 Joint probabilities

Long term wave statistics deals with the statistical properties of wave parameters like significant wave height, peak period etc. Recall that within short term wave statistics, these parameters were considered to be constant. This is therefore a change in viewpoint: We are considering the properties of the sea with respect to long term variations, i.e. seasonal or yearly variations.

A popular form of long term wave statistics is occurrence tables. Such tables describe the overall distribution of the wave parameters over, say, a year. The first table shows how  $H_{m0}$  varies over the year with a much rougher climate during the winter months. The next is a similar table for the mean wave period.

HP40	: 0.0	1.0	2.0: 3.0	3.0: 4.0:	4.0: 5.0:	5.0:	6.0: 7.0:	7.0: 8.0:	8.0: 9.0:	9.0: 10.0:	10.0: 11.0:	11.0: 12.0:	12.0: 13.0:	>	:	N :
JAN	:	6.9	20.2	23.0	20.2	13.0	7.4	5.1	2.1	1.3	0.8	0.2			:	625
FE8	:	11.3	25.9	24.1	16.9	9.2	6.6	3.0	1.9	0.6	0.4		0.2		:	532
MAR	3.0	27.2	28.8	17.9	13.0	6.5	2.5	1.0	0.1		1					-728
APR	1.5	27.1	27.9	23.3	11.0	5.4	2.4	0.7	0.3	0.2	0.1	0.1	:		:	989
MAY	12.9	57.9	22.9	3.9	1.7	0.8	i	÷	÷	ŧ	i	i	1		1	1164
JUN	17.8	55.2	23.3	3.0	0.3	0.4	:	:	1	:	Ī	i	-		1	1217
JUL	20.8	55.3	18.0	4.1	1.7	0.3	1	i	i	÷	i	i			1	1151
AUG	23.0	46.4	23.3	5.3	1.8	0.3	i	i		-		i			÷	797
SEP	2.2	40.5	37.6	12.4	4.6	1.2	0.7	0.5	0.1	0.1		į			1	947
OCT	1.5	14.3	29.3:	22.9	16.0	7.8	4.3	1.6	1.8	0.1	0.2	0.2	:		:	993
NOV	1.4	15.7	27.0	22.4	14.5	8.7	4.8	2.8	2.1	0.6	į				1	1178
DEC	-	10.4	30.2	24.6	14.5	9.3	5.1	3.6	1.1	0.6	0.4	0.1			i	936

MONTHLY DISTRIBUTION OF SIGNIFICANT WAVE HEIGHT, MMO AT HALTEN 1980-85

TH02	0.0:	1.0: 2.0:	2.0:	3.0:	4.0: 5.0:	5.0:	5.0:	8.0:	9.0:	9.0:	10.0:	12.0:	12.0:	>	
JAN :		:	:	:	1.1	10.6	23.4	29.0	22.6	9.6	3.5	0.3			625
FEB		÷	:		4.3	16.4	33.1	24.4	16.4	5.1	0.4				532
MAR			:	0.4	9.8	33.9	27.5	17.2	8.5	2.6	0.1				728
APR		:		0.3	11.7	23.9	34.5	21.7	5.7	1.4	0.8		i		989
MAY		÷	÷	1.2	31.4	40.3	20.4	6.2	0.4						1164
JUN	:	:	:	5.9	36.0	42.0	11.9	3.1	0.7	0.3					1217
วบเ :		:	:	4.5	39.2	38.2	14.1	3.3	0.5	0.2					1151
AUG		:	:	6.6	31.7	36.6	19.1	5.1	0.8						797
SEP :		-	1	0.2	15.6	38.9	26.0	13.3	5.1	1.0					947
OCT :			i	0.4	9.2	24.6	27.1	23.5	11.6	3.2	0.5				993
NOV		1		0.1	3.7	23.8	35.5	21.7	9.8	4.5	0.5	0.3			1170
DEC		i			3.1	15.4	29.9	27.9	16.1	5.6	1.5	0.4	0.1		936

.......

