

NL002 Solitons, a brief history of

In 1834, a young Scottish engineer named John Scott Russell was conducting experiments on the Union Canal (near Edinburgh) to measure the relationship between the speed of a boat and its propelling force, with the aim of finding design parameters for conversion from horse power to steam. One August day, a rope parted in his measurement apparatus and (Russell, 1844)

the boat suddenly stopped – not so the mass of water in the channel which it had put in motion; it accumulated round the prow of the vessel in a state of violent agitation, then suddenly leaving it behind, rolled forward with great velocity, assuming the form of a large solitary elevation, a rounded, smooth and well defined heap of water, which continued its course along the channel without change of form or diminution of speed.

Russell did not ignore this unexpected phenomenon, but “followed it on horseback, and overtook it still rolling on at a rate of some eight or nine miles an hour, preserving its original figure some thirty feet long and a foot to a foot and a half in height” until the wave became lost in the windings of the channel. He continued to study the solitary wave in tanks and canals over the following decade, finding it to be an independent dynamic entity moving with constant shape and speed. Using a wave tank he demonstrated four facts (Russell, 1844):

- (i) Solitary waves have the shape $h \operatorname{sech}^2[k(x - vt)]$;
- (ii) A sufficiently large initial mass of water produces two or more independent solitary waves;
- (iii) Solitary waves cross each other “without change of any kind”;
- (iv) A wave of height h and travelling in a channel of depth d has a velocity given by the expression

$$v = \sqrt{g(d + h)}, \quad (1)$$

(where g is the acceleration of gravity) implying that a large amplitude solitary wave travels faster than one of low amplitude.

Although confirmed by observations on the Canal de Bourgogne, near Dijon, most subsequent discussions of the hydrodynamic solitary wave missed the physical significance of Russell’s observations. Evidence that to the end of his life Russell maintained a much broader and deeper appreciation of the importance of his discovery is provided by a posthumous work where – among several provocative ideas – he correctly estimated the height of the earth’s atmosphere from Equation (1) and the fact that “the sound of a cannon travels faster than the command to fire it” (Russell, 1885).

In 1895, Korteweg and de Vries published a theory of shallow water waves that reduced Russell’s problem to its essential features. One of their results was the nonlinear partial differential equation

$$\frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} + \varepsilon \frac{\partial^3 u}{\partial x^3} + \gamma u \frac{\partial u}{\partial x} = 0, \quad (2)$$

which would play a key role in soliton theory (Korteweg & de Vries, 1895). In this equation, $u(x, t)$ is the wave amplitude, $c = \sqrt{gd}$ is the speed of small amplitude waves, $\varepsilon \equiv c(d^2/6 - T/2\rho g)$ is a dispersive parameter, $\gamma \equiv 3c/2d$ is a nonlinear parameter, and T and ρ are respectively the surface tension and the density of water. The authors showed that Equation (2) has a family of exact travelling wave solutions of the form $u(x, t) = \tilde{u}(x - vt)$, where $\tilde{u}(\cdot)$ is Russell's "rounded, smooth and well defined heap" and v is the wave speed. If the dispersive term (ε) and the nonlinear term (γ) in Equation (2) are both zero, then the Korteweg-de Vries (KdV) equation becomes linear ($u_t + c u_x = 0$) with a travelling wave solution for any pulse shape at the fixed speed $v = c = \sqrt{gd}$. In general Equation (2) is nonlinear with exact travelling wave solutions

$$u(x, t) = h \operatorname{sech}^2[k(x - vt)], \quad (3)$$

where $k = \sqrt{\gamma h/12\varepsilon}$, implying that higher amplitude waves are more narrow. With this shape, the effects of dispersion balance those of nonlinearity at an adjustable value of the pulse speed. Thus the solitary wave is recognized as an independent dynamic entity, maintaining a dynamic balance between these two influences. Interestingly, solitary wave velocities are related to amplitudes by

$$v = c + \gamma h/3 = \sqrt{gd}(1 + h/2d), \quad (4)$$

in accord with Russell's empirical results, given in Equation (1), to $O(h)$. Although unrecognized at the time, such an energy conserving solitary wave is related to the existence of a Bäcklund transform (BT), which was also studied during the late 19th century (Lamb, 1976). In such a transform, a known solution generates a new solution through a single integration, after which the new solution can be used to generate another new solution, and so on. It is straightforward to find a BT for any linear partial differential equation (PDE), which introduces a new eigenfunction into the total solution with each application of the transform. Only special nonlinear PDEs are found to have BTs, but late nineteenth mathematicians knew that these include

$$\frac{\partial^2 u}{\partial \xi \partial \tau} = \sin u, \quad (5)$$

which arose in research on the geometry of curved surfaces (Steuerwald, 1936).

