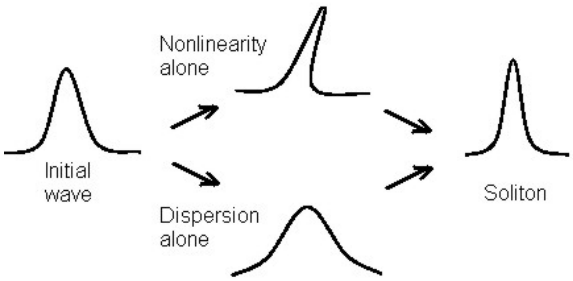
Solitons

Gabriele Filosofi

June 2005

Abstract

With the discovery of solitons (and chaos) over thirty years ago, our perception of how nonlinear systems behave has changed in a dramatic way. The amount of information on nonlinear wave phenomena obtained through the fruitful collaboration of mathematicians and physicists using this description makes the soliton concept one of the most significant developments in modern mathematical physics. There are now many text books, monographs and conference proceedings on soliton theory. The UW library catalog lists more than 70 books on solitons; in fact, we now have quite a deep mathematical understanding of soliton equations and there are many things concerning solitons which are well worth learning! I'll try to sketch a few of them in this paper.



Contents

1 Introduction

We will take a **solitary wave** to be a solution to a nonlinear differential equation that is non-constant and finite for some localized region and dies exponentially to a constant for large distances away from this region. A **soliton** can be viewed as a solitary wave displaying many properties of a particle. Nature provides us with many examples of solitary waves. They have been observed or assumed in a variety of phenomena, ranging from fluids hydrodynamics to plasma physics, from optics to protein models, from atmospheric events to magma flow, from high energy to solid state physics, from tornados to the Great Red Spot of Jupiter.

Solitons have become important items of research in diverse fields of physics and engineering. In recent years much effort has been expended on analyzing the properties of solitons for purposes such as high-speed communications and optical computing gates. Today, one of the most important application is the data transmission. The performance of a fibre strand is usually measured by two criteria: the transmission capacity and how far the light signal can travel from its source without having regenerated through an optical-electrical-optical conversion. On both counts fibre performance has been improved in recent years, thanks to the ability of solitons to propagate without dispersing over long distances with minimal losses (Marconi's Solstis project). Now it is possible to transmit 16 wavelengths at $10Gbps$ each over $3000km$ without regeneration, due to the utilization of dispersion-managed solitons.

Another important field of application is represented by nonlinear transmission lines (NLTL), a technology that has permitted the fabrication of broadband high efficiency frequency multipliers operating above $100GHz$.

Scientists are also trying to answer the question of whether effective computation can be performed by the interaction of solitons in a bulk medium, linear, planar, or three-dimensional, therefore potentially providing ultra-scale parallel processing. Various media are possible, including optical fiber, Josephson junctions and electrical transmission lines. An introduction to this topics can be found in this paper.

It has been suggested by several authors that solitary waves can have a general relevance for the study of energy transport in living systems. The most widely known theory in this context is probably that of *Davydov's soliton*, which is of essentially quantum nature. Davidov soliton describes the transmission of energy along protein chains. An entire paragraph is devoted to the hypothesis that solitons could play a fundamental functional role in the process of DNA transcription, effecting the opening of the double chain needed for RNA Polymerase to be able to copy the genetic code.

A close relationship between soliton equations and numerical algorithms has been pointed out. Among those numerical algorithms related to integrable systems: Matrix eigenvalue algorithms, Convergence acceleration algorithms, Continued fraction algorithms, Decoding algorithms. From these results, one may conjecture that a good

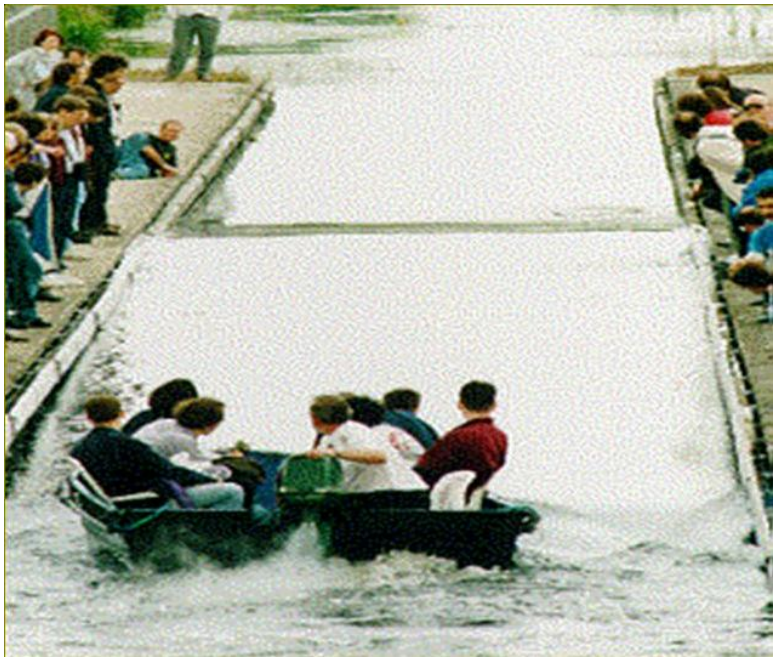
numerical algorithm is regarded as an integrable dynamical system.

Besides, solitons and nonlinear PDEs represent an exciting branch of mathematics. Not all nonlinear PDEs have soliton solutions. Those that do are generic and belong to a class for which the general initial-value problem can be solved by a technique called the inverse scattering transform (IST), a brilliant scheme developed by Kruskal and his coworkers in 1967. With this method, which can be viewed as a generalization of the Fourier transform to nonlinear equations, general solutions can be produced through a series of linear calculations. During the solution process it is possible to identify new nonlinear modes—generalized Fourier modes—that are the soliton components of the solution and, in addition, modes that are purely dispersive and therefore often called **radiation**.

The term *integrable*, referring to PDEs, is not used with perfect consistency throughout the literature. Here we use integrable to mean *solvable by the IST*. Nonintegrable equations, and integrable equations with arbitrary initial conditions, must in general be solved numerically.

