WATER WAVES, NONLINEAR SCHRÖDINGER EQUATIONS
AND THEIR SOLUTIONS

D. H. PEREGRINE

(Received 9 August 1982)

Abstract

Equations governing modulations of weakly nonlinear water waves are described. The modulations are coupled with wave-induced mean flows except in the case of water deeper than the modulation length scale. Equations suitable for water depths of the order the modulation length scale are deduced from those derived by Davey and Stewartson [5] and Dysthe [6]. A number of cases in which these equations reduce to a one dimensional nonlinear Schrödinger (NLS) equation are enumerated.

Several analytical solutions of NLS equations are presented, with discussion of some of their implications for describing the propagation of water waves. Some of the solutions have not been presented in detail, or in convenient form before. One is new, a “rational” solution describing an “amplitude peak” which is isolated in space-time. Ma’s [13] soliton is particularly relevant to the recurrence of uniform wave trains in the experiment of Lake et al. [10].

In further discussion it is pointed out that although water waves are unstable to three-dimensional disturbances, an effective description of weakly nonlinear two-dimensional waves would be a useful step towards describing ocean wave propagation.

1. Introduction

One of the remarkable developments in mathematics in the last 20 years has been the almost complete solution of certain types of nonlinear partial differential equations by the inverse scattering transform. See Ablowitz and Segur [2] for an up-to-date monograph on the topic which gives a wide perspective on the method. The method was first discovered in the solution of the Korteweg-de Vries (KdV) equation. The KdV equation is a canonical equation for weakly dispersive, weakly nonlinear waves which was first derived to describe shallow water waves (see Miles [14]).

1School of Mathematics, University of Bristol, Bristol BS8 1TW, England.
© Copyright Australian Mathematical Society 1983

https://doi.org/10.1017/S0334270000003891 Published online by Cambridge University Press
Once the inverse scattering method became known it was soon found to give solutions to other canonical equations, including the nonlinear Schrödinger (NLS) equations (Zakharov and Shabat [23, 24])

\[ iq_t + q_{xx} \pm 2 |q|^2 q = 0. \]  

(NLS ±)

These equations describe the evolution of modulations of dispersive waves with weak nonlinearity. They arise in the propagation of electromagnetic waves through matter and in view of the character of their solutions they are called the “self-focussing” and “defocussing” NLS equations for the + and − signs respectively. Here we use NLS + and NLS− to denote them.

For water wave modulations there is usually a coupling between the modulations and the wave-induced current so that it is only in certain cases that water wave modulations are described by an NLS equation. However, these include important cases such as (i) deep-water modulations for which the depth, \( h \), is such that the modulation wave number, \( K \), satisfies

\[ Kh \gg 1, \]  

(1.1)

and (ii) the modulations of a steady wave field which suffers small changes of direction on any depth for which \( kh \) is not very small (\( k \) is the water wave wavenumber).

The main aim of this paper is to review analytic solutions of the NLS equations in a water-wave context. Thus equations for weakly-nonlinear water-wave modulations are presented in the next section. Such equations have not previously been given for the case \( Kh = O(1) \). The results of Davey and Stewartson [5] for \( Kh \ll 1 \) and Dysthe’s [6] higher-order analysis for deep-water waves are combined to cover this case. The modulations are assumed to be long compared with a wavelength, \( i.e. \)

\[ K \ll k. \]  

(1.2)

For KdV and NLS equations the inverse scattering transform shows that, for an initial disturbance of finite extent, the solution is obtained from a discrete set of eigenvalues and a continuous spectrum. Each of the eigenvalues corresponds to a “soliton” and the continuous spectrum corresponds to an oscillatory dispersive wave. Certain conditions need to be satisfied in order that the set of eigenvalues be not empty. It is found that the asymptotic development of the solution with time leads to decay of the oscillatory part and thus the solitons asymptotically dominate the solution. This is relatively well-known, especially for the KdV equation, so we shall do no more than note that an individual soliton is effectively finite in extent, \( i.e. \) it decays exponentially to the undisturbed level; and that when solitons of different velocities meet each other they interact nonlinearly but eventually reemerge intact with only a phase shift due to the interaction.
It is natural with these results to concentrate on solitons. Unlike the KdV and defocussing NLS- equations, the self-focussing NLS+ equation has more than one soliton solution. It is thus useful to give the NLS solitons names as follows:

(a) An *isolated soliton* is that soliton of the NLS+ equation which decays to zero. This is the best known soliton (Zakharov and Shabat [23]), sometimes known as an “envelope” soliton to emphasis its modulational character.

(b) A *Ma soliton* is the soliton solution of the NLS+ equation, given by Ma [13], which decays to the uniform solution. This soliton is probably the most relevant to modulations of a wide-spread wave field.

(c) A *bi-soliton* is a solution of the NLS+ equation derived from two eigenvalues of equal real parts (Zakharov and Shabat [23]). The velocity of the soliton depends on the real part of the eigenvalue, so this solution corresponds to two isolated solitons which cannot separate and are thus “bound” solitons. The combination thus acts like a soliton. Multi-soliton solutions corresponding to more than two eigenvalues also exist.

(d) A *dark soliton* is a soliton of the defocussing NLS- equation which decays to a uniform solution (Zakharov and Shabat [24]). It is “dark” in the sense that its modulus is always less than that of the uniform solution in which it propagates. This soliton is similar to the KdV soliton.

Only the isolated soliton has received much attention in the water wave context. The Ma soliton and the bi-soliton are particularly interesting since they are oscillatory solutions, so more attention is given to them here. Less attention is given to the dark soliton since solutions of the NLS- equation are similar to solutions of the KdV equation, and also form part of Peregrine’s [18] discussion of jumps in water-wave properties, wave focussing and refraction.

The Ma soliton gives a nontrivial solution in the limit of zero amplitude. Apart from a simple exponential factor it is a rational function and hence is a typical example of such solutions (Ablowitz and Segur [2], Section 3.4). It describes an isolated “amplitude peak” in space-time arising out of the uniform solution. There is also a class of limiting bi-solitons, which are not rational functions. They describe distant equal isolated solitons drawing together and “bouncing” off each other.

