

On the Non-Dimensionalisation, Scaling and Resulting Interpretation of the Classical Governing Equations for Water Waves

Adrian CONSTANTIN^a and Robin Stanley JOHNSON^b

^a *School of Mathematics, Trinity College, Dublin 2, Ireland
E-mail: adrian@maths.tcd.ie*

^b *School of Mathematics & Statistics, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, United Kingdom
E-mail: r.s.johnson@ncl.ac.uk*

Received on April 6, 2008

Abstract

In this note we describe the underlying principles — and pitfalls — of the process of non-dimensionalising and scaling the equations that model the classical problem in water waves. In particular, we introduce the two fundamental parameters (associated with amplitude and with wave length) and show how they are used, independently, to represent different approximations (with corresponding different interpretations and applications). In addition, and most importantly, we analyse how these two parameters play a role in the derivation of the Korteweg-de Vries (KdV) equation, which then lead to predictions for the regions of physical space where solitons might be expected to appear. In particular, we address the issue of whether KdV theory can be used effectively to predict tsunamis. We argue that for tsunamis the propagation distances are much too short for KdV dynamics to develop.

1 The governing equations

We shall assume that the relevant equations and boundary conditions are: the Euler equation, the equation for incompressible fluids, surface and bottom kinematic conditions, and a dynamic condition — constant pressure — at the surface. All these together constitute the classical water-wave problem; we note, in particular, that the fluid is inviscid (but may be rotational) and that surface-tension effects are ignored. These assumptions reflect physical reality. Indeed, incompressibility is a physically reasonable assumption for water [22]. Experimental evidence shows that the length scales associated with an adjustment of the fluid velocity distribution due to laminar viscosity are long compared to typical wavelengths encountered in sea waves [10]. Also, for sea waves that are not near the breaking stage, the effects of friction due to turbulent mixing viscosity are also known to be negligible [2], so that we may reasonably neglect viscosity altogether. The effects of surface tension are negligible for wave lengths greater than a few centimeters [13, 22], the

typical wave length of a wind-generated wave at sea being about 150 m [2]. Therefore the balance between gravity and the inertia of the system is the major factor governing the evolution of sea waves from their initial profile.

We describe the governing equations in rectangular Cartesian coordinates, (x, y, z) , with the corresponding velocity field $\mathbf{u} \equiv (u, v, w)$, where each component depends on $\mathbf{x} \equiv (x, y, z)$ (with z measured vertically upwards) and time, t ; the pressure is P , ρ is the (constant) density of the fluid, and g is the constant acceleration of gravity. The governing equations are therefore [13]

$$\begin{cases} \frac{D\mathbf{u}}{Dt} = -\frac{1}{\rho} \nabla P + \mathbf{F} & (\text{where } \mathbf{F} \equiv (0, 0, -g)), \\ \nabla \cdot \mathbf{u} = 0 \end{cases} \quad (1.1)$$

and the corresponding boundary conditions are

$$\begin{cases} P = P_a = \text{constant}, \\ w = \frac{\partial H}{\partial t} + u \frac{\partial H}{\partial x} + v \frac{\partial H}{\partial y}, \end{cases} \quad \text{both on } z = H(x, y, t), \quad (1.2)$$

and

$$w = u \frac{\partial B}{\partial x} + v \frac{\partial B}{\partial y} \quad \text{on } z = B(x, y). \quad (1.3)$$

(We have used the familiar material-derivative operator: $\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla$.) Here, $H = H(x, y, t)$ is the free surface of the fluid, and $B = B(x, y)$ is the fixed, impermeable bed. (It is possible to include time in the function describing the bottom topography; this then enables, for example, the time evolution of marine quakes to be included in the model.) The second boundary condition in (1.2), and the boundary condition (1.3), ensure that both boundaries are interfaces: particles on these boundaries are confined to them at all times. The first boundary condition in (1.2) permits the motion of the water to decouple from that of the air above, with P_a representing the atmospheric pressure at the free surface, taken to be constant.

The vorticity of the water flow,

$$\boldsymbol{\omega} = \nabla \times \mathbf{u},$$

is a measure of the local spin of a fluid element [22]. For irrotational flows this local spin is completely absent:

$$\boldsymbol{\omega} = \left(\frac{\partial w}{\partial y} - \frac{\partial v}{\partial z}, \frac{\partial u}{\partial z} - \frac{\partial w}{\partial x}, \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) \equiv \mathbf{0}. \quad (1.4)$$

The idealization of irrotational flow is appropriate in the absence of non-uniform currents in the water (e.g. field data shows that for waves entering a region of still water the assumption of irrotational flow is realistic [22]), while vorticity is the hallmark of wave-current interaction. Kelvin's circulation theorem [13] guarantees that a water flow will be irrotational if the motion is started from rest by conservative forces. Also, a flow that is irrotational at some instant remains so at later times. Our presentation of the non-dimensionalisation/scaling does allow for arbitrary vorticity in the flow field.

2 Non-dimensionalisation of the governing equations

It is convenient to work with the perturbation of the pressure relative to the hydrostatic pressure distribution, and then this perturbation measures the change in pressure as a wave moves over the surface. Thus we introduce

$$P = P_a + \rho g(h_0 - z) + \rho g h_0 p, \quad (2.1)$$

where we have written p as a non-dimensional pressure and h_0 is suitably chosen. Indeed, we use h_0 to denote some average or typical depth of the water, and λ to be some average or typical wavelength of the wave; these two scale lengths are the basis for the generation of a non-dimensional version of the equations. To complete this, we need a speed scale and a time scale; to be consistent with our underlying aim (which is to produce a model for surface waves), we use $\sqrt{gh_0}$ as a speed scale and then $\lambda/\sqrt{gh_0}$ as a time scale. The process of non-dimensionalising then requires, for example, $x \mapsto \lambda x$, which is to be read as x (the dimensional, physical variable) is replaced by λx , where x is now a non-dimensional version of the original x . With this interpretation in mind, we transform according to

$$x \mapsto \lambda x, \quad y \mapsto \lambda y, \quad z \mapsto h_0 z, \quad t \mapsto \frac{\lambda}{\sqrt{gh_0}} t, \quad (2.2a)$$

