A Statistical Study on the Non-linear Properties of Shoaling Ocean Waves

by
Louise Cervantes

Master of Science Thesis in
Applied and Computational Mathematics

Department of Mathematics
University of Bergen

March 2021

Acknowledgements

First, I would like to thank my supervisors Henrik Kalisch and Francesco Lagona for their guidance throughout my time as a masters student. Your input and patience has been greatly appreciated. Furthermore, I want to thank Henrik Kalisch for handing me this project and express my gratitude for the time you have spent helping me out no matter the time or day of the week.

I would also like to thank associate professor Volker Roeber at the University of Pau and Pays de l'Adour and associate professor Francesco Fedele at Georgia Institute of Technology. Your input regarding the topics of this thesis has been both helpful and valuable.

Lastly, I'd like to thank my friends, family and classmates for their continuous support and encouragement. A special thanks to Haakon Ludvig Ervik for your help and friendship, chaos always makes more sense with you around. You have all been an enormous source of motivation when my own was lacking.

Abstract

Shoaling of ocean waves is studied numerically using a low-dimensional non-linear shoaling model coupled with Monte-Carlo simulations based on the statistical description of ocean waves and wave spectra. It is found that while non-linearity has a minor effect on the wave height, it has a major effect on the shape of the wave. In fact, in shallow water, the instantaneous surface elevation can be described using a Gram-Charlier distribution rather than a Gaussian distribution which is typical of waves in deep water. The positivity conditions of the Gram-Charlier expansion are enforced in a grid search to estimate the parameters of the distribution in a way that ensure a positive-definite distribution. The results are in line with field studies of coastal waves, such as the ARSLOE project [16]. An estimate of the wave spectra in shallow water is also presented for non-linear shoaling waves and results showed a slight shift in the peak frequency in favour of a lower frequency when considering shallow water sea states.

Contents

Notation

The following notations will be used throughout this thesis unless stated otherwise.

Underscored letters ($\underline{\eta}$): Random variables
Bold letters ($\mathbf{u}$): Vectors
∇: Gradient vector
$\mathcal{O}$: Big O notation
δ_{mn}: Kronecker-delta function

Abbreviations

ODE - Ordinary differential equation
PDF - Probability density function
CDF - Cumulative distribution function
SWH - Significant wave height
GC - Gram-Charlier
MLE - Maximum likelihood estimation
MC - Monte Carlo
P-M - Pierson-Moskowitz
LHS - Left hand side
RHS - Right hand side

Chapter 1

Introduction

1.1 Introduction

Various properties of wind-generated waves in coastal regions are significantly different from those in deep water regions. The differences are largely due to the influence of bathymetry, which is more pronounced in shallower water.

In general, deep water waves are considered a Gaussian random process with only minor discrepancies between the observed and theoretical probability density functions. The deviations from the Gaussian model are exhibited by that fact that high crests are observed more frequently than deep troughs [13]. In shallow water, these deviations are more pronounced due to the relative importance of non-linearity in these waves. Indeed irregularities in bathymetry, changes in wave height and wave steepness as the mean water depth decreases towards the shore affect wave properties and their probability distribution as a result. The steepening process near shore causes higher and sharper wave crests and shallower and flatter wave troughs. Under such conditions, the Gaussian model under such conditions is no longer sufficient for describing wave behaviour as it underestimates the higher values and overestimates the lower values of the observed surface elevation. Hence, a non-Gaussian probability density function has to be applied for representing shallow water wave profiles [17].

Previous statistical analyses on the non-Gaussian characteristics of coastal waves include the results of [16] and [17]. In these works, wave records were obtained at a location along the CERC Field Research Facility at Duck North Carolina. These wave records were taken during the growth stage of a storm in the ARSLOE project. The results show that the skewness of the distribution modelling the free surface elevation was the dominant parameter affecting the degree of deviation from the Gaussian model. To account for the skewness, a non-Gaussian probability density function was used to more accurately represent the distribution of the free surface elevation near the shore. The Gram-Charlier probability density function showed good agreement with the histograms of the surface elevation obtained near the shore in both studies.

While the studies mentioned above are based on measurements, the present study embodies a numerical framework for estimating the coastal surface elevation distribution. As will be elaborated on later in this paper, the combination of linear shoaling theory in deep water and non-linear cnoidal theory in shallow waters yields good agreement with the experimental results found in the above studies. In particular, with the approach used in the present paper, the distribution of the free surface elevation is also found to be non-Gaussian and well represented by a Gram-Charlier series.

The wave spectra of non-linear waves in shallow water will also be investigated in this thesis. Previous studies on the wave spectrum include those of [9] and [19]. In the former, it was shown that the crest and trough distributions follow the same Rayleigh distribution for a narrow spectrum if the free surface elevation can be considered a random Gaussian process. In [9], new analytical wave crest and trough distributions were derived to take into second-order effects of waves in deep water. The results were an extension to the work of Boccotti and are valid for the spectrum in deep water with frequencies of finite bandwidth. In this thesis, an estimation of the wave spectra in shallow water for frequencies of finite bandwidth will be presented. The free surface elevation in this case can no longer be considered a random Gaussian process due to non-linear effects and thus, the presented spectra is an estimate for waves approximated by the perturbed Gaussian distribution in the form of a Gram-Charlier expansion.

1.2 Thesis outline

Chapter 2
We begin with some basic wave theory and the formulation of the linear wave problem in terms of the Euler equations along with its periodic solution. The energy balance and wave height determination for linear shoaling processes is also presented. We proceed by presenting the non-linear wave problem again in terms of the Euler equations. The KdV equation, cnoidal wave solution and energy balance are also presented.

Chapter 3
Here we present the random-phase/amplitude model and wave spectrum when considering linear deep water gravity waves. For non-linear shallow water waves, the Gram-Charlier expansion is presented along with imposed conditions to ensure positivity.

Chapter 4
Chapter 4 is given in the form of our submitted paper. We investigate the shoaling of ocean waves numerically using a low-dimensional non-linear shoaling model coupled with Monte-Carlo simulations based on the statistical description of ocean waves and wave spectra.

Chapter 5
In Chapter 5 we carry out a zero-crossing analysis on real time series data and investigate statistical properties of the free surface elevation in shallow water. This is an extension to the experiments carried out in Case 1 of the submitted paper in the form of a comparison against real data.

Chapter 2

Wave theory

As mentioned in chapter 1, the study of deep water waves as they propagate shorewards into shallower regions has become a problem of interest in different fields. The degree to which wave height, as well as other properties, is affected during the shoaling process is, among other things, of particular importance in for example the maintenance of beaches and design of coastal structures [11].

2.1 Linear Theory

Linear wave theory is generally limited to small-slope, small amplitude surface gravity waves. This implies that $a/\lambda \ll 1$ and $a/h \ll 1$, respectively [12]. Here, a is the amplitude, λ is the wavelength and h is the depth. This section will comprise of the formulation of the linear wave problem along with its solution and the theory needed in order to obtain the wave height H of a shoaling wave.

2.1.1 Formulation of the linear wave problem

We begin by denoting the spatial coordinates of the two-dimensional position vector $\mathbf{x}$ to be (x, z) in agreement with a Eulerian description. Here, the x-axis is the direction of wave propagation and the z-axis points vertically. Then, the corresponding components of the velocity vector $\mathbf{u}(\mathbf{x}, t)$ are (u, w). Utilizing the conservation of mass property and also assuming an incompressible fluid layer with negligible changes in density leads to the well known continuity equation

$$\nabla \cdot \mathbf{u} = 0. \tag{2.1}$$

Moreover, assuming the fluid to be irrotational leads to the fluid vorticity $(\boldsymbol{\omega})$ being zero, namely

$$\boldsymbol{\omega} = \nabla \times \mathbf{u} = 0 \tag{2.2}$$

and the existence of a velocity potential $\phi(x, z, t)$ such that

$$\mathbf{u} = \nabla \phi. \tag{2.3}$$

In component form, this implies that

$$u = \frac{\partial \phi}{\partial x} \quad \text{and} \quad w = \frac{\partial \phi}{\partial z}. \tag{2.4}$$

An elliptic partial differential equation called the *Laplace equation* is a direct result of (2.1) and (2.3) and can be written mathematically as

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial z^2} = 0. \tag{2.5}$$

By also considering conservation of momentum, the linearized *Bernoulli equation* can be obtained and written as

$$\frac{\partial \phi}{\partial t} + \frac{P}{\rho} + gz = 0 \tag{2.6}$$

where P is the pressure, ρ is the water density and g is the gravitational acceleration. For the full derivation of (2.6), see [12]. To solve (2.5), boundary conditions need to be formulated. It is convenient to let $f(x, z, t) = 0$ describe the air-water interface and $z = \eta(x, t)$ denote the surface elevation from its undisturbed location $z = 0$. The equation for the surface is then $f(x, z, t) = z - \eta(x, t) = 0$. Three boundary conditions will now be formulated.

The first boundary condition is derived by noting that the bottom of the liquid layer is an impermeable surface. This implies that the velocity normal to the layer should be zero in a way that

$$w = \frac{\partial \phi}{\partial z} = 0 \quad \text{at} \quad z = -h \tag{2.7}$$

which is called the *free slip boundary condition.* At the free surface, particles near the surface should not leave the surface. Mathematically, this requires the fluid velocity normal to the surface be equal to the normal velocity of the surface itself:

$$(\mathbf{n} \cdot \mathbf{u})_{z=\eta} = \mathbf{n} \cdot \mathbf{u}_s \tag{2.8}$$

where $\mathbf{u}_s$ is the fluid velocity normal to the surface. Using the surface equation, the surface normal $\mathbf{n}$ can be written as

$$\mathbf{n} = \frac{\nabla f}{|\nabla f|} = \left(-\frac{\partial \eta}{\partial x}\mathbf{e}_x + \mathbf{e}_z\right) \frac{1}{\sqrt{\frac{\partial \eta}{\partial x}^2 + 1}} \tag{2.9}$$

and considering the velocity of the surface to be purely vertical gives

$$\mathbf{u}_s = \frac{\partial \eta}{\partial t}\mathbf{e}_z. \tag{2.10}$$

Multiplying (2.8) by $|\nabla f|$ leads to

$$\left(-u\frac{\partial \eta}{\partial x} + w\right)_{z=\eta} = \frac{\partial \eta}{\partial t} \tag{2.11}$$

where $\mathbf{u} = u\mathbf{e}_x + w\mathbf{e}_z$ has been used. Utilizing (2.4) and rearranging (2.11) gives

$$\left(\frac{\partial \phi}{\partial z}\right)_{z=\eta} = \frac{\partial \eta}{\partial t} + \frac{\partial \eta}{\partial x}\left(\frac{\partial \phi}{\partial x}\right)_{z=\eta} \tag{2.12}$$

and since we are limited to small-slope, small amplitude surface gravity waves, the non-linear term in (2.12) can be neglected as $\frac{\partial \eta}{\partial x}$ is sufficiently small in comparison. The LHS can then be Taylor expanded around $z = 0$ as an approximation for small slope waves:

$$\left(\frac{\partial \phi}{\partial z}\right)_{z=\eta} = \left(\frac{\partial \phi}{\partial z}\right)_{z=0} + \eta\left(\frac{\partial^2 \phi}{\partial z^2}\right)_{z=0} + ... = \frac{\partial \eta}{\partial t} \tag{2.13}$$

Neglecting the non-linear terms in a similar fashion leads to the linearized *kinematic boundary condition*

$$\frac{\partial \phi}{\partial z} = \frac{\partial \eta}{\partial t} \quad \text{at} \quad z = 0. \tag{2.14}$$

The third boundary conditions defines the water pressure to be equal to the atmospheric pressure so that $(P)_{z=\eta} = 0$, where P is the gauge pressure. Equation (2.6) evaluated at $z = \eta$ then reduces to

$$\left(\frac{\partial \phi}{\partial t}\right)_{z=\eta} = -g\eta. \tag{2.15}$$

Taylor expanding the first term in (2.15) in powers of η about $\eta = 0$ produces

$$\left(\frac{\partial \phi}{\partial t}\right)_{z=\eta} = \left(\frac{\partial \phi}{\partial t}\right)_{z=0} + \eta\left(\frac{\partial^2 \phi}{\partial t^2}\right)_{z=0} + ... = -g\eta \tag{2.16}$$

and neglecting the non-linear terms in the linear approximation as previously done for the kinematic boundary condition gives

$$\frac{\partial \phi}{\partial t} = -g\eta \quad \text{at} \quad z = 0 \tag{2.17}$$

which is the *dynamic boundary condition*. Equations (2.5), (2.7), (2.14) and (2.17) define the linear problem and its solution will be given in the next subsection.

2.1.2 Periodic wave solution

We begin by assuming the surface elevation η takes the form of a simple sinusoidal wave propagating in the positive x-direction

$$\eta(x,t) = a\cos(kx - \omega(k)t) \tag{2.18}$$

where k is the wave number and ω is the frequency. Equation (2.18) requires the velocity potential, ϕ, to be sine dependent and so a sought after solution for ϕ is of the form

$$\phi(x,z,t) = f(z)\sin(kx - \omega(k)t). \tag{2.19}$$

Determining the function $f(z)$ involves using the method of 'separation of variables' and will not be elaborated on here. For a detailed derivation, see [20]. The function $f(z)$ can be found from solving the differential equation

$$\frac{d^2 f(z)}{dz^2} - k^2 f(z) = 0 \tag{2.20}$$

and utilizing the characteristic equation, its solution is of the form

$$f(z) = Ae^{kz} + Be^{-kz}. \tag{2.21}$$

Applying the boundary conditions, A and B can be determined and the function f can be written as

$$f(z) = \frac{a\omega(k)}{k(1 - e^{-2kh})}e^{kz} + \frac{a\omega(k)e^{-2kh}}{k(1 - e^{-2kh})}. \tag{2.22}$$

From this, the following solution for the velocity potential is found

$$\phi(x, z, t) = \frac{a\omega(k)}{k}\frac{\cosh(k(z + h))}{\sinh(kh)}\sin(kx - \omega(k)t) \tag{2.23}$$

and the components of the velocity vector $\mathbf{u}$ can be readily determined:

$$\begin{aligned} u &= a\omega(k)\frac{\cosh(k(z + h)}{\sinh(kh)}\cos(kx - \omega(k)t) \\ v &= a\omega(k)\frac{\sinh(k(z + h)}{\sinh(kh)}\sin(kx - \omega(k)t). \end{aligned} \tag{2.24}$$

Determining the function $\omega(k)$ is done by differentiating (2.3) with respect to t and applying the dynamic boundary condition (2.17) so that

$$\left(\frac{\partial\phi}{\partial t}\right)_{z=0} = -\frac{a\omega(k)^2}{k}\frac{\cosh(kh)}{\sinh(kh)}\cos(kx - \omega(k)t) = -ag\cos(kx - \omega(k)t) \tag{2.25}$$

which, when solved for $\omega(k)$, leads to the well known *dispersion relation*

$$\omega(k) = \sqrt{gk\tanh(kh)} \tag{2.26}$$

that describes the relation between the wave frequency ω and wave number k. The phase speed is then

$$c = \frac{\omega}{k} = \sqrt{\frac{g}{k}\tanh(kh)}. \tag{2.27}$$

In deep water, $kh \to \infty$ s.t $\tanh(kh) \to 1$ and (2.27) becomes the deep water approximation

$$c = \sqrt{\frac{g}{k}} \tag{2.28}$$

In shallow water, $kh \ll 1$, so $\tanh(kh) \approx kh$ and (2.27) becomes the shallow water approximation

$$c = \sqrt{gh}. \tag{2.29}$$

2.1.3 Wave shoaling

In linear shoaling processes, the speed of wave propagation decreases. A consequence of this is the decrease in the kinetic energy of the wave. However, the total energy of a wave consists of both kinetic energy and potential energy which is conserved according to linear theory. A direct result of the decrease in the kinetic energy is then an increase in potential energy, which is found to be directly proportional to the wave height. This change in the wave height can be determined by utilizing the conservative property of the energy flux during the shoaling process. Consider first the energy per unit horizontal area

$$E = \frac{1}{\lambda}\int_0^\lambda \int_{-h}^0 \left[\frac{\rho}{2}|\nabla\phi|^2 + \rho g z\right] dz dx \tag{2.30}$$

and the group velocity c_g

$$c_g = \frac{d\omega}{dk} \tag{2.31}$$

which is the velocity with which the overall envelope shape of the wave propagates. Computing the integrals and substituting the velocity components (2.24) and dispersion relation (2.26) in (2.30) give the following expression for the total energy:

$$E = \frac{1}{8}\rho g H^2 \tag{2.32}$$

while the group velocity c_g can be written as

$$c_g = \frac{c}{2}\left[1 + \frac{2kh}{\sinh(2kh)}\right] \tag{2.33}$$

by differentiating the definition of the phase speed, (2.27). Conservation of the energy flux Ec_g then implies that the wave height H at a current depth is solely determined by the wave height at a previous depth and the respective group velocities at each depth. Namely,

$$H = H_0 \sqrt{\frac{c_{g,0}}{c_g}} \tag{2.34}$$

where the subscript '0' denotes the previous depth. To solve (2.34), the conservative property of the period T can be used in combination with the dispersion relation (2.26), leading to

$$\frac{2\pi}{T} - gk \tanh{(kh)} = 0 \tag{2.35}$$

which is a non-linear equation that can be solved for k. This in turn allows for the determination of H in (2.34).

2.2 Non-linear theory

When waves become too steep or the local depth becomes too shallow, the assumptions of linear theory are no longer satisfied and a new, higher-order framework is required. The Korteweg-de Vries equation is one example of such a framework and has been used with its cnoidal solution to describe wave behaviour during shoaling processes. Previous studies on the shoaling of non-linear waves are [18] and [24] among others. However, in these studies the wave energy density and energy flux are defined in terms of the linear framework. In the following, the non-linear wave problem and its solutions will be presented in terms of the KdV equation as well as an approximate energy balance which follows that given in [3].

2.2.1 Formulation of the non-linear wave problem

We begin, by again considering a 2-D system where (x, z) are chosen so that the x-axis is the direction of wave propagation and the z-axis points vertically upwards. The corresponding components of the velocity vector $\mathbf{u}(\mathbf{x}, t)$ are then (u, w). Letting $P(x, z, t)$ be the pressure and $\mathbf{g} = (0, -g)$ be the gravitational force, the surface water-wave problem can be given by the Euler equations

$$\nabla \cdot \mathbf{u} = 0, \tag{2.36}$$

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} = -\frac{1}{\rho}\nabla P - \mathbf{g} \tag{2.37}$$

By considering an incompressible and irrotational flow, the problem can be formulated in terms of the Laplace equation (2.5) for a velocity potential ϕ as shown in section 2.1. The boundary conditions to the problem are then the non-linearized forms of (2.7), (2.14) and (2.17). The complete problem is now given by

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial z^2} = 0 \quad \text{at} \quad -h_0 < z < \eta(x,t), \tag{2.38}$$

$$\frac{\partial \phi}{\partial t} + \frac{1}{2}\left(\left(\frac{\partial \phi}{\partial x}\right)^2 + \left(\frac{\partial \phi}{\partial z}\right)^2\right) + g\eta = 0 \quad \text{at} \quad z = \eta(x,t) \tag{2.39}$$

$$\frac{\partial \eta}{\partial t} + \frac{\partial \phi}{\partial x}\frac{\partial \eta}{\partial x} - \frac{\partial \phi}{\partial z} = 0 \quad \text{at} \quad z = \eta(x,t) \tag{2.40}$$

$$\frac{\partial \phi}{\partial z} = 0 \quad \text{at} \quad z = -h_0 \tag{2.41}$$

2.2.2 The KdV equation and cnoidal wave solutions

The Korteweg-de Vries (KdV) equation can be used to model weakly non-linear and dispersive waves travelling in one direction and is given in dimensional variables by

$$\eta_t + c_0\eta_x + \frac{3}{2}\frac{c_0}{h_0}\eta\eta_x + \frac{c_0 h_0^2}{6}\eta_{xxx} = 0 \tag{2.42}$$

where c_0 now denotes the shallow water approximation of the phase speed defined in (2.29). For waves to be accurately represented by solutions of the KdV equation, it is assumed the waves be of small amplitude and long wavelength relative to the undisturbed depth of the fluid layer. This requires that $\beta = \frac{h_0^2}{\lambda}$ and $\alpha = \frac{a}{h_0}$ are small parameters and $\frac{\alpha}{\beta} = \mathcal{O}(1)$. We begin by assuming the surface elevation η takes the form

$$\eta(x,t) = f(\xi(x,t)) = f(x - ct) \tag{2.43}$$

The KdV equation (2.42) then reduces to an ODE given by

$$\left(1-\frac{c}{c_0}\right)f'+\frac{3}{2}ff'+\frac{h_0^2}{6}f'''=0 \tag{2.44}$$

Integrating (2.44), multiplying with f' and integrating again leads to

$$-\frac{h_0^2}{3}\left(\frac{df}{d\xi}\right)^2=F(f)=f^3+2\left(1-\frac{c}{c_0}\right)f^2+Af+B \tag{2.45}$$

where $A, B \in \mathbb{R}$ are constants of integration. Several solutions to this problem exist and a complete derivation can be found in [20]. Considering only real solutions to the ODE, the function F can be written in terms of 3 distinct roots such that $f_3 < f_2 < f_1$ and

$$F(f)=(f-f_1)(f-f_2)(f-f_3) \tag{2.46}$$

Substitution into (2.45) leads to

$$\frac{df}{d\xi}=\pm\frac{\sqrt{3}}{h_0^2}\sqrt{(f-f_1)(f-f_2)(f-f_3)}. \tag{2.47}$$

from which the implicit solution can be written as

$$\int_{\xi_1}^{\xi}d\xi'=\pm\frac{h_0^2}{\sqrt{3}}=\int_{f_1}^{f(\xi)}\frac{dz}{\left\{(z-f_1)(z-f_2)(z-f_3)\right\}^{\frac{1}{2}}} \tag{2.48}$$