This paper does not review the whole area of NLS equations and water waves. A substantial review of deep water waves, with particular emphasis on the stability of steep waves, has recently appeared (Yuen and Lake [22]). A general discussion of wave packet evolution for water waves is given by Ablowitz and Segur [1]. This last account includes the effects of surface tension, which extend the variety of equations to be considered. The book by Ablowitz and Segur [2] should also be consulted for a more complete picture.
2. Modulation equations for water waves

A modulated water wave train can be described to a first, linear, approximation by the velocity potential

$$\phi(x,y,z,t) = A(x,y,t) \frac{\cosh k(z + h)}{\cosh kh} e^{i(kx - \omega t)} + \text{complex conjugate},$$

(2.1)

where the modulation is given by the "slowly-varying" complex-valued function $A(x,y,t)$; $\omega$ and $k$ are the frequency and wave number respectively of the modulated wave, or "carrier wave", which has an amplitude $a = 2\omega |A|/k$. This first approximation leads to the result that

$$A_t + c_g A_x = 0,$$

(2.2)

where $c_g = \omega'(k)$, the linear group velocity. That is, long modulations of linear waves travel at the group velocity.

The next approximation which includes both weakly nonlinear dispersive effects and the next order of terms in the modulation gradient gives

$$2i(A_t + c_g A_x) - b_1 A_{xx} + b_2 A_{yy} = B_1 |A|^2 A + B_2 A \Phi_x|_{z=0},$$

(2.3)

where $\Phi(x,y,z,t)$ is the velocity potential of the wave induced flow and satisfies

$$\nabla^2 \Phi = 0 \quad \text{in} \quad -h < z < 0,$$

(2.4a)

$$\Phi_x h_x + \Phi_y h_y = \Phi_z \quad \text{at} \quad z = -h(x,y),$$

(2.4b)

and

$$g \Phi_z + \Phi_{tt} = B_2 (|A|^2)_x \quad \text{at} \quad z = 0,$$

(2.4c)

where $z = 0$ is the undisturbed free surface and $z = -h$ is the bed. The derivation of these equations is briefly discussed in the Appendix where the coefficients $b_1, b_2, B_1$ and $B_2$ are also defined. The combined pair of terms from equation (2.2) which appear in equation (2.3) are, together, of the same order as each of the other terms. The second derivative terms are the higher-order modulation terms (or diffraction terms). The terms with $B_1$ and $B_2$ as coefficients are the first order nonlinear terms. In all but the first-mentioned pair of terms, the first approximation (2.2) allows $\partial/\partial t$ and $c_g \partial/\partial x$ to be interchanged.

There are two special cases of equations (2.3) and (2.4) which have been studied.

(i) The deep-water limit, in which $\Phi$ is of the order $(|A|^2)_x$ and may be neglected, giving

$$2i \left( \frac{1}{\omega} \frac{\partial A}{\partial t} + \frac{1}{2k} \frac{\partial A}{\partial x} \right) - \frac{1}{4k^2} \frac{\partial^2 A}{\partial x^2} + \frac{1}{2k^2} \frac{\partial^2 A}{\partial y^2} = \frac{4k^4}{\omega^2} |A|^2 A.$$

(2.5)
(ii) The case in which the modulation length-scale, say $1/K$, is very much greater than the depth $h$, that is

$$Kh \ll 1.$$  

The equations (2.4) for $\Phi$ can then be reduced to a long wave equation for $\Phi_0(x, y, t)$:

$$\left( gh - c^2 \right) \Phi_{xxx} + gh \Phi_{yyy} + B_2 \left( |A|^2 \right)_x = 0.$$  

Equations (2.3) and (2.7) are the Davey-Stewartson equations (Davey and Stewartson [5]).

3. NLS equations for water waves

In one space dimension there are two fundamentally different nonlinear Schrödinger equations. In the canonical forms used with the inverse scattering transform they are the "self-focussing" equation

$$iq_T + q_{xx} + 2|q|^2q = 0,$$  

(NLS+)

and the "de-focussing" equation

$$iq_T + q_{xx} - 2|q|^2q = 0.$$  

(NLS−)

The names represent the character of certain solutions of the equations which were first examined in the context of nonlinear optics (see Whitham [20], chapter 16). It is only the relative sign of $q_{xx}$ and $2|q|^2q$ which is significant; a change of sign of $iq_T$ simply gives the complex conjugate equation.

With more than one space dimension the term NLS equation is applied to any combination of second order space derivatives in place of $q_{xx}$. For example in two space dimensions the canonical forms are

$$iq_T + q_{xy} + 2|q|^2q = 0,$$  

(3.1)

and

$$iq_T + q_{xx} + q_{yy} \pm 2|q|^2q = 0.$$  

(3.2, 3.3)

Only a little work has been done with two dimensional problems, e.g. see Hui and Hamilton [7] and Yuen and Lake [22]. Here, only one-dimensional examples are considered.

The full modulation equations (2.3) and (2.4) are not an NLS equation, however NLS equations may be obtained from them in various ways.
(i) Steady waves, slow $x$ variation

For steady waves with much longer modulations in the $x$ direction than in the $y$ direction the terms $A_{xx}$ and $A_{xy}$ become negligible in equation (2.3), this decouples the wave-induced flow and gives

$$2ic_g A_x + b_2 A_{yy} = B_1 |A|^2 A.$$  \hspace{1cm} (3.4)

The transformation

$$T = x/2c_g, \quad X = b_2^{-1/2}y, \quad q = (\frac{1}{2}B_1)^{1/2}A,$$  \hspace{1cm} (3.5a, b, c)

transforms it into the NLS- equation.

Equation (3.4) is used by Yue and Mei [21] to study reflection of near-linear water waves, and by Peregrine [18] to discuss the focussing of near-linear water waves.

(ii) Transverse modulations

If there is no $x$ variation equation (2.4) becomes

$$2iA_t + b_2 A_{yy} = B_x A |A|^2 A,$$  \hspace{1cm} (3.6)

another example of the defocussing equation. Modulations at other specific angles do not give NLS equations except in the limiting cases considered below.

