Lecture 16: Solitary waves

Lecturer: Roger Grimshaw. Write-up: Daisuke Takagi, Yiping Ma and Andrew Stewart

June 24, 2009

1 Introduction

A solitary wave is a wave which propagates without any temporal evolution in shape or size when viewed in the reference frame moving with the group velocity of the wave. The envelope of the wave has one global peak and decays far away from the peak. Solitary waves arise in many contexts, including the elevation of the surface of water and the intensity of light in optical fibers. A soliton is a nonlinear solitary wave with the additional property that the wave retains its permanent structure, even after interacting with another soliton. For example, two solitons propagating in opposite directions effectively pass through each other without breaking.

Solitons form a special class of solutions of model equations, including the Korteweg de-Vries (KdV) and the Nonlinear Schrödinger (NLS) equations. These model equations are approximations, which hold under a restrictive set of conditions. The soliton solutions obtained from the model equations provide important insight into the dynamics of solitary waves. However, they are limited by the conditions under which the model equations hold. An alternative approach, which deals directly with the exact equations from which the model equations are derived, provides insight into a larger class of solitary waves than those obtained from the model equations.

In this lecture, we show that important properties of solitary waves can be determined directly from the exact equations governing a physical system by an asymptotic perturbation procedure. Information about the possible existence of certain types of solitary waves is obtained using a phase-plane formalism, a common technique in dynamical systems. In this framework, a solitary wave corresponds to a homoclinic orbit in a spatial dynamical system. Solitary waves exist in a number of different cases to be determined, given that they must decay far away from their global peak. A close examination of the tail regions, far away from the global peak, indicates that there are four possible cases of solitary waves. Three of the four possible cases are presented in detail. In the first case, the steady-state solution of the KdV equation is obtained. In the second case, a generalized solitary wave is obtained, which decays to non-zero oscillations of constant amplitude and wavenumber. In the third case, an envelope solitary wave is obtained, which satisfies the NLS equation.

2 Formulation of solitary waves as a dynamical system

Consider a solitary wave propagating with speed $c$ in the positive direction of the $x$ axis. The wave amplitude in the moving reference frame is the function $A(\xi)$ where $\xi \equiv x - ct$.  


Without loss of generality we can select $\xi$ such that the peak of $A$ is at $\xi = 0$, while $A$ decays in the tail regions, in the limits as $\xi \to \pm\infty$. In the tail regions, where a linearized analysis is assumed to hold, it is fruitful to seek solutions proportional to the real part of $\exp(ik\xi)$, where $k$ is a complex-valued wavenumber to be determined. The wavenumber must not be purely real, given that the wave amplitude must decay in the tail regions. This leads to the notion that solitary waves with a real phase speed $c$ only exist within a limited range of wavenumbers in the spectrum, called gaps.

A relationship between the phase speed $c$ and the wavenumber $k$ is obtained by linearizing the governing equations of a physical system of interest, which yields a linearized dispersion relation. For example, the linearized dispersion relation of surface water waves is given by

$$\frac{c^2}{gh} = f(q) = \left[(1 + Bq^2)/q\right] \tanh q,$$

where $g$ is the acceleration due to gravity, $h$ the constant depth of the water when it is unperturbed and $q = kh$ the dimensionless wavenumber. The dimensionless number $B = \sigma/\rho gh^2$ is the Bond number, which measures the relative importance of surface tension and gravity, where $\sigma$ is the coefficient of surface tension and $\rho$ is the density of water. A plot of the function $f(q)$ given by equation (1) for three representative values of the Bond number $B$ is shown in Figure 1.