MONTHLY DISTRIBUTION OF AVERAGE WAVE PERIOD, THO2 AT HALTEN 1980-85





For less quantitative relations a scatter plot is also convenient:

The scatter plot shows simultaneous observations of  $H_{m0}$  and  $T_p$ . The corresponding table given below is called a joint occurrence table:



Meteorological parameters are of great interest in ocean engineering, in particular the wind. Joint statistics of meteorological and wave parameters is therefore also common:

	Mont part	hly s pi	fre r. t	ien i	ncy thou	tabl sand	e fo lat	or H Hal	MÖ a teni	and 1 bank	Wind en f	Spe or (	ed Octo	in t ber	:егш 1980	s of 0-85.	
umati (141)	NG (H/   9.0	/S]   2.0	4.0	6.0	10.0	10.0	12.0	14.0		18.0	20.0 22.0	22.9 24.8	24.0	26 -8	38:8	e asc.	TRACTION
	25	61	61	25	12											15	<b>0.01</b>
1 0- 2.5	37	270	319	478	282	233	25									134	8.164
2.0- 3.0	61	355	441	417	515	931	306	98	12							256	0.314
3.0-4.9	49	74	135	221	245	306	613	221	98							160	9.196
4.0- 5.0		61	110	110	172	221	306	184	184	49						114	0.140
5.0- 6.0		12	25	25	49	49	196	233	123	74	12	12				66	0.001
6 0- 7.0					37	49	16	74	98	46	12	12				37	0.045
7.0- 8.0					12		37	37	12	25	25					12	0.015
a 0- 9.0				25			12	12	12	110	37					17	0.021
9.0-10.0													12			1	0.001
0 0-11 0				i					12					12		2	. 0.002
1 0-12.0	1				1							12	12			2	0.002
12 0-13.0				ļ <sup>.</sup>												•	6.000
13.0-14.0																0	0.000
14.0-15.0																0	0.000
A PEC		63	<b>1</b> 9	184	106	146	129	70	45	26	7	3	2	1	•	\$16	
FRACTION:	.017	.083	.109	.130	.132	.179	.154	.086	.055	.034	.909	.004	. 002	.001	. 900		

	OCTOBE	<b>1</b>												
	MDIR I	ÆG		•										
HM0 (M)	0. 30.	30. 60.	60. 90.	90. 120.	120. 150.	150. 180.	180. 210.	210. 240.	240. 270.	270. 300.	300. 330.	330. 360.	# REC.	FRACTION
0.0- 1.0								13	64	26	64	26	15	0.019
1.0- 2.0	357	89	38	26	38	13	115	204	179	217	77	217	123	0.157
2.0- 3.0	523	268	115	102	77	64	128	561	587	128	179	230	232	0.296
3.0- 4.0	306	64	38	140	38	38	191	612	765	140	128	89	200	0.255
4.0- 5.0	242	77		38	13	26	77	472	293	140	128	89	125	0.159
5.0- 6.0	115						38	242	64	13	38	128	50	0.064
6.0- 7.0	51							140	77	26	13	13	25	0.032
7.0- 8.0					1	İ		26	51	26			8	0.010
8.0- 9.0								13	64				6	0.008
9.0-10.0							Ì	}	[				0	0.000
0.0-11.0				}									0	0.000
1.0-12.0							· ·		1				0	0.000
2.0-13.0							}						0	0.000
3.0-14.0	1			1									0	0.000
4.0-15.0								Ì					C	0.000
REC. :	125	39	15	24	13	11	43	179	168	56	49	62	784	
FRACTION:	.159	.050	.019	.031	.017	.014	.055	.228	.214	.071	.063	.079		

A table involving the main wave direction for a certain month is shown in below

The graph on the following page shows the actual variation over a winter season for  $H_{m0}$ ,  $T_p$  and the wave direction at  $T_p$ ,  $\theta_p$ . The solid lines are measurements and the dotted line are daily numerical predictions carried out by a numerical wave model run by the Norwegian Meteorological Institute.

The purpose of long term statistics is to extract and compress the information in such curves in the best possible way!



# 5.2 Long term distributions of the significant wave height $H_{m0}$

The most important parameter is the significant wave height, and so far, the most studied long

term statistics has been for this parameter.

The following table shows a table of Hm0 from Utsira on the south west Norwegian coast.

Class	Interval[m]	Upper limit	ni	$\Sigma n_{\rm i}$	$P_{\rm j}\left(H_{\rm m0}\right) = \mathrm{P}(H_{\rm m0} < h)$
		[m]			-
1	0.0 - 0.49	0.49	29	29	0.00800
2	0.5 - 0.99	0.99	158	187	0.05150
3	1.0 - 1.49	1.49	830	1017	0.28055
4	1.5 - 1.99	1.99	746	1763	0.48634
5	2.0 - 2.49	2.49	626	2389	0.65903
6	2.5 - 2.99	2.99	415	2804	0.77352
7	3.0 - 3.49	3.49	298	3102	0.85572
8	3.5 - 3.99	3.99	189	3291	0.90786
9	4.0 - 4.49	4.49	152	3443	0.94979
10	4.5 - 4.99	4.99	57	3500	0.96552
11	5.0 - 5.49	5.49	44	3544	0.97766
12	5.5 - 5.99	5.99	36	3580	0.98759
13	6.0 - 6.49	6.49	26	3606	0.99476
14	6.5 - 6.99	6.99	11	3617	0.99779
15	7.0 - 7.49	7.49	1	3618	0.99807
16	7.5 - 7.99	7.99	3	3621	0.99890
17	8.0 - 8.49	8.49	0	3621	0.99890
18	8.5 - 8.99	8.99	2	3623	0.99945