(iii) Deep water modulations

Equation (2.5) is an NLS equation. The transformation

$$T = \frac{1}{2}\omega t, \quad X = kx - \frac{1}{2}\omega t + 2^{-1/2}ky, \quad Y = kx - \frac{1}{2}\omega t - 2^{-1/2}ky,$$

$$q = 2^{1/2}k^2A*/\omega,$$  \hspace{1cm} (3.7a, b, c, d)

gives the two-dimensional equation (3.1).

Cylindrical modulations at an angle $\alpha$ to the wave direction and the transformation

$$T = \frac{1}{2}\omega t \cos \alpha, \quad X = \frac{2\left[(kx - \frac{1}{2}\omega t)\cos \alpha + ky \sin \alpha\right]}{(\cos^2 \alpha - 2 \sin^2 \alpha)^{1/2}},$$

$$q = 2^{1/2}k^2A*/\omega = 2^{1/2}ka^*,$$  \hspace{1cm} (3.8a, b, c)

give the NLS+ equation for

$$\tan \alpha < \frac{1}{2};$$  \hspace{1cm} (3.9)

and similarly,

$$T = \frac{1}{2}\omega t, \quad X = \frac{2\left[(kx - \frac{1}{2}\omega t)\cos \alpha + ky \sin \alpha\right]}{(2 \sin^2 \alpha - \cos^2 \alpha)^{1/2}},$$

$$q = 2^{1/2}k^2A/\omega$$  \hspace{1cm} (3.10a, b, c)
gives the NLS– equation when
\[ \tan \alpha > \frac{1}{2}. \] (3.11)

See Hui and Hamilton [7] for discussion of the critical angle
\[ \alpha = \arctan 2^{-1/2} = 35.3^\circ. \] (3.12)

This is the angle of the modulation to the waves in the “bow wave” train of the Kelvin ship-wave pattern.

(iv) Shallow-water modulations

The Davey-Stewartson equations also transform into NLS equations when cylindrical modulations are considered in which the only spatial variation is in the coordinate
\[ Z = x \cos \alpha + y \sin \alpha. \] (3.13)

Equation (2.7) can then be integrated to give
\[ \Phi_0 = -\frac{B_2 |A|^2}{gh - c_s^2 \cos^2 \alpha}, \] (3.14)

and equation (2.4) becomes
\[ 2i(A_t + c_s \cos \alpha A_x) + (-b_1 \cos^2 \alpha + b_2 \sin^2 \alpha) A_{zz} \]
\[ = \left[ B_1 + \frac{B_2}{gh - c_s^2 \cos^2 \alpha} \right] |A|^2 A. \] (3.15)

This equation can correspond to either a self-focussing or defocussing NLS equation according to the sign of
\[ -b_1 \cos^2 \alpha + b_2 \sin^2 \alpha. \] (3.16)

Note, \( b_1 \) changes sign at \( kh = 1.36 \).

In the following sections it is impractical to refer back to all the above examples of water-wave NLS equations. The deep-water case of modulations in the wave direction is taken as a representative example for the self-focussing NLS+ equation,
\[ \frac{2i}{\omega} \frac{\partial A}{\partial t} + i \frac{A}{k} \frac{\partial A}{\partial x} - \frac{1}{4k^2} \frac{\partial^2 A}{\partial x^2} = \frac{4k^4}{\omega^2} |A|^2 A, \] (3.17)

with transformation to the NLS + equation effected by
\[ T = \frac{1}{2} \omega t, \quad X = 2kx - \omega t, \quad q = 2^{1/2}k^2A^*/\omega. \] (3.18)
4. Uniform solutions

The solutions of the self-focussing NLS+ equation which have constant amplitude \( q_0 \) are

\[
q = q_0 \exp i \left\{ (2q_0^2 - p_0^2)T + p_0X \right\},
\]

where \( p_0 \) is another constant. The corresponding solution of equation (3.17) is

\[
A = A_0 \exp i \left\{ 2p_0kx - \omega \left( p_0 - \frac{1}{2}p_0^2 + 2k^2A_0^2/\omega^2 \right)t \right\}.
\]

When this solution (4.2) is multiplied by the exponential from the carrier wave (2.1) it is seen that this solution corresponds to a plane wave of wave number \( k + 2p_0k \) with an attendant shift in frequency which is the appropriate approximation to the Stokes dispersion relation for \( p_0 \) and \( a_0k \ll 1 \), namely

\[
\omega_1 = \left[ g(k + 2p_0k)(1 + a_0^2k^2) \right]^{1/2}
\]

\[
= \omega \left[ 1 + p_0 - \frac{1}{2}p_0^2 + \frac{1}{2}a_0^2k^2 + O(p_0^3, p_0a_0^2k, a_0^4k^4) \right]
\]

(4.3)

where \( \omega = (gk)^{1/2} \). Thus, \( p_0 \) corresponds to a simple shift of carrier-wave wave number. Most such shifts are ignored in discussion of other solutions here.

Thus, the wave of constant wavenumber is

\[
q = q_0 e^{2i\omega_0T}
\]

(4.4)

and the wave of constant frequency is

\[
q = q_0 e^{i(2p_0T + p_0X)}
\]

(4.5)

where \( p_0 \) is one of the roots of

\[
p_0^2 + 2p_0 - 2q_0^2 = 0.
\]

(4.6)

However, only the smallest root (\( p_0 \approx q_0^2 \), when \( q_0 \ll 1 \)) is realistic for water waves.

The uniform solutions for the NLS− equation are obtained by changing the sign of \( q_0^2 \). If \( q_0^2 > 2^{-1/2} \) there is no real solution of equation (4.6), however this is not relevant to water waves.

Comparison of the Taylor series expansion (4.3) with the exact dispersion equation for deep water waves indicates that the series is a good approximation for a variation of 20% in \( k \) and for \( ak \) up to 0.2, that is \( q \) up to about 0.3.
5. The isolated soliton

The self-focussing NLS+ equation, with the boundary conditions $|q| \to 0$ as $|x| \to \infty$, has the isolated soliton solution

$$q = q_0 \text{sech} q_0(X - 2p_0 T) \exp i \{p_0 X + (q_0^2 - p_0^2) T\}. \quad (5.1)$$

As for the uniform solution, $p_0$ corresponds to a simple shift of carrier-wave wavenumber and hence it can often be set to zero without loss of generality. In that case, the soliton solution for equation (3.17) is

$$A = (\omega a_0/2k) \text{sech}^{1/2} a_0 k(kx - \frac{1}{2} \omega t) e^{-ia_0 k^2 \omega t/4}. \quad (5.2)$$

The nonlinear effect on wave frequency is only half that for the uniform wave train corresponding to the maximum amplitude. This difference is accounted for by the curvature, $q_{XX}$, of the wave envelope which has different signs for the high and low parts of the waves. The low parts travel faster and the high parts slower than the uniform solutions and thus the soliton maintains its integrity.