Solitary waves can exist provided that no real value of $k$ satisfies the dispersion relation, meaning that $k$ has a non-zero imaginary part. In the case of $B = 0$ (no surface tension), solitary waves can exist only when $c^2 > gh$. When effects due to surface tension are
significant, such that $B > 1/3$, $q$ is an increasing function of $c$ with $q \geq 0$ for $c^2/gh \geq 1$. An example with $B = 0.4$ is shown in figure 1 (blue curve). Thus, solitary waves with $B > 1/3$ can only exist when $c^2/gh < 1$. Finally, when $0 < B < 1/3$, a real-valued $q$ satisfies (1) for any speed $|c| > c_m$ for some positive $c_m$. Figure 1 shows a plot of (1) with $B = 0.2$, which attains a local minimum $c_m^2/gh$ at some non-zero wavenumber $k_m$. This indicates that solitary waves with $0 < B < 1/3$ can exist only when $c^2 < c_m^2$.

Solitary waves are found by reformulating the problem in the framework of a dynamical system. Let’s assume the system is conservative, which is the common and traditional scenario for solitary waves. The underlying physical system is Hamiltonian and reversible, meaning that the system conserves energy and is symmetric under the transformation $\xi \to -\xi$. This implies that if $k$ is a solution then $-k$ is also a solution due to reversibility. If $k$ has an imaginary part (as is required for a solitary wave), then $k^*$ is a solution because the dispersion relation for real-valued phase speed $c$ has real coefficients. The solutions generically form a quartet $(k, k^*, -k, -k^*)$ with an associated four-dimensional subspace for the corresponding wave mode. In the general case when the imaginary part of $k$ is non-zero, there are typically two roots in the limit as $\xi \to \infty$ and two other roots as $\xi \to -\infty$. In the limit as $\xi \to \infty$, the imaginary part of $k$ must be positive provided that the solitary wave decays in the tail regions. Likewise, the imaginary part of $k$ must be negative as $\xi \to -\infty$.

Let us now study the amplitude $A(\xi)$ in the framework of a dynamical system. Consider the phase plane where the $\xi$-derivative of the wave amplitude is plotted against the amplitude. Trajectories in this phase planes are the curves $(A(\xi), A_\xi(\xi))$. This can either represent the whole system (for a 2D problem) or a projection of the whole system (for a higher-dimensional problem). Given the form of $A$ at infinity discussed above, we deduce that the trajectory is a homoclinic orbit from the origin. Indeed, the origin corresponds to the tail regions as $\xi \to \pm\infty$, where both $A$ and $A_\xi$ tend to 0. Meanwhile, the peak of the wave ($\xi = 0$ point) lies somewhere on the $A_\xi = 0$ axis. If $k$ is pure imaginary then the trajectories satisfy $A_\xi = -|k|A$ near the origin, and approach it directly. If $k$ has a real part, then the trajectories spiral into the origin instead.

Consider now how the trajectory in the phase plane may change as parameters governing the original system are varied. Changes in the trajectory far from the origin of the phase plane mean that the overall shape of the wave $A(\xi)$ is qualitatively changed, but that its asymptotic behavior at $|\xi| \to \infty$ remains the same. However, much more dramatic changes in $A(\xi)$ occur if the behavior near the phase plane origin changes (for example when $k$ goes from being real to complex to pure imaginary, or vice-versa).

This leads us to consider how the quartet structure $(k, k^*, -k, -k^*)$ evolves as some global system parameter is varied. Clearly bifurcations arise when two solutions for $k$ coalesce, for which the necessary condition is that $\partial c/\partial k = 0$. This condition is equivalent to the condition that the phase speed $c$ is equal to the group velocity $c_g$ (i.e. the weak dispersion limit), when the bifurcation occurs at a real value of $k$. The equivalence follows immediately from the equation $c_g = c + k\partial c/\partial k$. For example, from Figure 1 we see that bifurcations arise in water waves when $(k, c^2)$ takes the values $(0, gh)$ or $(k_m, c_m^2)$.

Generically, there are two possible quartet structures at each of the bifurcation points:

- at $(k = 0, c^2 = gh)$: we either have the quartet $(0, 0, i\gamma, -i\gamma)$ or $(0, 0, \beta, -\beta)$, (two roots have coalesced at the origin, the other two are either real or pure imaginary)
• at \((k = k_m, c^2 = c_m^2)\): we either have \((\beta, \beta, -\beta, -\beta)\) (cc pairs coalesce on the real axis) or \((i\gamma, i\gamma, -i\gamma, -i\gamma)\) (pairs coalesce on the imaginary axis), where \(\beta\) and \(\gamma\) are real-valued.