$$u \mapsto \sqrt{gh_0} u, \quad v \mapsto \sqrt{gh_0} v, \quad w \mapsto \frac{h_0 \sqrt{gh_0}}{\lambda} w, \quad (2.2b)$$

$$\text{with } H = h_0 + h \quad \text{and} \quad B = h_0 b. \quad (2.2c)$$

Note that we must take care over the treatment of w , in order that the resulting non-dimensionalisation is consistent with the equation of mass conservation (and, equivalently, the existence of a stream function). So we see, for example, that in two spatial dimensions (x and z , say), we have $u_x + w_z = 0$ and so we may write $u = \psi_z$ and $w = -\psi_x$, (where ψ is the stream function, and we have used subscript notation for partial derivatives) which leads to the transformation introduced above. Also observe that we measure the surface disturbance relative to $z = h_0$ but we could elect to define this as $z = 0$ (or anything else appropriate). Upon introducing the parameter

$$\delta = \frac{h_0}{\lambda}, \quad (2.3)$$

the change of variables (2.2), which produces the non-dimensional version of the problem, gives from equations (1.1)-(1.3) the set

$$\left\{ \begin{array}{l} \frac{Du}{Dt} = -\frac{\partial p}{\partial x}, \quad \frac{Dv}{Dt} = -\frac{\partial p}{\partial y}, \quad \delta^2 \frac{Dw}{Dt} = -\frac{\partial p}{\partial z}, \\ \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0, \end{array} \right. \quad (2.4)$$

where

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} + w \frac{\partial}{\partial z},$$

with

$$p = \frac{h}{h_0} \quad \text{and} \quad w = \frac{1}{h_0} \left(\frac{\partial h}{\partial t} + u \frac{\partial h}{\partial x} + v \frac{\partial h}{\partial y} \right) \quad \text{on} \quad z = 1 + \frac{1}{h_0} h(x, y, t), \quad (2.5)$$

and

$$w = u \frac{\partial b}{\partial x} + v \frac{\partial b}{\partial y} \quad \text{on} \quad z = b(x, y). \quad (2.6)$$

For irrotational flows the validity of the equations

$$\delta^2 \frac{\partial w}{\partial y} - \frac{\partial v}{\partial z} = 0, \quad \frac{\partial u}{\partial z} - \delta^2 \frac{\partial w}{\partial x} = 0, \quad \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = 0, \quad (2.7)$$

additionally has to be ensured. (Observe that the combination h/h_0 is non-dimensional; this will be suitably rewritten and simplified in the next section.) These equations and boundary conditions are, of course, no different in their essential mathematical structure from the original, “physical” equations. However, they now contain our first parameter, δ , given by (2.3), which measures the relative size of the depth to wavelength. This parameter is the long wave, or shallowness parameter, and its magnitude can be used to identify different general types of wave problems — but more of this later.

3 Scaling

The main thrust of the principle and importance of scaling is driven by the role that the amplitude of the wave plays in the formulation of the water-wave problem. Let the surface wave be represented by $z = 1 + \epsilon \eta(x, y, t)$ in non-dimensional variables, so that $h = a \eta$, where

$$\epsilon = \frac{a}{h_0} \quad (3.1)$$

and a is a typical — perhaps maximum — amplitude of the wave (and so the choice of ϵ , together with δ , controls the type of water-wave problem that is under consideration). The surface boundary conditions then become

$$p = \epsilon \eta \quad \text{and} \quad w = \epsilon \left(\frac{\partial \eta}{\partial t} + u \frac{\partial \eta}{\partial x} + v \frac{\partial \eta}{\partial y} \right) \quad \text{on} \quad z = 1 + \epsilon \eta,$$

$$w = u \frac{\partial b}{\partial x} + v \frac{\partial b}{\partial y} \quad \text{on} \quad z = b(x, y),$$

which imply that w (at least near the surface) is proportional to ϵ , and that p is likewise. Indeed, in almost all water-wave problems, p decreases with depth and w is maximal at the free surface¹, so we may use ϵ as a measure of the maximum vertical velocity component,

¹See e.g. [13] for linear gravity water waves and [3, 4, 7] for nonlinear solutions to the governing equations representing irrotational gravity water waves of permanent shape propagating at constant speed at the free surface of water over a flat bed.

and of the maximum perturbation pressure, and scale accordingly. Then, for consistency, we require that u and v are similarly scaled; thus we now transform with

$$p \mapsto \epsilon p, \quad w \mapsto \epsilon w, \quad (u, v) \mapsto \epsilon(u, v), \quad (3.2)$$

so that, for example, the scaled w recovers the original (dimensional) w by constructing $\epsilon(h_0\sqrt{gh_0}/\lambda) \times$ (non-dimensional, scaled w). The full set of dimensionless, scaled equations and boundary conditions (from equations (2.4)-(2.6)) then become

$$\left\{ \begin{array}{l} \frac{Du}{Dt} = -\frac{\partial p}{\partial x}, \quad \frac{Dv}{Dt} = -\frac{\partial p}{\partial y}, \quad \delta^2 \frac{Dw}{Dt} = -\frac{\partial p}{\partial z}, \\ \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0, \end{array} \right. \quad (3.3)$$

where

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + \epsilon \left(u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} + w \frac{\partial}{\partial z} \right), \quad (3.4)$$

with

$$p = \eta \quad \text{and} \quad w = \frac{\partial \eta}{\partial t} + \epsilon \left(u \frac{\partial \eta}{\partial x} + v \frac{\partial \eta}{\partial y} \right) \quad \text{on} \quad z = 1 + \epsilon \eta(x, y, t), \quad (3.5)$$

and

$$w = u \frac{\partial b}{\partial x} + v \frac{\partial b}{\partial y} \quad \text{on} \quad z = b(x, y). \quad (3.6)$$

The requirement of irrotational flow (1.4) supplements the equations (3.3)-(3.6), giving

$$\delta^2 \frac{\partial w}{\partial y} - \frac{\partial v}{\partial z} = 0, \quad \frac{\partial u}{\partial z} - \delta^2 \frac{\partial w}{\partial x} = 0, \quad \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = 0. \quad (3.7)$$

We shall assume that suitable initial data is available to generate any particular solutions that might be implied by our various approximations to these equations. Satisfactory local well-posedness (existence, uniqueness, and continuous dependence on initial data for sufficiently small time intervals) results for the governing equations for water-waves, given a suitable initial wave profile and a suitable initial fluid velocity, were recently established [8, 21].