The velocity of solitons (5.1) depends only on $p_0$. Thus for fixed wavenumber, $p_0 = 0$, all solitons have the same velocity, zero in $(X, T)$ and $c_g$ in $(x, t)$ where $c_g$ is the linear group velocity for that wavenumber. For given frequency of the carrier wave

$$p_0^2 + 2p_0 - q_0^2 = 0, \quad (5.3)$$

there are two soliton solutions of amplitude $q_0$ corresponding to the two roots of equation (5.3), but again only the case with $p_0 = \frac{1}{2} q_0^2$ is relevant to water waves. The velocity is $\omega(1 + p_0)/2k$ in $(x, t)$ and is $2p_0$ in $(X, T)$.

The extent of the soliton in $X$ varies like $1/q_0$, so higher solitons are also shorter. See Figure 5 for examples. The solution (5.2) is a wave envelope, and the number of waves in the soliton can be determined for given maximum wave steepness if we assign a "length" to sech $x$. A convenient value is 3 since sech 1.5 = 0.425. In that case the number of deep-water waves in a soliton is $3/(4\pi a_0 k^2)$ at an instant and twice as many if the waves are counted as they pass a fixed point. This difference is due to the phase velocity being twice the group velocity, see Figure 1. For modulations at an angle to the wave direction transformation (3.8b) shows that the number of waves in a soliton of given steepness decreases like $(1 - 2\tan^2 \alpha)^{1/2}$. Similarly a soliton in water of finite depth has fewer waves for a given steepness, but the appropriate range of steepness is also reduced in this case.
Zakharov and Shabat [23] solved the NLS + equation with zero amplitude at infinity and demonstrated the soliton behaviour of these solutions. It is thus reasonable to consider the steepness versus modulation rate found here to be typical. From solution (5.2) we may identify \( K \) with \( 2^{1/2} a_0 k^2 \). Having done so we can now identify the region between shallow-water and deep-water modulations, as around

\[
Kh = 2^{1/2} a_0 k^2 h = 1. \tag{5.4}
\]

Lines corresponding to equation (5.4) are drawn in Figure 2, which covers the practical range of periods for ocean waves, and by suitable adjustment of units is appropriate for laboratory generated waves.

For given depth, the region well above the appropriate line in Figure 2 is where modulations are in deep-water; well below the line they are shallow-water modulations. In the region of the line the full equations (2.3) and (2.4) are needed. Although those equations do not simplify to an NLS equation it seems unlikely that their solutions differ greatly. Note that typical ocean waves (7–14 second period, steepness 0.05–0.2) are mostly in the intermediate region on continental shelves (50–200m deep). On the deep oceans, 4km deep, all significant waves have deep-water modulations.
6. The Ma soliton

Ma [13] extended the inverse scattering transform to solve the NLS+ equation with the uniform solution at infinity. That is,

\[ q(X, T) \to q_0 e^{2i\varphi T} \text{ as } |X| \to \infty. \]  \hfill (6.1)

The soliton solution that Ma derives, is, after some simplification and choice of the space and time origins,

\[ q = q_0 e^{2i\varphi T} \left[ 1 + \frac{2m(\cos 4mq_0^2 T + i\sin 4mq_0^2 T)}{n \cosh 2mq_0 X + \cos 4mq_0^2 T} \right], \]  \hfill (6.2)

where \( n^2 = 1 + m^2 \); an unsteady solution with period \( \pi/2mq_0^2 \).

For values of \( m \gg 1 \),

\[ q \sim q_0 e^{2i\varphi T} + 2mq_0 e^{4im^2q_0^2 T} \text{sech} 2mq_0 X, \]  \hfill (6.3)

which is the superposition of the uniform solution and a very much larger soliton of amplitude \( 2mq_0 \).

The modulus of a Ma soliton of moderate amplitude, \( m = 0.8 \), is illustrated for half a period in Figure 3.
Figure 3. One half period of the Ma soliton (6.2) for \( m = 0.8 \). The quantity \( |q/q_0| \) is given at
intervals of 30° in the variable \( 4mq_0^2T \).

For \( m \to 0 \), the uniform solution is not obtained. It is straightforward to verify that at

\[
4nq_0^2T = (2j + 1)\pi, \quad j = 0, \pm 1, \ldots,
\]

the solution has two zeros given by

\[
n \cosh 2mq_0X = n^2 + m^2.
\]

These always exist since \( n > 1 \). Midway between the zeros, at \( X = 0, \ |q/q_0| = (2n + 1) \) which approaches 3 as \( m \to 0 \). Figure 4 illustrates \( |q| \) for \( m = 0.1 \).

Notice particularly that there are two different amplitude scales used in the figure. The growth of the amplitude peak appears to be rapid. This is not so, the length of the soliton period increases indefinitely as \( m \to 0 \), the growth occurs on a time scale \( O(q_0^2) \).

Note that the amplitude can be scaled out of the Ma soliton by putting

\[
q' = q/q_0, \quad x' = q_0X, \quad t' = q_0^2T.
\]

A double Taylor series expansion about the amplitude peak gives a new solution of the NLS + equation:

\[
q' = e^{2\nu} \left[ 1 - \frac{4(1 + 4it')}{1 + 4x' + 16t'^2} \right].
\]

The method of deriving this solution is consistent with that for other rational solutions, see Ablowitz and Segur [2], Section 3.4.
Figure 4. One half period of the Ma soliton (6.2) for \( m = 0.1 \). The quantity \(|q/q_0|\) is given at intervals of 30° and 10° in \( 4mq_0^2 \). Note that the vertical scale of the lower 6 profiles is 100 times as great as that for the upper four.