2 Historical Background

In 1834, a young Scottish engineer named John Scott Russell (1808-1882) 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, first observed a wave formed when a boat which was rapidly drawn along a narrow channel by a pair of horses suddenly stopped. He wrote



"*I belive I shall best introduce this phaenomenon by descibing the circumstances of my first acquintance with it. I was observing the motion of a boat which was rapidly drawn along a narrow channel by a pair of horses, when 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 apparently without change of form or diminution of speed. I 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. Its height gradually diminished, and*

after a chase of one or two miles I lost it in the windings of the channel. Such, in the month of August 1834, was my first chance interview with that singular and beautiful phenomenon which I have called the Wave of Translation, a name which it now very generally bears." [J.Scott Russell, *Report on waves.* Fourteenth meeting of the British Association for the Advancement of Science, 1844].

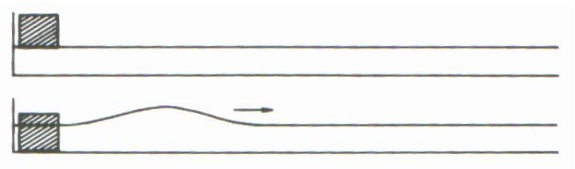


Figure 1: The method used by Russell to generate solitary waves.

He continued to study the solitary wave making extensive experiments in a laboratory scale wave tank over the following decade, in order to study this phenomenon more carefully. Russell observed a solitary wave or in Russell's terminology, the "*great primary wave of translation*", which is a long, shallow, water wave moving with constant shape and speed as an independent dynamic entity, and hence he deduced that it exist; this is his most significant result.

Figure 2: J.S.Russell

Russell produced solitary water waves letting a weight to fall in a little water channel. *A sufficiently large initial mass of water produced two or more independent solitary waves.*

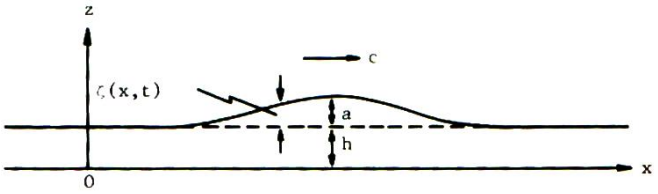


Figure 3: Configuration and parameters of a solitary wave

He empirically deduced that *the stable wave had a velocity c given by the expression*

$$c = \sqrt{g(h+a)} \tag{2.0.1}$$

where g is the gravity acceleration, a is the wave's height and h the depth of the channel. 2.0.1 implies that a large amplitude solitary wave travels faster than one of low amplitude. Russell found that *two solitary waves of different speed cross each other "without change of any kind"*.

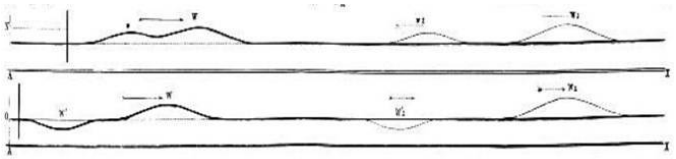


Figure 4:

Russell attempted also to produce depression waves by rising up a plunged body, but a dispersing wave train resulted.

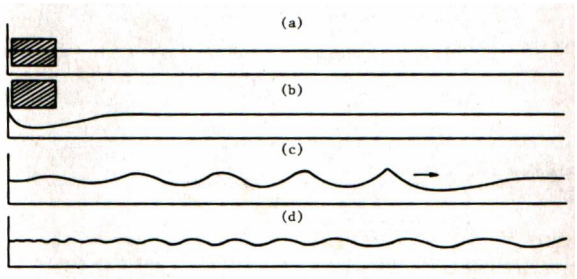


Figure 5: Failing method used by Russell to generate solitary waves.

Russell also determined *the shape of a solitary wave to be that of a sech()^2 function.* At the time there was no equation describing such water waves and having such a

solution. Thus he discovered the solution to an as yet unknown equation! The equation describing the unidirectional propagation of waves on the surface of a shallow channel was derived by Korteweg and de Vries in 1895.

3 Korteweg - de Vries Equation (KdV)

Despite Russell's detailed observations, it was many years before mathematicians formulated the relevant equation. This solitary wave describe by Russell, some thirty feet long and a foot high, moved along the channel at about eight miles per hour, maintaining its shape and speed for over a mile. The mathematicians Airy and Stokes made calculations which appeared to show that any such wave would be unstable and not persist for as long as Russell claimed.

Starting from the work of Boussinesq (1872) and Rayleigh (1876), in the 1895 Diederik Johannes Korteweg and Gustav de Vries, based on the thesis work of de Vries, developed a mathematical model for the motion of a long weakly nonlinear wave propagating along an incompressible fluid, i.e. the phenomenon described by Russell [D.J.Korteweg and G.de Vries, *On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves.* Philos. Mag. Ser. 5, **39**, (1895). 422-443]. So they pointed out errors in the analysis of Airy and Stokes and vindicated Russell's conclusions.

Figure 6: D.J.Korteweg.

Korteweg was a well-known Dutch mathematician. He was born in 1848, received his doctorate in 1878 from Uni- versity of Amsterdam, became a professor there in 1881, retired in 1918, and died in 1941. De Vries wrote his doctoral thesis under Korteweg, defended it in 1894, served as a high school teacher until 1931, and died in 1934. An extensive obituary of Korteweg was published in the 1945/1946 Annals of the Royal Dutch Academy of Arts and Science, but there was no mention of the KdV there.

It is a non-linear Partial Differential Equation (PDE) of the 3^{rd} order in two dimensions (i.e. one space and time dimension)

$$\eta_t = \frac{3}{2}\sqrt{\frac{g}{h}}\left(\eta\eta_x + \frac{2h}{3}\eta_x + \frac{1}{3}\epsilon\eta_{xxx}\right) \tag{3.0.2}$$

where $\eta(x,t)$ is the wave amplitude (i.e. the fluid surface elevation relative to the undisturbed depth), h is the channel height, g is the gravitational acceleration, $\epsilon = \frac{h^3}{3} - \frac{Th}{\rho g}$ is a dispersive parameter, with $T = \rho\sigma$ the surface tension of the fluid with density ρ. The subscripts denote partial differentiation. The wave propagation is directed as x, but a change in the sign of the η_t term reverses the direction, because it is equivalent to a sign inversion for the space variable x. The KdV is nonlinear because of the product $\eta\eta_x$. This nonlinearity is quadratic ($\eta^2\eta_x$ would be a cubic nonlinearity). The coefficient of the term η_x is $c = \sqrt{gh}$, the linear long wave phase speed. It should be remarked that water waves (and similarly for other applications) are only approximately described by the KdV equation, in the limit of small amplitude and long waves (λ large enough).