4 Approximate equations

The problem described by equations (3.3)-(3.6) contains two parameters, and it is their interpretation and role that we now examine. It is clear that these two represent the contributions, in the governing equations, of the amplitude (ϵ) and the wavelength (δ) of the wave under consideration; in each case, this is accomplished by measuring a typical amplitude and a typical wavelength relative to the average (or typical) depth of the water. It is immediately evident that the three quantities (amplitude, wavelength and depth) are independent, and so it is misleading to associate any functional relation between them. (Of

course, it is always possible to construct, in the laboratory, a wave maker that generates waves with a specific relation between amplitude and wavelength — and even to allow such a wave to move over a depth that is chosen to be a function of the amplitude, say — but this would be, at the very least, perverse.) This observation has an important consequence: no general meaning can be attached to statements such as $\delta^2 = O(\epsilon)$ or even $\delta^2 = o(\epsilon)$ as $\epsilon \rightarrow 0$, for to do so is to imply a relation between these parameters — albeit it a weak one. We must therefore recognise that we are working with a problem that contains two independent parameters, ϵ and δ , which must be treated as such (although we can allow each to approach zero, provided that this is through independent processes).

The two most obvious approximate problems — and arguably the most useful in a general discussion of water-wave properties — are to let either one or the other parameter tend to zero, keeping the other fixed. (This is equivalent to treating the one that is fixed as $O(1)$ as the limit in the other proceeds.) Then, being specific, we have the two standard approximations:

- (a) $\epsilon \rightarrow 0$ (δ fixed) — the linearised problem;
- (b) $\delta \rightarrow 0$ (ϵ fixed) — the long-wave or shallow-water problem.

The first case, (a), represents the problem of waves of small amplitude, and so we obtain a set of linear equations, to leading order (as $\epsilon \rightarrow 0$):

$$\left\{ \begin{array}{l} \frac{\partial u}{\partial t} = -\frac{\partial p}{\partial x}, \quad \frac{\partial v}{\partial t} = -\frac{\partial p}{\partial y}, \quad \delta^2 \frac{\partial w}{\partial t} = -\frac{\partial p}{\partial z}, \\ \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0, \end{array} \right. \quad (4.1)$$

with

$$p = \eta \quad \text{and} \quad w = \frac{\partial \eta}{\partial t} \quad \text{on} \quad z = 1, \quad (4.2)$$

and

$$w = u \frac{\partial b}{\partial x} + v \frac{\partial b}{\partial y} \quad \text{on} \quad z = b(x, y). \quad (4.3)$$

An important property of this set is that the evaluation at the surface now occurs on $z = 1$ (which is therefore known), yet the free surface — which is often the primary unknown that we seek — still appears in the problem via the boundary conditions on this surface. (The equivalent mapping from the unknown surface, $z = 1 + \epsilon \eta$, to $z = 1$, exists in these type of problems, essentially by virtue of the validity of a Taylor expansion about $z = 1$, the underlying solution structure being polynomial in z .) The equations (4.1)-(4.3) then describe the general linear water-wave problem, with dispersion (by virtue of the terms generated by δ^2) and over a general topography, $z = b(x, y)$. Irrotational flows require the additional equations

$$\delta^2 \frac{\partial w}{\partial y} - \frac{\partial v}{\partial z} = 0, \quad \frac{\partial u}{\partial z} - \delta^2 \frac{\partial w}{\partial x} = 0, \quad \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = 0. \quad (4.4)$$

The corresponding set for $\delta \rightarrow 0$, with ϵ fixed, is

$$\left\{ \begin{array}{l} \frac{Du}{Dt} = -\frac{\partial p}{\partial x}, \quad \frac{Dv}{Dt} = -\frac{\partial p}{\partial y}, \quad \frac{\partial p}{\partial z} = 0, \\ \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0, \end{array} \right. \quad (4.5)$$

where

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + \epsilon \left(u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} + w \frac{\partial}{\partial z} \right), \quad (4.6)$$

with

$$p = \eta \quad \text{and} \quad w = \frac{\partial \eta}{\partial t} + \epsilon \left(u \frac{\partial \eta}{\partial x} + v \frac{\partial \eta}{\partial y} \right) \quad \text{on} \quad z = 1 + \epsilon \eta(x, y, t), \quad (4.7)$$

and

$$w = u \frac{\partial b}{\partial x} + v \frac{\partial b}{\partial y} \quad \text{on} \quad z = b(x, y). \quad (4.8)$$

These constitute a general version of the shallow-water equations, the crucial simplification being that the pressure perturbation p (due to the passage of the wave) is now independent of z , resulting in the dispersive effects being absent here. Although this approximation, based on neglecting higher-order terms in δ , must be justified — and this uses a different basis as compared with that for the small amplitude approximation — this is readily available. Note that in the limit $\delta \rightarrow 0$, irrotational flows are characterized by the additional equations

$$\frac{\partial u}{\partial z} = \frac{\partial v}{\partial z} = 0, \quad \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = 0. \quad (4.9)$$

The two sets, equations (4.1)-(4.3) and (4.5)-(4.8), are the leading-order problems for either $\epsilon \rightarrow 0$ or $\delta \rightarrow 0$, the other parameter being fixed, i.e. $O(1)$, as the limit is imposed. However, if both ϵ and δ are to be regarded as small (although under independent limiting processes), then the leading-order equations become

$$\left\{ \begin{array}{l} \frac{\partial u}{\partial t} = -\frac{\partial p}{\partial x}, \quad \frac{\partial v}{\partial t} = -\frac{\partial p}{\partial y}, \quad \frac{\partial p}{\partial z} = 0, \\ \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0, \end{array} \right. \quad (4.10)$$

with

$$p = \eta \quad \text{and} \quad w = \frac{\partial \eta}{\partial t} \quad \text{on} \quad z = 1, \quad (4.11)$$

and

$$w = u \frac{\partial b}{\partial x} + v \frac{\partial b}{\partial y} \quad \text{on} \quad z = b(x, y). \quad (4.12)$$