The Ma solitons for small \( m \) are illustrative of the instability of the uniform solution. A disturbance with maximum value \( m^2q_0 \), when \( T = 0 \) in equation (6.2) grows in a time \( \pi/4mq_0^2 \) to have maximum amplitude of \( 3q_0 \). The disturbance then decays again, a behaviour which is similar to the water wave experiments of Lake et al. [10] where modulated waves return to a uniform condition.

7. Bound solitons

The solitons of the self-focussing NLS+ equation differ from those of the defocussing NLS− and KdV equations. Different solitons may have the same velocity. Thus in circumstances where more than one soliton exists it is possible
for two or more solitons to remain close to each other and to interact indefinitely. The inverse scattering transform may be used to find these “bound soliton” solutions. Each of these “multi-solitons” has properties like a single soliton in that they interact with other solitons, or multi-solitons, of different velocities but eventually regain their identity apart from some displacement and phase shift. Here only the bi-solitons, two bound solitons, with \(|q| \to 0\) as \(|x| \to \infty\) are considered.

To find the most general form of bi-soliton solution of the \(NLS^+\) equation the inverse-scattering solutions of Zakharov and Shabart [23] may be used. From their equations (17’) and (18’) the solution for \(n\) solitons is

\[ q = 2 \sum_{k=1}^{n} \lambda_k^2 \phi_k, \tag{7.1} \]

where \(\phi_k = \psi_{2k}^*\) in their notation, and is found from the set of equations

\[ \phi_j + \sum_{k=1}^{n} \sum_{l=1}^{n} \frac{\lambda_j^* \lambda_k^* \lambda_l^*}{(\xi_j^* - \xi_k^*)(\xi_k^* - \xi_l^*)} \phi_l = \lambda_j^*. \tag{7.2} \]

The \(\xi_j\) are the eigenvalues of the scattering problem, and

\[ \lambda_j = c_j^{1/2} \exp \left( i \xi_j X + 4i \xi_j^2 T \right), \tag{7.3} \]

where \(c_j\) is a constant.

These expressions involve four complex constants for \(n = 2\); they are \(c_j\) and \(\xi_j\). For bound solitons these simplify a little since \(\xi_1 = \xi_2\), where

\[ \xi_j = \xi_j + i \eta_j. \tag{7.4} \]

In this development we are not concerned with particular initial conditions for the equation but are looking for general solutions for bi-solitons. This means we can simplify expressions by choosing the origins of \(X\) and \(T\) appropriately and eliminating constants that correspond to a uniform change in the carrier-wave frequency and wavenumber. The carrier wave wavenumber is changed by the constant \(\xi_j\) and we may put \(\xi_j = 0\) without loss of generality.

With

\[ \lambda_j = c_j^{1/2} \exp \left( -\eta_j X - 4i \eta_j^2 T \right), \tag{7.5} \]

it may be seen that changing the value of \(c_j\) corresponds to changing the origin of \(X\) and \(T\), thus for the present purposes it is only the relative change of origin between \(\lambda_1\) and \(\lambda_2\) that is important. This leaves \(\eta_1, \eta_2\) and the relative displacement of the solitons in space as the only significant parameters. The bi-soliton solution is periodic in time so that even the relative shift in time is unimportant. As may be noted later only the ratio \(\eta_1/\eta_2\) gives qualitatively different solutions for different \(\eta_1\) and \(\eta_2\).
The following choice of constants leads to some simplification in the final solution

\[ \lambda_1 = (LM)^{1/2} \exp\left(-\frac{1}{2}MX_1 - \frac{1}{2}iM^2T\right), \]  
(7.6)

\[ \lambda_2 = (LN)^{1/2} \exp\left(-\frac{1}{2}NX_2 - \frac{1}{2}iN^2T\right), \]  
(7.7)

where

\[ X_2 - X_1 = \text{constant} = B, \]  
(7.8)

\( M \) and \( N \) are positive constants and it is convenient to introduce

\[ L = \frac{M - N}{M + N}, \quad K = \frac{2(MN)^{1/2}}{M + N}, \quad \text{with} \quad K^2 + L^2 = 1. \]  
(7.9)

There are various ways of writing the bi-soliton solution which is eventually found. Three forms which have been found to be convenient are

\[ q = \frac{2L\left(e^{iM^2T}M \cosh NX_2 - e^{iN^2T}N \cosh MX_1\right)}{L^2 \cosh(MX_1 + NX_2) + \cosh(MX_1 - NX_2) - K^2 \cos(M^2 - N^2)T}, \]  
(7.10)

\[ q = \frac{2Le^{iM^2T}(M \sech MX_1 - e^{iSN} \sech NX_2)}{1 + L^2 - K^2 \tanh MX_1 \tanh NX_2 - K^2 \cos S \sech MX_1 \sech NX_2}, \]  
(7.11)

and

\[ q = \frac{e^{iM^2T}M \sech MX_1 - e^{iN^2T}N \sech NX_2}{\cosh J - \sinh J(\tanh MX_1 \tanh NX_2 + \cos S \sech MX_1 \sech NX_2)}, \]  
(7.12)

where

\[ S = (M^2 - N^2)T \quad \text{and} \quad \tanh J = K^2/(1 + L^2) = 2MN/(M^2 + N^2). \]  
(7.13)

In the form (7.12) the individual solitons are displayed in the numerator and the interaction effects are thus entirely in the denominator. For example if \( B \) is very large, the denominator becomes approximately

\[ \cosh(MX_1 - J) \sech MX_1 \]  
(7.14)

in the neighbourhood of the origin of \( X_1 \), showing that the distant soliton of amplitude \( N \) causes a shift of origin of \( J/M \). To the next approximation the term in \( \cos S \) gives a periodic oscillation in each soliton's position, a result obtained by
Karpman and Solov'ev [9] who solve perturbation differential equations for this type of interaction between solitons.

The bi-soliton solution in the form (7.11) is convenient for examining some of its properties since except for the factor $e^{iM^2T}$, all the time variation is in terms of $S$. Thus $|q|$ has period $2\pi/(M^2 - N^2)$.