In 1939 Frenkel and Kontorova introduced a seemingly unrelated problem arising in solid state physics to model the relationship between dislocation dynamics and plastic deformation of a crystal (Frenkel & Kontorova, 1939). From this study, an equation describing dislocation motion is

$$\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial t^2} = \sin u, \quad (6)$$

where $u(x, t)$ is atomic displacement in the x -direction and the "sin" function represents periodicity of the crystal lattice. A travelling wave solution of Equation (6), corresponding to the propagation of a dislocation, is

$$u(x, t) = 4 \arctan \left[\exp \left(\frac{x - vt}{\sqrt{1 - v^2}} \right) \right], \quad (7)$$

with velocity v in the range $(-1, +1)$. Since Equation (6) is identical to Equation (5) after an independent variable transformation [$\xi = (x-t)/2$ and $\tau = (x+t)/2$], exact solutions of Equation (6) involving arbitrary numbers of dislocation components as in Equation (7) can be generated through a succession of Bäcklund transforms, but this was not known to Frenkel and Kontorova.

In the late 1940s, Fermi and Ulam suggested one of the first scientific problems to be assigned to the Los Alamos MANIAC computing machine: the dynamics of energy equipartition in a slightly nonlinear crystal lattice, which is related to thermal conductivity. The system they chose was a chain of 64 equal mass particles connected by slightly nonlinear springs, so from a linear perspective there were 64 normal modes of oscillation in the system. It was expected that if all the initial energy were put into a single vibrational mode, the small nonlinearity would cause a gradual progress toward equal distribution of the energy among all modes (thermalization), but the numerical results were surprising. If all the energy is originally in the mode of lowest frequency, it returns almost entirely to that mode after a period of interaction among a few other low frequency modes. In the course of several numerical refinements, no thermalization was observed (Fermi *et al.*, 1955). Pursuit of an explanation for this “FPU recurrence” led Zabusky and Kruskal to approximate the nonlinear spring-mass system by the KdV equation. In 1965 they reported numerical observations that KdV solitary waves pass through each other with no change in shape or speed, and coined the term “soliton” to suggest this property (Zabusky & Kruskal, 1965). Zabusky and Kruskal were not the first to observe nondestructive interactions of energy conserving solitary waves. Apart from Russell’s tank measurements (Russell, 1844), Perring and Skyrme, had studied solutions of Equation (6) comprising two solutions as in Equation (7) undergoing a collision. In 1962 they published numerical results showing perfect recovery of shapes and speeds after a collision and went on to discover an exact analytical description of this phenomenon (Perring & Skyrme, 1962). This result would not have surprised nineteenth century mathematicians; it is merely the second member of the hierarchy of solutions generated by a Bäcklund transform. Nor would it have been unexpected by Seeger and his colleagues, who had noted in 1953 the connections between the nineteenth century work (Steuerwald, 1936) and that of Frenkel and Kontorova (Seeger *et al.*, 1953). Since Perring and Skyrme were interested in Equation (6) as a nonlinear model for elementary particles of matter, however, the complete absence of scattering may have been disappointing. Throughout the 1960s, Equation (6) arose in a variety of problems including the propagation of ferromagnetic domain walls, self-induced transparency in nonlinear optics, and the propagation of magnetic flux quanta in long Josephson transmission lines. Eventually Equation (6) became known as the “sine-Gordon” (SG) equation (a nonlinear version of the Klein-Gordon equation: $u_{xx} - u_{tt} = u$). Perhaps the most important contribution made by Zabusky and Kruskal in their 1965 paper was to recognize the relation between nondestructive soliton collisions and the riddle of FPU recurrence. Viewing KdV solitons as independent and localized dynamic entities, they

explained the FPU observations as follows. The initial condition generates a family of solitons with different speeds, moving apart in the x - t plane. Since the system studied was of finite length with perfect reflections at both ends, the solitons could not move infinitely far apart; instead they eventually reassembled in the x - t plane, approximately recreating the initial condition after a surprisingly short “recurrence time.” By 1967, this insight had led Kruskal and his colleagues to devise a nonlinear generalization of the Fourier transform method for constructing solutions of the KdV emerging from arbitrary initial conditions (Gardner *et al.*, 1967). Called the inverse scattering (or inverse spectral) method (ISM), this approach proceeds in three steps.

- (i) The nonlinear KdV dynamics are mapped onto an associated linear problem, where each eigenvalue of the linear problem corresponds to the speed of a particular KdV soliton.
- (ii) Since the associated problem is linear, the time evolution of its solution is readily computed.
- (iii) An inverse calculation then determines the time evolved KdV dynamics from the evolved solution of the linear associated problem. Thus the solution of a nonlinear problem is found from a series of linear computations.

Another development of the 1960s was Toda’s discovery of exact two-soliton interactions on a nonlinear spring-mass system (Toda, 1967). As in the FPU system, equal masses were assumed to be interconnected with nonlinear springs, but Toda chose the potential

$$\left(\frac{a}{b}\right) [e^{-bu_j} - 1] + au_j, \quad (8)$$

where $u_j(t)$ is the longitudinal extension of the j -th spring from its equilibrium value and both a and b are adjustable parameters. (In the limit $a \rightarrow \infty$ and $b \rightarrow 0$ with ab finite, this reduces to the quadratic potential of a linear spring. In the limit $a \rightarrow 0$ and $b \rightarrow \infty$ with ab finite, it describes the interaction between hard spheres.) Thus by the late 1960s it was established – although not widely known – that solitons were not limited to PDEs (KdV and SG). Local solutions of difference differential equations could also exhibit the unexpected properties of unchanging shapes and speeds after collisions.