The first three of the four possible cases will be examined in detail later. The fourth case, for the quartet \((i\gamma, i\gamma, -i\gamma, -i\gamma)\), has only rarely been studied and will not be considered here.

Consider a projection of the full system onto the appropriate four-dimensional subspace. The resulting bifurcation is analyzed within the framework of this subspace. A 4-vector \(W(\xi)\), representing the structure of the subspace, satisfies a first-order differential equation of the form

\[
dW/d\xi = L(W; \epsilon) + N(W),
\]

where \(L(W; \epsilon)\) is a linear operator and \(N(W)\) contains all nonlinear terms. The parameter \(\epsilon\) represents the distance from the bifurcation point. Near the bifurcation (which takes place at \(\epsilon = 0\)) the linear operator \(L\) can be written as \(L(W; \epsilon) = L_0(W) + \epsilon L_1(W) + \ldots\), where the eigenvalues \(\lambda = ik\) of the operator \(L_0\) reproduce one of the four possible quartet structures described above. In each of the first three cases, the structural form of the small-amplitude solutions \(W\) is examined to identify and describe the nature of corresponding solitary waves.

3 The three cases

The three types of bifurcations may now be studied with the standard techniques of center manifold reduction and normal form analysis in bifurcation theory. For more details about these techniques, we refer the reader to the review article by Crawford [3].

3.1 Case (1)

Let us first consider case (1). At the bifurcation point \((\epsilon = 0)\) the linearized system (2) has eigenvalues \((0, 0, \mp \gamma)\). Since the \(\lambda = 0\) eigenvalue is degenerate, there will be a corresponding single eigenvector \(V_0\), and a single generalized eigenvector \(V_1\) such that \(L_0 V_1 = \lambda V_1 + V_0 = V_0\). Small-amplitude solutions are then sought in the form

\[
W = A(\xi)V_0 + B(\xi)V_1 + W^{(2)}. \tag{3}
\]

Here \(A, B\) are real variables of \(O(\alpha), \alpha \ll 1\), where \(\alpha\) measures the wave amplitude. The leading terms form a two-dimensional subspace \((A, B)\), while \(W^{(2)}\) is a small error term of \(O(\alpha^2, \alpha \epsilon)\), where \(\epsilon, \alpha \ll 1\) are both small parameters. Note that the two remaining eigenvalues \(\mp \gamma\) play no role at the leading order here, since they correspond to strong exponential decay at infinity, and their effects are included in the small error term \(W^{(2)}\).

Substituting (3) into (2) yields

\[
A_\xi V_0 + B_\xi V_1 + \frac{d}{d\xi} W^{(2)} = L_0 \left( A(\xi)V_0 + B(\xi)V_1 + W^{(2)} \right) + \epsilon L_1 \left( A(\xi)V_0 + B(\xi)V_1 + W^{(2)} \right) + N \left( A(\xi)V_0 + B(\xi)V_1 + W^{(2)} \right) \tag{4}
\]
Projection onto the two-dimensional subspace \((\mathbf{V}_0, \mathbf{V}_1)\) yields, to lowest order in \(\epsilon\) and \(\alpha\),

\[
A_\xi = B + O(\epsilon, \alpha^2, \ldots) \\
B_\xi = \mathbf{V}_1 \cdot \left[ \epsilon A L_1(\mathbf{V}_0) + \epsilon B L_1(\mathbf{V}_1) + L_0(\mathbf{W}^{(2)}) + N(\mathbf{A}\mathbf{V}_0 + \mathbf{B}\mathbf{V}_1) \right] + O(\alpha^2, \epsilon \alpha^2, \epsilon^2, \alpha^3, \ldots)
\]