Another property of all bi-solitons is that $q = 0$ when both $S = 0 \mod 2\pi$ and

$$M \text{ sech } MX_1 = N \text{ sech } NX_2.$$  \hfill (7.15)

Equation (7.15) always has two real solutions since the side of the equation which has the largest maximum value also decays most rapidly to zero as $|X| \to \infty$. These zeros would occur even with a simple superposition of solutions since solutions of different amplitudes have different frequencies.

The zeros of $q$ are a striking feature of some diagrams of $|q|$ and of any plot of $\arg q$. Some examples of particular solutions are given in Figures 5, 6 and 7; $N$ is always taken equal to unity since the solution for some other value, with the same ratio $M/N$, is easily found by the transformation (6.6).

Figure 5 shows the amplitude for the case where two solitons of disparate size are at the same point, i.e. $X_1 = X_2$. It may be noted that the outskirts of the longer soliton appear to be unaffected by the higher soliton.

Figure 6 shows two solitons of similar size at the same point. Now the difference wavenumber, $M - N$, is much smaller and hence this solution has a greater spatial extent than either individual soliton, note the difference in the scale of $X$ from Figure 5. The interaction between the solitons is also greater. For most of the period there are two equal symmetrically displaced "solitons" and these come together at the centre for only a small portion of the period. It should be noted that the period becomes longer as $M \to N$, and the amplitude peak that occurs between the two zeros still grows on a time scale of $O(\frac{1}{4}N^2)$ which is short compared with the period.

A solution with solitons displaced, $X_2 - X_1 = 0.4$, is shown in Figure 7. The same values of $M$ and $N$ are used in both Figures 6 and 7.

Although in all the above examples two solitons that make up the bi-soliton can be identified from the diagrams, this is not always the case. For example $M = 3, N = 1$, $X_1 = X_2$, is an intermediate case between those of Figures 5 and 6. It is the bi-soliton illustrated by Miles [15] (also Satsuma and Yajima [19]) and has

$$|q| = 2 \text{ sech } X \text{ at } S = \pi,$$  \hfill (7.16)

as may readily be found from solution (7.10). The transition from a maximum at $X = 0, S = \pi$ occurs at $M/N = 2.618$. 

https://doi.org/10.1017/S0334270000003891 Published online by Cambridge University Press
Figure 5. One half period of the bi-soliton (7.11) for $M = 5$, $N = 1$, and $X_1 = X_2$. The quantity $|q|$ is plotted at intervals of $30^\circ$ in $S = (M^2 - N^2)T$. The two individual solitons are illustrated above.

In the limit $M \to N$ the bi-soliton becomes aperiodic. In deriving the limiting solution from expression (7.10) one finds that the displacement between solitons must also tend to zero. In the joint limit

$$M - N \to 0, \quad \frac{X_2 - X_1}{M - N} \to b, \quad (7.17)$$

the solution for $M = N = 1$ is

$$q = \frac{4e^{iT}[(1 + 2iT) \cosh x - (X - b) \sinh X]}{1 + 2(X - b)^2 + 8T^2 + \cosh 2X}. \quad (7.18)$$
Figure 6. One half period of the bi-soliton (7.11) for \( M = 1.2, N = 1, \) and \( X_1 = X_2. \) The quantity \(| q |\) is plotted at intervals of 30° in \( S = (M^2 - N^2)T. \) A supplementary profile is shown by a broken line.

This solution is not very different from the examples of Figures 6 and 7 in the region around the origin.

At large distances from the origin,

\[
q \approx \frac{8ie^{iT}Te^{i|X|}}{16T^2 + e^{-2|X|}} = \frac{ie^{iT}}{\cosh(|X| - 2\log|2T|)}. \tag{7.19}
\]

That is, the solution represents a pair of equal solitons with the distance between them proportional to \( 4\log|2T|, \) a result deduced by Zakharov and Shabat [23]. It is interesting to note that the solution is only symmetrical in \( X \) when \( b = 0. \) For the asymmetric case the solitons do not “pass” through each other at \( T = 0 \) but “bounce” off each other. See Figure 7 and note that \(| q |\) is symmetrical in \( T.\)

Multiple bound solitons (multi-solitons) are not in general periodic. They are only periodic when all the difference frequencies, \( \eta_1^2 - \eta_j^2, \) are integer multiples of a single number. Swenson (private communication) reports that computation of the tri-soliton in the family of solutions examined by Miles [15] reveals the approach to its maximum amplitude is very similar to that of the bi-soliton formed from the two largest eigenvalues.
There are no solutions of the defocussing NLS− equation corresponding to the isolated soliton (5.1) of the NLS+ equation. However, it is straightforward to find solutions of the form

\[ q = Q(X - CT)\exp\{-2i\beta_0^2 T - iF(X - CT)\}, \quad (8.1) \]

which tend to the uniform solution \( q_0 e^{-2i\beta_0^2 T} \) as \( |X - CT| \to \infty \). The solution is

\[ Q^2(z) = q_0^2\{1 - \sin^2 B \sech^2(q_0 z \sin B)\}, \quad (8.2a) \]

and

\[ F(z) = \arctan\{\tan B \tanh(q_0 z \sin B)\} \quad (8.2b) \]

where

\[ C = \pm \frac{1}{2} q_0 \cos B. \quad (8.2c) \]

In these expressions \( B \) is a constant such that \( q_0^2 \sin^2 B \) is the maximum deviation of \( |q|^2 \) below the uniform level \( q_0^2 \).
The solution (8.2) has all the usual soliton features as is shown by Zakharov and Shabat [24]. Their arrangement of the solution is

\[ q = \frac{q_0 \left[ e^{tB} + \exp \left\{ 2q_0 (X - CT) \sin B \right\} \right] e^{-2q_0^2 t - 1a}}{1 + \exp \left\{ 2q_0 (X - CT) \sin B \right\}}, \]  

(8.3)

in our notation.

These solutions have a maximum amplitude, when \( B = \frac{1}{2} \pi \). The limiting wave is

\[ q = q_0 \tanh q_0 X e^{-2q_0^2 T}, \]  

(8.4)

and is stationary in \((X, T)\).