For irrotational flows we further require

$$\frac{\partial u}{\partial z} = \frac{\partial v}{\partial z} = 0, \quad \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = 0. \quad (4.13)$$

These equations therefore constitute the leading-order approximation for small amplitude, long waves (or shallow water), and these are consistent with both sets (4.1)-(4.3) and (4.5)-(4.8); in the former, let $\delta \rightarrow 0$ and, in the latter, $\epsilon \rightarrow 0$. It should be noted, however, that

to extend this approximation by including higher-order terms, requires the introduction of a double asymptotic expansion, of the form

$$q(x, y, z, t; \epsilon, \delta) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \epsilon^n \delta^{2m} q_{nm}(x, y, z, t) \quad \text{for } \epsilon \rightarrow 0, \delta \rightarrow 0,$$

where q (and correspondingly q_{nm}) stands for each of u, v, w and p ; the expansion for η follows the same form, but without the dependence on z . Terms at each order $\epsilon^n \delta^{2m}$ are collected and solved sequentially; this formulation does not admit the combining of terms from different orders e.g. $n = 1, m = 0$ with $n = 0, m = 1$ (a manoeuvre that is sometimes attempted!). Furthermore, no meaning can be attached to requirements such as “ ϵ is smaller than δ^2 ”, for to do so is to imply some (weak) functional relation between ϵ and δ . Nevertheless, the apparent “balance” between ϵ and δ^2 has excited a lot of interest — particularly in the context of Korteweg-de Vries type approximations (see [9, 11, 20, 24, 25, 26]) — and so we shall explore this aspect of the formulation.

5 Small-amplitude, nonlinear, long-wave approximation

The Korteweg-de Vries (KdV) equation — the archetypal equation of soliton theory — first appeared in the context of water waves in 1895 [17] (although the soliton connection did not arise for another 60 years or more), so to examine the position of this equation in our presentation is altogether appropriate. The significant aspect of this equation is that it describes the balance between the effects of nonlinearity and of dispersion. Now these two properties are represented by the parameters ϵ and δ , respectively, but they appear in the equations in the forms ϵ and δ^2 . This suggests — and a simple calculation confirms this — that the relevant effects will be retained (to leading order) by including terms as far as $O(\epsilon)$ and $O(\delta^2)$. However, this does not imply that they will “balance” i.e. be of equal importance in the resulting equation. In the early days of soliton theory, when attempts to put the KdV equation in a fairly robust physical context were enthusiastically pursued, it was thought that the choice “ $\delta^2 = O(\epsilon)$ ” was necessary (see e.g. [1] and the comments in [26]) even though, as we have mentioned earlier, this can have no basis in a consistent theory. Indeed, if this were to be the case, then observations of solitons would be very rare! Fortunately, there is no requirement for this condition to hold for the KdV equation to be the appropriate leading-order approximation to the description of the surface waves. As we shall now demonstrate, there is a balance (somewhere and sometime) for any δ , provided only that $\epsilon \rightarrow 0$.

We consider a two-dimensional irrotational wave propagating in the positive x -direction (so that $v = 0$ and there is no dependence on y); although it is possible to include the effects of vorticity and of a variable depth — and so to derive equations which are generalisations of the classical KdV equation (for examples of these problems see [14, 15, 16]) — but for simplicity we shall consider only constant depth, $b = 0$, and irrotational flow. The equations (3.3)-(3.7) then become

$$\begin{cases} u_t + \epsilon(uu_x + uw_z) = -p_x & , & \delta^2 \{w_t + \epsilon(uw_x + ww_z)\} = -p_z, \\ u_x + w_z = 0 & , & u_z - \delta^2 w_x = 0, \end{cases} \quad (5.1)$$

with

$$p = \eta \quad \text{and} \quad w = \eta_t + \epsilon u \eta_x \quad \text{on} \quad z = 1 + \epsilon \eta(x, t), \quad (5.2)$$

and

$$w = 0 \quad \text{on} \quad z = 0. \quad (5.3)$$

(We have used subscripts to denote partial derivatives.) The two parameters are evident here, as is the obvious, simple special case where they are both the “same size”. Now, these equations possess an important and far-reaching property: the parameter δ can be scaled out, in favour of ϵ . This, in turn, leads to a precise prescription — in asymptotic terms — of the region where the KdV equation will be valid (and so where solitons might be expected to appear). We transform according to

$$x \mapsto \frac{\delta}{\sqrt{\epsilon}} x, \quad z \mapsto z, \quad t \mapsto \frac{\delta}{\sqrt{\epsilon}} t, \quad (5.4a)$$

$$p \mapsto p, \quad \eta \mapsto \eta, \quad u \mapsto u, \quad w \mapsto \frac{\sqrt{\epsilon}}{\delta} w, \quad (5.4b)$$

the last scaling being required to ensure the existence of a stream function. (We note, in passing, that if we were to allow the mathematical solecism $\delta^2 = O(\epsilon)$, then this would be no transformation at all.) The resulting equations, from (5.1)-(5.3), are

$$\begin{cases} u_t + \epsilon (uu_x + wu_z) = -p_x & , & \epsilon \{w_t + \epsilon (uw_x + ww_z)\} = -p_z, \\ u_x + w_z = 0 & , & u_z - \epsilon w_x = 0, \end{cases} \quad (5.5)$$

with

$$p = \eta \quad \text{and} \quad w = \eta_t + \epsilon u \eta_x \quad \text{on} \quad z = 1 + \epsilon \eta(x, t), \quad (5.6)$$