Each interval is half a meter and the numbers of observations are seen to vary from class to class.  $\Sigma n_i$  signifies the number of observations up to and including the present class. We can read from the table that 3580 observations out of 3624 are smaller than 6 m. Thus, a fraction of .988 of the observations are smaller than 6 m.

The function

 $F_e(h) = \frac{number of data \le h}{total number of data}$ 

is called the *empirical cumulative distribution function*.

In the books of probability theory it is discussed how this function approaches the cumulative distribution function in the case of a large number of data.

Instead of a table, it is more convenient to have an analytical function fitted to  $F_e$ .

A much used form is the Weibull distribution function.

This function gives the fraction of the cases with a significant wave height smaller than a certain level h:

$$P(H_{m0} \le h) = 1 - \exp\left(-\left(\frac{h - H_0}{H_c - H_0}\right)^{\gamma}\right), h \ge H_0$$

(The notation  $P(H_{m0} \le h)$  means the probability that  $H_{m0}$  is less or equal to *h*.)

 $H_c$ ,  $H_0$  and  $\gamma$  are parameters in the distribution.

By specifying  $H_c$ ,  $H_0$  and  $\gamma$ , we specify the whole function, and this is of course much more convenient than a long table.

WARNING: There is no standard notation here: all parameters may have other names or being mixed up. Always check the definition in the text you are reading.

Let us now consider how we can fit a data set to the function above.

We shall limit ourselves to the case where  $H_0$  is equal to a constant. In practice, several values are tried and the one fitting the data best is used.

Below we assume that  $H_0 = 0$ . Then, if we set  $F(h) = P(H_{m0} \le h)$ ,

$$F(h) = 1 - \exp\left(-\left(\frac{h}{H_c}\right)^{\gamma}\right)$$
$$\log(1 - F(h)) = -\left(\frac{h}{H_c}\right)^{\gamma}$$

$$\log(-\log(1-F(h))) = \gamma(\log h - \log H_c)$$

By plotting  $\log(-\log(1 - F(h)))$  vs  $\log(h)$  this gives a straight line:



The slope of the line is  $\gamma$  and the ordinate axis crossing is -  $\gamma \log(H_c)$ .

For a set of data,  $\log(-\log(1 - F_e)))$  is plotted vs  $\log(h)$  and a suitable straight line is found by regression:



You find a copy of a Weibull scaled diagram in Appendix B.



A series of examples of long term distribution of  $H_{m0}$  are given below.

Exercise 5.1:

Assume that we know  $F_e(h_1)$  and  $F_e(h_2)$ . How can we fit a Weibull distribution with  $H_0 = 0$  through these two points?

Review questions:

- 1) What is long term statistics?
- 2) What is a joint occurrence table and what is a scatter plot?
- 3) What is the Weibull distribution function and how can we express the distribution of significant wave height by the Weibull distribution?

# 6. EXTREME VALUE ANALYSIS

For an offshore structure, a fundamental question is how large waves it will experience during its lifetime. What is the chance it will experience a wave 15 m high, 20 m high, 25 m high etc? Such and similar questions (maximum water levels, largest floods, greatest drought, strongest wind) may be attacked by the methods of extreme value analysis.

This is a large field with may different techniques, but we shall limit ourselves to the question of the maximum wave.

This problem may be split in two:

- 1) What is the largest wave experienced for a given sea sate?
- 2) How do we use the answer of 1) when the sea state varies?

None of the questions are simple to answer. We first need a result from probability theory. Consider the stochastic variable X and N independent outcomes of X:  $X_1,...,X_N$ . Let  $X_{max}$  be largest of  $X_1,...,X_N$ ,  $X_{max} = \max_{i=1,N} X_i$ . The statement that  $X_{max} \le x$  is equivalent to that  $X_1 \le x$ ,  $X_2 \le x,...,X_N \le x$ . By the assumption of independence (see a book on probability!),

$$P(X_{\max} \le x) = P(1 \le x, X_2 \le x, \dots, X_N \le x)$$

$$= P(X_1 \le X) \cdot P(X_2 \le x) \dots P(X_N \le x) = (F_X(x))^N$$

where  $F_X$  is the cumulative distribution function of *X*.