For small values of \(|q - q_0|\) there is a mathematical analogy between the defocussing NLS– and the Boussinesq equations for shallow water waves for which surface tension dominates the dispersive effects. The above soliton solutions hence correspond to the shallow water solitary wave and the solitons of the KdV equation which is derived from the Boussinesq equations. Details are in Peregrine [18] which, among other things, discusses wave focussing as described by the NLS equation (3.6). Since the Boussinesq equations are only appropriate for near-linear waves (see Peregrine [17]) this correspondence with the NLS–equation may be useful in studying their solutions.

9. Periodic solutions

Corresponding to all of the solitons discussed above (isolated soliton, Ma soliton, bi-soliton and dark soliton) there are solutions periodic in space. For the two solitons with steady profiles the periodic solutions are known and can be expressed in terms of elliptic functions. For the others periodic solutions of long wavelength are readily obtained by “matching” solitons with intervening stretches of uniform or zero solution as appropriate. It is likely that shorter wavelength solutions also exist.

A spatially periodic solution corresponding to the Ma soliton will be similar to the solution obtained from the initial conditions of a uniform wave train with small sinusoidal modulation. Such solutions have been numerically computed and are illustrated in Figure 12, cases 1 and 1a of Yuen and Lake [22].

https://doi.org/10.1017/S0334270000003891 Published online by Cambridge University Press
It is doubtful if the aperiodic amplitude "peak" solution has a spatially periodic counterpart since on the "outskirts" of such waves the linearized equation

\[ iq_t + q_{xx} = 0 \]  \hspace{1cm} (9.1)

should be satisfied. If \( q \) is periodic in \( x \) then consideration of a Fourier component shows that it is also periodic in time.

The period in time of the oscillating solitons is likely to be changed by the presence of neighbouring solitons as is readily seen by considering the sketch in Figure 8 of a space periodic solution that might be obtained by combining equal-bi-soliton solutions (7.18) with \( B = 0 \). The lines correspond to the local maxima of \( |q| \). Such a solution is periodic in time, although solution (7.18) is aperiodic.

\[ \text{Figure 8. Sketch of a space-periodic solution of the NLS+ equation which may be obtained by matching an array of the equal bi-soliton solutions (7.18) for } b = 0. \text{ The lines correspond to the local maxima of } |q|. \]

By incorporating solitons of differing velocities so that they can pass through each other, much more intricate patterns than Figure 8 can be conceived which are periodic in both space and time. It is doubtful whether it is worth pursuing these solutions rather than examining the basic elements of such patterns as is done here.

A solution which appears to be equivalent to a periodic bi-soliton is described by Bryant [4]. It is computed from the evolution of a set of Fourier modes.
10. Conclusion

This paper has reviewed explicit solutions of one-dimensional NLS equations. These equations are obtained by considering restricted classes of modulations of water waves. Section 3 lists a range of such classes but discussion is mainly confined to the uni-directional unsteady modulations of deep water waves. It is well known that such waves are unstable to long modulations, the Benjamin-Feir instability. The Ma soliton illustrates this instability.

The periodicity of the Ma solitons reflect the long period return to uniformity that Lake et al. [10] found in experiments on weakly modulated deep-water wave trains. In the experiments the steeper waves had a lower frequency and wavenumber when uniform conditions returned. This shift in the carrier wave is not modelled by the NLS computations of Lake et al. [10] or by the solutions discussed here. However, the Ma soliton, or the amplitude peak (6.7), may provide a suitable starting point for a more accurate theoretical investigation.

The amplitude-peak solution (6.7) shows how the self-focussing of an undisturbed uniform wave, for \( T \to -\infty \), can grow into a disturbance of three times the original wave amplitude. It would be interesting to have more details of experimental results in order to see whether this triple amplification is typical of the growth of instabilities.

The steady one-dimensional solutions \( q(X, T) \) of the two-dimensional NLS equation

\[
iq_T + q_{XX} - q_{YY} + 2|q|^2q = 0, \tag{10.1}
\]

are unstable to two-dimensional disturbances. This equation is equivalent to equation (3.1) and governs weakly nonlinear modulations of deep water waves. The modulations are three-dimensional in physical space. Yuen and Lake [22] give a substantial review of deep-water instabilities. See also Ablowitz and Segur [1] and Larsen [11]. Water waves in sufficiently narrow channels are immune to these instabilities.

The discovery of Benjamin-Feir instability and the NLS isolated solitons stimulated study of wave groups among ocean waves (Mollo-Christensen and Ramamonjiarisoa [16]), with the implicit hope that some improved representation of the ocean surface and its statistics might emerge with the inclusion of these nonlinear phenomena.

The three-dimensional instabilities of deep water waves imply that such an approach is unlikely to be sufficient. Despite this, the two-dimensional case described by the one-dimensional NLS equation still merits study. If a suitable wave representation is found it may indicate how the three-dimensional problem...
should be tackled; there are experiments in narrow channels to be interpreted, and there are applications to waves in channels and similar restricted waters. The discussion here is based on this view, and not on any belief that the NLS+ equation is entirely appropriate.

One possible avenue of interpreting wave behaviour arises from the occurrence of peaks of amplitude in numerical solutions of the NLS equation and of similar calculations with sets of Fourier components, e.g. see Figure 12 of Yuen and Lake [22] (see also Figure 18 which shows peaks in $(x, y, t)$ for the two-dimensional NLS equation (10.1)). It appears from the diagrams that many such peaks are similar to those found here in that there is a zero on each side of a peak.

The peaks represent the steepest water waves. However a note of caution: the similarity that appears may be rather superficial as can be seen from the solutions in this paper. The peak of a Ma soliton differs very considerably from the peak of a symmetrical bi-soliton. Not only are the analytic forms different, e.g. compare (6.7) with (7.19) for $B = 0$, but the pattern of phase variation is also different. The two cases are sketched in Figure 9, in each case the exponential factor has been excluded to simplify the diagram.

The Ma-peak is probably more relevant to the propagation of a wave field since it sits in a uniform background; whereas the bi-soliton has a zero background. In view of the frequency shift of the carrier wave that occurs in water wave experiments it is interesting to note that the gradient of phase between the zeros corresponds to an increased frequency. Perhaps finite amplitude or high-order modulation effects in some way negate this increase of frequency without affecting the subsequent decrease.