3.1 Cnoidal waves

3.0.2 is the original form of the **Korteweg-de Vries Equation** (KdV). By scaling η, t and x, i.e. by multiplying and summing them with some constants, it is possible to change the coefficient in front of each of the terms of 3.0.2 at will. This is often done in the literature. For instance, with

$$u = \frac{1}{2\epsilon^{1/3}}(\eta + \frac{2h}{3}); \quad \xi = -\frac{x}{\epsilon^{1/3}}; \quad \tau = \frac{t}{2}\sqrt{\frac{g}{h}}$$

we obtain a nondimensionalized version of the 3.0.2, the so-called "standard form", expressed by

$$u_\tau + 6uu_\xi + u_{\xi\xi\xi} = 0 \quad \xi \in \mathbf{R}, \quad \tau > 0 \tag{3.1.1}$$

or, by renaming variables,

$$u_t + 6uu_x + u_{xxx} = 0 \quad x \in \mathbf{R}, \quad t > 0 \tag{3.1.2}$$

Note that now the variable u it is not a real displacement, indeed it is adimensional. Replacing coefficient 6 with -6 amounts to replacing u by $-u$. Replacing coefficient of u_t with -1 amounts to replacing x by $-x$, so that reversing the direction of propagation of the solutions. Sometimes KdV appears in these forms.

Although equation 3.1.2 now bears the name KdV, it was apparently first obtained by Boussinesq (1877). Korteweg and de Vries found a two-parameter family of periodic travelling wave solutions, described by elliptic functions and commonly called **cnoidal waves**,

$$u(x,t) = b + 2m\gamma^2 \; cn^2(\gamma(x - vt), m), \qquad v = 6b + 4(2m-1)\gamma^2 \tag{3.1.3}$$

where $cn(x,m)$ is the Jacobian elliptic function of modulus m, defined as

$$\begin{cases} x = \int_0^\phi \frac{d\vartheta}{\sqrt{1 - m\sin(\vartheta)^2}}, & 0 \le m \le 1 \\ cn(x,m) = \cos(\phi) \end{cases}$$

For $0 < m < 1$, 3.1.3 are periodic solutions. As Korteweg and de Vries noted, the period may be expressed as an elliptic integral. Graphically, cnoidal waves look much like cosine waves, but close examination shows they have sharper crests and flatter troughs.

As $m \to 1$ the period goes to infinity, $cn(x,m) \to sech(x)$, and then the cnoidal wave 3.1.3 becomes the famous "one crested soliton" or "1-soliton" solution 3.6.11 (derived below), riding on a background level b. On the other hand, as $m \to 0$, $cn(x,m) \to \cos(2x)$, and the cnoidal wave collapses to a linear sinusoidal wave.

This solitary wave solution found by Korteweg and de Vries had earlier been obtained directly from the governing equations (in the absence of surface tension T) independently by Boussinesq (1871, 1877) and Rayleigh (1876) who were motivated to explain the experiments of Russell. Curiously, it was not until quite recently that it was recognized that the KdV equation is not strictly valid if surface tension is taken into account and $\epsilon < 0$, as then there is a resonance between the solitary wave and very short capillary waves. This condition resembles the *Lighthill criterion* valid for NSE (see below). After this ground-breaking work of Korteweg and de Vries, interest in solitary water waves and the KdV equation declined until the dramatic discovery of the soliton by Zabusky and Kruskal in 1965.

3.2 The FPU Problem

No one could have predicted that the KdV would be a famous equation some 70 years later.

The MANIAC computer was designed to carry out some computations needed at the Los Alamos laboratories for the design of the first hydrogen bombs, and of course it was a marvel for its day. But it is worth noting that it was very weak by today's standards even when compared with modest desktop machines.

In the late 1940s, E.Fermi J.Pasta and S.Ulam suggested one of the first scientific problems to be assigned to MANIAC: the dynamics of energy equipartition in a slightly nonlinear crystal lattice, now called the **FPU problem**.

Fermi, Pasta and Ulam make it clear that the problem that they want to simulate is the vibrations of a "one-dimensional continuum" or "string" with fixed end-points and nonlinear elastic restoring forces, but that "*for the purposes of numerical work this continuum is replaced by a finite number of points ... so that the PDE describing the motion of the string is replaced by a finite number of ODE.*" To rephrase this in the current jargon, FPU study a one-dimensional lattice of N oscillators with nearest neighbor interactions and zero boundary conditions. The system they choose was a chain of $N = 64$ equal mass particles connected by slightly nonlinear springs, so from a linear perspective there were 64 normal modes of oscillation in the system.

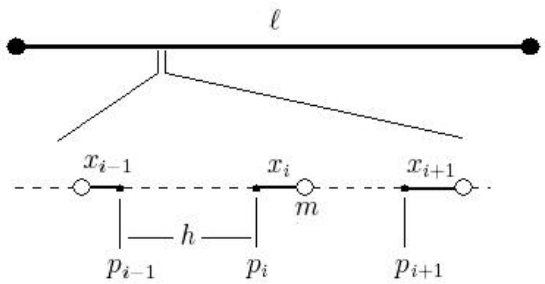


Figure 7: The discrete model for a continous elastic wire

We imagine the original string to be stretched along the x-axis from 0 to its length ℓ. The N oscillators have equilibrium positions $p_i = ih, i = 0, \ldots, N-1$, where $h = \frac{\ell}{N-1}$ is the lattice spacing, so their positions at time t are $X_i(t) = p_i + x_i(t)$, (where the x_i represent the displacements of the oscillators from equilibrium). The force attracting any oscillator to one of its neighbors is taken as $k(\delta + \alpha\delta^2)$, δ denoting the "strain", i.e., the deviation of the distance separating these two oscillators from their equilibrium separation h. Note that when $\alpha = 0$ this is just a linear Hooke's law force with spring constant k. The force acting on the i-th oscillator due to its right neighbor is $F(x)_i^+ = k[(x_{i+1} - x_i) + \alpha(x_{i+1} - x_i)^2]$, while the force acting on it due to its left neighbor is $F(x)_i^- = -k[(x_i - x_{i-1}) + \alpha(x_i - x_{i-1})^2]$. Thus the total force acting on the i-th oscillator will be the sum of these two forces, and assuming that all of the oscillators have the same mass, m, Newton's equations of motion read:

$$m\ddot{x}_i = F_i^+ + F_i^- = k(x_{i+1} - 2x_i + x_{i-1}) + k\alpha[(x_{i+1} - x_i)^2 - (x_i - x_{i-1})^2], \quad i = 0, 1 \ldots N-1 \tag{3.2.1}$$

or

$$m\ddot{x}_i = k(x_{i+1} - 2x_i + x_{i-1})[1 + \alpha(x_{i+1} - x_{i-1})], \qquad i = 0, 1, \ldots N-1 \tag{3.2.2}$$

with the boundary conditions $x_0(t) = x_{N-1}(t) = 0$. In addition, Fermi, Pasta and Ulam looked at motions of the lattice that start from rest, i.e., they assumed that $\dot{x}_i(0) = 0$, so the motion of the lattice is completely specified by giving the $N-2$ initial displacements $x_i(0), i = 1, \ldots, N-2$. We shall call this the *FPU initial value problem* (with initial condition $x_i(0)$).

It will be convenient to rewrite Newton's equations in terms of parameters that refer more directly to the original string that we are trying to model. Namely, if ρ denotes the density of the string, then $m = \rho h$, while if κ denotes the Young's modulus for the string (i.e., the spring constant for a piece of unit length), then $k = \frac{\kappa}{h}$ will be the spring constant for a piece of length h. Defining $c = \sqrt{\frac{\kappa}{\rho}}$ we can now rewrite Newton's equations as

$$m\ddot{x}_i = c^2\big(\frac{x_{i+1} - 2x_i + x_{i-1}}{h^2}\big)[1 + \alpha(x_{i+1} - x_{i-1})] \tag{3.2.3}$$

We can now pass to the continuum limit, i.e., by letting $N \to \infty$ (so $h \to 0$) we can attempt to derive a PDE for the function $u(x,t)$ that measures the displacement at time t of the particle of string with equilibrium position x. We shall leave the nonlinear case for later, and here restrict our attention to the linear case, $\alpha = 0$.

$$u_{tt}(x,t) = c^2\big(\frac{u(x+h,t) - 2u(x,t) + u(x-h,t)}{h^2}\big) \tag{3.2.4}$$

Applying Taylor's formula

$$f(x \pm h) = f(x) \pm hf'(x) + \frac{h^2}{2!}f''(x) \pm \frac{h^3}{3!}f'''(x) + \frac{h^4}{4!}f''''(x) + O(h^5) \tag{3.2.5}$$

to $u(x,t)$, we get

$$\frac{u(x+h,t) - 2u(x,t) + u(x-h,t)}{h^2} = u_{xx}(x,t) + \frac{h^2}{12}u_{xxxx}(x,t) + O(h^3) \tag{3.2.6}$$

so letting $h \to 0$, we find $u_{tt} = c^2 u_{xx}$, i.e., u satisfies the linear wave equation, with propagation speed c (and of course the boundary conditions $u(0,t) = u(\ell,t) = 0$, and initial conditions $u_t(x,0) = 0$, $u(x,0) = u_0(x)$. This is surely one of the most famous initial value problems of mathematical physics, and nearly every mathematician sees a derivation of both the d'Alembert and Fourier version of its solution early in their careers. For each positive integer k there is a normal mode or "standing wave" solution

$$u^{(k)}(x,t) = \cos\big(\frac{k\pi ct}{\ell}\big)\ \sin\big(\frac{k\pi x}{\ell}\big) \tag{3.2.7}$$

and the solution to the initial value problem is $u(x,t) = \sum_{k=1}^{\infty} a_k u^{(k)}(x,t)$ where the a_k are the Fourier coefficients of u_0

$$a_k = \frac{2}{\ell}\int_0^{\ell} u_0(x)\sin\big(\frac{k\pi x}{\ell}\big)dx \tag{3.2.8}$$

In order to find the solution to the FPU initial value problem we have to replace x by ih in $u^{(k)}(x,t)$ (and using $\ell = (N-1)h$). So we get the normal modes for the FPU initial value problem (linear case of course)

$$x_i^{(k)}(t) = \cos\frac{k\pi ct}{(N-1)h}\ \sin\big(\frac{k\pi i}{N-1}\big), \qquad i = 0, 1, \ldots N-1 \tag{3.2.9}$$

It follows that, in the linearized case, the solution to the FPU initial value problem with initial conditions $x_i(0)$ is given explicitly by $x_i(t) = \sum_{k=1}^{N-2} a_k x_i^{(k)}(t)$, where the Fourier coefficients a_k are determined from the formula

$$a_k = \sum_{i=1}^{N-2} x_i(0)\sin\big(\frac{k\pi i}{N-1}\big) \tag{3.2.10}$$

Of course, when $\alpha = 0$ and the interactions are linear, we are in effect dealing with $N-2$ uncoupled harmonic oscillators (the above normal modes) and there is no thermalization.

On the contrary, the sum of the kinetic and potential energy of each of the normal modes is a constant of the motion!

But if α is small but non-zero, Fermi, Pasta and Ulam expected (on the basis of then generally accepted statistical mechanics arguments) that the energy would gradually shift between modes so as to eventually roughly equalize the total of potential and kinetic energy in each of the $N-2$ normal modes $x^{(k)}$. To test this they started the lattice in the fundamental mode $x^{(1)}$, with various values of α, and integrated Newton's equations numerically for a long time interval, interrupting the evolution from time to time to compute the total of kinetic plus potential energy in each mode.