We recall that for the wave height,

$$F_{H}(h) = 1 - \exp\left(-2\left(\frac{h}{Hm0}\right)^{2}\right),$$

(The Rayleigh distribution)

For a given  $H_{m0}$ , the largest of, say, N waves has then the following cumulative distribution function

$$P(\max_{i=1,N} H_i < h) = \left[1 - \exp\left(-2\left(\frac{h}{H_{m0}}\right)^2\right)\right]^N$$

Some computations show that the expectation of the highest wave out of N is

$$E(H_{\max} - \max_{i=1,N} H_i) \approx Hm \left(\sqrt{(\log N)/2} + \frac{0.57}{\sqrt{8\log N}}\right)$$

These relations are approximate; based on a somewhat more accurate distribution for  $H_{\text{H}}$ . The graph in Fig 6.1 shows the variation of  $H_{\text{max}}$  as N varies.





For a sea state lasting A seconds the number of waves is approximately

$$N = \frac{A}{T_{m02}}$$

since  $T_{m02}$  is the average wave period. The probability distribution for the largest wave height during this period is then (approximately)

$$P(H_{\max} < h) = \left[1 - \exp(-2\left(\frac{h}{H_{m0}}\right)^2\right]^N$$

This solves question No 1).

Consider now a sea state "1" and a sea state "2". Exactly as before

$$P(H_{max} < h \text{ during both "1" and "2"})$$
$$= P(H_{max} < h \text{ during "1"}) \cdot P(H_{max} < h \text{ during "2"})$$

$$= \left[1 - \exp\left(-2\left(\frac{h}{H_{m0_1}}\right)^2\right)\right]^{A_1/T_{m2_1}} \cdot \left[1 - \exp\left(-2\left(\frac{h}{H_{m0_2}}\right)^2\right)\right]^{A_2/T_{m02_2}}$$

It is obvious how this generalises as a product involving several different sea states.

Consider finally a joint ( $H_{m0}$ ,  $T_{m02}$ ) long term statistics. Each entry in the table expresses the fraction of the time  $H_{m0}$  and  $T_{m02}$  is in that particular class, i.e.

$$H_{m0} \approx H_{m0}$$
 and  $T_{m02} \approx T_{m02}$ 

for a fraction  $p_{ij}$  of the time. For a time span of A (e.g. 100 years) the sea state is in this class  $p_{ij} \cdot A = a_{ij}$  years. The probability distribution for the highest wave during all these years is therefore

$$P(H_{\max} < h) = \prod_{ij} \left[ 1 - \exp\left(-2\left(\frac{h}{H_{m0_{ij}}}\right)^2 \right]^{a_{ij}/T_{m02_{ij}}} \right]^{a_{ij}/T_{m02_{ij}}}$$

(The symbol  $\Pi_{ij}$  simply means the product over all classes in the joint occurrence table.)

The table below shows a  $H_{m0}$  -  $T_{m02}$  table expressed in 0.001%.

Mean zerocrossing period,  $T_{\pi}$  (s) Significant wave height, 6-7 7-8 0-2 2-3 3-4 4-5 5-6 8-9 9-10 10-11  $H_{S}(\mathbf{m})$ 0.0-0.5 0.5-1.0 5 05 1 1.0-1.5 10 288 115 1.5-2.0 2 2 2 1 2.0-2.5 127 53 94 **36**0 2.5-3.0 3.0-3.5 3.5-4.0 4.0-4.5 65 112 43 4.5-5.0 5.5 - 6.06.0-6 6.5-7.0 7.0 - 7.57 5-8 0 27 49 8.0-8.5 8.5-9.0 9.0-9.5 9.5 - 10.010.0-10.5 10.5-11.0 11.0-11.5 11.5-12.0

Normalized joint occurrence table of  $H_s$  and  $T_z$ . Data from Haltenbanken. 1974–80.

Using the formula above, with a slightly different short term distribution, we obtain the following result:



## Exercise 6.1:

Find the expected maximum wave at a point in the sea if  $H_{m0}$  has been 1 m and  $T_{m02} = 6$  s from the creation of the earth ( $\approx 4.10^9$  years).

## Exercise 6.2:

Two sea states,  $(H_{m0} = 4 \text{ m}, T_{m02} = 10 \text{ s})$  and  $(H_{m0} = 8 \text{ m}, T_{m02} = 12 \text{ s})$ , have both lasted for 12 hours. Determine the expected maximum wave for the heaviest sea state and show that the probability that a larger wave should have occurred during the first sea state is vanishingly small.