Computing the integral and substituting $z=f_1+(f_2-f_1)\sin^2\theta$ with the Jacobian $\frac{dz}{d\theta}=2(f_2-f_1)\sin\theta\cos\theta$ gives the following expression for ξ

$$\xi=\xi_1\pm\frac{2h_0^2}{\sqrt{3(f_1-f_3)}}\int_0^{\phi(\xi)}\frac{\mathrm{d}\theta}{\left\{1-m\sin^2\theta\right\}^{\frac{1}{2}}} \tag{2.49}$$

where $m=\frac{f_1-f_2}{f_1-f_3}$. The elliptic integral (2.49) has a known solution and satisfies the relation

$$\cos\phi=\mathrm{cn}\left(\frac{\xi-\xi_1\sqrt{3(f_1-f_3)}}{2};m\right) \tag{2.50}$$

in a way that f can be given in terms of ϕ:

$$
\begin{aligned}
f &= f_1 + (f_2 - f_1)\sin^2\phi \\
&= f_1 + (f_2 - f_1)(1 - \cos^2\phi) \\
&= f_2 + (f_1 - f_2)\mathrm{cn}^2\left(\frac{\xi - \xi_1\sqrt{3(f_1 - f_3)}}{2}\right)
\end{aligned} \tag{2.51}
$$

where where cn is the Jacobian elliptic function that gives periodic waves for the modulus $m \in [0, 1)$. This expression comes from utilizing the transformation $z = f_1 + (f_2 - f_1)\sin^2\theta$ and some trigonometric identies. The wave speed c and wave length λ can now be defined as

$$
c = c_0\left(1 + \frac{f_1 + f_2 + f_3}{2h_0}\right) \quad \text{and} \quad \lambda = K(m)\sqrt{\frac{16h_0^3}{3(f_1 - f_3)}} \tag{2.52}
$$

where $K(m)$ is the complete elliptic integral of the first kind. The KdV equation (2.42) can now be given in terms of its stationary solution

$$
\eta(x, t) = f_2 + (f_1 - f_2)\mathrm{cn}^2\left(\sqrt{\frac{3(f_1 - f_3)}{4h_0^3}}(x - ct); m\right) \tag{2.53}
$$

2.2.3 Energy balance

Recall that for waves to be accurately represented by solutions of the KdV equation it is assumed that $\frac{\alpha}{\beta} = \mathcal{O}(1)$. To ensure that the energy conservation is valid to the same order as α and β we consider the following change of variables presented in [11]:

$$
\tilde{x} = \frac{x}{\lambda}, \quad \tilde{z} = \frac{z + h_0}{h_0}, \quad \tilde{\eta} = \frac{\eta}{a}, \quad \tilde{t} = \frac{c_0 t}{\lambda}, \quad \tilde{\phi} = \frac{c_0}{ga\lambda}\phi.
$$

The KdV equation (2.42) in non-dimensional form is then

$$
\tilde{\eta}_{\tilde{t}} + \tilde{\eta}_{\tilde{x}} + \frac{3}{2}\alpha\tilde{\eta}\tilde{\eta}_{\tilde{x}} + \frac{1}{6}\beta\tilde{\eta}_{\tilde{x}\tilde{x}\tilde{x}} = \mathcal{O}\left(\alpha^2, \alpha\beta, \beta^2\right) \tag{2.54}
$$

and the corresponding non-dimensional potential velocity field is given by

$$
\tilde{\phi}_{\tilde{x}}(\tilde{x}, \tilde{z}, \tilde{t}) = \tilde{\eta} + \frac{1}{4}\alpha\tilde{\eta}^2 + \beta\left(\frac{1}{3} - \frac{\tilde{z}^2}{2}\right)\tilde{\eta}_{\tilde{x}\tilde{x}} + \mathcal{O}\left(\alpha^2, \alpha\beta, \beta^2\right) \tag{2.55}
$$

$$\tilde{\phi}_{\tilde{z}}(\tilde{x}, \tilde{z}, \tilde{t}) = -\beta \tilde{z} \tilde{\eta}_{\tilde{x}} + \mathcal{O}\left(\alpha\beta, \beta^2\right). \tag{2.56}$$

By considering the Bernoulli equation within the fluid domain, the dynamic pressure P' can be written as

$$P' = P - P_{atm} + \rho g z = -\rho_t - \frac{\rho}{2}|\nabla\phi|^2 \tag{2.57}$$

or in non-dimensional variables by using the scaling $P' = \rho g a \tilde{P}'$ as

$$\tilde{P}' = \tilde{\eta} + \frac{1}{2}\beta\left(\tilde{z}^2 - 1\right)\tilde{w}_{\tilde{x}\tilde{t}} + \mathcal{O}\left(\alpha\beta, \beta^2\right). \tag{2.58}$$

The total energy balance can be written as

$$\frac{\partial}{\partial t}\int_{-h_0}^{\eta}\left(\frac{1}{2}|\nabla\phi|^2 + g(z + h_0)\right)dz + \frac{\partial}{\partial x}\int_{-h_0}^{\eta}\left(\frac{1}{2}|\nabla\phi|^2 + g(z + h_0) + P\right)\phi_x dz = 0$$

and by assuming the potential energy to be zero when there is no wave motion, the expression above simplifies to

$$\frac{\partial}{\partial t}\left(\int_{-h_0}^{\eta}\frac{1}{2}|\nabla\phi|^2\, dz + \int_{0}^{\eta} gz dz\right) + \frac{\partial}{\partial x}\int_{-h_0}^{\eta}\left(\frac{1}{2}|\nabla\phi|^2 + gz + P\right)\phi_x dz = 0. \tag{2.59}$$

Using non-dimensional variables and integrating with respect to $\tilde{z}$ gives

$$\frac{\partial}{\partial \tilde{t}}\left(\alpha^2\tilde{\eta}^2 + \frac{\alpha^3}{4}\tilde{\eta}^3 + \frac{\alpha^2\beta}{6}\tilde{\eta}\tilde{\eta}_{\tilde{x}\tilde{x}} + \frac{\alpha^2\beta}{6}\tilde{\eta}_{\tilde{x}}^2\right) + \frac{\partial}{\partial \tilde{x}}\left(\alpha^2\tilde{\eta}^2 + \frac{5}{4}\alpha^3\tilde{\eta}^3 + \frac{\alpha^2\beta}{2}\tilde{\eta}\tilde{\eta}_{\tilde{x}\tilde{x}}\right) = \mathcal{O}\left(\alpha^4, \alpha^3\beta, \alpha^2\beta^2\right).$$

Now, in order to be of the same order as α and β, the energy density E in non-dimensional variables must be given by

$$\tilde{E} = \alpha^2\tilde{\eta}^2 + \frac{\alpha^3}{4}\tilde{\eta}^3 + \frac{\alpha^2\beta}{6}\tilde{\eta}\tilde{\eta}_{\tilde{x}\tilde{x}} + \frac{\alpha^2\beta}{6}\tilde{\eta}_{\tilde{x}}^2 \tag{2.60}$$

and thus, the energy flux is given by

$$\tilde{q}_E = \alpha^2\tilde{\eta}^2 + \frac{5}{4}\alpha^3\tilde{\eta}^3 + \frac{\alpha^2\beta}{2}\tilde{\eta}\tilde{\eta}_{\tilde{x}\tilde{x}}. \tag{2.61}$$

Transforming back to dimensional variables by using the scaling $E = c_0^2 h_0 \tilde{E}$ and $q_E = c_0^3 h_0 q_{\tilde{E}}$ gives

$$E = c_0^2\left(\frac{1}{h_0\eta^2} + \frac{1}{4h_0^2}\eta^3 + \frac{h_0}{6}\eta\eta_{xx} + \frac{h_0}{6}\eta_x^2\right) \tag{2.62}$$

and

$$q_E = c_0^3\left(\frac{1}{h_0}\eta^2 + \frac{5}{4h_0^2}\eta^3 + \frac{h_0}{2}\eta\eta_{xx}\right) \tag{2.63}$$

as found in [2].

2.2.4 Wave shoaling

The wave height of a shoaling wave can now be determined by imposing preservation of wave frequency, conservation of mass and conservation of energy [20]. Thus, if the wave motion at a certain water depth h_A is given, the wave height at water depth h was found in [11] to be given by the following equations:

$$\begin{aligned} \frac{c_A}{\lambda_A} &= \frac{c}{\lambda}, \\ \int_0^T q_{E_A} dt &= \int_0^T q_E dt, \\ \int_0^\lambda \eta_A dx &= \int_0^\lambda \eta dx \end{aligned} \tag{2.64}$$

Using the stationary solution of the KdV equation (2.53) with wave speed and wavelength given in (2.52) and also utilizing the energy flux (2.63), a system of three non-linear equations that can be solved for f_1, f_2 and f_3 and the height of a wave at depth h can be determined.

Chapter 3

Statistical theory

Short-term statistical theory to characterize deep water gravity waves is largely based on the assumption that the surface elevation is a stationary Gaussian process. This assumption is usually accurate for wave records of duration between 15 and 30 minutes. In shallow, coastal waters this is however not always the case due to the non-linear nature of the waves described in the previous chapter. This chapter will provide some statistical preliminaries for linear waves before moving on to the more complex, non-linear case.

3.1 Deep water waves

One of the most descriptive and important ways of characterizing the sea surface is in terms of the *wave spectrum*. The aim is to describe the sea surface as a stochastic process (see Appendix A) by characterizing all possible time records that could have been made under the conditions of the actual time record [13]. First, an introduction to the random-phase/amplitude model will be presented before moving onto the wave spectrum approach of characterizing deep water waves.

3.1.1 The random-phase/amplitude model

Consider the surface elevation η as a function of time t at one location. Using Fourier theory, the wave record can be reproduced by considering the surface elevation to be a sum of harmonic wave components and can therefore be approximated mathematically in terms of its *Fourier series* as

$$\eta(t) = \sum_{i=1}^{N} a_i \cos(2\pi f_i t + \alpha_i) \tag{3.1}$$

where a_i, f_i and α_i are the Fourier amplitude, frequency and phase respectively.

If we now consider the surface elevation, amplitude and phase as random variables chosen for each realisation of the wave record, (3.1) becomes

$$\underline{\eta}(t) = \sum_{i=1}^{N} \underline{a}_i \cos(2\pi f_i t + \underline{\alpha}_i) \tag{3.2}$$

and is more formally known as the *random-phase/amplitude model*. Random variables are fully characterized by their respective probability density functions. Here, the phase α_i is uniformly distributed between 0 and 2π at each frequency so that

$$p(\alpha_i) = \frac{1}{2\pi}, \quad 0 < \alpha_i \leq 2\pi \tag{3.3}$$

and the amplitude $\underline{a}_i$ is Rayleigh distributed at each frequency:

$$p(a_i) = \frac{\pi}{2} \frac{a_i}{\mu_i^2} \exp\left(-\frac{\pi a_i^2}{4\mu_i^2}\right) \tag{3.4}$$

where μ_i is the expected values of the amplitude $\mu_i = E\{\underline{a}_i\}$ (see Appendix A). A large set of realizations of $\eta(t)$ can then be constructed for a given amplitude spectrum by drawing a random amplitude a_i and phase α_i from their probability density functions at each frequency and inserting into (3.2). This approach is generally accurate when the problem is linear so that interactions between the wave components are weak and can be neglected.

3.1.2 The wave spectrum

Since wave energy can be proven to be proportional to the variance in linear wave theory [13], the variance provides a useful link between statistical and physical properties. Therefore, it is often beneficial to consider the variance $E\{\frac{1}{2}\underline{a}_i^2\}$ rather than the expectation of the amplitude $E\{\underline{a}_i\}$ as in the random-phase/amplitude model described above.

By distributing the variance of the surface elevation $E\{\frac{1}{2}\underline{a}_i^2\}$ over the frequency interval Δf_i at frequency f_i, we obtain the following definition of the variance density spectrum:

$$S(f_i) = \frac{1}{\Delta f_i} E\left\{\frac{1}{2}\underline{a}_i^2\right\}, \quad \forall f_i. \tag{3.5}$$

A continuous version of (3.5) is obtained by letting the width of the frequency interval Δf_i approach zero and can be written mathematically as

$$S(f) = \lim_{\Delta f \to 0} \frac{1}{\Delta f} E \left\{ \frac{1}{2} \underline{a}^2 \right\}. \tag{3.6}$$

The variance density spectrum given in (3.5) can be directly related to the scaling parameter σ of the Rayleigh distributed Fourier amplitudes and is shown in Appendix A.

The Pierson-Moskowitz spectrum

The Pierson-Moskowitz (P-M) spectrum is a unidirectional spectrum describing waves in fully developed seas and is often used in applications. The underlying assumption is that waves reach a point of equilibrium with the wind if the wind blows steadily over a large area for a sufficient period of time [21]. Now, the spectral formulation for fully developed seas can be given by the following:

$$S(\omega) = \frac{A}{\omega^5} \exp\left(-\frac{B}{\omega^4}\right) \tag{3.7}$$

where ω is the *circular frequency* and A and B are constants that can be defined as

$$A = 0.0081 g^2 \quad \text{and} \quad B = 0.74 \left(\frac{g}{U}\right)^4 \tag{3.8}$$

for the P-M spectrum, where U is the mean wind speed 19.5m above the sea surface and g is the gravitational acceleration. It is often more convenient to consider the significant wave height rather than the wind speed and the following relation for B can be used in place of that in (3.8) [21]:

$$B = \frac{4A}{H_s^2}. \tag{3.9}$$

Figure 3.1 accordingly shows the Pierson-Moskowitz spectrum as a function of significant wave height.

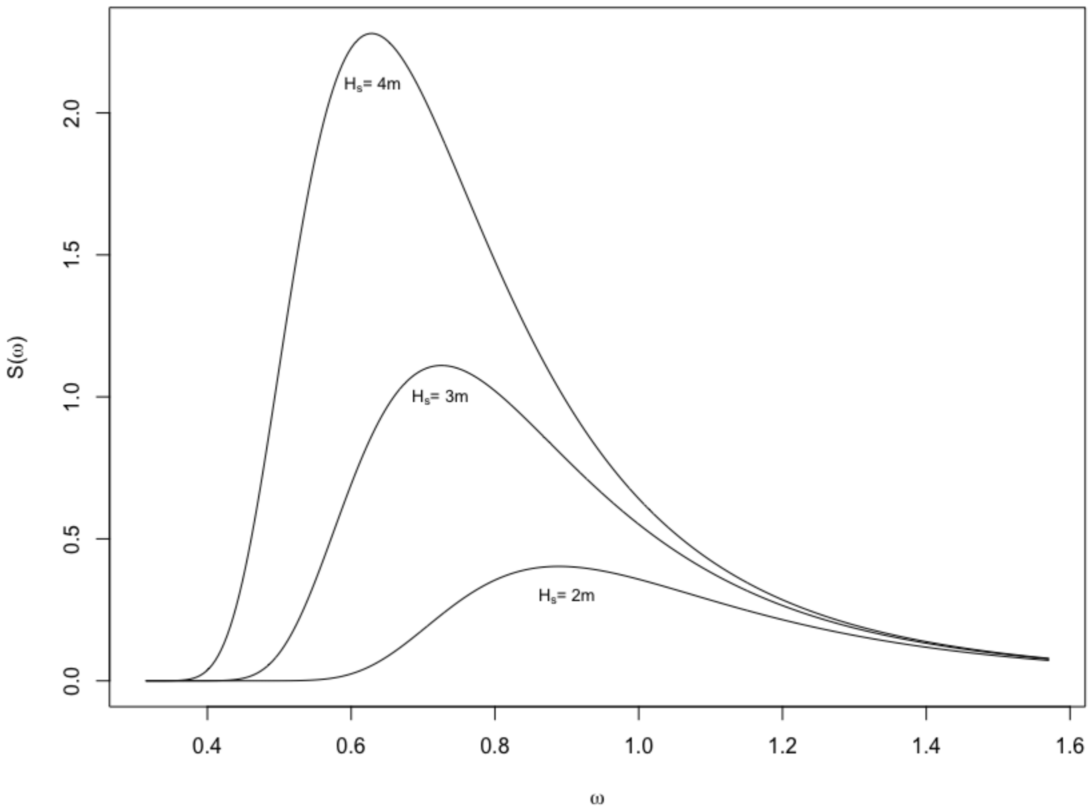


Figure 3.1: Pierson-Moskowitz spectrum as a function of significant wave height H_s

.

3.2 Shallow water waves

Wind-generated waves in coastal regions have been observed to be significantly different from those in deep water regions. The differences are largely due to the effects of shoaling as deep water waves enter shallower water. Therefore, a non-Gaussian probability density function has to be applied to represent shallow water wave profiles. The Gram-Charlier series expansion is a popular approach when it comes to the PDF estimation of wave profiles obtained in coastal regions [17]. The approach is based on the orthogonality of Hermite polynomials with respect to the normal distribution and will be treated in the following.

3.2.1 The Gram-Charlier expansion

Consider first the normal PDF with arbitrary mean μ and variance σ^2 given by

$$p(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) = \frac{1}{\sigma}p\left(\frac{x-\mu}{\sigma}\right). \tag{3.10}$$

For the standardized, zero mean, unit variance normal PDF, (3.10) reduces to

$$p(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right). \tag{3.11}$$

The nth-order Hermite polynomial $H_n(x)$ can then be written in terms of the derivatives of (3.11) as follows:

$$H_n(x) = (-1)^n \frac{d^n p}{dx^n}\frac{1}{p(x)}. \tag{3.12}$$

The polynomials are mutually orthogonal (see Appendix B) with respect to the normal PDF so that

$$\int_{-\infty}^{\infty} H_m(x)H_n(x)p(x)dx = n!\delta_{mn}. \tag{3.13}$$

Now consider a random variable z with unknown PDF $f(z)$. The unknown function $f(z)$ can be approximated in terms of Hermite polynomials, i.e.

$$f(z) = g_n(z)p(z) = \sum_{n=0}^{\infty} c_n H_n(z)p(z) \tag{3.14}$$

where $g_n(z) = \sum_{n=0}^{\infty} c_n H_n(z)$. To find these coefficients, we begin by multiplying both sides of (3.14) by $H_m(z)$ to get

$$f(z)H_m(z) = \sum_{n=0}^{\infty} c_n H_m(z)H_n(z)p(z). \tag{3.15}$$

Integrating from $-\infty$ to ∞ and using the property of orthogonality from (3.13) yields

$$\int_{-\infty}^{\infty} f(z)H_m(z)dz = \int_{-\infty}^{\infty} \sum_{n=0}^{\infty} c_n H_m(z)H_n(z)p(z)dz = c_n n! \tag{3.16}$$

and rearranging to solve for c_n leads to the final expression of

$$c_n = \frac{1}{n!}\int_{-\infty}^{\infty} f(z)H_n(z)dz. \tag{3.17}$$

Letting β_1 and β_2 be the skewness and kurtosis of (3.14), the values

$$c_0 = 1, \quad c_1 = 0 = c_2, \quad c_3 = \frac{\sqrt{\beta_1}}{3!}, \quad c_4 = \frac{(\beta_2 - 3)}{4!} \tag{3.18}$$

are the first 5 coefficients in (3.17), [7]. When z is standardized (zero mean and unit variance), th 4th-order approximation of $g_n(z)$ is

$$g_4(z) = 1 + \frac{\gamma_1}{6}H_3(z) + \frac{\gamma_2}{24}H_4(z) \tag{3.19}$$

where $\gamma_1 = \sqrt{\beta_1}$ and $\gamma_2 = \beta_2 - 3$. This expression is more formally known as the Gram-Charlier Type-A expansion. The Edgeworth expansion is another popular representation [10] and is given by

$$g_6(z) = 1 + \frac{\gamma_1}{6}H_3(z) + \frac{\gamma_2}{24}H_4(z) + \frac{\gamma_1^2}{72}H_6(z). \tag{3.20}$$

Observe that the Edgeworth expansion requires one more Hermite polynomial while keeping the number of parameters constant. Also, when $\gamma_1 = 0 = \gamma_2$, (3.14) reduces to a standard normal distribution.

3.2.2 Positivity conditions of the GC Type A series

One of the major drawbacks of a polynomial approximation is that certain parameters can lead to negative values which is undesirable when considering probability density functions. It is therefore necessary to define the conditions for which the function $f(z)$ is positive-definite.

In general, $f(z)$ in (3.14) is positive-definite when

$$g_n(z) = \sum_{n=0}^{\infty} c_n H_n(z) \geq 0 \quad, \forall z. \tag{3.21}$$

By considering $P = (c_0, ..., c_n)$ to be a point in n-dimensional space, (3.21) requires that P lies on the same side as $(0, ..., 0)$ of the hyperplane

$$\sum_{n=0}^{\infty} c_n H_n(z) = 0, \quad \forall n$$

as described in [7]. This implies that for each $z \in [-\infty, \infty]$, P should lie within the envelope given parametrically by

$$\sum_{n=0}^{\infty} c_n H_n(z) = 0 = \sum_{n=0}^{\infty} c_n n H_{n-1}(z) = 0 \tag{3.22}$$

where (B.4) has been used in the above equality.