What did they find? Here is a quotation from their report

> "*Let us say here that the results of our computations show features which were, from the beginning, surprising to us. Instead of a gradual, continuous flow of energy from the first mode to the higher modes, all of the problems showed an entirely different behavior. Starting in one problem with a quadratic force and a pure sine wave as the initial position of the string, we did indeed observe initially a gradual increase of energy in the*

higher modes as predicted (e.g., by Rayleigh in an infinitesimal analysis). Mode 2 starts increasing first, followed by mode 3, and so on. Later on, however, this gradual sharing of energy among the successive modes ceases. Instead, it is one or the other mode that predominates. For example, mode 2 decides, as it were, to increase rather rapidly at the cost of the others. At one time it has more energy than all the others put together! Then mode 3 undertakes this role. It is only the first few modes which exchange energy among themselves, and they do this in a rather regular fashion. Finally, at a later time, mode 1 comes back to within one percent of its initial value, so that the system seems to be almost periodic." [E.Fermi, J.Pasta, S.Ulam, *Studies of non linear problems*, Report No LA-1940, Los Alamos National Laboratory, May 1955].

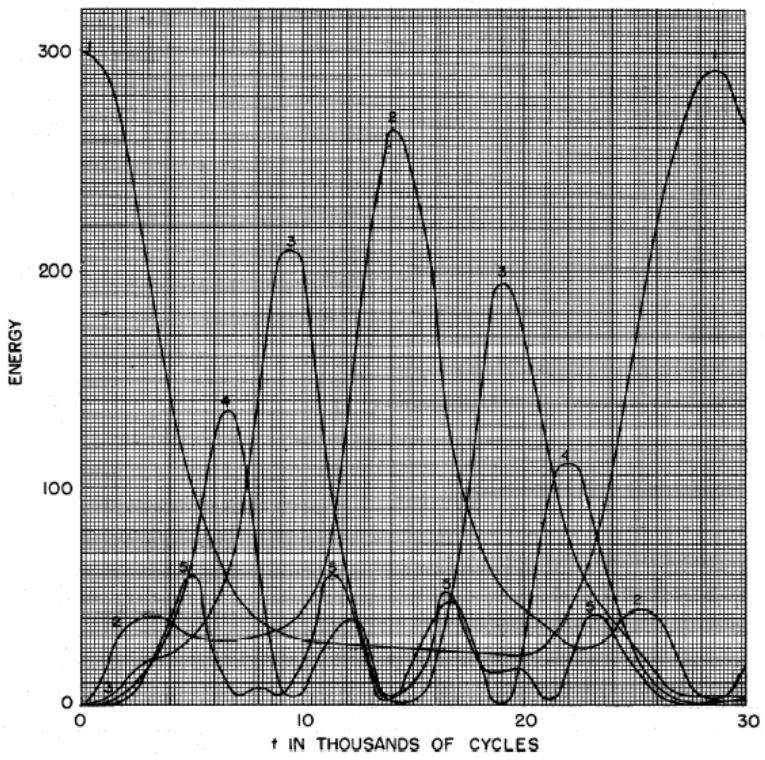


Figure 8: Energy (kinetic plus potential) in each of the first five modes, in arbitrary units. $N = 64$, $\alpha = 0.25$, $\delta t^2 = 0.125$. The initial form of the string was a sine wave. Note the recurrence of the original wave after about 29000 cycles. The little loss was certainly due to numerical approximations.

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 (a process known as *thermalization*), but the numerical results were surprising. *If all the energy is originally in the mode of lowest frequency, after flowing back and forth amongst the low order modes, it returns almost entirely to that mode to within an accuracy of one or two percent, and from there on the process approximately repeated.* This phenomenon was called **recurrence**. Unfortunately, Fermi died in November 1954, even before the paper cited above was published. Pasta and Ulam completed their last few computational examples.

3.3 The work of Zabusky and Kruskal

The FPU puzzle was solved in 1965 by Zabusky and Kruskal, who explained it in terms of solitary-wave solutions to the KdV. In 1965 Norman Zabusky and Martin Kruskal resumed the FPU system in studying the heat conductivity of solids. The fact that only the lowest order (long-wavelength) modes of the discrete FPU lattice were "active" led them to find a connection with the KdV equation in a continuum approximation.

Finding a good continuum limit for the nonlinear FPU lattice is a lot more sophisticated than one might at first expect after the easy time we had with the linear case. Following a similar computation as previously done for 3.2.6 we obtain

$$\alpha(x_{i+1} - x_{i-1}) = 2\alpha h u_x + \frac{\alpha h^3}{3} u_{xxx} + O(h^5) \tag{3.3.1}$$

so substituting in 3.2.3 and neglecting terms of order greater than $O(h^2)$ we obtain

$$\frac{1}{c^2} u_{tt} = u_{xx} + 2\alpha h u_x u_{xx} + \frac{h^2}{12} u_{xxxx} \tag{3.3.2}$$

Remark 1. If we differentiate this equation with respect to x and make the substitution $v = u_x$, we see that it reduces to the more familiar Boussinesq equation

$$\frac{1}{c^2} v_{tt} = v_{xx} + \alpha \frac{\partial^2 (v^2)}{\partial x^2} + \frac{h^2}{12} v_{xxxx} \tag{3.3.3}$$

The general solution of the linear wave equation $u_{tt} - c^2 u_{xx} = 0$ is of course $u(x,t) = w(x+ct) + z(x-ct)$, i.e., the sum of an arbitrary left moving traveling wave and an arbitrary right moving traveling wave, both moving with speed c. It is customary to simplify the analysis in the linear case by treating each kind of wave separately. Then, we would like to look for solutions $u(x,t)$ that behave more and more like, say, right

moving traveling waves of velocity c and for longer and longer periods of time as α and h tend to zero. With this aim in mind we define new variables $\xi = x - ct$, $\tau = \alpha hct$. Let $u(x,t) = z(\xi,\tau)$ be such a solution, the following relations hold

$$\frac{\partial^k}{\partial \xi^k} = \frac{\partial^k}{\partial x^k}$$
$$\frac{\partial}{\partial t} = \frac{\partial}{\partial \xi}\frac{\partial \xi}{\partial t} + \frac{\partial}{\partial \tau}\frac{\partial \tau}{\partial t} = -c\frac{\partial}{\partial \xi} + \alpha hc\frac{\partial}{\partial \tau}$$
$$\frac{\partial^2}{\partial t^2} = c^2\frac{\partial^2}{\partial \xi^2} - 2\alpha hc^2\frac{\partial^2}{\partial \xi \partial \tau} + c^2(\alpha h)^2\frac{\partial^2}{\partial \tau^2}$$