## Review questions: EXTREME VALUE ANALYSIS

1) Derive the cumulative distribution function for  $X_{max} = \max(X_1, ..., X_N)$ ,  $P(X_{max} \le x)$ , when the cumulative distribution of all  $X_i$ -s is  $F_X(x)$  and the  $X_i$ -s are independent.

2) How can you compute the expected maximum wave for a sea state with a certain significant wave height and mean period if the sea state persists for "*A*" hours?

# 7 **REFERENCES PART B**

Wetherill, G.B.: Elementary Statistical Methods, 3rd ed. Chapman and Hall, 1982.

Bendat, J.S. and Piersol, A.G.: Random Data: Analysis and Measurement Procedures, Wiley 1971.

Marin Hydrodynamikk (Compendium in Norwegian), Dept. of Structural Engineering, NTNU, Trondheim, Norway, 1982.

### APPENDIX A. CENTRAL CONCEPT IN PROBABILITY THEORY

Warning: This very short introduction is just for refreshing the concepts. Consult a textbook for a much more complete introduction.

The most central concept in probability theory is that of a *random* (or *stochastic*) *variable*. A random variable is actually a function, but contrary to an ordinary function where we put in an argument and compute an output value, we have no knowledge about the input (Some like to think of a *god* putting in the value). All we can see is the output and we have to deduce all properties of the function by observing the output. It turns out that what we have to determine is the *distribution* of the outcomes. For a real number valued random function the distribution function is defined

$$F_X(x) = \operatorname{Prob}(X \le x)$$

where X symbolizes both the function and the values, and "Prob" means the probability (or the fraction of times) that the outcome is less or equal to x. If the outcome is an angle between 0 and  $2\pi$  the

definition is similar. The derivative of distribution function is called the *probability density*,  $f_X(x) = dF_X(x)/dx$ . Note that  $f_X(x) \ge 0$  and  $\int_{x=-\infty}^{\infty} f_X(x)dx = 1$ . For a real number valued random function X the *expectation* is by

$$E(X) = \int_{x=-\infty}^{\infty} x f_X(x) dx$$

and the variance of X,

$$Var(X) = \int_{x = -\infty}^{\infty} (x - E(X))^2 f_X(x) dx = E\left\{ \left( X - E(X) \right)^2 \right\}.$$

If g(X) is a function of X then  $E(g(X)) = \int_{x=-\infty}^{\infty} g(x) f_X(x) dx$ . The square root of the variance is called the standard deviation.

Two stochastic variables have a *joint* distribution function  $F_{X,Y}(x,y) = \operatorname{Prob}(X \le x \cap Y \le y)$  and a joint probability density  $f_{X,Y}(x,y) = \partial^2 F_{X,Y}(x,y) / \partial x \partial y$ . The covariance between two stochastic variables *X* and *Y* is defined as

$$Cov(X,Y) = \int_{x,y=-\infty}^{\infty} (x - E(X))(y - E(Y))f_{X,Y}(x,y)dxdy = E\{(X - E(X))(Y - E(Y))\}.$$

*X* and *Y* are independent if  $f_{X,Y}(x, y) = f_X(x)f_Y(y)$ . Then Cov(X, Y) = 0 and we say that *X* and *Y* are *uncorrelated*. Note also that Cov(X, X) = Var(X).

The most important rules concerning the sums of stochastic variables are

$$E(a_1X_1 + a_2X_2 + \dots + a_nX_n) = a_1E(X_1) + a_2E(X_2) + \dots + a_nE(X_n)$$

and, if the variables are uncorrelated,

$$Var(a_1X_1 + a_2X_2 + \dots + a_nX_n) = a_1^2 Var(X_1) + a_2^2 Var(X_2) + \dots + a_n^2 Var(X_n)$$

Below we shall consider a stochastic phase angle  $\Theta$  uniformly distributed on  $[0,2\pi]$ . This means that its probability density is  $f_{\Theta}(\theta) = 1/(2\pi), 0 \le \theta \le 2\pi$ . Consider now the random function  $g(\Theta) = \cos(\theta_0 + \Theta)$ , By the rules we just have introduced we have

$$E(g(\Theta)) = \frac{1}{2\pi} \int_{\theta=0}^{2\pi} \cos(\theta_0 + \theta) d\theta = 0$$
  
$$Var(g(\Theta)) = \frac{1}{2\pi} \int_{\theta=0}^{2\pi} (\cos(\theta_0 + \theta) - 0)^2 d\theta = \frac{1}{2\pi} \int_{\theta=0}^{2\pi} \cos^2(\theta) d\theta = \frac{1}{2\pi}$$

These expressions were used in Section. 2B



## Appendix B A Wibull scaled diagram