For the Gram-Charlier Type-A expansion, we begin by considering $\mathscr{D}$ to be the region in the (γ_1, γ_2)-plane for which $f(z)$ in (3.14) is positive definite. Mathematically, this entails that

$$g_4(z) = 1 + \frac{\gamma_1}{6} H_3(z) + \frac{\gamma_2}{24} H_4(z) \geq 0, \quad \forall z. \tag{3.23}$$

Furthermore, for each value of z the equation

$$g_4(z) = 1 + \frac{\gamma_1}{6} H_3(z) + \frac{\gamma_2}{24} H_4(z) = 0 \tag{3.24}$$

defines a hyperplane in (γ_1, γ_2)-space in the form of a 1-dimensional line. By also considering the derivative of (3.24) given by

$$g_4'(z) = \frac{\gamma_1}{2} H_2(z) + \frac{\gamma_2}{6} H_3(z) = 0, \tag{3.25}$$

it is possible to determine the set of (γ_1, γ_2) as a function of z that satisfies $\mathscr{D}$. The set that satisfies (3.24) and (3.25) simultaneously is called the *envelope* of $p_4(z)$ [10].

Straightforward computations give

$$\gamma_1(z) = -24 \frac{H_3(z)}{h(z)} \quad \text{and} \quad \gamma_2(z) = 72 \frac{H_2(z)}{h(z)}, \tag{3.26}$$

where (3.24) and (3.25) has been solved simultaneously for γ_1 and γ_2 and $h(z) = 4H_3^2(z) - 3H_2(z)H_4(z)$.

To determine the set (γ_2, γ_2), we begin by rewriting $h(z)$ as $h(z) = z^6 - 3z^4 + 9z^2 + 9$ (see Appendix B). Since $h(z) \geq 0$ for all z, the signs of γ_1 and γ_2 depend solely on $H_3(z)$ and $H_2(z)$, respectively. Now, $H_3(z) = z^3 - 3z$ so $\gamma_1 \geq 0$ for $z \in (-\infty, -\sqrt{3}]$ and $z \in [0, \sqrt{3}]$. Similarly, $H_2(z) = z^2 - 1$ so $\gamma_2 \geq 0$ for $z \in (-\infty, -1]$ and $z \in [1, \infty)$.

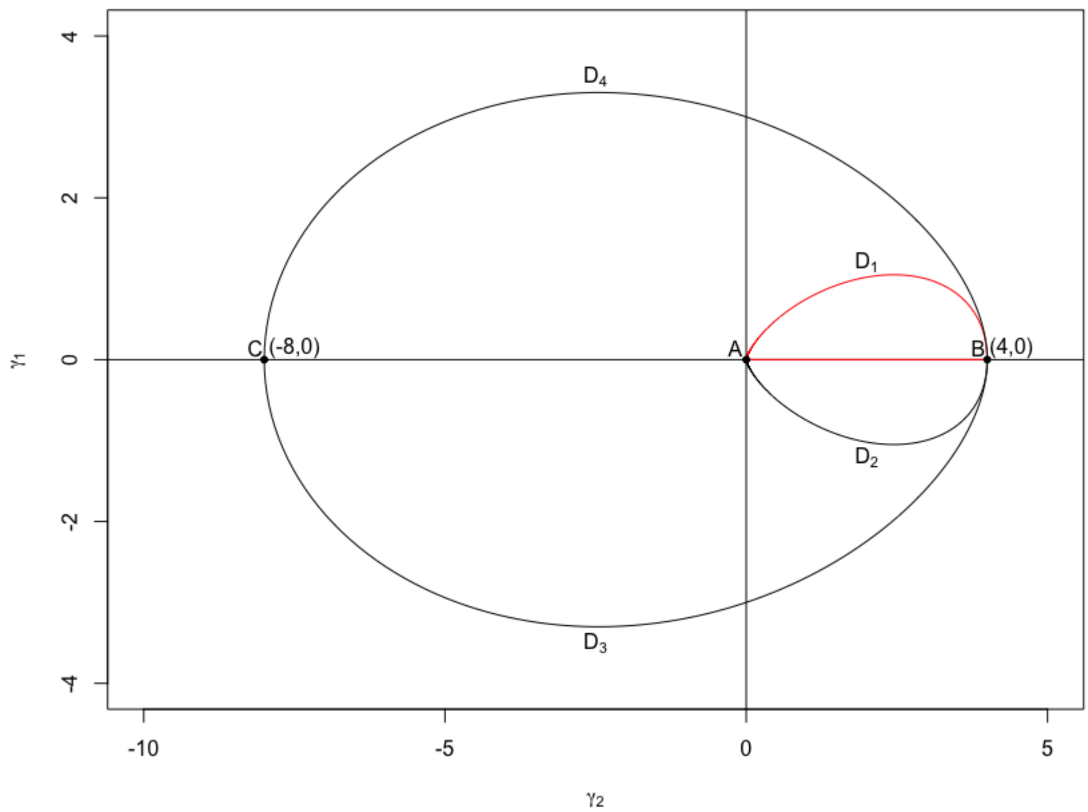


Figure 3.2: Global plot of the envelope of $p_4(z)$. The red lines define the boundary of $\mathscr{D}$.

Figure 3.2 presents the global envelope of $p_4(z)$. The curve AD$_1$B represents the region of the (γ_1, γ_2)-plane when $z \in (-\infty, -\sqrt{3})$. For $z \in [-\sqrt{3}, 0]$ and $z \in [0, \sqrt{3}]$, the curves BD$_3$C and BD$_4$C are obtained. Lastly, the curve AD$_2$B represents the region of the (γ_1, γ_2)-plane when $z \in [\sqrt{3}, \infty)$. It is now clear that $f(z) \geq 0$ for all z when both γ_1 and γ_2 are positive. For the kurtosis, this means that $\gamma_2 \in [0, 4]$. The maximum skewness is obtained when $\left|\gamma_1'(z)\right| = \left|z^4 - 6z^3 + 6z^2 - 18z + 9\right| = 0$ and can be found numerically to be at the point $(2.451, 1.051)$ in the (γ_1, γ_2)-plane. This implies that $\gamma_1 \in [0, 1.051]$. The region $\mathscr{D}$ is then the envelope obtained when

$\gamma_1 \in [0, 1.051]$ and $\gamma_2 \in [0, 4]$ and is shown in Figure 3.2 where the red lines define the boundary of $\mathscr{D}$.

3.3 Maximum likelihood estimation

Maximum likelihood estimation (MLE) is a well known method used to estimate the parameters of a statistical model. The procedure is based on maximizing what is known as the *likelihood function* of the model.

Consider the random sample $\underline{Z} = (z_1, z_2, ..., z_n)$ generated from the PDF $f(z_i; \boldsymbol{\theta})$ where $\boldsymbol{\theta}$ is a k-dimensional parameter vector in the parameter space Ω. Then, the joint probability density of the sample $\underline{Z}$ is

$$f(z_1, z_2, ..., z_n; \boldsymbol{\theta}) = f(z_1; \boldsymbol{\theta}) \cdot f(z_2; \boldsymbol{\theta}) \cdot ... \cdot f(z_n; \boldsymbol{\theta}) = \prod_{i=1}^{n} f(z_i; \boldsymbol{\theta}) \tag{3.27}$$

If we view the joint PDF as a function of $\boldsymbol{\theta}$, (3.27) can be written as

$$L(\boldsymbol{\theta}; z_1, z_2, ..., z_n) = \prod_{i=1}^{n} f(z_i; \boldsymbol{\theta}) \tag{3.28}$$

where $L(\boldsymbol{\theta})$ is the likelihood function. The goal of MLE is to find the values of the model parameters that maximize the likelihood function (3.28) over the parameter space Ω in a way that makes the observed data most probable. We can now define the *maximum likelihood estimator*

$$\hat{\boldsymbol{\theta}} = \underset{\theta \in \Omega}{\operatorname{argmax}} \, L(\boldsymbol{\theta}) \tag{3.29}$$

to be the parameter values that maximize $L(\boldsymbol{\theta})$. Candidates for the maximum likelihood estimator are then all points θ_j such that

$$\frac{\partial L}{\partial \theta_j} = 0, \quad j = 1, ..., k \tag{3.30}$$

By taking the natural logarithm of the likelihood function, the product in the joint density (3.27) can be written as a sum which is more convenient when differentiating. The *log-likelihood* is then readily defined as

$$l(\boldsymbol{\theta}; \underline{Z}) = \log\left(\prod_{i=1}^{n} L(\boldsymbol{\theta}; \underline{Z})\right) = \sum_{i=1}^{n} \log L(\boldsymbol{\theta}; \underline{Z}) \tag{3.31}$$

and candidates for the maximum likelihood estimator are now all points θ_j such that

$$\frac{\partial l}{\partial \theta_j} = 0, \quad j = 1, ..., k \tag{3.32}$$

The monotonic behaviour of the logarithm function ensures that the maximum of $l(\boldsymbol{\theta}))$ occurs at the same values of θ_j as for $L(\boldsymbol{\theta})$ which is a desirable result of the transformation.

3.4 Monte Carlo methods

Monte Carlo (MC) methods are a class of computational algorithms based on repeated stochastic sampling. The underlying concept is to use randomness to solve problems that might be deterministic in principle. Some typical uses of MC methods are estimation, optimization and sampling as described in [8]. This thesis will focus on the latter. MC sampling methods are entirely random in a way that all simulated samples fall within the support of the distribution used. This is achieved by using a pseudo-random number generator which is repeatedly called and returns a real number in $[0, 1]$. The results are then used to generate a distribution of samples that is an accurate representation of the desired probability distribution [1].

For this kind of technique to be effective in representing a random variable with a given distribution, a sufficient number of iterations should be performed. When the sample size is not sufficiently large, the problem of clustering can arise. This is due to the samples tendency to take high probability values when a low number of iterations is used, leaving values in the outer ranges of the distribution unrepresented. Increasing the number of iterations ensures that a larger range of values is covered belonging to both high and low probability occurrences. The effects of both high and low probability outcomes are then accounted for in the simulation and the representation of the desired probability distribution is more accurate.

Since random wave heights can be shown to follow a Rayleigh distribution in deep water [14], one MC simulation results in one realization of the wave height distribution at a specific position in time or space. Each run then yields different results leading to an ensemble of realizations. This ensemble is then a set of plausible realizations of the wave height distribution under the conditions of an actual observation.

Chapter 4

Wave Spectra in Shallow Water Using Cnoidal Theory

In this chapter, our submitted paper is presented.

Wave Spectra in Shallow Water Using Cnoidal Theory

Louise Cervantes[1], Henrik Kalisch [1*], Francesco Lagona[1,2], Martin O. Paulsen[1] and Volker Roeber[1,3]

[1] *Department of Mathematics, University of Bergen, Postbox 7800, 5020 Bergen, Norway*
[2] *Department of Political Sciences, University of Roma Tre, Italy*
[3] *Université de Pau et des Pays de l'Adour, Anglet campus, France*

March 2021

Abstract

Shoaling of ocean waves is studied numerically using a low-dimensional non-linear shoaling model coupled with Monte-Carlo simulations based on the statistical description of ocean waves and wave spectra. It is found that while non-linearity has a minor effect on the wave height, it has a major effect on the shape of the wave. In fact, in shallow water, the instantaneous surface elevation can be described using a Gram-Charlier distribution rather than a Gaussian distribution which is typical of waves in deep water. The positivity conditions of the Gram-Charlier expansion are enforced in a grid search to estimate the parameters of the distribution in a way that ensure a positive-definite distribution and the results are in line with field studies of coastal waves, such as the ARSLOE project [10].

Contents

1 Introduction

Various properties of wind-generated waves in coastal regions are significantly different from those in deep water regions. The differences are largely due to the influence of bathymetry, which is more pronounced in shallower water.

In general, deep water waves are considered a Gaussian random process with only minor discrepancies between the observed and theoretical probability density functions. The deviations from the Gaussian model are exhibited by that fact that high crests are observed more frequently than deep troughs [9]. In shallow water, these deviations are more pronounced due to the relative importance of non-linearity in these waves. Indeed irregularities in bathymetry, changes in wave height and wave steepness as the mean water depth decreases towards the shore affect wave properties and their probability distribution as a result. The steepening process near shore causes higher and sharper wave crests and shallower and flatter wave troughs. Under such conditions, the Gaussian model under such conditions is no longer sufficient for describing wave behaviour as it underestimates the higher values and overestimates the lower values of the observed surface elevation. Hence, a non-Gaussian probability density function has to be applied for representing shallow water wave profiles [11].

Previous statistical analyses on the non-Gaussian characteristics of coastal waves include the results of [10] and [11]. In these works, wave records were obtained at a location along the CERC Field Research Facility at Duck North Carolina. These wave records were taken during the growth stage of a storm in the ARSLOE project. The results show that the skewness of the distribution modelling the free surface elevation was the dominant parameter affecting the degree of deviation from the Gaussian model. To account for the skewness, a non-Gaussian probability density function was used to more accurately represent the distribution of the free surface elevation near the shore. The Gram-Charlier probability density function showed good agreement with the histograms of the surface elevation obtained near the shore in both studies.

While the studies mentioned above are based on measurements, the first part of the present study embodies a numerical framework for estimating the coastal surface elevation distribution. As will be elaborated on later in this paper, the combination of linear shoaling theory in deep water and non-linear cnoidal theory in shallow waters yields good agreement with the experimental results found in the above studies. In particular, with the approach used in the present paper, the distribution of the free surface elevation is also found to be non-Gaussian and well represented by a Gram-Charlier series.

The second part of this paper concerns the wave spectra of non-linear waves in shallow water. This may sound like a stretch since the superposition principle can not be applied to non-linear waves. However, for the shoaling of long swells, the time scale of the shoaling process may be short enough that non-linear interactions between the different wave components can not play out completely. In particular, in the present study we are concerned with the range of the shoaling curve between where the linear theory ceases to be valid and waves begin to break as show in [14]. Previous studies on the wave spectrum include those of [5] and [13]. In the former, it was shown that the crest and trough distributions follow the same Rayleigh distribution for a narrow spectrum if the free surface elevation can be considered a random Gaussian process. In [5], new analytical wave crest and trough distributions were derived to take into consideration second-order effects for waves in deep water. The results were an extension to the work of Boccotti and are valid for the spectrum in deep water with frequencies of finite bandwidth. In the present study, an estimation of the wave spectra in shallow water for frequencies of finite bandwidth is presented. The free surface elevation in this case can no longer be considered a random Gaussian process due to non-linear effects and thus, the presented spectra is an estimate for waves approximated by the perturbed Gaussian distribution in the form of a Gram-Charlier expansion.

2 Wave theory

Waves convey mass, momentum and energy and in shoaling processes, wave energy is generally conserved while wave momentum may vary. The linear theory of wave shoaling imposes utilizes energy conservation to obtain the wave height of a shoaling wave. For the nonlinear case, momentum and energy balances are described using the KdV equation together with periodic cnoidal wave solutions.

Linear theory

Linear wave theory is generally limited to small-slope, small amplitude surface gravity waves. This implies that $a/\lambda \ll 1$ and $a/h \ll 1$, respectively [8]. Here, a is the amplitude, λ is the wavelength and h is the depth.

The solution to the linear problem is found by assuming the surface elevation η takes the form of a simple sinusoidal wave propagating in the positive x-direction

$$\eta(x,t) = a\cos(kx - \omega(k)t), \tag{2.1}$$

where k is the wave number and ω is the circular frequency. The velocity potential is given by

$$\phi(x,z,t) = \frac{a\omega(k)}{k}\frac{\cosh(k(z+h))}{\sinh(kh)}\sin(kx - \omega(k)t) \tag{2.2}$$

and ω is given by the dispersion relation

$$\omega(k) = \sqrt{gk\tanh kh}. \tag{2.3}$$

In linear shoaling processes, the speed of wave propagation decreases. A consequence of this is the decrease in the kinetic energy of the wave. However, the total energy of a wave consists of both kinetic energy and potential energy which is conserved according to linear theory. A direct result of the decrease in the kinetic energy is then an increase in potential energy which is found to be directly proportional to the wave height. This change in the wave height can be determined by utilizing the conservative property of the energy flux during the shoaling process. Consider first the energy per unit horizontal area

$$E = \frac{1}{\lambda}\int_0^{\lambda}\int_{-h}^{0}\left[\frac{\rho}{2}|\nabla\phi|^2 + \rho g z\right] dz dx. \tag{2.4}$$

Substituting the solution of the velocity potential (2.2), the dispersion relation (2.3) and computing the integrals gives the following expression for the total energy:

$$E = \frac{1}{8}\rho g H^2. \tag{2.5}$$

Now, the phase speed c is defined as $c = \frac{\omega}{k} = \sqrt{\frac{g}{k}\tanh kh}$ and so the group velocity (the velocity with which the overall envelope shape of the wave propagates) is

$$c_g = \frac{d\omega}{dk} = \frac{c}{2}\left[1 + \frac{2kh}{\sinh(2kh)}\right] \tag{2.6}$$

Conservation of the energy flux Ec_g then implies that the wave height H at a current depth is solely determined by the wave height at a previous depth and the respective group velocities at each depth. Namely,

$$H = H_0\sqrt{\frac{c_{g0}}{c_g}} \tag{2.7}$$

where the subscript '0' denotes the previous depth [17]. To solve (2.7), the conservative property of the period T can be used in combination with the dispersion relation (2.3), leading to

$$\frac{2\pi}{T} - gk\tanh(kh) = 0 \tag{2.8}$$

which is a non-linear equation that can be solved for k. This in turn allows for the determination of H in (2.7).

Non-Linear theory

When waves become too steep or the local depth becomes too shallow, the assumptions of linear theory are no longer satisfied and a new, higher-order framework is required. The Korteweg-de Vries equation is one example of such a framework and has been used with its cnoidal solution to describe wave behaviour during shoaling processes. Previous studies on the shoaling of non-linear waves are [12] and [18] among others.

The Korteweg-de Vries (KdV) equation is a weakly non-linear dispersive model equation given in dimensional variables by

$$\eta_t + c_0\eta_x + \frac{3}{2}\frac{c_0}{h_0}\eta\eta_x + \frac{c_0 h_0^2}{6}\eta_{xxx} = 0 \tag{2.9}$$

where c_0 denotes the shallow water approximation of the phase speed and h_0 denotes the local water depth. The KdV equation has an exact travelling wave solution given by

$$\eta(x,t) = f_2 + (f_1 - f_2)\mathrm{cn}^2\left(\sqrt{\frac{3(f_1-f_3)}{4h_0^3}}(x-ct);m\right). \tag{2.10}$$

where f_1 is the wave crest, f_2 is the wave trough, m is the elliptic parameter, cn is the Jacobian elliptic function and $f_3 = f_1 - \frac{1}{m}(f_1 - f_2)$. The wave speed c and wave length λ can be defined as

$$c = c_0\left(1 + \frac{f_1+f_2+f_3}{2h_0}\right) \quad \text{and} \quad \lambda = K(m)\sqrt{\frac{16h_0^3}{3(f_1-f_3)}} \tag{2.11}$$

where $K(m)$ is the complete elliptic integral of the first kind. It has been shown in [3], [2] and [1] that the energy balance in the KdV equation is given by

$$\frac{\partial}{\partial t}E + \frac{\partial}{\partial x}q_E = 0 \tag{2.12}$$

to the second order, where

$$E = c_0^2\left(\frac{1}{h_0\eta^2} + \frac{1}{4h_0^2}\eta^3 + \frac{h_0}{6}\eta\eta_{xx} + \frac{h_0}{6}\eta_x^2\right) \tag{2.13}$$

and

$$q_E = c_0^3\left(\frac{1}{h_0}\eta^2 + \frac{5}{4h_0^2}\eta^3 + \frac{h_0}{2}\eta\eta_{xx}\right). \tag{2.14}$$

The wave height of a shoaling wave can now be determined by imposing preservation of wave frequency, conservation of mass and conservation of energy. Thus, if the wave motion at a certain water depth h_A is given, the wave height at water depth h was found in [7] to be given by the following equations:

$$\begin{aligned} \frac{c_A}{\lambda_A} &= \frac{c}{\lambda}, \\ \int_0^T q_{E_A}dt &= \int_0^T q_E dt, \\ \int_0^\lambda \eta_A dx &= \int_0^\lambda \eta dx. \end{aligned} \tag{2.15}$$

Using the stationary solution of the KdV equation (2.10) with wave speed and wavelength given in (2.11) and also utilizing the energy flux (2.14), a system of three non-linear equations that can be solved for f_1, f_2 and f_3 and the height of a wave at depth h can be determined. For more details on the numerical procedure see [14].

The theory in this chapter is the foundation upon which the non-linear transfer function presented in [15] is built upon for individual waves. When considering sea states consisting of a wave spectrum, some statistical preliminaries are necessary and will be presented in section 3. The applications of the non-linear transfer function in this case are then presented accordingly in section 5.

3 Statistical theory

Short-term statistical theory to characterize deep water gravity waves is largely based on the assumption that the surface elevation is a stationary Gaussian process. In shallow, coastal waters this is however not always the case due to the non-linear nature of the waves described in the Introduction. Therefore, a non-Gaussian probability density function has to be applied to represent shallow water wave profiles. The Gram-Charlier series expansion is a popular approach for estimating the distribution of the free surface elevation in coastal regions [11]. Another popular distribution is the Tayfun distribution [19]. This section will provide some statistical preliminaries for deep and shallow water as well as presenting the method of Maximum Likelihood Estimation (MLE).