and

$$w = 0 \quad \text{on} \quad z = 0, \quad (5.7)$$

which are the same as (5.1)-(5.3), but with δ^2 replaced by ϵ , for arbitrary δ . These equations, (5.5)-(5.7), constitute the classical representation that leads to the KdV equation to leading order as $\epsilon \rightarrow 0$, at fixed δ , in the region where

$$\tau = \epsilon t = O(1) \quad \text{and} \quad \xi = x - t = O(1).$$

To see this, it is convenient first to simplify the system (5.5)-(5.7) by reducing the number of unknown functions. The fourth equation in (5.5) ensures the existence of a function, the velocity potential $\phi(x, z, t)$, such that

$$u = \phi_x, \quad \epsilon w = \phi_z.$$

Notice also that the first two equations in (5.5) merely express the fact that the expression

$$E(t) = \phi_t + \frac{\epsilon}{2} \phi_x^2 + \frac{1}{2} \phi_z^2 + p$$

is, at any fixed t , constant throughout the fluid domain $\{(x, z) \in \mathbb{R}^2 : 0 < z < 1 + \epsilon \eta(x, t)\}$, this being a version of the “pressure equation”; see [13]. Since ϕ is uniquely defined up to an additive function of t , we can define ϕ such that $E(t) = 0$ for all t . The equations (5.5)-(5.7) then become

$$\begin{cases} \epsilon \phi_{xx} + \phi_{zz} = 0 & \text{in } 0 < z < 1 + \epsilon \eta(x, t), \\ \phi_z = \epsilon (\eta_t + \epsilon \phi_x \eta_x) & \text{on } z = 1 + \epsilon \eta(x, t), \\ \phi_z = 0 & \text{on } z = 0, \\ \phi_t + \frac{\epsilon}{2} \phi_x^2 + \frac{1}{2} \phi_z^2 + \eta = 0 & \text{on } z = 1 + \epsilon \eta(x, t). \end{cases} \quad (5.8)$$

Let us now expand ϕ and η in powers of ϵ : we assume that (5.8) has an asymptotic solution of the form

$$\phi(x, z, t) \sim \sum_{k=0}^{\infty} \epsilon^k \phi_k(x, z, t), \quad \eta(x, t) \sim \sum_{k=0}^{\infty} \epsilon^k \eta_k(x, t).$$

The leading order approximation (ϵ^0) to (5.8) is simply

$$\begin{cases} \partial_z^2 \phi_0 = 0 & \text{in } 0 < z < 1, \\ \partial_z \phi_0 = 0 & \text{on } z = 1, \\ \partial_z \phi_0 = 0 & \text{on } z = 0, \\ \partial_t \phi_0 + \frac{1}{2} (\partial_z \phi_0)^2 + \eta_0 = 0 & \text{on } z = 1. \end{cases} \quad (5.9)$$

(We have used the notation e.g. ∂_z for the partial derivative with respect to z .) The first three equations in (5.9) imply that ϕ_0 is independent of z so that

$$\phi_0(x, z, t) = F(x, t) \quad (5.10)$$

for some function F . The last equation in (5.9) becomes

$$F_t + \eta_0 = 0 \quad \text{on } z = 1. \quad (5.11)$$

At the next level of approximation (ϵ^1) of (5.8) we obtain the system

$$\begin{cases} \partial_x^2 \phi_0 + \partial_z^2 \phi_1 = 0 & \text{in } 0 < y < 1, \\ \partial_z \phi_1 = \partial_t \eta_0 & \text{on } z = 1, \\ \partial_z \phi_1 = 0 & \text{on } z = 0, \\ \partial_t \phi_1 + \frac{1}{2} (\partial_x \phi_0)^2 + \eta_1 + \partial_z \phi_1 \cdot \partial_z \phi_0 = 0 & \text{on } z = 1. \end{cases} \quad (5.12)$$

From (5.10), and the first and third equation in (5.12), we deduce that

$$\partial_z \phi_1 = -z F_{xx} \quad \text{for } 0 \leq z \leq 1.$$

The second equation in (5.12) subsequently becomes

$$F_{xx} = -\partial_t \eta_0 \quad \text{on } z = 1.$$

Combined with (5.11) this yields the linear wave equation

$$\partial_x^2 \eta_0 - \partial_t^2 \eta_0 = 0.$$

The general solution being $\eta_0(x, t) = f(x - t) + \tilde{f}(x + t)$, where the sign \mp refers to a wave of profile f (respectively \tilde{f}) translated to the right/left at constant unit speed; our choice of following waves propagating in the positive x -direction leads us to

$$\eta_0(x, t) = f(x - t). \quad (5.13)$$

This suggests that in a particular neighborhood of the (x, t) -space it should be possible to obtain precise information about the evolution of the water's free surface. In order to do this we look at the equations (5.8) in a frame of reference which is moving with a speed of unity to the right. To gain insight, we also introduce the slow time scale $\tau = \epsilon t$. If

$$\xi = x - t, \quad \tau = \epsilon t,$$

then $x = \xi + \frac{\tau}{\epsilon}$ and $t = \frac{\tau}{\epsilon}$ so that

$$\partial_x = \partial_\xi \frac{\partial \xi}{\partial x} + \partial_\tau \frac{\partial \tau}{\partial x} = \partial_\xi, \quad \partial_t = \partial_\xi \frac{\partial \xi}{\partial t} + \partial_\tau \frac{\partial \tau}{\partial t} = -\partial_\xi + \epsilon \partial_\tau.$$

The system (5.8) in (ξ, τ) -coordinates then becomes

$$\begin{cases} \epsilon \phi_{\xi\xi} + \phi_{zz} = 0 & \text{in } 0 < z < 1 + \epsilon\eta, \\ \phi_z = \epsilon(-\eta_\xi + \epsilon\eta_\tau + \epsilon\phi_\xi \eta_\xi) & \text{on } z = 1 + \epsilon\eta, \\ \phi_z = 0 & \text{on } z = 0, \\ \epsilon \partial_\tau \phi - \partial_\xi \phi + \frac{\epsilon}{2} \phi_\xi^2 + \frac{1}{2} \phi_z^2 + \eta = 0 & \text{on } z = 1 + \epsilon\eta. \end{cases} \quad (5.14)$$