These events are only the salient features of a growing panorama of nonlinear wave activities that became gradually less parochial during the 1960s. Solid state physicists began to see relationships between their solitary waves (magnetic domain walls, self-shaping pulses of light, quanta of magnetic flux, polarons, etc.), and those from classical hydrodynamics and oceanography, while applied mathematicians began to suspect that the ISM might be used for a broader class of nonlinear wave equations. It was amid this intellectual ferment that Newell and his colleagues organized the first soliton research workshop during the summer of 1972 (Newell, 1974). Interestingly, one of the most significant contributions to this conference came by post. From the Soviet Union arrived

a paper by Zakharov and Shabat formulating Kruskal's ISM for the nonlinear PDE (Zakharov & Shabat, 1972)

$$i\frac{\partial u}{\partial t} + \frac{\partial^2 u}{\partial x^2} + 2|u|^2u = 0. \quad (9)$$

In contrast to KdV, SG and the Toda Lattice, the dependent variable in this equation is complex rather than real, so the evolutions of two quantities (magnitude and phase of u) are governed by the equation. This reflects the fact that Equation (9) is a nonlinear generalization of a linear equation $iu_t + u_{xx} + u = 0$, solutions of which comprise both an envelope and a carrier wave. Since this linear equation is a Schrödinger equation for the quantum mechanical probability amplitude of a particle (like an electron) moving through a region of uniform potential, it is natural to call Equation (9) the nonlinear Schrödinger (NLS) equation. When the NLS equation is used to model wave packets in such fields as hydrodynamics, nonlinear optics, nonlinear acoustics, plasma waves and biomolecular dynamics, however, its solutions are devoid of quantum character. Upon appreciating the Zakharov and Shabat paper, many left the 1972 conference convinced that four nonlinear equations (KdV, SG, NLS, and the Toda lattice) display solitary wave behavior with the special properties that led Zabusky and Kruskal to coin the term soliton (Newell, 1974). Within two years, ISM formulations had been constructed for the SG equation and also for the Toda lattice.

Since the mid-1970s, the soliton concept has become established in several areas of applied science, and dozens of dynamic systems are now known to be integrable through the ISM. Even if a system is not exactly integrable, additionally, it may be close to an integrable system, allowing analytic insight to be gleaned from perturbation theory. Thus one is no longer surprised to find stable spatially localized regions of energy, balancing the opposing effects of nonlinearity and dispersion and displaying the essential properties of objects. Current soliton studies are focused on working out the details of such object-like behavior in a wide range of research areas (Scott, 1999).

ALWYN SCOTT

Further Reading

- Fermi, E., Pasta, J.R. & Ulam, S.M. 1955. Studies of nonlinear problems, Los Alamos Scientific Laboratory Report No. LA-1940
- Frenkel, J. & Kontorova, T. 1939. On the theory of plastic deformation and twinning. *Phys. Z. Sowjetunion*, 1: 137–49
- Gardner, C.S., Greene, J.M., Kruskal, M.D. & Miura, R.M. 1967. Method for solving the Korteweg-de Vries equation. *Physical Review Letters*, 19: 1095–97

- Korteweg, D.J. & de Vries, H. 1895. On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves. *Philosophical Magazine*, 39: 422–43
- Lamb, G.L. Jr. 1976. Bäcklund transforms at the turn of the century. In *Bäcklund Transforms*, edited by R.M. Miura, Berlin and New York: Springer
- Newell, A.C. (editor) 1974. *Nonlinear Wave Motion*, Providence, R.I.: American Mathematical Society
- Perring, J.K. & Skyrme, T.R.H. 1962. A model unified field equation. *Nuclear Physics*, 31: 550–55
- Russell, J.S. 1844. *Report on Waves*, 14th meeting of the British Association for the Advancement of Science, London: BAAS
- Russell, J.S. 1885. *The Wave of Translation in the Oceans of Water, Air and Ether*, London: Trübner
- Scott, A.C. 1999. *Nonlinear Science: Emergence and Dynamics of Coherent Structures*, Oxford and New York: Oxford University Press
- Seeger, A., Donth, H. & Kochendörfer, A. 1953. Theorie der Versetzungen in eindimensionalen Atomreihen. *Zeitschrift für Physik*, 134: 173–93
- Steuerwald, R. 1936. Über Enneper'sche Flächen und Bäcklund'sche Transformation. *Abhandlungen der Bayerischen Akademie der Wissenschaften München* 1–105
- Toda, M. 1967. Vibration of a chain with nonlinear interactions. *Journal of the Physical Society of Japan*, 22: 431–36; Wave propagation in anharmonic lattices. *Journal of the Physical Society of Japan*, 23: 501–06
- Zabusky, N.J. & Kruskal, M.D. 1965. Interactions of solitons in a collisionless plasma and the recurrence of initial states. *Physical Review Letters*, 15: 240–43
- Zakharov, V.E. & Shabat, A.B. 1972. Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media. *Soviet Physics, JETP*, 34: 62–69