The random-phase/amplitude model

The random-phase/amplitude model is generally accurate when the waves are not too steep and are in sufficiently deep waters so that interactions between the wave components are weak and can be neglected. We begin by considering the surface elevation η as a function of time t at one location. Using Fourier theory, the wave record can be reproduced by considering the surface elevation to be a sum of harmonic wave components and can therefore be approximated mathematically in terms of its *Fourier series* as

$$\underline{\eta}(t) = \sum_{i=1}^{N} \underline{a}_i \cos(2\pi f_i t + \underline{\alpha}_i) \tag{3.1}$$

where $\underline{a}_i$, f_i and $\underline{\alpha}_i$ are the amplitude, frequency and phase respectively and the underscore indicates that they are random variables. Equation (3.1) is more formally known as the *random-phase/amplitude model.* Random variables are fully characterized by their respective probability density functions and here the phase α_i is uniformly distributed between 0 and 2π at each frequency so that

$$p(\alpha_i) = \frac{1}{2\pi}, \quad 0 < \alpha_i \leq 2\pi. \tag{3.2}$$

The amplitude $\underline{a}_i$ is Rayleigh distributed at each frequency so that

$$p(a_i) = \frac{\pi}{2} \frac{a_i}{\mu_i^2} \exp\left(-\frac{\pi a_i^2}{4\mu_i^2}\right) \tag{3.3}$$

where μ_i is the expected values of the amplitude $\mu_i = E\{\underline{a}_i\}$ [9]. A large set of realizations of $\eta(t)$ can then be constructed for a given amplitude spectrum by drawing a random amplitude a_i and phase α_i from their probability density functions at each frequency and inserting into (3.1).

The wave spectrum

Since wave energy can be proven to be proportional to the variance in linear wave theory [9], the variance provides a useful link between statistical and physical properties.

Therefore, it is often beneficial to consider the variance $E\{\frac{1}{2}\underline{a}_i^2\}$ rather than the expectation of the amplitude $E\{\underline{a}_i\}$ as in the random-phase/amplitude model described above [9].

By distributing the variance of the surface elevation $E\{\frac{1}{2}\underline{a}_i^2\}$ over the frequency interval Δf_i at frequency f_i, we obtain the following definition of the variance density spectrum:

$$S(f_i) = \frac{1}{\Delta f_i} E\left\{\frac{1}{2}\underline{a}_i^2\right\}, \quad \forall f_i. \tag{3.4}$$

A continuous version of (3.4) is readily obtained by letting the width of the frequency interval Δf_i approach zero and can be written mathematically as

$$S(f) = \lim_{\Delta f \to 0} \frac{1}{\Delta f} E\left\{\frac{1}{2}\underline{a}^2\right\}. \tag{3.5}$$

Now, the random surface elevation $\underline{\eta}(t)$ given in (3.1) is the sum of a large number of harmonic waves. The variance of a single harmonic wave with amplitude a is given by $\overline{\eta^2} = \frac{1}{2}a^2$ so that the variance of the sum is given by summing the individual variance contributions of each harmonic [9], i.e.

$$\overline{\underline{\eta}^2} = E\{\underline{\eta}^2\} = \sum_{i=1}^{N} E\{\frac{1}{2}\underline{a}_i^2\} \tag{3.6}$$

when the overbar indicates averaging and $E\{\underline{\eta}\} = 0$. The variance density spectrum given in (3.4) can be directly related to the scaling parameter σ of the Rayleigh distributed Fourier amplitudes and is presented in Section 4.

The Pierson-Moskowitz spectrum

The Pierson-Moskowitz (P-M) spectrum is a unidirectional spectrum describing waves in fully developed seas and is often used in applications. The underlying assumption is that waves reach a point of equilibrium with the wind if the wind blows steadily over a large area for a sufficient period of time [16]. Now, the spectral formulation for fully developed seas can be given by the following:

$$\hat{S}(\omega) = \frac{Ag^2}{\omega^5} \exp -\frac{5}{4}\left(\frac{\omega_0}{\omega}\right)^4 \tag{3.7}$$

where $\omega = 2\pi f$ is the *circular frequency* in Hertz and $A = 8.1 \cdot 10^{-3}$.

The Gram-Charlier type-A expansion

When a random variable z has an unknown probability density function (PDF), the unknown function $f(z)$ can be approximated in terms of Hermite polynomials, i.e.

$$f(z) = g_n(z)p(z) = \sum_{n=0}^{\infty} c_n H_n(z)p(z) \tag{3.8}$$

where $p(z)$ is the standardized (zero mean and unit variance) normal PDF and $H_n(z) = (-1)^n \frac{d^n p}{dz^n} \frac{1}{p(z)}$ is the nth-order Hermite polynomial. Letting β_1 and β_2 be the skewness and kurtosis of (3.8), the values

$$c_0 = 1, \quad c_1 = 0 = c_2, \quad c_3 = \frac{\sqrt{\beta_1}}{3!}, \quad c_4 = \frac{(\beta_2 - 3)}{4!} \tag{3.9}$$

are the first 5 coefficients inside the sum [4]. When z is standardized (zero mean and unit variance), a 4th-order approximation of $g_n(z)$ is then

$$g_4(z) = 1 + \frac{\gamma_1}{6} H_3(z) + \frac{\gamma_2}{24} H_4(z) \tag{3.10}$$

where $\gamma_1 = \sqrt{\beta_1}$ and $\gamma_2 = \beta_2 - 3$. This expression is more formally known as the Gram-Charlier Type-A expansion. Note that when $\gamma_1 = 0 = \gamma_2$, (3.8) reduces to a standard normal distribution.

Postivity conditions

One of the major drawbacks of a polynomial approximation is that certain parameters can lead to negative values which is undesirable when considering probability density functions [6]. It is therefore necessary to define the conditions for which the function $f(z)$ is positive-definite.

We begin by considering $\mathscr{D}$ to be the region in the (γ_1, γ_2)-plane for which $f(z)$ in (3.8) is positive definite. Mathematically, this entails that

$$g_4(z) = 1 + \frac{\gamma_1}{6} H_3(z) + \frac{\gamma_2}{24} H_4(z) \geq 0, \quad \forall z. \tag{3.11}$$

Now, the set (γ_1, γ_2) that satisfies

$$g_4(z) = g_4'(z) = 0 \tag{3.12}$$

where $g_4'(z) = \frac{\gamma_1}{2} H_2(z) + \frac{\gamma_2}{6} H_3(z)$ is called the *envelope* of $p_4(z)$ [6]. Straightforward computations give

$$\gamma_1(z) = -24 \frac{H_3(z)}{h(z)} \quad \text{and} \quad \gamma_2(z) = 72 \frac{H_2(z)}{h(z)}, \tag{3.13}$$

where (3.12) has been solved simultaneously for γ_1 and γ_2 and $h(z) = 4H_3^2(z) - 3H_2(z)H_4(z)$.

To determine the set (γ_2, γ_2), we begin by rewriting $h(z)$ as $h(z) = z^6 - 3z^4 + 9z^2 + 9$. Since $h(z) \geq 0$ for all z, the signs of γ_1 and γ_2 depend solely on $H_3(z)$ and $H_2(z)$, respectively. Now, $H_3(z) = z^3 - 3z$ so $\gamma_1 \geq 0$ for $z \in (-\infty, -\sqrt{3}]$ and $z \in [0, \sqrt{3}]$. Similarly, $H_2(z) = z^2 - 1$ so $\gamma_2 \geq 0$ for $z \in (-\infty, -1]$ and $z \in [1, \infty)$.

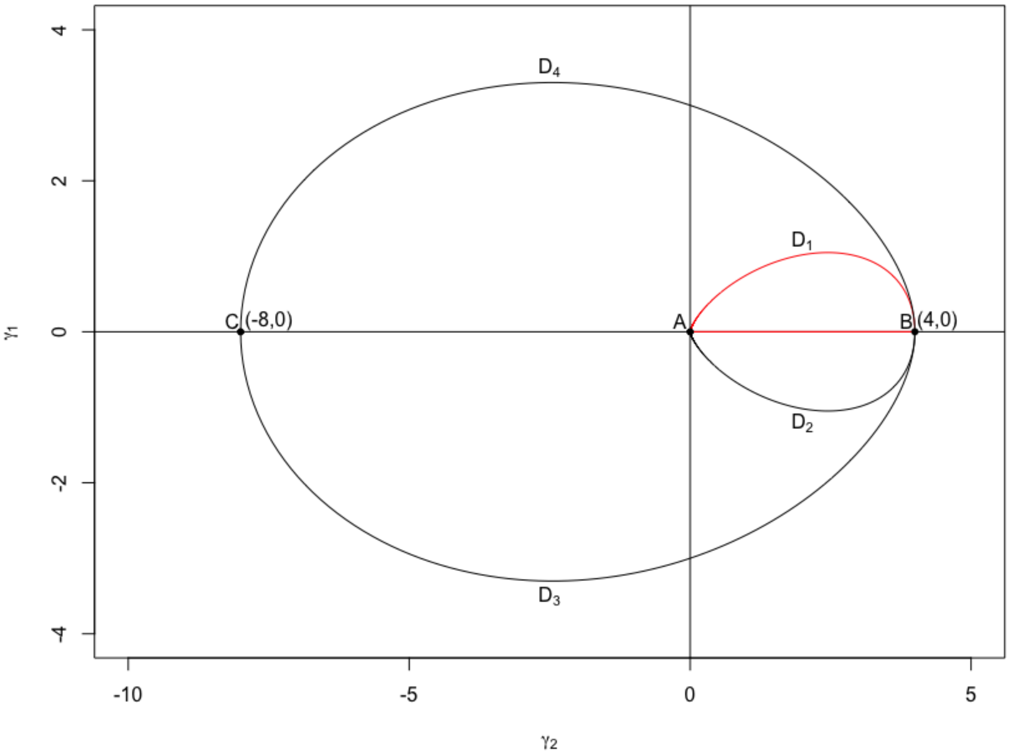


Figure 1: Global plot of the envelope of $p_4(z)$. The red lines define the boundary of $\mathscr{D}$.

Figure 1 presents the global envelope of $p_4(z)$. The curve $\mathrm{AD_1B}$ represents the region of the (γ_1, γ_2)-plane when $z \in (-\infty, -\sqrt{3})$. For $z \in [-\sqrt{3}, 0]$ and $z \in [0, \sqrt{3}]$, the curves $\mathrm{BD_3C}$ and $\mathrm{BD_4C}$ are obtained. Lastly, the curve $\mathrm{AD_2B}$ represents the region of the (γ_1, γ_2)-plane when $z \in [\sqrt{3}, \infty)$. It is now clear that $f(z) \geq 0$ for all z when both γ_1 and γ_2 are positive. For the kurtosis, this means that $\gamma_2 \in [0, 4]$. The maximum skewness is obtained when $\left|\gamma_1'(z)\right| = \left|z^4 - 6z^3 + 6z^2 - 18z + 9\right| = 0$ and can be found numerically to be at the point $(2.451, 1.051)$ in the (γ_1, γ_2)-plane. This implies that $\gamma_1 \in [0, 1.051]$. The region $\mathscr{D}$ is then the envelope obtained when $\gamma_1 \in [0, 1.051]$ and $\gamma_2 \in [0, 4]$ and is shown in Figure 1 where the red lines define the boundary of $\mathscr{D}$.

Maximum likelihood estimation

Maximum likelihood estimation (MLE) is a well known method used to estimate the parameters of a statistical model. The procedure is based on maximizing what is known as the *likelihood function* of the model. Consider the random sample $\underset{\sim}{Z} = (z_1, z_2, ..., z_n)$ generated from the PDF $f(z_i; \boldsymbol{\theta})$ where $\boldsymbol{\theta}$ is a k-dimensional parameter vector in the parameter space $\boldsymbol{\Omega}$. Then, the joint probability density of the sample $\underline{Z}$ is

$$f(z_1, z_2, ..., z_n; \boldsymbol{\theta}) = f(z_1; \boldsymbol{\theta}) \cdot f(z_2; \boldsymbol{\theta}) \cdot ... \cdot f(z_n; \boldsymbol{\theta}) = \prod_{i=1}^{n} f(z_i; \boldsymbol{\theta}) \tag{3.14}$$

If we view the joint PDF as a function of $\boldsymbol{\theta}$, (3.14) can be written as

$$L(\boldsymbol{\theta}; z_1, z_2, ..., z_n) = \prod_{i=1}^{n} f(z_i; \boldsymbol{\theta}) \tag{3.15}$$

where $L(\boldsymbol{\theta})$ is the likelihood function. The goal of MLE is to find the values of the model parameters that maximize the likelihood function (3.15) over the parameter space $\boldsymbol{\Omega}$ in a way that makes the observed data most probable. We can now define the *maximum likelihood estimator*

$$\hat{\boldsymbol{\theta}} = \underset{\theta \in \Omega}{\operatorname{argmax}} \, L(\boldsymbol{\theta}) \tag{3.16}$$

to be the parameter values that maximize $L(\boldsymbol{\theta})$. Candidates for the maximum likelihood estimator are then all points θ_j such that

$$\frac{\partial L}{\partial \theta_j} = 0, \quad j = 1, ..., k \tag{3.17}$$

By taking the natural logarithm of the likelihood function, the product in the joint density (3.14) can be written as a sum which is more convenient when differentiating. The *log-likelihood* is then readily defined as

$$l(\boldsymbol{\theta}; \underline{Z}) = \log \left(\prod_{i=1}^{n} L(\boldsymbol{\theta}; \underline{Z}) \right) = \sum_{i=1}^{n} \log L(\boldsymbol{\theta}; \underline{Z}) \tag{3.18}$$

and candidates for the maximum likelihood estimator are now all points θ_j such that

$$\frac{\partial l}{\partial \theta_j} = 0, \quad j = 1, ..., k \tag{3.19}$$

The monotonic behaviour of the logarithm function ensures that the maximum of $l(\boldsymbol{\theta})$) occurs at the same values of θ_j as for $L(\boldsymbol{\theta})$ which is a desirable result of the transformation.

4 Methodology

Case 1: Sea states with waves of single frequency

The standard form of the Rayleigh probability density function of the random variable $\underline{z}$ is given by

$$p(z) = \frac{z}{\sigma^2} \exp \left(-\frac{z^2}{2\sigma^2} \right), \quad z \geq 0 \tag{4.1}$$

where σ is the scaling parameter of the distribution. Now, we begin by considering singular frequency gravity waves in deep water that are Rayleigh distributed with their probability density function given by

$$p(H) = \frac{H}{4m_0} \exp\left(-\frac{H^2}{8m_0}\right). \tag{4.2}$$

From this, the relationship between the zeroth-order moment m_0 and the scaling parameter σ of the Rayleigh distribution can be defined as $\sigma = 2\sqrt{m_0}$. In deep water, the following approximation of the significant wave height can be used [13]:

$$H_s \approx 4\sqrt{m_0} \tag{4.3}$$

so that $m_0 \approx \frac{H_s^2}{16}$ and the scaling parameter σ of the Rayleigh distribution can be computed.

Now, for a given sea state with significant wave height $1\text{m} \leq H_s \leq 3\text{m}$ and peak period $8\text{s} \leq T_p \leq 12\text{s}$, Rayleigh distributed wave heights were randomly sampled using σ (~ 500) from Monte Carlo simulations and stored in the matrix $\mathbf{H}$. Each simulated value of H_i in $\mathbf{H}$ then corresponds to one realization of the wave height under the conditions of an actual observation. The non-linear transfer function implemented in [15] was then readily applied to each realization with their corresponding frequency $f = 1/T$ to acquire the local wave heights, wave lengths, modulus m and root solutions f_2 in shallow water stored in the matrices $\mathbf{H}^*$, $\boldsymbol{\lambda}^*$, $\mathbf{m}$ and $\mathbf{f_2}$ respectively for later use.

To compute the surface elevation η in both deep and shallow water, the parameter m was used as a switch. Using each m_i to calculate $K(m_i)$ which is the complete elliptic integral of first kind, the Jacobian elliptic function cn was computed for each m_i. As mentioned in section 2, m gives periodic waves for $0 \leq m < 1$. For the case $m = 0$, the cnoidal solution given in (2.10) reduces to the linear solution given in (2.1). The surface elevation of each individual wave was then computed at 100 uniformly spaced grid points x_i so that $-\frac{\lambda_i}{2} \leq x_i \leq \frac{\lambda_i}{2}$, using either the linear or non-linear solution depending on the nature of the wave. The results were stored in the matrices $\boldsymbol{\eta}$ and $\boldsymbol{\eta^*}$ for deep and shallow water, respectively.

Once the surface elevation was computed, the parameter grid search described at the end of this section was used to estimate the parameter vector $\boldsymbol{\theta}$ of the surface elevation distribution by the method of MLE as described in section 3 so that

$$\hat{\boldsymbol{\theta}} = \underset{\theta \in \Omega}{\operatorname{argmax}} \, L(\boldsymbol{\theta}) \tag{4.4}$$

A statistical analysis of both the wave height and the surface elevation was then carried out. First, the question of whether the wave heights are Rayleigh distributed in shallow water was addressed by fitting a Rayleigh distribution to the data and hypothesis testing. Similarly, to determine if the the deep water free surface elevation is in fact normally distributed, a Gaussian distribution was fit to the data by method of maximum likelihood estimation along with visual inspections in the form of histograms and Q-Q plots. For the free surface elevation in shallow water, the Gram-Charlier expansion is used in place of the Gaussian distribution and a comparison was carried out.

Case 2: Sea states with waves of several frequencies

Now we begin by considering a wave spectrum defined in terms of its significant wave height H_s and peak period T_p. The Pierson-Moksowitz spectrum dependent on H_s is then computed using (3.7). Namely,

$$S(f_i) = \frac{H_{rms}^2}{8 \int \hat{S}(f_i) df_i} \hat{S}(f_i) \tag{4.5}$$

where H_{rms} is the root-mean-square wave height. The scaling parameters σ_i used to simulate were then calculated using the relation

$$\sigma_i = \sqrt{S(f_i)\Delta f_i}. \tag{4.6}$$

Each σ_i was then used to randomly sample Rayleigh distributed Fourier amplitudes ($\sim$ 100) at each frequency f_i with Monte Carlo simulations. The result is the 100-by-100 matrix $\mathbf{A}$ where each column $\mathbf{A}_{*,i}$ represents 100 realizations of the random Fourier amplitude $\underline{a}_i$ at the frequency f_i in deep water. Each column then has an expected value given by $E\{\underline{a}_i\} = \sigma_i \sqrt{\frac{\pi}{2}}$.

The non-linear transfer function ([15]) requires that the amplitudes given as the input are the physical amplitudes rather than the Fourier amplitudes. To approximate the physical amplitudes we propose the following scaling:

$$E\{\underline{\tilde{a}}_i\} = \kappa_i E\{\underline{a}_i\} \tag{4.7}$$

where κ_i is chosen to be

$$\kappa_i = \frac{H_{rms}}{E\{\underline{a}_i\}} = \frac{H_{rms}}{\sigma_i \sqrt{\frac{\pi}{2}}}, \tag{4.8}$$

where $E\{\underline{\tilde{a}}_i\}$ is the expectation of the physical amplitudes. The transformed scaling parameter to be simulated with is then given by

$$\tilde{\sigma}_i = \frac{E\{\underline{\tilde{a}}_i\}}{\sqrt{\frac{\pi}{2}}} = \sqrt{\frac{2}{\pi}} \kappa_i E\{\underline{a}_i\} = \kappa_i \sigma_i. \tag{4.9}$$

Using $\tilde{\sigma}_i$ to simulate with in the same manner as before results in the transformed matrix $\tilde{\mathbf{A}}$ now consisting of the physical amplitudes at each frequency f_i in deep water. Our choice of κ_i ensures that each column of the transformed matrix $\tilde{\mathbf{A}}_{*,i}$ has an expected value $E\{\tilde{a}_i\} = H_{rms}$ in agreement with the significant wave height of the original deep water spectrum. The non-linear transfer function was then applied to each column $\tilde{\mathbf{A}}_{*,i}$ along with its corresponding frequency f_i to acquire the local amplitude of each harmonic in shallow water and was stored in the matrix $\tilde{\mathbf{A}}^*$. Each column of $\tilde{\mathbf{A}}^*$ now represents realizations of the physical amplitudes at each frequency f_i in shallow water. Analogous to the deep water case, the columns $\tilde{\mathbf{A}}^*_{*,i}$ each have an expected value given by $E\{\underline{\tilde{a}}_i^*\} = \tilde{\sigma}_i^*\sqrt{\frac{\pi}{2}}$ where $\tilde{\sigma}_i^*$ is the scaling parameter of the Rayleigh distributed amplitudes in shallow water and is estimated for each column by fitting a Rayleigh distribution to each $\tilde{\mathbf{A}}^*_{*,i}$. The scaling parameters belonging to the Rayleigh distributed Fourier amplitudes in shallow water were then calculated using the relation given in (4.9) and rearranging. So,

$$\sigma_i^* = \frac{1}{\kappa_i}\tilde{\sigma}_i^*. \tag{4.10}$$

The spectrum in shallow water $S^*(f_i)$ can now be approximated using (4.6) so that

$$S^*(f_i) = \frac{{\sigma_i^*}^2}{\Delta f}. \tag{4.11}$$

Implementation of parameter grid search

A description of the methodology used in the parameter grid search will now be given. First, we denote n as the number of possible values for γ_1, γ_2 and σ such that the total number of grid points to search is n^3.

Step 1: Standardize data

We begin by standardizing the surface elevation data so

$$z_\eta = \eta - \mu_\eta,$$

where z_n is now the standardized surface elevation and μ_η is the surface elevation mean.