Thus in these new coordinates the 3.3.2 transforms to

$$z_{\xi\tau} - \frac{\alpha h}{2} z_{\tau\tau} = -z_\xi z_{\xi\xi} - \frac{h}{24\alpha} z_{\xi\xi\xi\xi} \tag{3.3.4}$$

and, at last, we are prepared to pass to the continuum limit. We assume that α and h tend to zero at the same rate, i.e., that as h tends to zero, the quotient h/α tends to a positive limit, and we define $\delta = \lim_{h\to 0}\sqrt{\frac{h}{24\alpha}}$. Then $\alpha h = O(h^2)$, so letting h approach zero gives $z_{\xi\tau} + z_\xi z_{\xi\xi} + \delta^2 z_{\xi\xi\xi\xi} = 0$. Finally, making the substitution $v = z_x$ we arrive at the KdV equation

$$v_\tau + vv_\xi + \delta^2 v_{\xi\xi\xi} = 0 \tag{3.3.5}$$

Let us recapitulate the relationship between the FPU Lattice and the KdV equation. Given a solution $x_i(t)$ of the FPU Lattice, we get a function $u(x,t)$ by interpolation, i.e., $u(ih,t) = x_i(t)$, $i = 0,\ldots,N-1$. For small lattice spacing h and nonlinearity parameter α there will be solutions $x_i(t)$ so that the corresponding $u(x,t)$ will be an approximate right moving traveling wave with slowly varying shape, i.e., it will be of the form $u(x,t) = z(x - ct, \alpha hct) \equiv z(\xi,\tau)$ for some smooth function z, whose derivative $v = z_\xi$ will satisfy the KdV equation 3.3.5, where $\delta^2 = \frac{h}{24\alpha}$. Having found this relationship between the FPU Lattice and the KdV equation, Kruskal and Zabusky made some numerical experiments, solving the KdV initial value problem for various initial data.

The KdV model considered by Zabusky and Kruskal was (let's rename symbols for convenience)

$$u_t + uu_x + \delta^2 u_{xxx} = 0 \tag{3.3.6}$$

For numerical reasons, they choose to deal with the case of periodic boundary conditions, studying the equation on the circle instead of on the line. For their published report, they choose

$$
\begin{aligned}
\delta &= 0.022 \\
u(2,t) &= u(0,t),\ u_x(2,t) = u_x(0,t),\ u_{xx}(2,t) = u_{xx}(0,t), \quad \forall\ t > 0 \\
u(x,0) &= \cos(\pi x) \quad for \quad 0 \le x \le 2
\end{aligned}
$$

The boundary conditions above are periodic (representing essentially a ring of coupled nonlinear springs). Analyzing their KdV model numerically, Zabusky and Kruskal observed that a simple initial profile $u(x,0)$ broke up into a train of solitary waves with the tallest ones moving the fastest. They found that when two of these pulses collide, the pulses passed through each other yet preserving their size, shapes and speeds after the collision. *The only effect produced in wave interaction was a phase change* (representing a displacement ahead of their centres). The evidence of the nonlinear character of the interaction was that during the collision *the resulting amplitude was less than the sum of the two separated waves.*

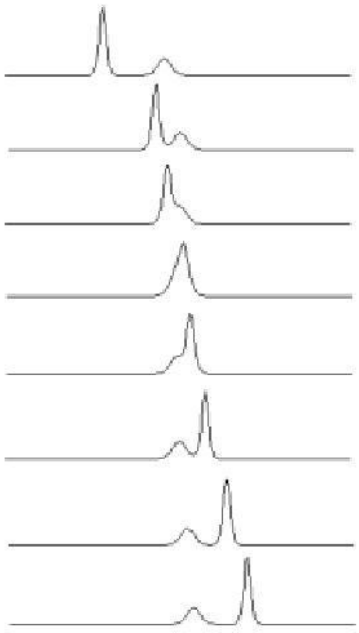


Figure 9: KdV 2-soliton. The larger, faster soliton will start to overtake the smaller. For most dispersive evolution equations these solitary waves would scatter inelastically and lose energy to radiation. Not so for the KdV equation. these solitary waves interact nonlinearly when they collide but soon reappear unchanged. Except, that is, for one very curious fact: the fast wave which has overtaken the slower one will have jumped ahead!

Since these pulses behaved more like particles than waves, they were given the name **solitons** (Zabusky & Kruskal, 1965). At that time Zabusky and Kruskal were unaware that solitons were discovered in 1834 by John Scott-Russell. The first success of the soliton concept was explaining the recurrence in the Fermi-Pasta-Ulam system.

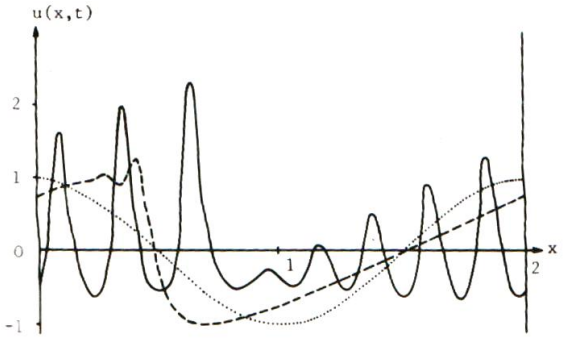


Figure 10: Solution to the KdV problem posed by Zabusky e Kruskal in 1965. The three wave profiles refer to three different time instants.

The figure above shows the solution for different t. Initially the first two terms in the equation dominate the evolution with a progressive arising of a discontinuity. After that dispersive term u_{xxx} starts to stabilizing the wave, preventing the emergence of the singularity. In the same time eight $sech(x)^2$-like pulses came into existence and move along x interacting one each other, but conserving their individuality.