Expanding ϕ and η in powers of ϵ ,

$$\phi(t, x, z) \sim \sum_{k=0}^{\infty} \epsilon^k \Phi_k(\xi, z, \tau), \quad \eta(t, x) \sim \sum_{k=0}^{\infty} \epsilon^k H_k(\xi, \tau),$$

the leading order (ϵ^0) approximation of (5.14) is

$$\begin{cases} \partial_z^2 \Phi_0 = 0 & \text{in } 0 < z < 1, \\ \partial_z \Phi_0 = 0 & \text{on } z = 1, \\ \partial_z \Phi_0 = 0 & \text{on } z = 0, \\ -\partial_\xi \Phi_0 + \frac{1}{2} (\partial_z \Phi_0)^2 + H_0 = 0 & \text{on } z = 1. \end{cases} \quad (5.15)$$

From the first three equations in (5.15) we get $\partial_z \Phi_0 \equiv 0$. Therefore

$$\Phi_0(x, z, \tau) = \theta(\xi, \tau) \quad \text{in } 0 \leq z \leq 1, \quad (5.16)$$

for some function θ , and the last equation in (5.15) yields

$$H_0(\xi, \tau) = \partial_\xi \theta(\xi, \tau) \quad \text{on } z = 1. \quad (5.17)$$

At the next level of approximation (ϵ^1) of (5.14) we get

$$\begin{cases} \partial_z^2 \Phi_1 + \partial_\xi^2 \Phi_0 = 0 & \text{in } 0 < z < 1, \\ \partial_z \Phi_1 = -\partial_\xi H_0 & \text{on } z = 1, \\ \partial_z \Phi_1 = 0 & \text{on } z = 0, \\ \partial_\tau \Phi_0 + \partial_z \Phi_0 \cdot \partial_z \Phi_1 + \frac{1}{2} (\partial_\xi \Phi_0)^2 + H_1 = 0 & \text{on } z = 1. \end{cases} \quad (5.18)$$

From the first three equations in (5.18), in combination with (5.16), we deduce that

$$\partial_z \Phi_1(x, z, \tau) = -z \theta_{\xi\xi}(\xi, \tau) \quad \text{in } 0 \leq z \leq 1.$$

For some function of integration $\alpha(\xi, \tau)$, independent of z , we then get

$$\Phi_1(x, z, \tau) = -\frac{1}{2} z^2 \theta_{\xi\xi}(\xi, \tau) + \alpha(\xi, \tau) \quad \text{in } 0 \leq z \leq 1. \quad (5.19)$$

From the last equation in (5.18) we now obtain

$$H_1(\xi, \tau) = -\theta_\tau(\xi, \tau) - \frac{1}{2} \theta_{\xi\xi\xi}(\xi, \tau) + \alpha_\xi(\xi, \tau) - \frac{1}{2} \theta_\xi^2(\xi, \tau) \quad (5.20)$$

in view of (5.16).

At the third level of approximation (ϵ^2) to (5.14) we obtain

$$\begin{cases} \partial_z^2 \Phi_2 + \partial_\xi^2 \Phi_1 = 0 & \text{in } 0 < z < 1, \\ \partial_z \Phi_2 - H_0 \theta_{\xi\xi} = -\partial_\xi H_1 + \partial_\tau H_0 + \partial_\xi \Phi_0 \cdot \partial_\xi H_0 & \text{on } z = 1, \\ \partial_z \Phi_2 = 0 & \text{on } z = 0, \\ \partial_\tau \Phi_1 - \partial_\xi \Phi_2 + \partial_\xi \Phi_0 \cdot \partial_\xi \Phi_1 + \frac{1}{2} (\partial_z \Phi_1)^2 + \partial_z \Phi_0 \cdot \partial_z \Phi_2 + H_2 = 0 & \text{on } z = 1. \end{cases} \quad (5.21)$$

The presence of the term $(-H_0 \theta_{\xi\xi})$ in the second equation of (5.21) is the first time in our expansion that a lower order term contributes towards the subsequent order of approximation: by (5.19),

$$\partial_z \Phi_1 = -(1 + \epsilon H_0) \theta_{\xi\xi} + O(\epsilon^2) \quad \text{on } z = 1 + \epsilon \eta.$$

From the first and third equation in (5.21), in combination with (5.19), we get

$$\Phi_2(x, z, \tau) = \frac{1}{24} z^4 \partial_\xi^4 \theta(\xi, \tau) - \frac{1}{2} z^2 \partial_\xi^2 \alpha(\xi, \tau) + \beta(\xi, \tau)$$

for some new function of integration $\beta(\xi, \tau)$. Using this, in combination with (5.17), (5.19) and (5.20) in the second equation from (5.21), we obtain

$$\frac{1}{3} \theta_{\xi\xi\xi\xi} + 3 \theta_\xi \theta_{\xi\xi} + 2 \theta_{\tau\xi} = 0.$$

In view of (5.17), the above equation is precisely the Kortweg-de Vries (KdV) equation for the surface wave $\eta \sim H_0(\xi, \tau)$:

$$2 \partial_\tau H_0 + 3 H_0 \partial_\xi H_0 + \frac{1}{3} \partial_\xi^3 H_0 = 0. \quad (5.22)$$

The balance between nonlinearity and dispersion, as embodied by the KdV equation, therefore occurs in non-dimensional scaled variables in that region of time and space defined by $t = O(\epsilon^{-1})$ and $x = O(\epsilon^{-1})$ (because $x-t = O(1)$ here), respectively, and for any δ . This region of space is therefore given, from (5.4), by $x = O(\delta\epsilon^{-3/2})$ and $t = O(\delta\epsilon^{-3/2})$ in non-dimensional variables which, in original physical variables, can be estimated by the distance $x = O(\lambda\delta\epsilon^{-3/2})$ and the corresponding time $t = O(\frac{\lambda}{\sqrt{gh_0}}\delta\epsilon^{-3/2})$ in accordance to (2.2); these two provide estimates for the time and place where the balance occurs (for any δ , but for $\epsilon \rightarrow 0$). Of course, we assume that the times and distances involved here are such that other physical properties do not intervene. If the wave travels for a significant distance/time, then the effects of, for example, viscosity may have to be included, although this property is represented by a new, independent, parameter, so the formal asymptotic argument already described is unaltered.