Step 2: Define grid

The grid vertices can be defined from the conditions imposed on γ_1 and γ_2 found in section 3. We let $\boldsymbol{\gamma_1}$ and $\boldsymbol{\gamma_2}$ be the equally spaced n-by-1 column vectors where each $\gamma_{1,i} \in [0, 1.051]$ and $\gamma_{2,j} \in [0, 4]$, respectively. Similarly, we let $\boldsymbol{\sigma}$ be the equally spaced n-by-1 column vector centered around the sampling standard deviation so that each $\sigma_k \in [\sigma_{z_\eta} - 2, \sigma_{z_\eta} + 2]$.

Step 3: Define probability density and negative log-likelihood functions

The PDF function $f(z_\eta)$ and negative log-likelihood functions were defined in R as the following:

Algorithm 1: Unknown PDF $f(z_\eta)$ function

Input: Standardized surface elevation z_η, parameter vector $\boldsymbol{\theta}$

Output: Values of PDF at each surface elevation point

$$f(z_\eta) = \left(1 + \frac{\gamma_1}{6} H_3(z_\eta) + \frac{\gamma_2}{24} H_4(z_\eta)\right) \left(\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{z_\eta}{\sigma_\eta}\right)^2}\right)$$

Algorithm 2: Negative log-likelihood function

Input: Standardized surface elevation z_η, parameter vector $\boldsymbol{\theta}$

Output: Negative log-likelihood value of $f(z_\eta)$

$$l = -\left(\sum \log f(z_{\eta)}\right)$$

Step 4: Parameter grid search

Then a grid search was implemented in R in the following way:

Algorithm 3: Parameter grid search

Input: Standardized surface elevation z_η, initial parameter vector $\boldsymbol{\theta}$

Output: Updated parameter vector $\boldsymbol{\theta}$

for $\gamma_{1,i}$ **do**
    for $\gamma_{2,j}$ **do**
        for σ_k **do**
            Calculate minimum of $f(z_\eta)$
            if *minimum of* $f(z_\eta) > 0$ **then**
                Compute negative log-likelihood l of $f(z_\eta)$ at that grid point
                if $l < l_0$ **then**
                    // l_0 denotes the negative log-likelihood value from the previous iteration
                    Parameter vector $\boldsymbol{\theta} = [\gamma_{1,i}, \gamma_{2,k}, \sigma_k]$
                end
            end
        end
    end
end

5 Case Studies and Analysis

In this section, the results obtained when implementing the methodology proposed in section 4 for two different cases will be shown.

Case 1: Sea states with waves of single frequency

In all experiments the deep water depth is defined as 70m and the coastal depth as 5m. Experiments 1, 2 and 3 are carried out with a period of $T = 8s, T = 10s$ and $T = 12s$ respectively and values of the significant wave height $H_{s,0}$ were chosen so that $H_{s,0} \in [1, 2, 3]$m in deep water. Tables 1, 2 and 3 show the estimated values of the skewness (γ_1) and kurtosis (γ_2) in the Gram-Charlier type-A expansion given in (3.10) as well as the standard deviation (σ) of the normal distribution $p(z)$. The parameters β_1 and β_2 are calculated given that $\gamma_1 = \sqrt{\beta_1}$ and $\gamma_2 = \beta_2 - 3$ as presented in section 3. The parameter $H_{s,0}$ defines the significant wave height in deep water whereas H_s denotes the shallow water significant wave height and is calculated for each experiment.

Experiment 1:

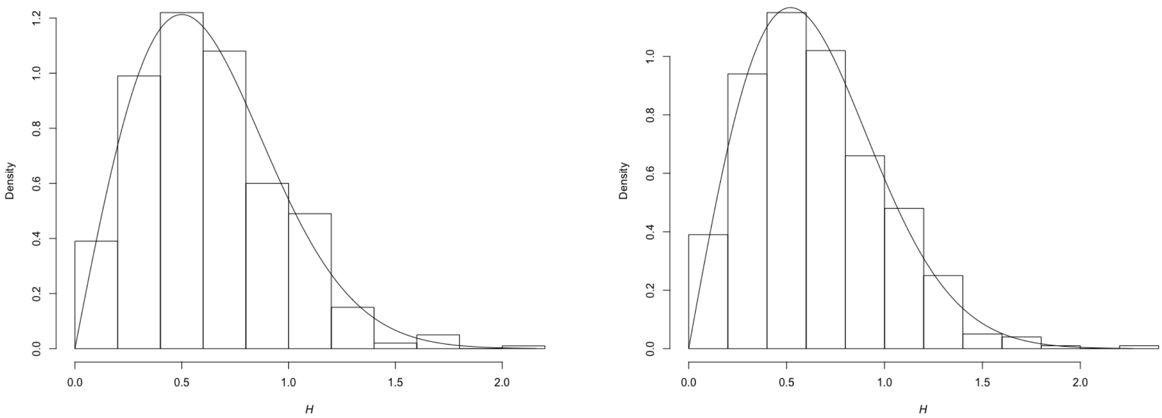
Table 1: Estimated values of γ_1, γ_2 and σ for simulated sea states with $T = 8s$.

$H_{s,0}$	H_s	γ_1	γ_2	σ	β_1	β_2
1m	1.05m	0.517	0.552	0.249	0.267	3.552
2m	2.05m	0.879	1.379	0.512	0.773	4.379
3m	3.18m	0.983	1.793	0.731	0.966	4.793

Figure 2 shows a Rayleigh distribution fit to histograms of the wave height at 70m and 5m depth, respectively for the case $T = 8s$ and initial significant wave height $H_{s,0} = 1$m. In both subfigures, the Rayleigh distribution fits the data very well. A Kolmogorov-Smirnov (K-S) test was also carried out to test the null hypothesis that the wave height data in shallow water comes from a Rayleigh distribution. The p-value obtained for this case was statistically significant ($p = 0.7725$) indicating that the null hypothesis can not be rejected and we conclude that the wave heights in shallow water are Rayleigh distributed. The same result was obtained for all 3 experiments and 9 cases, i.e. the p-values obtained from the K-S test were all greater than the significance level $(\alpha) = 0.05$. We therefore conclude that wave heights obtained at a depth of 5m can still be considered Rayleigh distributed.

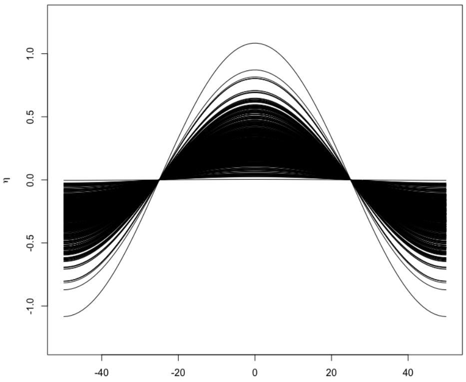
Figure 3 shows plots of each individual wave profile (~ 500) over its respective wave length λ. An increase in wave height and decrease in wavelength can be observed while the frequency remains constant in each case. This is due to the group velocity changing with water depth. A decrease in the group velocity is analogous to a decrease in the wave-energy transport velocity and must be compensated for. Since wave energy is conserved, a decrease in the kinetic energy leads accordingly to an increase in the potential energy and thus an increase in wave height as described in section 2. Similar results were observed in experiments 2 and 3.

Figure 5 presents Q-Q plots of the free surface elevation data in both deep and shallow water. In all 3 cases, the surface elevation in deep water follows the normal line reasonably well with only small deviations at the end points. This indicates that the free surface elevation in deep water has a ”thin-tailed” distribution. In these cases, the Q-Q plot of the distribution has small or negligible deviations at the ends. Thus, the surface elevation in deep water can still be classed as normally distributed. In shallow water however, the deviations from the normal line are much greater. The degree of deviation is more pronounced in the right end point than the left which indicates that the surface elevation in shallow water follows a distribution that is positively skewed. Again, similar results were obtained for experiments 2 and 3.

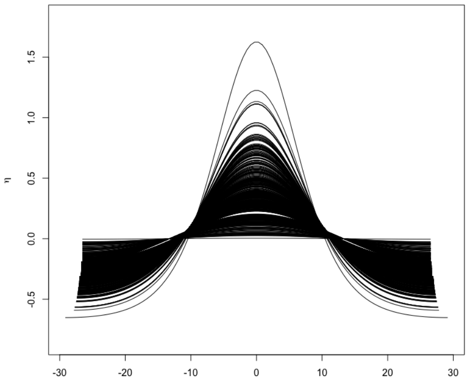


(a) Waveheight at 70m, $H_{s,0} = 1$m (b) Wave height at 5m, $H_s = 1.05$m

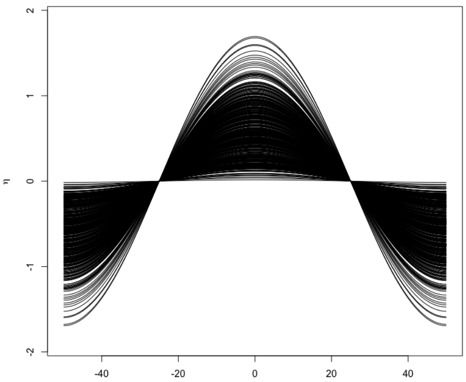
Figure 2: Histograms of the wave height H at 70m and at 5m depth (after the non-linear transfer function has been applied) for waves with $T = 8s$. A Rayleigh distribution has been fit to the waves at both depths.



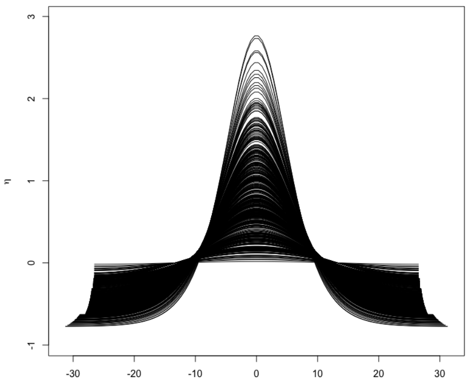
(a) Surface elevation at 70m for $H_{s,0} = 1$m



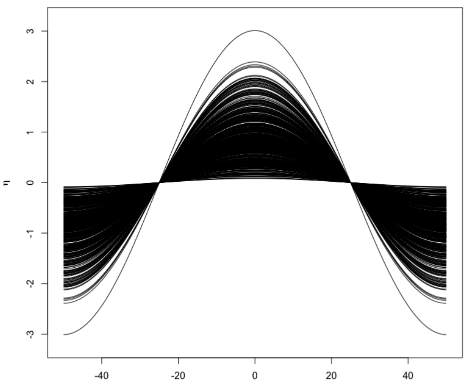
(b) Surface elevation at 5m for $H_s = 1.05$m



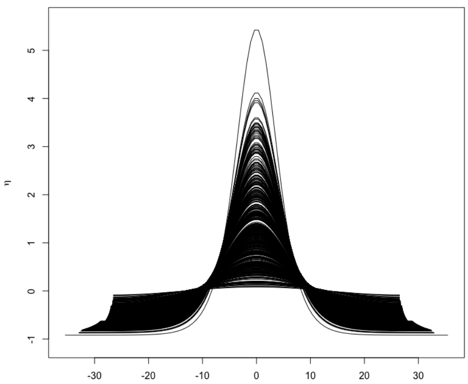
(c) Surface elevation at 70m for $H_{s,0} = 2$m



(d) Surface elevation at 5m for $H_s = 2.05$m

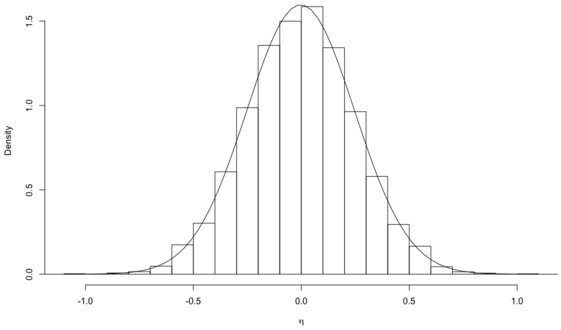


(e) Surface elevation at 70m for $H_{s,0} = 3$m

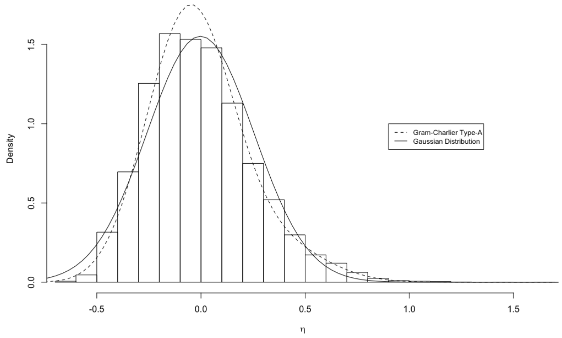


(f) Surface elevation at 5m for $H_s = 3.18$m

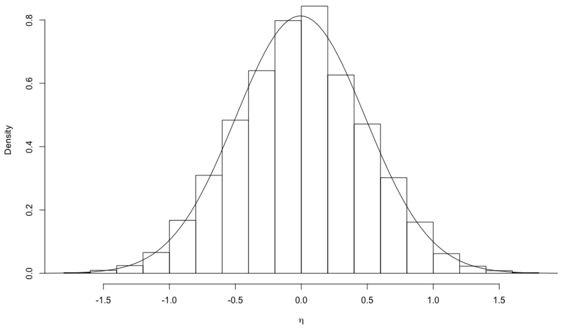
Figure 3: Plots of surface elevation η at 70m and at 5m depth (after the non-linear transfer function has been applied) for waves with $T = 8s$ over each of their respective wavelengths λ.



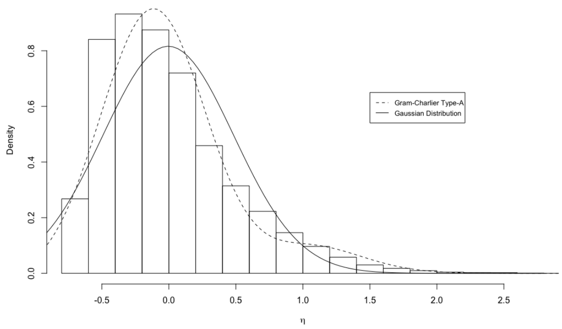
(a) Surface elevation at 70m for $H_{s,0} = 1$m



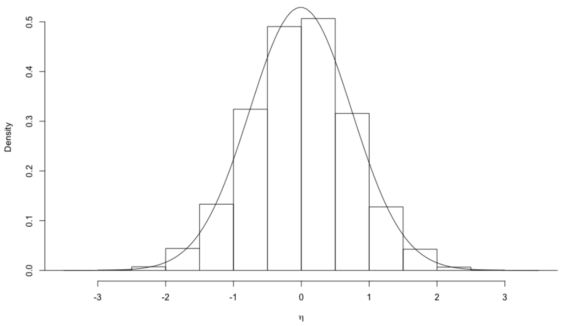
(b) Surface elevation at 5m for $H_s = 1.05$m



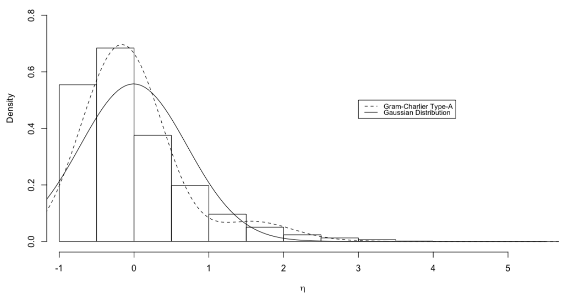
(c) Surface elevation at 70m for $H_{s,0} = 2$m



(d) Surface elevation at 5m for $H_s = 2.05$m

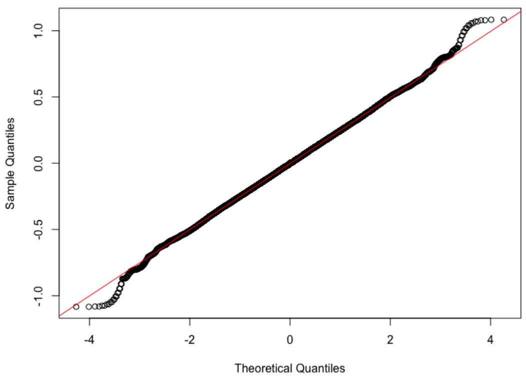


(e) Surface elevation at 70m for $H_{s,0} = 3$m

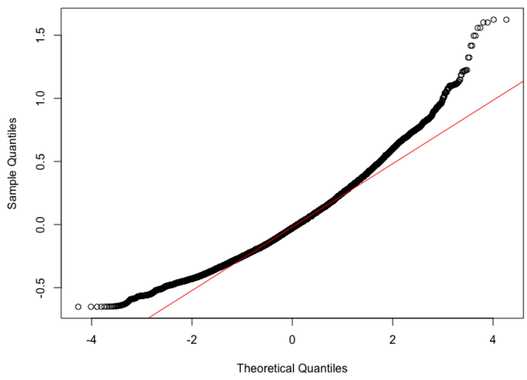


(f) Surface elevation at 5m for $H_s = 3.18$m

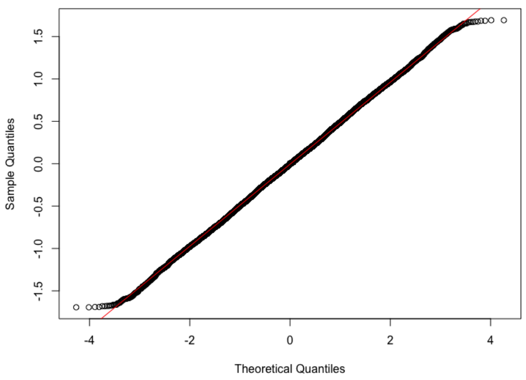
Figure 4: Histograms of surface elevation η at 70m and at 5m depth (after the non-linear transfer function has been applied) for waves with $T = 8s$. A Gaussian distribution has been fit to the waves at 70m depth whereas both the Gaussian and Gram-Charlier densities have been fit to the waves at 5m depth.



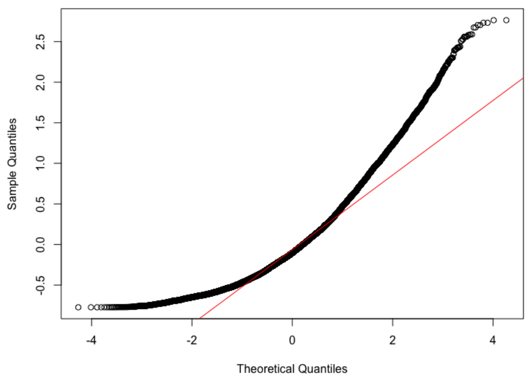
(a) Q-Q plot of surface elevation at 70m for $H_{s,0}$ = 1m



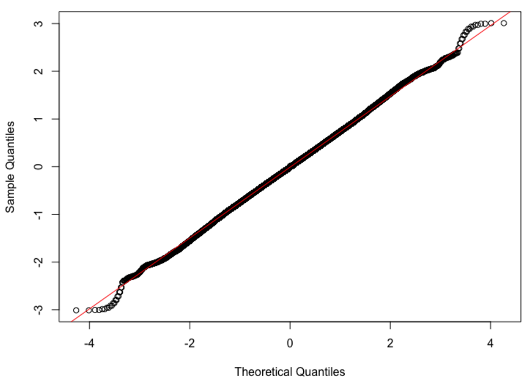
(b) Q-Q plot of surface elevation at 5m for H_s = 1.05m



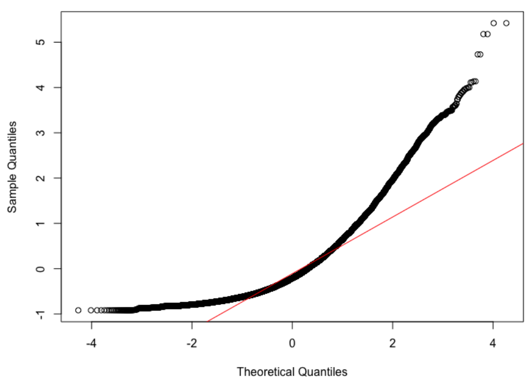
(c) Q-Q plot of surface elevation at 70m for $H_{s,0}$ = 2m



(d) Q-Q plot of surface elevation at 5m for H_s = 2.05m

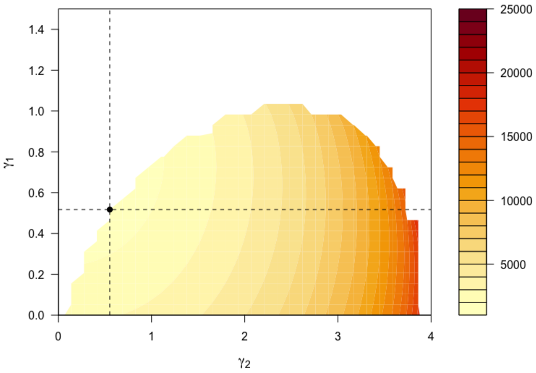


(e) Q-Q plot of surface elevation at 70m for $H_{s,0}$ = 3m

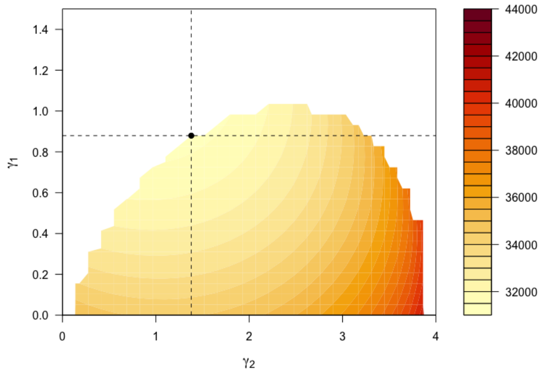


(f) Q-Q plot of surface elevation at 5m for H_s = 3.18m

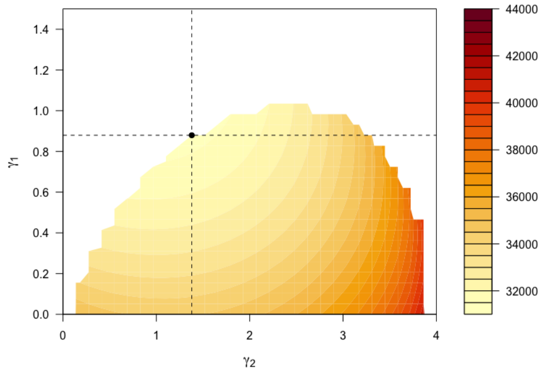
Figure 5: Q-Q plots of surface elevation η at 70m and 5m depth for waves with $T = 8s$.