If you play around with different profiles $u(x, 0)$, you'll find that the number of solitons and their spacing and heights depends on the starting profile in some nonobvious ways, but there should generally appear some finite number of solitons. Then each one collides with the next peak at left, etc. But, they don't split up into zillions of different solitons: *the number of solitons is conserved.*

Next: *recurrence. After a not very long time, something similar to the initial profile results.* This is absolutely amazing! Roughly speaking, we have zillions of oscillators, so from general principles from ergodic theory we'd expect a very long recurrence time, i.e. we'd never expect to see the initial sine profile reappear in our lifetime. But it does! This was one of the great surprises of FPU. It turns out that this amazing recurrence happens because the original oscillators aren't important, only the number of solitons (eight) is important. And eight phases lining up the same way again after a not very long period of time is not as surprising as 100 zillions phases suddenly lining up the same way again.

Let us explain the reason why the recurrence in the FPU Lattice is so surprising. The lattice is made up of a great many identical oscillators. Initially the relative phases of these oscillators are highly correlated by the imposed cosine initial condition. If the

interactions are linear ($\alpha = 0$), then the oscillators are harmonic and their relative phases remain constant. But, when α is positive, the anharmonic forces between the oscillators cause their phases to start drifting relative to each other in an apparently uncorrelated manner. The expected time before the phases of all of the oscillators will be simultaneously close to their initial phases is enormous, and increases rapidly with the total number N. But, from the point of view of the KdV solitons, an entirely different picture appears. As mentioned in the above paragraph, if $\delta = 0$ in the KdV equation, it reduces to the so-called *non-viscous Burgers Equation*, which exhibits steepening and breaking of a negatively sloped wave front in a finite time T_B. (For the above initial conditions $T_B = 1/\pi$). However, when $\delta > 0$, just before breaking would occur, a small number of solitons emerge (eight in the case of the above initial wave shape, $\cos(\pi x)$, *and this number depends only on the initial wave shape, not on the number of oscillators*. The expected time for their respective centers of gravity to all eventually "focus" at approximately the same point of the circle is of course much smaller than the expected time for the much larger number of oscillators to all return approximately to their original phases. In fact, the recurrence time T_R for the solitons turns out to be approximately equal to $30.4T_B$, and at this time the wave shape $u(x, T_R)$ is uniformly very close to the initial wave form $u(x, 0) = \cos(\pi x)$. There is a second (somewhat weaker) focusing at time $t = 2T_R$, etc.

Notice that, as Zabusky and Kruskal emphasize, it is the persistence or shape conservation of the solitons that provides the explanation of recurrence. If the shapes of the solitons were not preserved when they interacted, there would be no way for them to all get back together and approximately reconstitute the initial condition at some later time. A full understanding of FPU recurrence requires that we comprehend the reasons behind the remarkable new phenomenon of solitonic behavior, and in particular *why solitons preserve their shape.*

The unusual nonlinear phenomena described by Zabusky and Kruskal created a lot of excitement, but no one knew how to solve the KdV, except numerically. In 1967, Gardner, Greene, Kruskal, and Miura showed that the KdV equation was integrable. They presented a method, known as the **Inverse Scattering Transform** (IST), to solve the initial-value problem for the KdV. They showed that $u(x, t)$ can be obtained from $u(x, 0)$ with the help of the solution to the inverse scattering problem for the linear 1-D Schrödinger equation. They also explained that soliton solutions to the KdV correspond to a zero reflection coefficient in the scattering data.

Basically IST is a nonlinear generalization of the Fourier transform method for constructing solutions of the KdV emerging from arbitrary initial conditions (Gardner et al., 1967). Further extension of the IST have lead to the solution of other problems and to the discovery of other types of solitons in a variety of fields.

3.4 How can be a soliton defined?

"*Each such 'solitary-wave pulse' or 'soliton' begins to move uniformly at a rate.. which is linearly proportional to its amplitude*", [1965 Zabusky & Kruskal in Physical Rev. Lett. XV. 240/1].

"*The solitons exhibit a remarkable stability in that their identity is preserved through nonlinear interactions. This property of solitons.. was discovered numerically and justifies the name suggestive of particles*", [1967 Zabusky & Kruskal in Physical Rev. Lett. XIX. 1096/1].

"*Steep waves in shallow water have nonlinear properties similar to those exhibited by interacting 'solitons', nonlinear dispersive wave entities that arise in solutions of the Korteweg-de-Vries.. equation*", [1968 Trans. Amer. Geophysical Union XLIX. 209/2].

It is not easy to define precisely what a soliton is. Drazin and Johnson (1989) describe solitons as solutions of nonlinear differential equations which represent waves of permanent form; are localised, so that they decay or approach a constant at infinity; can interact strongly with other solitons, but they emerge from the collision unchanged apart from a phase shift; they are the non-linear modes of an integrable system. More formal definitions exist, but they require substantial mathematics. On the other hand, some scientists use the term soliton for phenomena that do not quite have these properties (for instance, the 'light bullets' of nonlinear optics are often called solitons despite losing energy during interaction). In some cases, the term soliton has been used even for dissipative systems.

Now, we can give a couple of operative definitions

Definizione 3.4.1. A **solitary wave** can be defined a solution $u(x,t) \in \mathbf{R}$ of a nonlinear PDE, such that its dependency from $x \in \mathbf{R}$ and $t \geq 0$ is in the form $u(x,t) = f(x - ct)$ and

$$\forall t \geq 0, \qquad \lim_{x \to -\infty} u(x,t) = A \qquad \lim_{x \to \infty} u(x,t) = B$$

with constant A and B.

In other words, a solitary wave is a solution that for two specific variables only depends on their difference. Thus, it is translated and keeps its shape in those directions. Note that A and B should not necessarily vanishing, as demonstrated by the SGE soliton solutions.

Definizione 3.4.2. An N-soliton solution is a solution that is asymptotic to a nontrivial sum of N solitary waves $\sum_{i=1}^{N} f_i(x - c_i t)$ as $t \to -\infty$ and to the sum of the same waves $\sum_{i=1}^{N} f_i(x - c_i t + r_i)$ with some nonzero phase shifts r_i as $t \to +\infty$.

In other words, solitons are special solitary waves which after nonlinear interaction don't "radiate away energy", but do pass through each other, keeping their direction, velocity and shape, but with phase shifts, i.e. they are displaced from the position they would occupy if the collision had never happened.