The solitons are only part of the solution to an NLS problem. An initial-value problem also gives rise to a continuous spectrum in the inverse scattering method. For initial disturbances of finite extent this part of the solution eventually decays like $t^{-1/2}$. For some ocean wave propagation the initial generation area may be so large that even a trans-oceanic distance may be insufficient for solitons to reach their asymptotic dominance. Generally, after a sufficiently great distance of propagation only solitons with closely similar velocities would contribute to the waves at any one place.

An important aspect of ocean-wave propagation is the change in the waves' characteristics as they propagate into coastal waters. Figure 3 demonstrates that as typical ocean swell passes onto a continental shelf it passes through a depth range in which the modulation length is comparable with the depth and the coupled equation for wave-induced flows does not simplify. Although this is a region where an NLS equation is not directly applicable it would be surprising if the solutions of these equations differed in character from the NLS solutions (e.g.
Figure 9. (a) Sketch of the lines of constant phase of $q e^{-2iT}$ near the peak of the limiting Ma soliton (6.7). (b) Sketch of the lines of constant phase of $q e^{-2iT}$ near the peak of the limiting bi-soliton (7.19).
long wave propagation in shallow water is described by the KdV equation which can be solved by the inverse scattering method; however physically relevant problems are equally well modelled by a whole range of equations, see Broer [3]).

A change of more significance is that from deep-water waves to shallow-water waves. The type of governing NLS equation changes from self-focussing to defocussing. There does not appear to be the same wealth of analytic solutions for the latter equation. Others have studied this change (e.g. Johnson [8] and Larsen [12]). For the present we note just two things. (i) The periodic modulations (and other solutions) show “peaky” modulations for the self focussing equation and “flat-topped” modulations for the defocussing equation (see Figure 10). This might show up in wave statistics in the appropriate circumstances. (ii) The time scale of the solutions discussed here is long. It may in some circumstances mean that the region over which propagation conditions vary is too short for weakly nonlinear effects to have $O(1)$ effects. Typical evolution time scales are $O(1/a^2k^2\omega)$ and in that time waves propagate a distance $O(c_g/a^2k^2\omega)$. For example, for 14 second waves with $ak = 0.05$ this gives a distance of about 10km. Thus it is only for very long gentle waves or steep continental slopes that the evolution distance is longer than the topographic scale.

![Figure 10. Typical periodic modulation solutions, (a) deep water, (b) for waves with $kh < 1.3$.](https://doi.org/10.1017/S0334270000003891)

**Acknowledgement**

This work was prepared while the author was visiting the Institute of Geophysics and Planetary Physics, University of California, San Diego. Receipt of a Cecil and Ida Green Scholarship in partial support of the visit is gratefully acknowledged.
Appendix

Full expressions for equations (2.3) and (2.4) are

\[ 2i\omega (A_t + c_g A_x) - \left[ c_g^2 - gh(1 - p^2)(1 - khp) \right] A_{xx} + \left( \omega c_g/k \right) A_{yy} \]
\[ = (k^4/2p^2)(9 - 12p^2 + 13p^4 - 2p^6) |A|^2 A \]
\[ + k^2 [2c + c_g(1 - p^2)] A_{\Phi x} |z=0, \quad (A.1) \]
\[ \nabla^2 \Phi = 0 \quad \text{in} \quad -h < z < 0, \quad (A.2a) \]
\[ \Phi_x h_x + \Phi_y h_y = \Phi_z \quad \text{at} \quad z = -h, \quad (A.2b) \]
\[ g \Phi_z + \Phi_{tt} = k^2 [2c + c_g(1 - p^2)] (|A|^2)_x \quad \text{at} \quad z = 0, \quad (A.2c) \]

where

\[ c = \omega/k \quad \text{and} \quad p = \tanh kh = \omega^2/gk. \quad (A.3) \]

These equations do not appear to have been stated before. They are readily derived from the results presented by Dayey and Stewartson [5] and Dysthe [6]. Dysthe’s equation for \( \Phi \) only needs the bed boundary conditions (A.2b) in order to be appropriate for a large, finite depth of water. It only needs the relation

\[ \Phi(x, y, z, t) = \Phi_0(x, y, t) - \frac{1}{2}(z + h)^2 \Phi_{0xx} + \cdots \quad (A.4) \]

in the Davey and Stewartson case, to be noted together with the form of the forcing of \( \Phi \) given by Dysthe to deduce the boundary condition (A.2c).

In cases where (A.1) and (A.2) cannot be simplified to the deep-water limit or the Davey-Stewartson equation, i.e. when \( Kh = O(1) \), the coefficients in the equations can be simplified since \( K \ll k \) and hence the carrier wave is a deep-water wave train, \( p = 1, c_g = \frac{1}{2}c \).

The order of magnitude of \( \Phi \) is not immediately clear. Suppose, for the purpose of estimating it, that

\[ |A| = A_0 \cos K(x - c_g t). \quad (A.5) \]

Then

\[ \Phi = B \frac{\cosh 2K(z + h)}{\cosh 2Kh} \sin 2K(x - c_g t) \quad (A.6) \]

where

\[ B = \frac{k^2 [2c + c_g(1 - p^2)] A_0^2}{2g \tanh 2Kh - 4Kc_g^2}. \quad (A.7) \]
Thus, for $K_h$ large, $c_g^2 = \frac{1}{2} g / k$, and

$$B = \left( \frac{\omega k}{g} \right) A_0^2$$

since it is assumed *ab initio* that $K \ll k$. Thus the term $A \Phi_x$ is smaller than other terms in (A.1) since the derivative is of order $K$. Hence in deep water wave-induced flows are only significant at the next order of approximation, given by Dysthe [6].

The case of shallow-water modulations, $K_h \ll 1$ gives

$$B = \frac{k^2 \left[ 2c + c_g(1 - p^2) \right] A_0^2}{4K \left( gh - c_g^2 \right)}$$

with the order of magnitude of $\Phi$ larger than it is in deep water by a factor $1/K$. Davey and Stewartson [5] comment on the nonuniformity of the approximation for shallow-water waves as $c_g^2 \to gh$.

References