The first two terms in the \(B\) equation are linear terms of \(O(\epsilon \alpha)\), and their projection onto \(\mathbf{V}_1\) yields \(\epsilon (c_1 A + c_2 B)\), where \(c_1\) and \(c_2\) are the projection coefficients. Similarly, the linear and nonlinear terms of \(O(\alpha^2)\) yield contributions of the kind \(A^2, AB\) or \(B^2\), so that

\[
A_\xi = B \\
B_\xi = \epsilon (c_1 A + c_2 B) + (d_1 A^2 + d_2 AB + d_3 B^2) + ... 
\]

where the omitted terms are \(O(\alpha \epsilon^2, \epsilon^2, \alpha^3)\). Finally, a normal form analysis reveals the normal form of the system near the bifurcation:

\[
A_\xi = B, \\
B_\xi = \epsilon A + \mu A^2 + ...
\] (5)

where \(\mu\) is a real-valued coefficient, specific to the system being considered.

The eigenvalues \(\lambda = i k\) of this system are \(\pm \epsilon^{1/2}\) if \(\epsilon > 0\), and \(\pm i |\epsilon|^{1/2}\) if \(\epsilon < 0\). As discussed earlier, only the former case, \(\epsilon > 0\), yields the solitary wave solution, with \(k = \pm i \epsilon^{1/2}\). Finally, comparison with the dispersion relation expressed near \(k = 0\)

\[
c(k) = c(0) + \frac{k^2}{2} \frac{d^2 c}{dk^2} \bigg|_{k=0} + ... 
\] (6)

leads to the identification of \(\epsilon\) as

\[
\epsilon = -\frac{2(c - c(0))}{c_{kk}(0)}. 
\] (7)

Hence, for solitary waves to exist \((\epsilon > 0)\), we either need \(c > c(0)\) if \(c_{kk}(0) < 0\), or \(c < c(0)\) if \(c_{kk}(0) > 0\). These inequalities recover the graphical argument discussed in the previous section.

When the error terms in (5) are omitted, we can eliminate \(B\), and the resulting ODE for \(A\) can be recognized as the steady-state KdV equation, which has the well-known “sech^2” solution. It is then a delicate and intricate task to establish that this solitary wave solution persists when the error terms are restored.

This dynamical systems approach to the problem has therefore established the steady-state KdV equation as the normal form for weakly nonlinear, weakly dispersive solitary waves (see also Lecture 6) whenever the dispersion relation satisfies \(dc/dk \to 0\) when \(k \to 0\), and has provided mathematical conditions on their propagation speed. The KdV is thus seen to be strictly correct for any solitary wave of the kind described above, far from the wave peak (i.e. in the low-amplitude tails). Whether this statement still holds nearer the peak \(\xi = 0\) depends on the peak amplitude \(A(0)\).
3.2 Case (2)

Next consider case (2), where the linearized system (2) at the bifurcation point ($\epsilon = 0$) has eigenvalues $(0, 0, \pm i\beta)$. Again, the degeneracy of $\lambda = 0$ implies that there is a single eigenvector $V_0$, and a single generalized eigenvector $V_1$. However, account must now be taken of the other two eigenvalues $\pm i\beta$, with their associated eigenvectors $V_2$, $V_2^*$, since they do not now lead to decaying solutions at infinity. Small-amplitude solutions are sought in the form

$$ W = A(\xi)V_0 + B(\xi)V_1 + C(\xi)V_2 + C^*(\xi)V_2^* + W^{(2)}. $$

Here $C$ is a complex-valued variable, and hence the leading terms form a four-dimensional subspace $(A, B, C)$, while $W^{(2)}$ is a small error term. Projection onto this four-dimensional subspace, followed by normal form analysis now reveals that $(A, B, C)$ satisfy the system

$$
\begin{align*}
A_{\xi} &= B, \\
B_{\xi} &= \epsilon A + \mu A^2 + \nu |C|^2 + \cdots, \\
C_{\xi} &= i\gamma(1 + \delta A)C + \cdots.
\end{align*}
$$

Here $\mu, \nu, \delta$ are real-valued coefficients specific to the system being considered, and the omitted terms are small error terms as above.