(a) $\gamma_1 = 0.517$ and $\gamma_2 = 0.552$



(b) $\gamma_1 = 0.879$ and $\gamma_2 = 1.379$



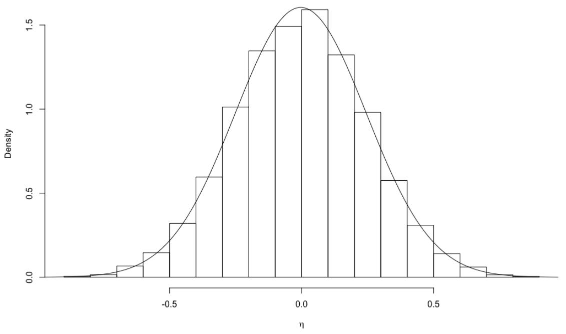
(c) $\gamma_1 = 0.983$ and $\gamma_2 = 1.793$

Figure 6: γ_1, γ_2 plane showing estimated parameters of the Gram-Charlier type-A expansion for waves with $T = 8s$.

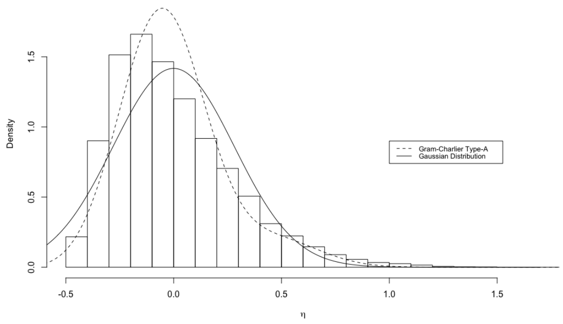
Experiment 2:

Table 2: Estimated values of γ_1, γ_2 and σ for simulated sea states with $T = 10s$.

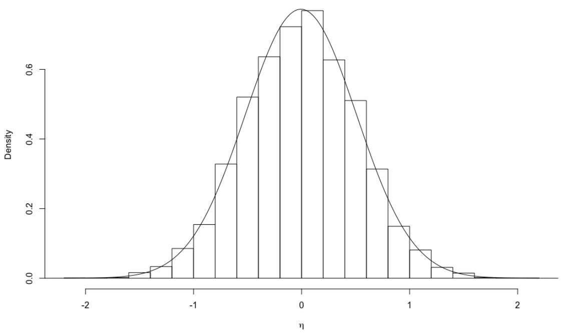
$H_{s,0}$	H_s	γ_1	γ_2	σ	β_1	β_2
1m	1.17m	0.724	0.965	0.250	0.524	3.965
2m	2.46m	0.983	1.793	0.522	0.966	4.793
3m	3.68m	1.034	2.207	0.649	1.069	5.207



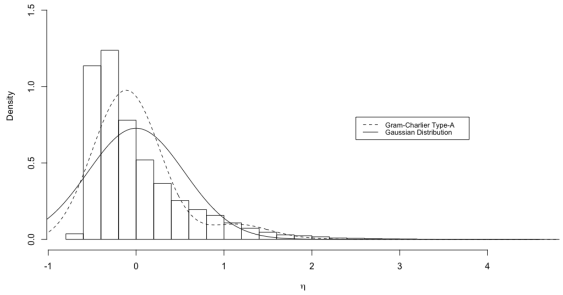
(a) Surface elevation at 70m for $H_{s,0} = 1\text{m}$



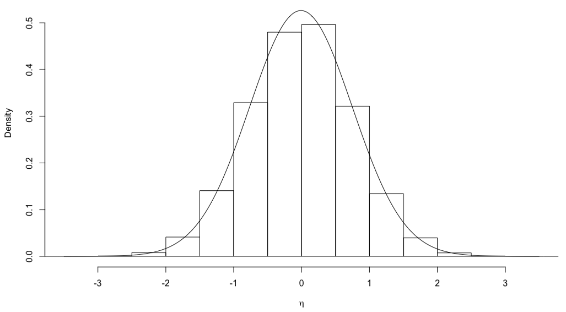
(b) Surface elevation at 5m for $H_s = 1.17\text{m}$



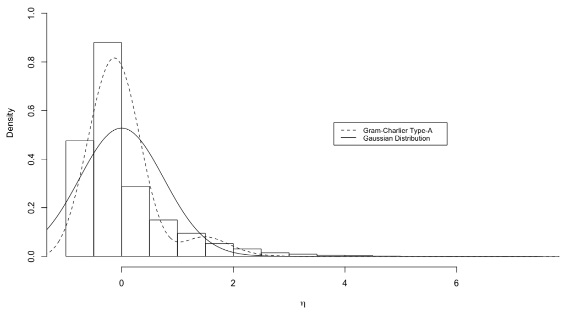
(c) Surface elevation at 70m for $H_{s,0} = 2\text{m}$



(d) Surface elevation at 5m for $H_s = 2.46\text{m}$

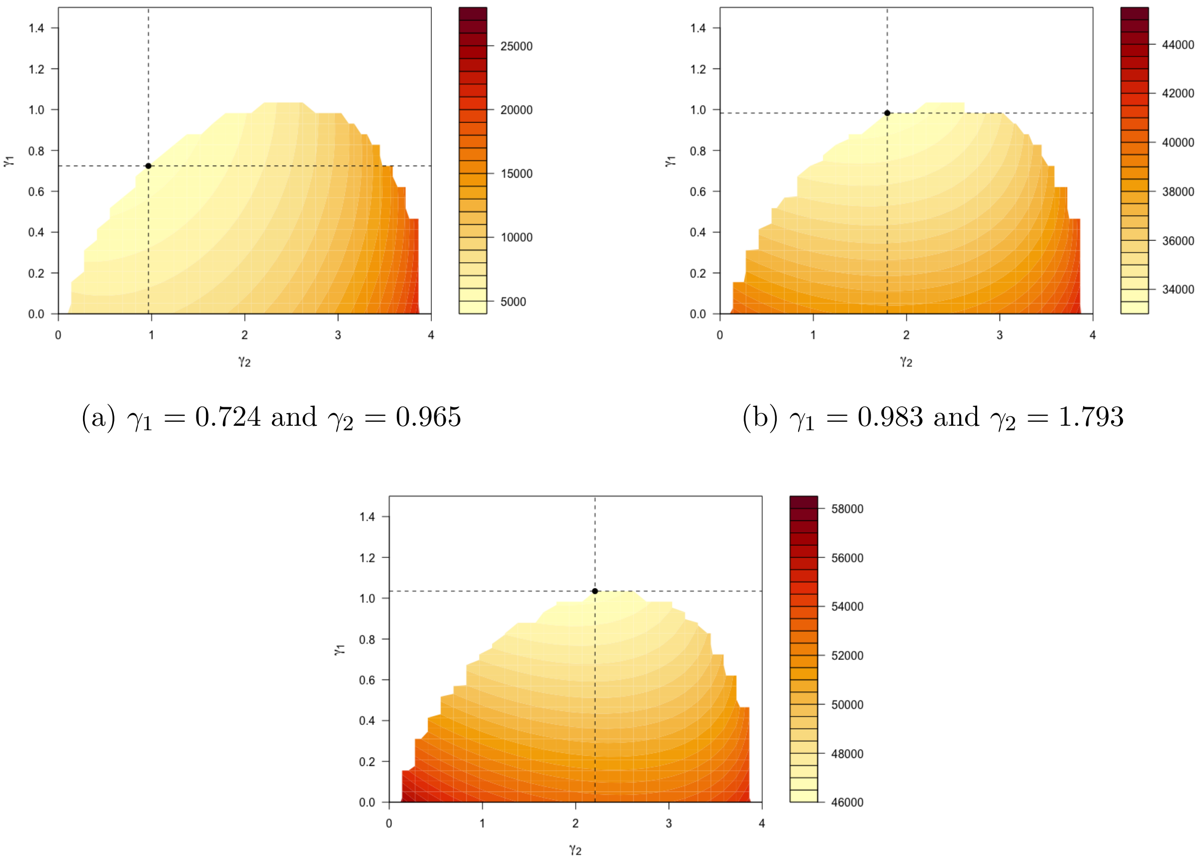


(e) Surface elevation at 70m for $H_{s,0} = 3\text{m}$



(f) Surface elevation at 5m for $H_s = 3.68\text{m}$

Figure 7: Histograms of surface elevation η at 70m and at 5m depth (after the non-linear transfer function has been applied) for waves with $T = 10s$. A Gaussian distribution has been fit to the waves at 70m depth whereas both the Gaussian and Gram-Charlier densities have been fit to the waves at 5m depth.



(a) $\gamma_1 = 0.724$ and $\gamma_2 = 0.965$

(b) $\gamma_1 = 0.983$ and $\gamma_2 = 1.793$

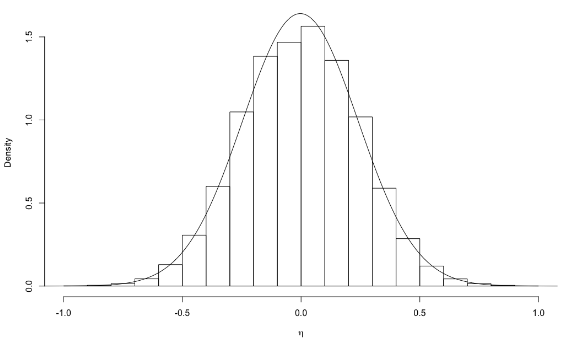
(c) $\gamma_1 = 1.034$ and $\gamma_2 = 2.207$

Figure 8: γ_1, γ_2 plane showing estimated parameters of the Gram-Charlier type-A expansion for waves with $T = 10s$.

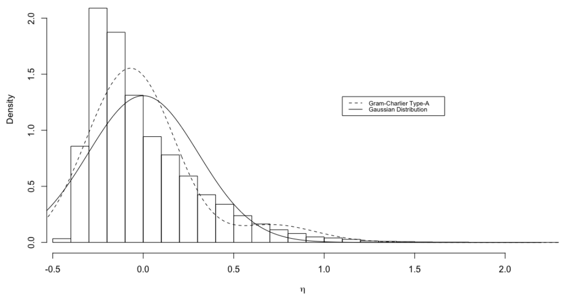
Experiment 3

Table 3: Estimated values of γ_1, γ_2 and σ for simulated sea states with $T = 12s$.

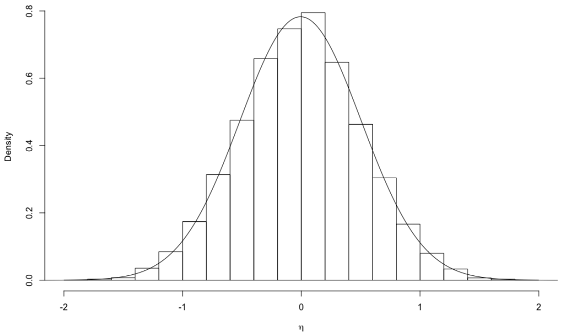
$H_{s,0}$	H_s	γ_1	γ_2	σ	β_1	β_2
1m	1.32m	0.983	1.793	0.328	0.966	4.793
2m	2.94m	1.034	2.207	0.527	1.069	5.207
3m	4.51m	1.034	2.621	0.753	1.069	5.621



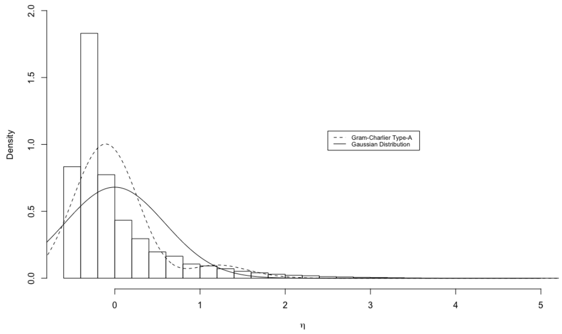
(a) Surface elevation at 70m for $H_{s,0} = 1$m



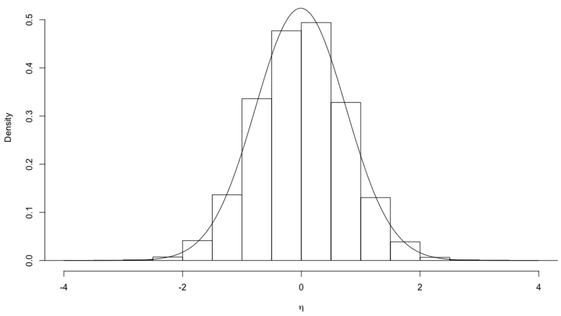
(b) Surface elevation at 5m for $H_s = 1.31$m



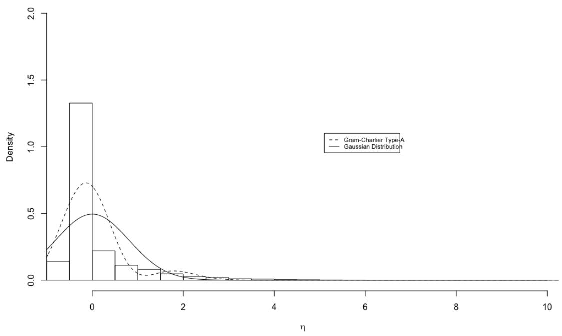
(c) Surface elevation at 70m for $H_{s,0} = 2$m



(d) Surface elevation at 5m for $H_s = 2.94$m

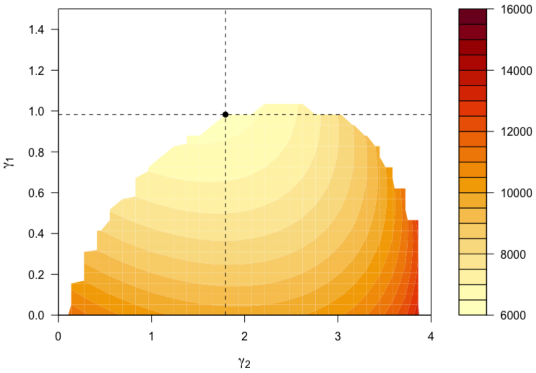


(e) Surface elevation at 70m for $H_{s,0} = 3$m

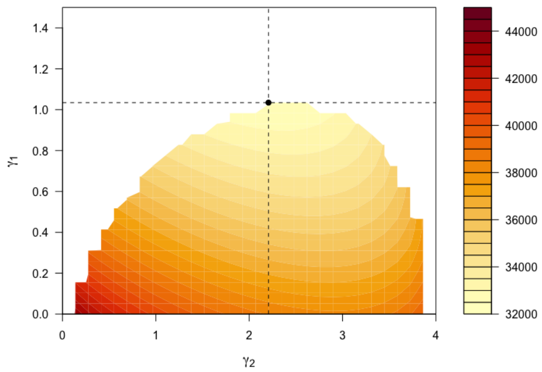


(f) Surface elevation at 5m for $H_s = 4.51$m

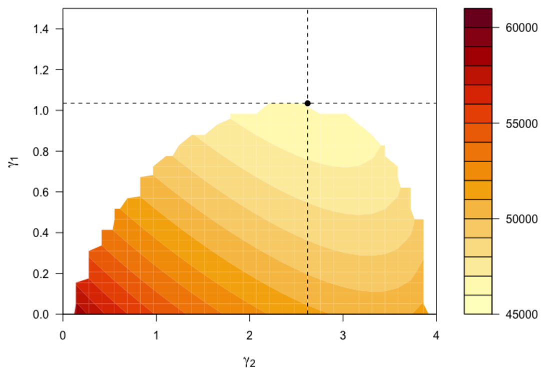
Figure 9: Histograms of surface elevation at 70m and 5m depth for waves with $T = 12s$.



(a) $\gamma_1 = 0.983$ and $\gamma_2 = 1.793$



(b) $\gamma_1 = 1.034$ and $\gamma_2 = 2.207$



(c) $\gamma_1 = 1.034$ and $\gamma_2 = 2.621$

Figure 10: γ_1, γ_2 plane showing estimated parameters of the Gram-Charlier type-A expansion for waves with $T = 12s$.

Figures 4, 7 and 9 show histograms of the free surface elevation. In each figure, subfigures (a), (c) and (e) present the histogram with a Gaussian distribution fit to the data. Similarly, subfigures (b), (d) and (f) present the histogram along with comparisons between a Gaussian distribution and Gram-Charlier type-A distribution fit to the data. In all 3 experiments, the Gaussian distribution (solid line) fits the data well in deep water as anticipated from the Q-Q plots in Figure 5. Regarding the surface elevation in shallow water, the results vary depending on the significant wave height. As can be observed, sea states with a smaller significant wave height are in general better approximated by a Gram-Charlier type-A series. As the significant wave height increases, the surface elevation data becomes excessively skewed which can possibly be explained by the non-linearity of the waves. Recall that the modulus $m \in [0, 1)$ gives periodic waves. For $m = 0$, the solution to the problem is given in terms of (2.1). When the non-linear terms are more dominant however, the parameter m increases and causes a surface deformation in the form of sharper crests and flatter troughs which can be seen in the histograms. In general, the non-linear terms seem more dominant in the sea states with an original significant wave height $H_{s,0} = 3$m. Figure 11 presents the parameter γ_2 as a function of γ_1. As we can see, some scatter can be observed when considering each pair of (γ_1, γ_2) values corresponding to each experiment with a significant wave height H_s. In general however, γ_2 seems to increase with γ_1. For all experiments considered here in Case 1, the wave spectrum (the focus of Case 2) reduces to a delta-function at the frequency under consideration. The results of Case 2 presented in the next subsection are then each an extension to this situation where sea states consisting of several frequencies are investigated.

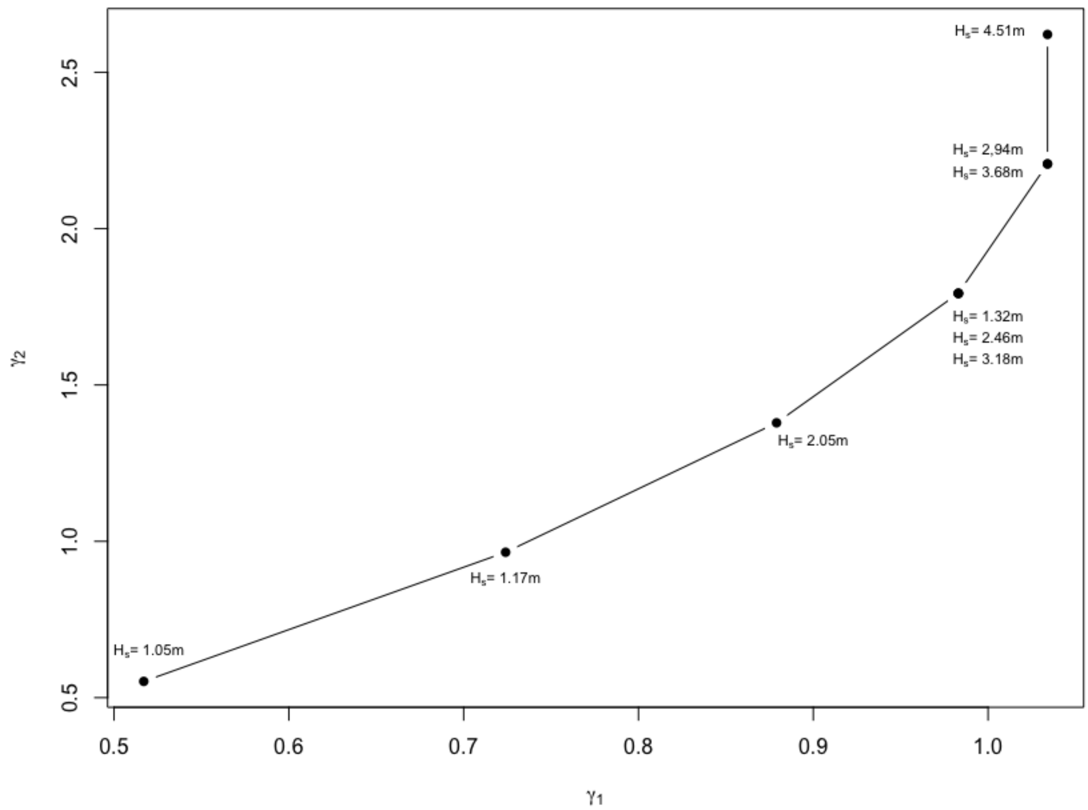
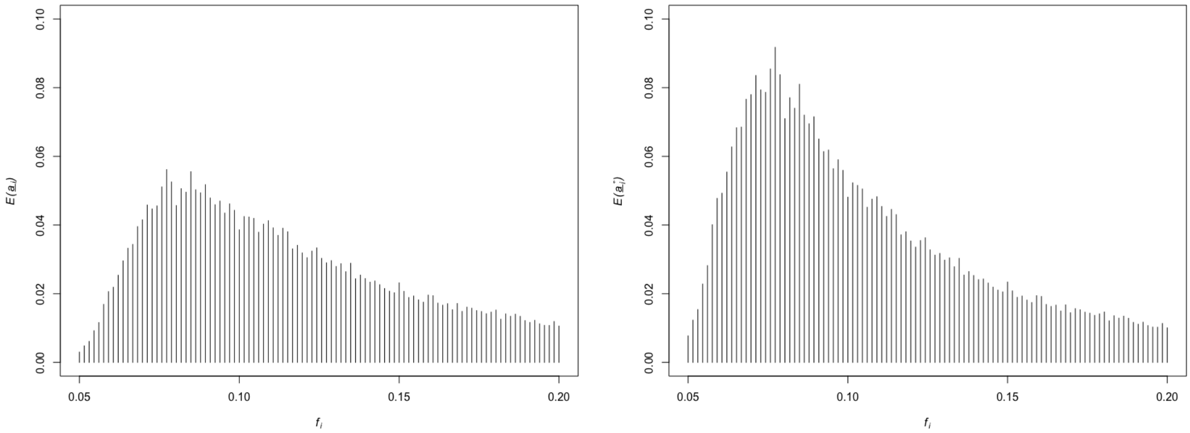


Figure 11: Parameter γ_2 as a function of parameter γ_1.