It is instructive to present some estimates of the distance that waves will have to travel before the nonlinear-dispersion balance, epitomised by the KdV equation, can occur. Thus we compute values of

$$\lambda\delta\epsilon^{-3/2} = h_0 \left(\frac{a}{h_0} \right)^{-\frac{3}{2}},$$

which is independent of λ , demonstrating that the scaling (5.4), is equivalent to using the single scale length h_0 for all non-dimensionalisations; the results are given in the table below (but note that, in the cases when $\epsilon = a/h_0 \geq 1$, the estimate is not applicable).

amplitude	0.1 m	0.5 m	1 m	5 m	10 m
depth 5 m	1.77 km	0.16 km	0.056 km	N/A	N/A
depth 10 m	10 km	0.89 km	0.32 km	0.028 km	N/A
depth 100 m	3162 km	283 km	100 km	8.9 km	3.2 km
depth 1 km	10^6 km	89442 km	31622 km	2828 km	1000 km
depth 4 km	$32 \cdot 10^6$ km	$28 \cdot 10^6$ km	10^6 km	$9 \cdot 10^4$ km	$32 \cdot 10^3$ km

These figures demonstrate that, as John Scott Russell so eloquently described in his 1844 “Report on Waves” [23], a solitary wave of moderate amplitude in a river could evolve and be followed on horseback — the distances needed for a KdV balance are no more than a few kilometres. On the other hand, a wave generated over the deep ocean — the devastating tsunami of December 2004, for example — would have to travel many thousands of kilometres before the KdV balance would be achieved. The Boxing Day Tsunami 2004 was generated on December 26, 2004, by a major earthquake with epicentre located about 250 km off the west coast of the island of Sumatra [28]. The earthquake caused a rupture of the ocean floor of approximate length 1200 km essentially to the north-northwest; the slip along it acted as a line source generating tsunami waves which propagated through the Indian Ocean to the west, and through the Andaman Basin to the east [9]. In the basin of the Indian Ocean/Bay of Bengal the parameters $a = 1$ m, $h_0 = 4$ km and $\lambda = 180$ km are advocated in [9], based on accurate measurements of the tsunami wave height over a distance of 1800 km provided by a radar altimeter on board a satellite along a track traversing the Indian Ocean/Bay of Bengal from north to south,

about 2 h after the main earthquake shock [18, 19]. These yield the values

$$\delta = \frac{h_0}{\lambda} \approx 2 \cdot 10^{-2}, \quad \epsilon = \frac{a}{h_0} \approx 25 \cdot 10^{-5},$$

and the apparent approximate balance “ $\delta^2 = O(\epsilon)$ ” is then taken in [9] to be indicative for the suitability of KdV as a model for the evolution of these tsunami waves. (This is in contrast to the Andaman Basin where the parameters $a = 1$ m, $h_0 = 1$ km and $\lambda = 180$ km do not characterize values of δ and ϵ considered to be appropriate for KdV dynamics [9].) However, for these values of the parameters, our considerations show that the distance required for a balance between nonlinearity and dispersion characteristic of the KdV equation is in excess of 10^6 km in the Indian Ocean and of $3 \cdot 10^4$ km in the Andaman Basin cf. the above table. If we allow the wave amplitude, based on estimates of the wave just a few seconds after the earthquake and so applicable to the initial-value problem for waves propagating in one direction only, to be closer to 1 m (or even less, e.g. 0.6 m as advocated in [11]) the distances involved are even greater. This implies that it is quite unrealistic to suppose that a balance of nonlinearity and dispersion can occur over oceans that extend, at most, about 10^4 km: we must conclude that any tsunamis initiated in the deep oceans are unlikely to be a manifestation of soliton theory. Our considerations rely on the assumption that a large wave length λ is appropriate for tsunamis but, of course, the initial data might contain a range of wavelengths so, perhaps, smaller wavelengths might be relevant. However, a short time after the initial disturbance, out in the open sea, the tsunami waves are generally of small amplitude (records show that 1 m is a good upper estimate) and therefore over moderate distances the dynamics of the tsunami is likely to be captured by linear theory (whereas in coastal regions where the typical water depth changes drastically, nonlinear theory is needed to describe the effects of the bottom topography on the shape of the surface wave). The dispersion relation

$$c = \sqrt{\frac{g\lambda}{2\pi} \tanh\left(\frac{2\pi h_0}{\lambda}\right)},$$

written in the original physical variables, gives the approximate propagation speed of irrotational small amplitude waves [7, 13]. The monotonicity of c as a function of λ shows that longer waves travel faster, with the long wave limit $\delta = h_0/\lambda \rightarrow 0$ providing the upper bound $c \approx \sqrt{gh_0}$. (For example, the average depth of the Indian Ocean being $h_0 = 4$ km, taking $g = 9.8 \text{ m/s}^2$, linear shallow water theory predicts that the 2004 Boxing Day tsunami would propagate at approximate speed $\sqrt{gh_0} \approx 720$ km/h, thus needing about 2 h 10 min to travel the distance of about 1550 km from the epicentre of the earthquake to the affected coast of Sri Lanka — which is about right [9, 28]. A similar calculation for the Andaman Basin yields a speed of about 360 km/h, which is consistent with the fact that the tsunami waves impacted upon the west coast of Thailand about 1 h after the earthquake [9].) The dispersive effect in linear water waves is exemplified by the dispersion relation, the fundamental feature being that waves of different wavelength (i.e. different λ) travel at different speeds. Thus an initial profile which comprises a number of different wavelengths will suitably evolve. Therefore, after initiation, provided that the typical depth h_0 remains roughly constant, we see that, in time, the longer waves move to the front and the short wave lengths are then confined to the back of the tsunami profile.