When the error terms are omitted the system is integrable. In that limit, it is easy to verify that $|C|^2_{\xi} = 0$ by constructing the quantity $C^*C_{\xi} + C^*_{\xi}C$. Hence, $|C| = C_0$ is a constant. By a change of origin from $A \rightarrow A + A_0$ with $\epsilon A_0 + \mu A_0^2 + \nu C_0^2 = 0$, the system reduces to the same form as (5) in case (1). Thus, for the case $\epsilon > 0$ (when case (1) is a KdV-type solitary wave), the solution is a one-parameter family of homoclinic-to-periodic solutions, with $|C| = C_0$ constant and $(A, B) \rightarrow (A_0, 0)$ as $\xi \rightarrow \pm \infty$. The solution is a generalized solitary wave which typically has a “sech^2” core, and decays at infinity to non-zero oscillations of constant amplitude $C_0$ and wavenumber $\gamma$ (Figure 2). A delicate analysis of the full system (2) with the small error terms shows that at least two of these solutions persist; the minimal amplitude $C_0$ being exponentially small, that is $O(\exp (-K/|\epsilon|^{1/2}))$ where $K$ is a positive real constant. Although such waves are permissible as solutions of the steady-state equations, they have infinite energy and their associated group velocity is inevitably inward at one end and outward at the other end. Hence, they cannot be realized in a physical system from any localized initial condition because energy is finite. Instead localized initial conditions will typically generate a one-sided generalized solitary wave, whose central core is accompanied by small-amplitude outgoing waves on one side only. Such waves cannot be steady, and instead will slowly decay with time as energy is radiated away by outgoing waves from the core.

3.3 Case (3)

Finally we consider case (3), when there is a double eigenvalue $\lambda = i\beta$ with generically a corresponding single eigenvector $V_0$, and a single generalized eigenvector $V_1$, while the complex conjugate double eigenvalue $\lambda = -i\beta$ has corresponding complex conjugate eigenvectors. Small-amplitude solutions are now sought in the form

$$ W = A(\xi)V_0 + B(\xi)V_1 + A^*(\xi)V_0^* + B^*(\xi)V_1^* + W^{(2)}. $$

163
Here $A$ and $B$ are complex-valued variables, forming a four-dimensional subspace while $W^{(2)}$ is again a small error term. Projection onto this subspace and a normal form analysis reveal that

$$A_\xi = i\beta A + B + iAP(\epsilon, |A|^2, K) + \cdots,$$
$$B_\xi = i\beta B + iBP(\epsilon, |A|^2, K) + AQ(\epsilon, |A|^2, K) + \cdots,$$  \hspace{1cm} (11)

where $K = i(AB^* - A^*B)$ (note that $K$ is real). Here $P$ and $Q$ are real-valued polynomials of degree 1, which take the form

$$P(\epsilon, |A|^2, K) = \epsilon + \nu_1|A|^2 + \nu_2K,$$
$$Q(\epsilon, |A|^2, K) = 2\epsilon\beta + \mu_1|A|^2 + \mu_2K,$$  \hspace{1cm} (12)

where all coefficients are real-valued. When the error terms in (11) are omitted, the resulting system is integrable. There are two constants of motion. The first one is $K$, which is verified by a direct construction of $K_\xi$. The second one is $H$, where

$$H = |B|^2 - (2\epsilon\beta|A|^2 + \frac{H_1}{2}|A|^4 + \mu_2K|A|^2).$$  \hspace{1cm} (13)

To prove that $H$ is constant, note that $(|A|^2)_\xi = A^*B + B^*A$ and that $(|B|^2)_\xi = Q(|A|^2)_\xi$. Hence, $(|B|^2)_\xi - Q(|A|^2)_\xi = 0$, so that the integral based on this quantity is constant. This integral is $H$. 