Case 2: Sea states with waves of several frequencies

In both experiments the deep water depth was again defined as 70m and the coastal depth as 5m. Experiments 1 and 2 were carried out with significant wave heights $H_{s,0} = 1$m and $H_{s,0} = 2$m in deep water, respectively. In both cases, 100 uniformly distributed frequencies f_i were defined such that 0.05 Hz $\leq f_i \leq 0.2$ Hz and $\Delta f_i = 0.001$. The results of both experiments are presented below.

Experiment 1



(a) Amplitude spectrum at 70m, $H_{s,0} = 1$m (b) Amplitude spectrum at 5m, $H_s = 1.41$m

Figure 12: The amplitude spectrum in deep vs. shallow water, i.e. the expected values of the Fourier amplitudes as a function of frequency.

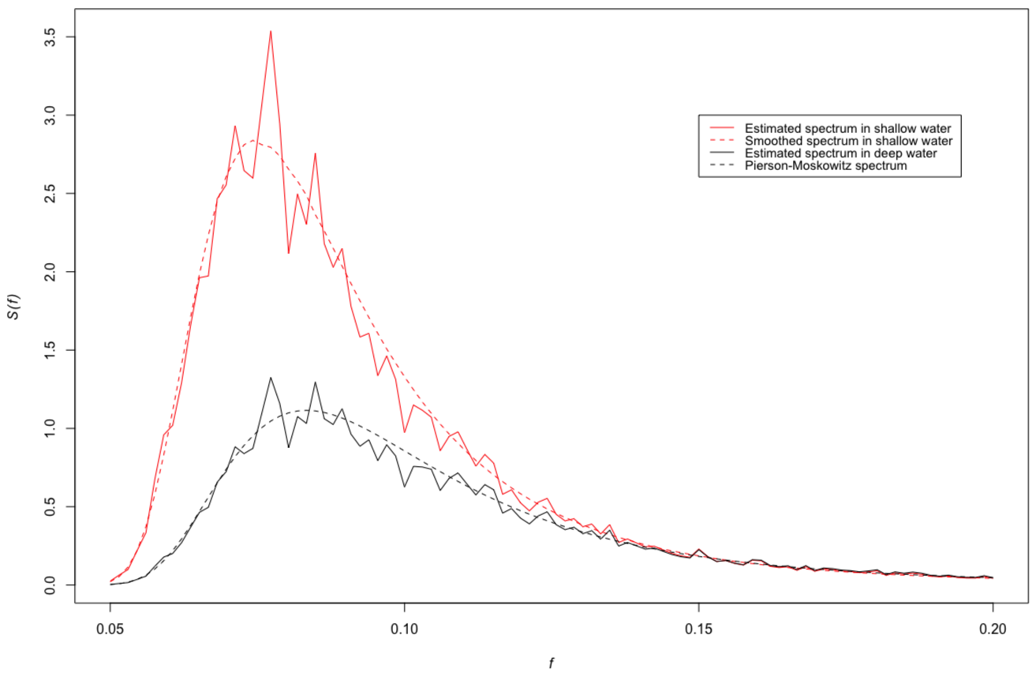
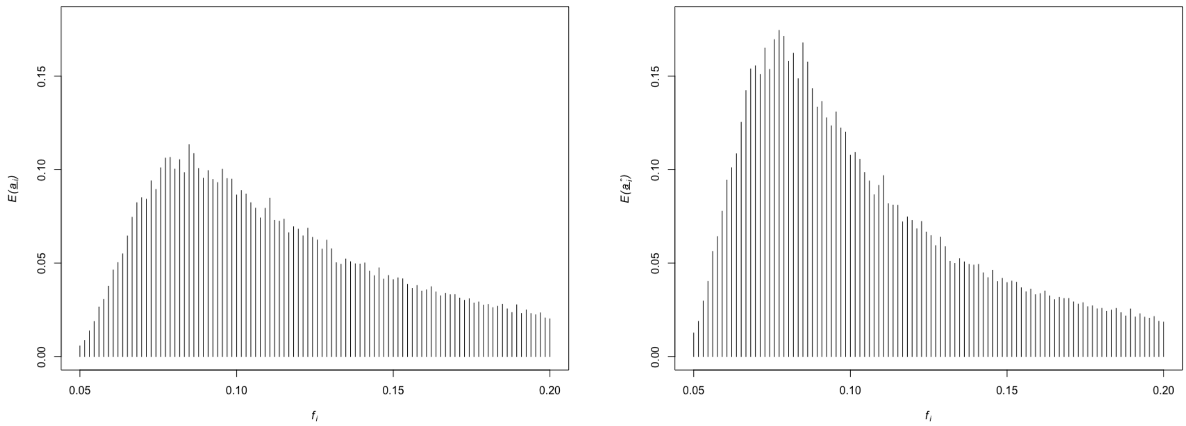


Figure 13: Comparison between P-M spectrum and estimated spectrum in deep water as well as the estimated spectrum spectrum in shallow water for $H_{s,0} = 1$m and $T_p = 12$s.

Experiment 2



(a) Amplitude spectrum at 70m, $H_{s,0} = 2$m (b) Amplitude spectrum at 5m, $H_s = 2.77$m

Figure 14: The amplitude spectrum in deep vs. shallow water, i.e. the expected values of the Fourier amplitudes as a function of frequency.

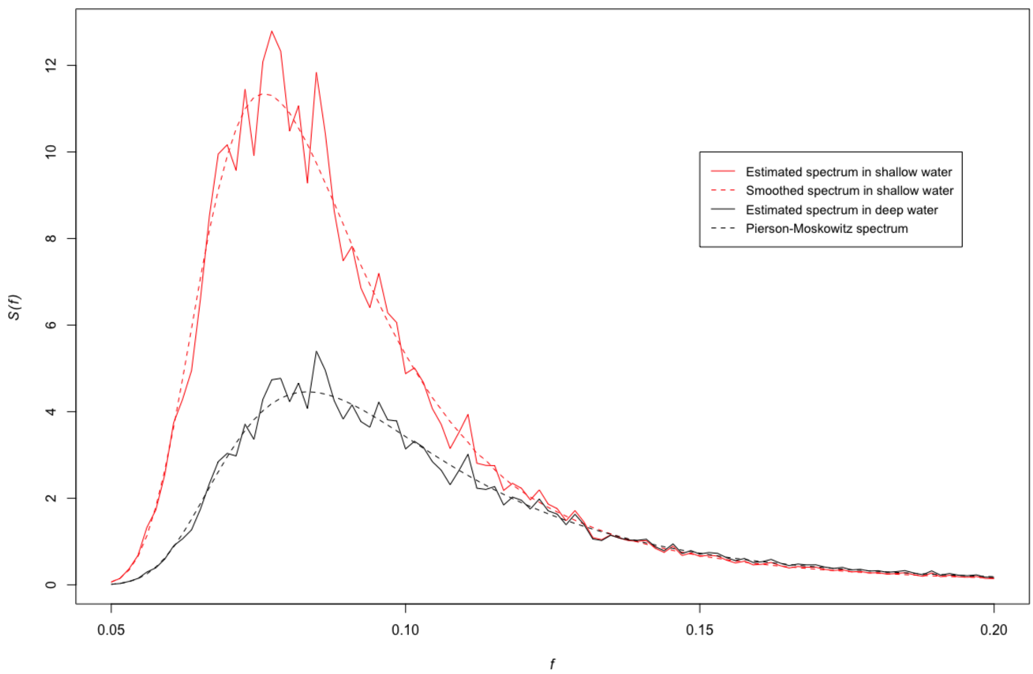


Figure 15: Comparison between P-M spectrum and estimated spectrum in deep water as well as the estimated spectrum spectrum in shallow water for $H_{s,0} = 2$m and $T_p = 12s$.

Figures 12 and 14 show the computed amplitude spectrum in both deep and shallow water for both of the experiments that were carried out, respectively. In both experiments, the expected values of the amplitudes in shallow water $E\{a_i^*\}$ corresponding to the lower frequencies ($f_i \in [0.065, 0.09]$) were amplified to a greater extent than for those corresponding to the higher frequencies. In experiment 1, the significant wave height in shallow water was calculated to be $H_s = 1.41$m, whereas in experiment 2 a significant wave height $H_s = 2.77$m was obtained.

Figures 13 and 15 show the initial P-M spectrum along with its estimate by the procedure in section 4. The estimated and smoothed estimate of the spectrum in shallow water is also presented. As can be seen in both figures the peak frequency T_p^* is slightly shifted to a lower frequency in shallow water in both experiments. This indicates that waves with lower frequencies have a greater contribution to the total variance $\overline{\eta^2}$ in shallow water compared to that of deep water. From a physical perspective, this indicates that waves with lower frequencies have a greater contribution to the total *energy* in shallow water compared to that of deep water.

6 Discussion and Further Work

In the first part of this work, the non-Gaussian characteristics of the free surface elevation in shallow water was investigated for sea states consisting of waves with a single frequency. The wave heights obtained at 5m depth could still be considered Rayleigh distributed and the Gram-Charlier series fit the computed surface elevation data at 5m depth to a satisfactory degree. However, the histograms of the surface elevation became excessively skewed for the sea state considering an initial significant wave height $H_{s,0} = 3$m and period $T = 12s$. A natural extension to these experiments would be the investigation of the limiting sea severity above which the Gram-Charlier series is no longer accurate in describing the distribution of the free surface elevation in shallow waters. A comparison between the Gram-Charlier series and the Tayfun distribution could then be carried out to identify which distribution is most accurate depending on the sea severity. It was also observed that the significant wave height did not change significantly after the non-linear transfer function was applied to the wave height data, although the wave shape had a noticeable change in the form of sharper crests and flatter troughs. Since the model used in this work does not take into account wave breaking, further studies could involve the investigation of a region between the linear region and region dominated by non-linear effects where the waves haven't yet reached breaking point but the significant wave height of the sea state undergoes a noticeable change during the shoaling process.

In the second part of this paper, a scaling of the Fourier amplitudes was proposed to approximate the physical amplitudes of waves belonging to the Pierson-Moskowitz spectrum in deep water. The wave spectrum in shallow water was then estimated using the relation between $S^*(f_i)$ and the scaling parameter σ_i^* of the Fourier amplitudes for 2 cases. A slight shift in the peak frequency was observed in both cases in favour of a lower frequency than that of the peak frequency of the deep water spectrum. In practice, it could be more practical to use real time series data in both deep and shallow water as a means of comparison. A Fourier analysis could be carried out on the deep water time series to identify the Fourier amplitudes before using our proposed scaling to estimate the physical amplitudes. The estimated physical amplitudes could then be used as the input to the non-linear transfer function and thus, the shallow water spectrum estimate can be computed from the output.

Acknowledgments

The authors would like to thank Francesco Fedele for helpful discussions.

References

[1] A. Ali and H. Kalisch. Energy balance for undular bores. *Compt. Rend. Mecanique 338*, pages 67–70, 2010.

[2] A. Ali and H. Kalisch. Mechanical balance laws for boussinesq models of surface water waves. 2012.

[3] A. Ali and H. Kalisch. On the formulation of mass, momentum and energy conservation in the kdv equation. 2013.

[4] B. D. E and K. E. Dennis. The conditions under which gram-charlier and edgeworth curves are positive definite and unimodal. 1952.

[5] F. Fedele and F. Arena. Weakly non-linear statistics of high random waves. 2005.

[6] E. Jondeau and M. Rockinger. Gram-charlier densities. 1999.

[7] Z. Khorsand and H. Kalisch. On the shoaling of solitary waves in the kdv equation. 2014.

[8] Kundu, Pijush K. *Fluid Mechanics*. Elsevier Inc., 2016.

[9] Leo H. Holthuijsen. *Waves in Oceanic and Coastal Waters*. Cambridge University Press, 2007.

[10] M. K. Ochi, S. Malakar, and W.-C. Wang. Statistical analysis of coastal waves observed during the arsloe project. 1982.

[11] M. K. Ochi and W.-C. Wang. Non-gaussian characteristics of coastal waves. 1984.

[12] L. Ostrovskiy and Pelinovskiy. *Wave transformation on the surface of a fluid on variable depth*, pages 552–555. E.N., 1970.

[13] Paolo Boccotti. *Wave Mechanics for Ocean Engineering*. Elsevier Science, 2000.

[14] M. Paulsen and H. Kalisch. A nonlinear formulation of radiation stress and applications to cnoidal shoaling. 2021.

[15] M. O. Paulsen. On the relation between wave conditions and mathematical properties of some asymptotic water wave models. 2020.

[16] M. Pawlowski. Sea spectra revisited. 2011.

[17] Robert M. Sorensen. *Basic wave mechanics: for coastal and ocean engineers*. Wiley, 1993.

[18] I. Svendsen and Brink-Kjær. *Shoaling of cnoidal waves*, pages 365–383. O., 1972.

[19] M. A. Tayfun. Narrow-band non-linear sea waves. *Journal of Geophysical Research: Oceans*, 85(C3):1548–1552, 1980.

Chapter 5

Time Series Analysis

In this chapter, a description of a zero-crossing analysis is presented before applying said analysis to real time series data and carrying out a statistical analysis on par with the analysis carried out in Case 1 of our submitted paper in Chapter 4.

5.1 Zero-Crossing Analysis

In a time record, the surface elevation is the instantaneous elevation of the sea surface at an arbitrary moment in time relative to a reference level [13]. Individual waves can then be defined as the as profile of the surface elevation between two consecutive upward zero-crossings (when zero is the reference level under consideration). Figure 5.1 shows a time record of the surface elevation with the arrows indicating upward zero-crossings. The surface elevation profile between the 2 points marking the upward zero-crossings is then accordingly one wave in the time series.

Approximating the surface elevation at a specific moment in time t_i is then done by first defining the height of the wave as the distance between the wave crest H_{crest} and wave trough H_{trough} so that

$$a_i = \frac{H_{\text{crest},i} - H_{\text{trough},i}}{2}. \tag{5.1}$$

The surface elevation at time t is then given by

$$\eta(t_i) = a_i \cos\left(2\pi f_i t_i + \alpha_i\right) \tag{5.2}$$

where f_i can be found by computing the wave period T_i, i.e. the time between 2 consecutive upward zero-crossing points so that $f_i = 1/T_i$.

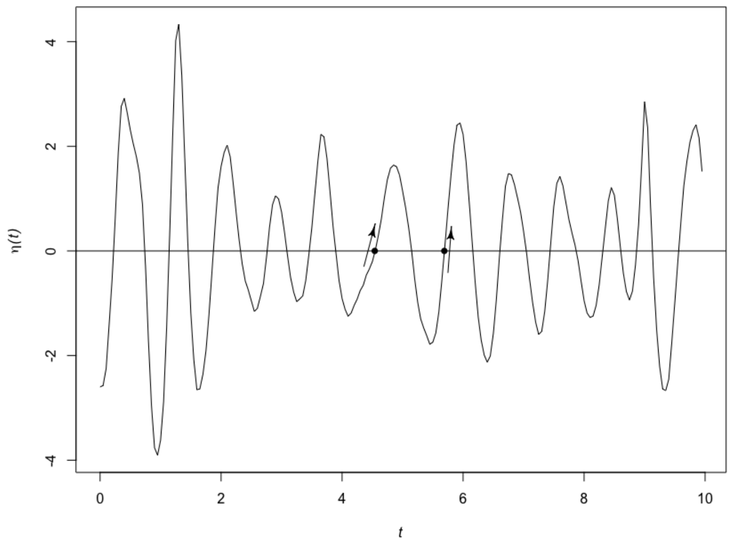


Figure 5.1: Time record of the surface elevation $\eta(t)$ with upward zero-crossings.

5.1.1 Application in MATLAB

An overview of how the zero-crossing analysis was carried out is now shown by the pseudo codes at the end of this chapter. First, the start and end points of the analysis were defined in Algorithm 1 to ensure that the first and last wave in the time series both undergo a full oscillation. Further, Algorithm 2 shows the determination of upward and and downward zero-crossing as well as the wave crest and trough of each wave in the time series. Each wave period T_i was then determined by calculating the time difference between 2 consecutive upward zero-crossings. Figure 5.2 shows the results of the application for the first 100 data points in the time series obtained at WG.1 for resolution purposes.

5.2 Experiment and results

The data under investigation was obtained by measuring the water surface elevation at 2 wave gauges and corresponds to wave gauge 1 and wave gauge 7 (WG.1 $\sim$ WG.7) in the study carried out in [15]. The measurements at WG.1 and WG.7 are at water depths of 47cm and 15cm over a 1/20 plane beach, respectively. Figure 5.3 shows histograms of the measured surface elevation η at both gauges, where a Guassian distribution has been fit to the data. As can be observed, the Gaussian distribution fits the surface elevation data at 47cm (WG.1) better than that of WG.7 at 15cm.

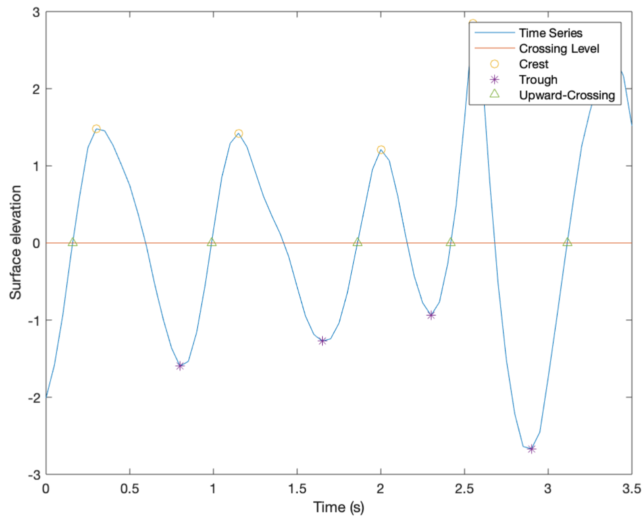
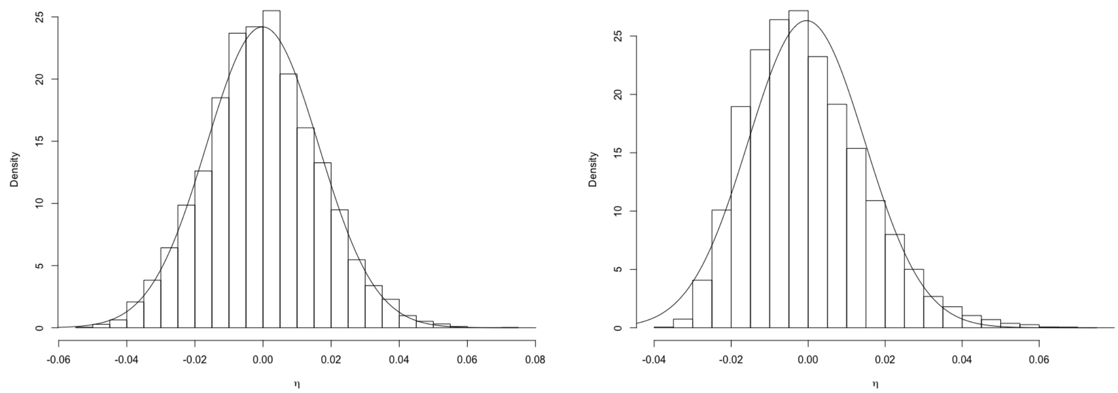


Figure 5.2: Results of application of upward-zero-crossing analysis on surface elevation time series at WG.1 in MATLAB.

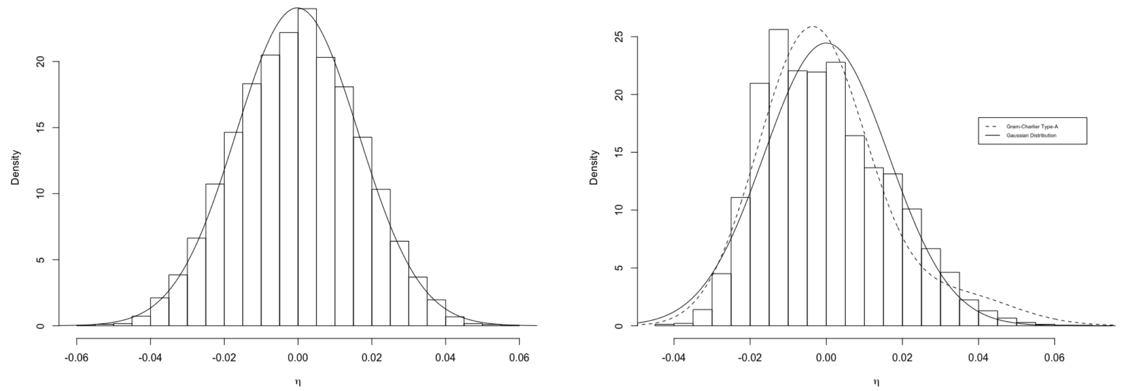


(a) Histogram of measured surface elevation at 47cm.