Thus, even if short wave lengths were relevant in the initial stage, after some time the thrust of the tsunami problem is confined to long waves — a feature confirmed by the satellite measurements of the 2004 Boxing Day tsunami mentioned before. Behind these waves a dispersive effect might be noticeable but dispersion cannot be essential for the leading waves in a tsunami. (Of course, the background state of the sea — prior to the arrival of the tsunami — could contain short waves but such background states would then most likely exhibit vorticity and the irrotational flow setting is no longer adequate.) A more extensive discussion of the tsunamis, including a formulation and solution of the problem which does not rest on soliton theory, can be found in [5] (see also the brief discussion in [6]).

6 Conclusion

This note has attempted to describe the background to a careful mathematical statement of the classical water-wave problem, with the emphasis on the scaling issues that must be addressed. In particular, we have seen that a systematic formulation of the problem requires some care as to the choice of scales and, most especially, the appropriate choice of variables relevant to the KdV equation. One of the important observations is the demonstration that a “KdV balance” is available for any value of the long-wave (shallowness) parameter δ , requiring only that the waves be of small amplitude (in the sense that ϵ is small). This, in turn, has led us to some estimates for the distances over which waves will have to travel in order to achieve this balance and this is likely to be unattainable for tsunami waves on a planet with the dimensions of the Earth.

Acknowledgments. *The authors are grateful for the support of the Hamilton Mathematics Institute, Trinity College Dublin. The authors would like to thank the referee for helpful comments.*

References

- [1] ABLOWITZ M J and SEGUR H, *Solitons and the inverse scattering transform*, SIAM Studies in Applied Mathematics, Philadelphia, Pa., 1981.
- [2] BARNETT T P and KENYON K E, Recent advances in the study of wind waves, *Rep. Prog. Phys.* **38** (1975), 667–729.
- [3] CONSTANTIN A, The trajectories of particles in Stokes waves, *Invent. Math.* **166** (2006), 523–535.
- [4] CONSTANTIN A and ESCHER J, Particle trajectories in solitary water waves, *Bull. Amer. Math. Soc.* **44** (2007), 423–431.
- [5] CONSTANTIN A and JOHNSON R S, Propagation of very long water waves, with vorticity, over variable depth, with applications to tsunamis, *Fluid Dynamics Research* **40** (2008), 175–211.
- [6] CONSTANTIN A and JOHNSON R S, Modelling tsunamis, *J. Phys. A* **39** (2006), L215–L217.
- [7] CONSTANTIN A and STRAUSS W A, Exact steady periodic water waves with vorticity, *Comm. Pure Appl. Math.* **57** (2004), 481–527.

-
- [8] COUTAND D and SHKOLLER S, Well-posedness of the free-surface incompressible Euler equations with or without surface tension, *J. Amer. Math. Soc.* **20** (2007), 829–930.
- [9] CRAIG W, Surface water waves and tsunamis, *J. Dynam. Differential Equations* **18** (2006), 525–549.
- [10] DA SILVA A F T and PEREGRINE D H, Steep, steady surface waves on water of finite depth with constant vorticity, *J. Fluid Mech.* **195** (1988), 281–302.
- [11] DIAS F and DUTYKH D, Dynamics of tsunami waves, in *Extreme man-made and natural hazards in dynamics of structures*, pp. 1–25, Springer, Berlin, 2006.
- [12] DRAZIN P G and JOHNSON R S, *Solitons: an introduction*, Cambridge University Press, Cambridge, 1989.
- [13] JOHNSON R S, *A modern introduction to the mathematical theory of water waves*, Cambridge University Press, Cambridge, 1997.
- [14] JOHNSON R S, Water waves and Korteweg- de Vries equations, *J. Fluid Mech.* **97** (1980), 701–719.
- [15] JOHNSON R S, Nonlinear gravity waves on the surface of an arbitrary shear flow with variable depth, in *Nonlinear instability analysis*, pp. 221–243, Adv. Fluid Mech., Southampton, 1997.
- [16] JOHNSON R S, The classical problem of water waves: a reservoir of integrable and nearly-integrable equations, *J. Nonlinear Math. Phys.* **10** (2003), 72–92.
- [17] KORTEWEG D J and DE VRIES G, On the change of form of long waves advancing in a rectangular canal and on a new type of long stationary waves, *Phil. Mag.* **39** (1895), p. 422.
- [18] KULIKOV E A, MEDVEDEV P P and LAPPO S S, Satellite recording of the Indian Ocean tsunami on December 26, 2004, *Dokl. Akad. Nauk* **401** (2005), 537–542.
- [19] KULIKOV E A, Dispersion of the Sumatra Tsunami waves in the Indian Ocean detected by satellite altimetry, *Russian J. Earth Sci.* **8** (2006), ES4004, doi:10.2205/2006ES000214
- [20] LAKSHMANAN M, Integrable nonlinear wave equations and possible connections to tsunami dynamics, in *Tsunami and nonlinear waves*, pp. 31–49, Springer, Berlin, 2007.
- [21] LANNES D, Well-posedness of the water-waves equations, *J. Amer. Math. Soc.* **18** (2005), 605–654.
- [22] LIGHTHILL J, *Waves in fluids*, Cambridge University Press, Cambridge, 1978.
- [23] SCOTT RUSSELL J, Report on waves, in *Report on the Fourteenth Meeting*, British Association for the Advancement of Science, ed. John Murray, pp. 311–390, London, 1844.
- [24] SATTINGER D H, Tsunamis and barge canals, *J. Math. Fluid Mech.* **9** (2007), 226–234.
- [25] SCHNEIDER G and WAYNE C E, On the validity of 2D-surface water wave models, *GAMM Mitt. Ges. Angew. Math. Mech.* **25** (2002), 127–151.
- [26] SEGUR H, Waves in shallow water, with emphasis on the tsunami of 2004, in *Tsunami and nonlinear waves*, pp. 3–29, Springer, Berlin, 2007.
- [27] SEGUR H, Integrable models of waves in shallow water, *MSRI Publ.* **55** (2007), 345–371.
- [28] TITOV V, RABINOVICH A B, MOFJELD H O, THOMSON R E, and GONZÁLEZ F I, The global reach of the 26 December Sumatra tsunami, *Science* **309** (2005), 2045–2048.