Figure 2: Generalized solitary wave.
For a solitary wave solution, conditions at infinity require $K = H = 0$. It then follows that

$$|B|^2 = 2\epsilon|A|^2 + \frac{\mu_1}{2}|A|^4$$

using the fact that $\langle |A|^2 \rangle_\xi = A^*B + B^*A$ and that $A^*B = B^*A$ is real. Thus solitary wave solutions exist provided that $\epsilon > 0$, and that the nonlinear coefficient $\mu_1 < 0$. The condition $\epsilon > 0$ implies that the perturbed eigenvalues, $\lambda \approx i\beta \pm (2\epsilon\beta)^{1/2}$ have split off the imaginary axis, and so provide the conditions needed for exponential decay at infinity. The normal form coefficient $\mu_1$ must be computed from the physical parameters through a reduction procedure [3], so the condition $\mu_1 < 0$ is problem-specific.

The normal form coefficient $\mu_1$ is determined by the physical parameters through a reduction procedure [3], so the condition $\mu_1 < 0$ is problem-specific. The solution of the truncated system governed by (14) is

$$A = a \exp (i[\beta + \epsilon]\xi) \text{sech}(\gamma \xi), \quad (15)$$

where $\gamma = (2\epsilon\beta)^{1/2}$, $|a|^2 = -4\epsilon\beta/\mu_1$. This solution describes an envelope solitary wave, with a carrier wavenumber $\beta + \epsilon$ and an envelope described by the “sech”-function. As we saw in Lecture 14, these solitary waves can also be obtained from the soliton solutions of the NLS equation, in the special case when the phase velocity equals the group velocity, $c = c_g$, or more precisely when $c + \Omega/K = c_g + V$, where $V$ is the soliton speed, $\Omega$ the frequency and $K$ the wavenumber correction. Note that the solution (15) contains an arbitrary phase in the complex amplitude $a$, meaning that the location of the crests of the carrier wave vis-a-vis the maximum of the envelope (here located at $\xi = 0$) is arbitrary.

However, restoration of the error terms in (11) leads to the result that only two of these solutions persist, namely, those for which a carrier wave crest or trough is placed exactly at $\xi = 0$, so that the resulting solitary wave is either one of elevation or depression. This result requires very delicate analysis, but could be anticipated by noting that these are the only two solutions which persist under the symmetry transformation $\xi \to -\xi$.

4 Applications to water waves

The linearized dispersion relation holds the key to finding solitary waves. For water waves, for which the dispersion relation is (1), these two cases (1) and (2) imply that pure solitary waves of elevation exist for $B = 0$, and of depression for $B > 1/3$, while generalized solitary waves arise whenever $0 < B < 1/3$. For the case of generalized solitary waves, there is always the possibility that the amplitude of the oscillations is zero, and the solution then reduces to a pure solitary wave, called an “embedded” solitary wave. There are now many examples of such embedded solitary waves arising in various physical systems, notably for internal waves. This “dynamical-systems” approach to finding solitary waves has also been applied to interfacial waves, where again the linear dispersion relation holds the key to

---

1 This may be understood by regarding $|A|$ as space, $\xi$ as time, and $|A|_\xi$ as velocity in (14). Then (14) describes a velocity field on the positive real line. In order for an orbit starting from and returning to the origin as $\xi \to \mp \infty$ to exist, the velocity field must have a stagnation point, which occurs only when $\mu_1 < 0$. 

165
determining where solitary waves can be found. However, various numerical and analytical studies suggest that embedded solitary waves do not arise in the context of water waves. Instead, case (3) implies that envelope solitary waves arise for capillary-gravity waves with $0 < B < 1/3$, where it can be shown that the coefficient $\mu_1$ in (12) is negative as required.

Finally, we remark that the method of treating solitary waves as homoclinic orbits in a spatial dynamical system has many applications (cf. [2]) in general evolution equations, most of which are not integrable. This point of view, coupled with numerical continuation techniques, turns out to be fruitful in studying spatially localized states in both one (cf. [1]) and two (cf. [7]) dimensional pattern forming systems.

**References**