(b) Histogram of measured surface elevation at 15cm

Figure 5.3: Histograms of surface elevation measurements fit with a Gaussian distribution at 47cm (WG.1) and 15cm (WG.7) depth, respectively.

The non-linear transfer function in [20] was then applied to the wave heights obtained from the upward-zero-crossing analysis and the surface elevation was computed in the same manner as described in section 4 of our submitted paper. Namely, computing the surface elevation using either (2.18) or (2.53) depending on the nature of the wave. The results are show in Figure 5.4. Subfigure (a) presents the histogram of the surface elevation with a Gaussian distribution fit to the data. Similarly, subfigure (b) presents the histogram along with comparisons between a Gaussian distribution and Gram-Charlier type-A distribution fit to the data. As can be observed in subfigure (a), the Gaussian distribution (solid line) fits the data reasonably well in at 47cm depth as can also be seen from the Q-Q plots in Figure 5.5. Regarding the surface elevation in shallow water, the parameter vector $\boldsymbol{\theta}$ in the Gram-Charlier expansion (see section 4 of submitted paper) was found to be $\boldsymbol{\theta} = [0.7105, 1.0526, 0.1]$. Both histograms represent the original measured data reasonably well, although some deviations are to be expected due to our assumption that each wave is a single harmonic.

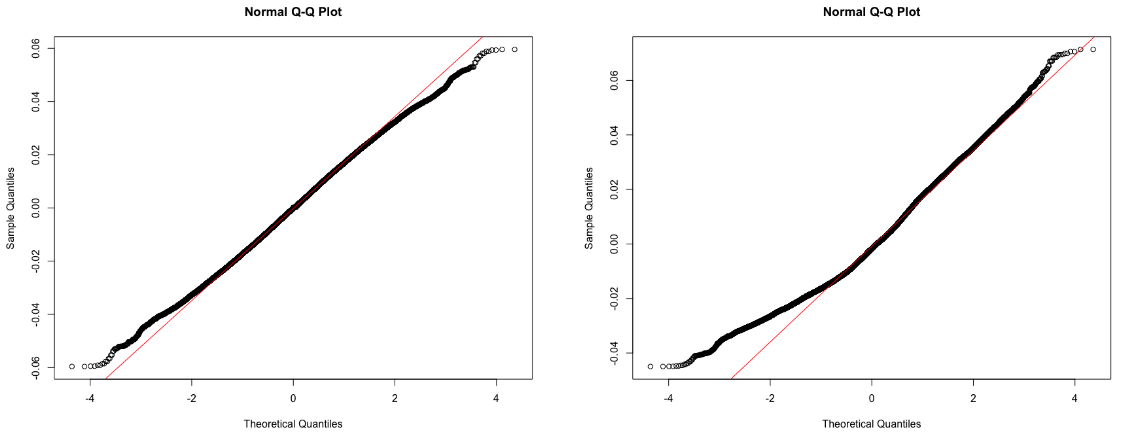


(a) Surface elevation η at 47cm.

(b) Surface elevation η at 15cm.

Figure 5.4: Histograms of surface elevation η at 47cm and at 15cm depth (after the non-linear transfer function has been applied) for waves of several frequencies. A Gaussian distribution has been fit to the waves at 47cm depth whereas both the Gaussian and Gram-Charlier densities have been fit to the waves at 15cm depth.

As mentioned, Figure 5.5 presents Q-Q plots of the free surface elevation data at both gauge depths. The surface elevation at 47cm depth follows the normal line reasonably well with only small deviations at the end points. This again indicates that the free surface elevation in deep water has a "thin-tailed" distribution. In these cases, the Q-Q plot of the distribution has small or negligible deviations at the ends. Thus, the surface elevation at 47cm depth can still be represented by a Gaussian distribution to a satisfactory degree. At 15cm depth, the deviations from the normal line are larger at the end points indicating a distribution at 15cm that is more skewed. This extra skewness can be accounted for by using the Gram-Charlier expansion to fit the data rather than a Gaussian distribution as shown in Figure 5.4.



(a) Q-Q plot of surface elevation at 47cm. (b) Q-Q plot of surface elevation at 15cm.

Figure 5.5: Q-Q plots of surface elevation η at 47cm and 15cm depth.

5.3 Discussion and Further Work

In this chapter we presented the method of an upward zero-crossing analysis and applied it to real time series data. The wave heights obtained were then used as the input to the non-linear transfer function and the resulting histograms yielded good agreement with the original measured surface elevation data. A Gram-Charlier expansion was then fit to the surface elevation at WG.7 (15cm depth), taking into account the skewing of the surface elevation data during the shoaling process and the results were shown in Figure 5.4.

An extension to this work could be the Fourier analysis of the time series obtained at both wave gauges to obtain the Fourier amplitudes. In this way, the wave spectra at both depths could be computed and a comparison could be carried out to identify any major differences like the downshifting/up-shifting of the wave energy, changes in peak frequency etc. Since we only considered 2 gauges in this chapter, the analysis considered here could be applied to data obtained at several of the other wave gauges to see the range of depths for which the Gram-Charlier expansion is accurate in representing the distribution of the free surface elevation. However, it should be noted that at Gauge 8 and beyond, the wave experiments carried out in [15] feature wave breaking. In principle, it would at least be possible to incorporate wave breaking in the non-linear shoaling code using the approach detailed in [6] and [4].

Algorithm 1: Upward zero-crossing starting points

Input: Surface elevation time series data η, sampling frequency f_s
Output: Wave height and wave period data
if $\eta_{start} == 0$ *and* $\eta_{start+1} > 0$ **then**
    Denote first upward zero-crossing at this point in time, t_{start}
end
if $\eta_{start} == 0$ *and* $\eta_{start+1} < 0$ **then**
    for *i = 1,...,end-1* **do**
        if $\eta_i < 0$ *and* $\eta_{i+1} > 0$ **then**
            Denote first upward zero-crossing at this point in time, t_i
        end
    end
end
if $\eta_{end} == 0$ *and* $\eta_{end-1} < 0$ **then**
    Denote last upward zero-crossing at this point in time, t_{end}
end
if $\eta_{end} == 0$ *and* $\eta_{end-1} > 0$ **then**
    for *i = end - 1,...,1* **do**
        if $\eta_i < 0$ *and* $\eta_{i+1} > 0$ **then**
            Denote last upward zero-crossing at this point in time, t_i
        end
    end
end

Algorithm 2: Upward zero-crossing

Input: Surface elevation time series data η, sampling frequency f_s
Output: Wave height and wave period data
for *i = index of first upward crossing point,..., index of last upward crossing point* **do**
    if $\eta_i < 0$ *and* $\eta_{i+1} > 0$ **then**
        Store values and index of upward zero-crossing point
    end
    if $\eta_i > 0$ *and* $\eta_{i+1} < 0$ **then**
        Store values and index of downward zero-crossing point
    end
end
for *i = indexes of upward crossing points* **do**
    if $\eta_i > \eta_{i-1}$ *and* $\eta_i > \eta_{i+1}$ *and* $\eta_i > 0$ **then**
        Denote this point as the crest of the wave
    end
    if $\eta_i < \eta_{i-1}$ *and* $\eta_i < \eta_{i+1}$ *and* $\eta_i < 0$ **then**
        Denote this point as the trough of the wave
    end
end

Bibliography

[1] Adrian Barbu and Song-Chun Zhu. *Monte Carlo Methods*. Springer, 2020.

[2] A. Ali and H. Kalisch. Energy balance for undular bores. *Compt. Rend. Mecanique 338*, pages 67–70, 2010.

[3] A. Ali and H. Kalisch. On the formulation of mass, momentum and energy conservation in the kdv equation. 2013.

[4] M. Bjørkavåg and H. Kalisch. Wave breaking in boussinesq models for undular bores. *Physics Letters A, 375(14)*, pages 1570–1578, 2011.

[5] T. Brenn and S. N. Anfinsen. A revisit of the gram-charlier and edgeworth series expansions. 2017.

[6] M. K. Brun and H. Kalisch. Convective wave breaking in the kdv equation. *Analysis and Mathematical Physics, 8(1)*, pages 57–75, 2018.

[7] B. D. E and K. E. Dennis. The conditions under which gram-charlier and edgeworth curves are positive definite and unimodal. 1952.

[8] T. T. Dirk P. Kroese, Tim Brereton and Z. I. Botev. Why the monte carlo method is so important today. 2014.

[9] F. Fedele and F. Arena. Weakly non-linear statistics of high random waves. 2005.

[10] E. Jondeau and M. Rockinger. Gram-charlier densities. 1999.

[11] Z. Khorsand and H. Kalisch. On the shoaling of solitary waves in the kdv equation. 2014.

[12] Kundu, Pijush K. *Fluid Mechanics*. Elsevier Inc., 2016.

[13] Leo H. Holthuijsen. *Waves in Oceanic and Coastal Waters*. Cambridge University Press, 2007.

[14] M. S. Longuet-Higgins. On the statistical distribution of the heights of sea waves. 1953.

[15] H. Mase and J. T. Kirby. Hybrid frequency-domain kdv equation for random wave transformation. 1992.

[16] M. K. Ochi, S. Malakar, and W.-C. Wang. Statistical analysis of coastal waves observed during the arsloe project. 1982.

[17] M. K. Ochi and W.-C. Wang. Non-gaussian characteristics of coastal waves. 1984.

[18] L. Ostrovskiy and Pelinovskiy. *Wave transformation on the surface of a fluid on variable depth*, pages 552–555. E.N., 1970.

[19] Paolo Boccotti. *Wave Mechanics for Ocean Engineering.* Elsevier Science, 2000.

[20] M. O. Paulsen. On the relation between wave conditions and mathematical properties of some asymptotic water wave models. 2020.

[21] M. Pawlowski. Sea spectra revisited. 2011.

[22] Pinsky, M. A. and Karlin, S. *An Introduction to Stochastic Modeling.* Elsevier Inc., 2011.

[23] L. R. Shenton. Efficiency of the method of moments and the gram-charlier type a distribution. 1951.

[24] I. Svendsen and Brink-Kjær. *Shoaling of cnoidal waves*, pages 365–383. O., 1972.

[25] Y. J. Y. Young Myung Choi and S. H. Kwon. Validity of ocean wave spectrum using rayleigh probability density function. 2012.

Appendices

Appendix A

Random variables

A.1 Random variables

Random variables are defined to be any variable whose value cannot be predicted and they are fully characterised by their PDF $p(z)$. Consider a random variable $\underline{z}$. The probability of $\underline{z}$ acquiring a value between z and $z + dz$ is given by

$$Pr\{z < \underline{z} \leq z + \mathrm{d}z\} = \int_{z}^{z+\mathrm{d}z} p(z)\mathrm{d}z \tag{A.1}$$

The probability of $\underline{z}$ taking on a value less than or equal to z is then defined using its CDF $P(z)$ and can be written mathematically as

$$P(z) = \Pr\{\underline{z} \leq z\} = \int_{-\infty}^{z} p(z)dz \tag{A.2}$$

A.1.1 Estimation

Considering a set of sample values (an *ensemble*) of the random variable $\underline{z}$ and the notation $\langle . \rangle$ denoting the ensemble average, the following holds for the mean and standard deviation of $\underline{z}$ respectively:

$$\mu_z \approx \langle \underline{z} \rangle = \frac{1}{N} \sum_{i=1}^{N} \underline{z}_i \tag{A.3}$$

$$\sigma_z^2 \approx \langle (\underline{z} - \langle \underline{z} \rangle)^2 \rangle = \frac{1}{N} \sum_{i=1}^{N} (\underline{z}_i - \langle \underline{z} \rangle)^2 = \frac{1}{N} \sum_{i=1}^{N} \langle \underline{z}_i \rangle^2 - \langle \underline{z} \rangle^2 \tag{A.4}$$

A.1.2 Moments

The moments of a function are quantitative measures related to the shape of the function's graph. The nth-order moment, m_n, of $p(z)$ about a value c can be defined as

$$m_n = \int_{-\infty}^{\infty} (z-c)^n p(z) dz. \tag{A.5}$$

If $c = 0$, the nth moment is called a *raw* moment and a *central* moment if $c = \mu$. Then, the following statements can be made about the zeroth, first-, and second-order moments:

The zeroth raw moment m_0 of any PDF is 1 since

$$\Pr\{\underline{z} \leq \infty\} = P(\infty) = \int_{-\infty}^{\infty} p(z) dz = 1$$

The first raw moment m_1 is known as the *mean* or the *expected value* of $\underline{z}$ and can be written as

$$m_1 = \mu_z = E\{\underline{z}\} = \int_{-\infty}^{\infty} z p(z) dz$$

The second central moment, m_2, is the variance σ^2 of $\underline{z}$. It can be defined as

$$m_2 = \sigma_z^2 = E\{(\underline{z} - \mu_z)^2\} = \int_{-\infty}^{\infty} (z - \mu_z)^2 p(z) dz = E\{\underline{z}^2\} - \mu_z^2 = m_2 - m_1^2$$

In addition to these definitions, the third- and fourth-order moments are used to describe the skewness and kurtosis of a probability density function, respectively.

A.1.3 Relationship between the wave spectrum and scaling parameter of Rayleigh distributed Fourier amplitudes

We begin by considering the standard form of the Rayleigh probability density function of the random variable $\underline{z}$ is given by

$$p(z, \sigma) = \frac{z}{\sigma^2} \exp\left(-\frac{z^2}{2\sigma^2}\right), \quad z \geq 0 \tag{A.6}$$

where σ is the scaling parameter of the distribution. Now, the probability distribution of a Rayleigh distributed Fourier amplitude at a specific frequency f_i was given in (3.4). Namely,

$$p(a_i) = \frac{\pi}{2}\frac{a_i}{\mu_i^2}\exp\left(-\frac{\pi a_i^2}{4\mu_i^2}\right). \tag{A.7}$$

The following expression for σ_i^2 is then easily determined as

$$\sigma_i^2 = \frac{2\mu_i^2}{\pi}. \tag{A.8}$$

Now, the variance of $\underline{a}_i$ is given by

$$Var(\underline{a}_i) = \frac{4-\pi}{2}\sigma_i^2 = \frac{4-\pi}{2}\left(\frac{2\mu_i^2}{\pi}\right) = E\{\underline{a}_i^2\} - (E\{\underline{a}_i\})^2. \tag{A.9}$$

where $(E\{\underline{a}_i\})^2$ and $E\{\underline{a}_i^2\}$ are the first- and second raw moments, respectively. Rearranging for the second raw moment gives

$$E\{a_i^2\} = Var(a_i) + (E\{a_i\})^2 = \frac{2\mu_i^2}{\pi}\left(\frac{4-\pi}{2} + \frac{\pi}{2}\right) = \frac{4\mu_i^2}{\pi}. \tag{A.10}$$

Recall the variance density spectrum (3.5):

$$S(f_i) = \frac{1}{\Delta f_i}E\left\{\frac{1}{2}\underline{a}_i^2\right\}, \quad \forall f_i. \tag{A.11}$$

By substituting the expression for $E\{\underline{a}_i^2\}$ and rearranging, the following for the scale parameter σ holds:

$$\frac{4\mu_i^2}{\pi} = 2S(f_i)\Delta f_i = 2\sigma_i^2 \tag{A.12}$$

so,

$$\sigma_i = \sqrt{S(f_i)\Delta f_i}. \tag{A.13}$$

A.2 Stochastic processes

A ***stochastic process*** is a family of random variables $\underline{z}_{t_i}$, where t is a parameter running over a suitable index set t. Often, the index t corresponds to discrete units of time so that the index set t is $t = \{0, 1, 2, ...\}$ [22]. Then $\underline{z}_{t_1}$ corresponds to z at $t = 0$. It can also be convenient to write $\underline{z}(t_1)$ or z_1 to denote the same variable.

A fitting example of a stochastic process in one-dimension can be visualized by considering wind-generated surface waves. Let the index t start at $t = 0$ and the set of surface elevations $\underline{\eta}$ be observed at a location O over a period of time. The random variable $\underline{\eta}$ at time t_1 has a different value than $\underline{\eta}$ at t_2 and $\underline{\eta}$ at t_3 etc. since the values are random. This set $\underline{\eta}(t_1), \underline{\eta}(t_2), \underline{\eta}(t_3), ..., \underline{\eta}(t_i)$, is one realization of the stochastic process and can be repeated to obtain several realizations.

Appendix B

Hermite polynomials

Like other orthogonal polynomials, Hermite polynomials can be defined in several different ways. Here, the following definition of the Probabilists' Hermite polynomial is considered and is given by

$$H_n(x) = (-1)^n e^{\frac{x^2}{2}} \frac{d^n}{dx^n} e^{-\frac{x^2}{2}} \tag{B.1}$$

The first eleven Hermite polynomials can be found by straightforward computations and are:

$$\begin{aligned}
H_0(x) &= 1, \\
H_1(x) &= x, \\
H_2(x) &= x^2 - 1, \\
H_3(x) &= x^3 - 3x, \\
H_4(x) &= x^4 - 6x^2 + 3, \\
H_5(x) &= x^5 - 10x^3 + 15x \\
H_6(x) &= x^6 - 15x^4 + 45x^2 - 15, \\
H_7(x) &= x^7 - 21x^5 + 105x^3 - 105x, \\
H_8(x) &= x^8 - 28x^6 + 210x^4 - 430x^2 + 105, \\
H_9(x) &= x^9 - 36x^7 + 378x^5 - 1260x^3 + 945x, \\
H_{10}(x) &= x^{10} - 45x^8 + 630x^6 - 3150x^4 + 4725x^2 - 945
\end{aligned} \tag{B.2}$$

In general,

$$H_n(x) = x^n - \frac{n(n-1)}{1!}x^{n-2} + \frac{n(n-1)(n-2)(n-3)}{2!}x^{n-4} - +... \tag{B.3}$$

Differentiating (B.3) gives

$$H_n'(x) = n\left[x^{n-1} - \frac{(n-1)(n-2)}{1!}x^{n-3} + \frac{(n-1)(n-2)(n-3)(n-4)}{2!}x^{n-5} - +...\right]$$

i.e. (B.1) obeys the differentiation rule

$$H_n'(x) = nH_{n-1}(x). \tag{B.4}$$

B.1 Orthogonality

In this section we will prove that Hermite polynomials form an orthogonal set with respect to the weight function

$$w(x) = e^{-\frac{x^2}{2}}. \tag{B.5}$$

We begin showing this by defining the Kronecker delta function

$$\delta_{mn} = \begin{cases} 0, & m \neq n \\ 1, & m = n. \end{cases} \tag{B.6}$$

The n-th order polynomial is then orthogonal with respect to $w(x)$ so that

$$\int_{-\infty}^{\infty} H_m(x)H_n(x)w(x)dx = \sqrt{2\pi}n!\delta_{mn} \tag{B.7}$$

Substituting (B.1) for $H_n(x)$ gives

$$\int_{-\infty}^{\infty} H_m(x)H_n(x)w(x)dx = (-1)^n \int_{-\infty}^{\infty} H_m(x)\frac{d^n e^{-\frac{x^2}{2}}}{dx^2}dx$$

and integration by parts for $m \neq n$ yields

$$(-1)^n \int_{-\infty}^{\infty} H_m(x)\frac{d^n e^{-\frac{x^2}{2}}}{dx^2}dx = (-1)^n \left[H_m(x)\frac{d^{n-1}e^{-\frac{x^2}{2}}}{dx^{n-1}}\Bigg|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} H_m'(x)\frac{d^{n-1}e^{-\frac{x^2}{2}}}{dx^{n-1}}dx \right].$$

Observe that the first term after the equality is zero because $e^{-\frac{x^2}{2}}$ and its derivatives are zero at $\pm\infty$. Using (B.4) to rewrite the second term after the equality gives

$$(-1)^{n+1}\int_{-\infty}^{\infty} H'_m(x)\frac{d^{n-1}e^{-\frac{x^2}{2}}}{dx^{n-1}}dx = m\int_{-\infty}^{\infty} H_{m-1}(x)\frac{d^{n-1}e^{-\frac{x^2}{2}}}{dx^{n-1}}dx.$$

Integrating by parts a second time yields

$$(-1)^{n+2}m(m-1)\int_{-\infty}^{\infty} H_{m-2}(x)\frac{d^{n-2}}{dx^{n-2}}.e^{-\frac{x^2}{2}}\,dx$$

This implies that after integrating by parts m times we get

$$(-1)^{n+m}m!\int_{-\infty}^{\infty} H_0(x)\frac{d^{n-m}}{dx^{n-m}}e^{-\frac{x^2}{2}}\,dx.$$

Recall from (B.2) that $H_0(x) = 1$. Substituting and evaluating the integral leads to

$$(-1)^{n+m}m!\left[1\cdot\frac{d^{n-m-1}}{dx^{n-m-1}}e^{-\frac{x^2}{2}}\Bigg|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} 0\cdot\frac{d^{n-m-1}}{dx^{n-m-1}}e^{-\frac{-x^2}{2}}\,dx\right] = 0$$

For $m = n$ we follow the same procedure and obtain

$$\int_{-\infty}^{\infty}\big(H_n(x)\big)^2 e^{-\frac{x^2}{2}}\,dx = (-1)^{2n}n!\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}\,dx = \sqrt{2\pi}n!$$

and (B.7) is satisfied.