There are many equations which have solitary wave solutions, but these are not usually soliton solutions, in that they undergo inelastic scattering.

For instance, there is an alternative shallow water wave model proposed by Peregrine and Benjamin, Bona and Mahoney,

$$u_t + u_x + 6uu_x - u_{xxt} = 0 \tag{3.4.1}$$

which also has a solitary wave solution

$$u(x,t) = \frac{2\kappa^2}{1-4\kappa^2} sech^2[\kappa(x - ct - x_0)], \qquad c^2 = \frac{1}{1-4\kappa^2} \tag{3.4.2}$$

But they are not solitons, because after a collision they loose their initial profile. This equation also differs from the KdV equation in that the KdV equation is asymptotically 'unique' in the sense that all small parameters are removed via the analysis leading to the equation. An asymptotic reduction of equation 3.4.1 yields the KdV equation.

Further, solitons exhibit quasiperiodicity, fast recurrence, persistence (or stability), speed depending upon amplitude (linearly, at leaast for KdV solitons), but don't exhibit dispersion or dissipation of the wavecrest.

Soliton equations seem to generally arise as "continuum limits" of interesting discrete models, often in highly unexpected ways. Equations with multisoliton solutions are very rare (they occur nearly always in one space dimension); these equations are called soliton equations. An example is the KdV discussed above.

3.5 Qualitative characterization of the KdV Equation

In this paragraph we want to gain some insight into why the KdV equation can generate solitons. Let's come back to the KdV equation in nondimensional standard form

$$u_t + 6uu_x + u_{xxx} = 0 \quad x \in \mathbf{R}, \quad t > 0 \tag{3.5.1}$$

The second and the third terms in the equation are the nonlinear and the dispersive terms, respectively.

3.5.1 Dispersion

Dispersion arise from the fact that waves with different wavelengths travel with different phase velocities. An initial well shaped wave is composed by many small waves with different wavelengths, so it will soon spread out (or disperse) in the many components, resulting in signal distorsion.

Let us consider a linear wave equation of the form

$$u_t + P(\tfrac{\partial}{\partial x})\ u = 0 \tag{3.5.2}$$

where P is a polynomial. Recall that a solution $u(x,t)$ of the form $e^{j(kx-\omega t)}$ is called a plane-wave solution, where k is the *wave number* (waves per unit length) and ω is the *angular frequency*. Rewriting this in the form $e^{jk(x-\frac{\omega}{k}t)}$, we recognize that this is a traveling wave of velocity $\frac{\omega}{k}$.

If we substitute this plane-wave solution into our wave equation 3.5.2, we get a formula determining a unique frequency $\omega(k)$ associated to any wave number k

$$\omega(k) = -jP(jk) \tag{3.5.3}$$

This is called the **dispersion relation** for our wave equation. For the plane-wave solution with wave number k it expresses the *phase velocity* $v \equiv \frac{\omega}{k}$, and also the *group velocity* $v_g \equiv \frac{d\omega}{dk}$. The phase velocity measures how fast a point of constant phase is moving, while the group velocity measures how fast the energy of the wave moves.

For example, for the *linear advection equation* $u_t + cu_x = 0$ we obtain $\frac{\omega}{k} = \frac{d\omega}{dk} = c$ so that all plane-wave solutions travel at the same phase velocity c, and we say that we have trivial dispersion in this case. To appreciate the dispersion effect, let's consider the KdV equation under small amplitude waves ($u << 1$). This linearizes the KdV equation as follows

$$u_t + u_{xxx} = 0 \tag{3.5.4}$$

The dispersion relation now reads

$$\omega(k) = -k^3 \tag{3.5.5}$$

The *phase velocity* is $v \equiv \frac{\omega}{k} = -k^2$, while the *group velocity* is $v_g \equiv \frac{d\omega}{dk} = -3k^2$.

For a larger wavelength $\lambda = \frac{2\pi}{k}$, v is smaller. In this case, plane waves of large wave-number (and hence high frequency) are traveling much faster than low-frequency waves.

Therefore, a wave composed of a superposition of elementary components with different wavelengths will disperse, or change its form as it propagates.

That is, suppose our initial condition is $u_0(x)$. We can use the Fourier Transform to write u_0 in the form $u_0(x) = \int \hat{u}_0(k)e^{jkx}dk$, and then, by superposition, the solution

to our wave equation will be $u(x,t) = \int \hat{u}_0(k) e^{jk(x-\frac{\omega}{k}t)} dk$. Suppose for example our initial wave form is a highly peaked Gaussian. Then in the case of the *linear advection equation* $u_t + cu_x = 0$ all the Fourier modes travel together at the same speed and the Gaussian lump remains highly peaked over time. On the other hand, for the linearized KdV equation the various Fourier modes all travel at different velocities, so after a short time they start cancelling each other by destructive interference, and the originally sharp Gaussian quickly broadens.

In an optical fiber, for example, dispersion originates from the frequency dependence of the refractive index of the fiber, i.e. $n = n(\omega)$. Instead in a NLTL, dispersion is due to the periodicity of nonlinear loads.

3.5.2 Nonlinearity

Nonlinearity makes the top of the wave moving faster than the low sides and this causes the wave to shock in the same way as the waves we see on the beach.

Several times already we have referred to the phenomenon of steepening and breaking of negatively sloped wave-fronts for certain wave equations. What does it means?

Let's consider the equation

$$u_t + f(u)u_x = 0 \tag{3.5.6}$$

where $f : \mathbf{R} \to \mathbf{R}$ is some smooth function. This is the *nonlinear advection equation* or NLA.

Theorem 3.5.1. *Let $u(x,t)$ be a smooth solution of 3.5.6 for $x \in \mathbf{R}$ and $t \in [0, t_0]$, and with initial condition $u_0(x) \equiv u(x,0)$. Then for $0 < t < t_0$ the graph of $u(x,t)$ as a function of x can be constructed from the graph of u_0 by translating each point $(x, u_0(x))$ horizontally by an amount $f(u_0(x))t$.*

Proof. The proof is by the **method of characteristics**, i.e., we look for parametric curves $(x(s), t(s))$ along which $u(x,t) = c$ with a constant c, in such a way 3.5.6 is satisfied. Differentiating $u(x(s), t(s)) = c$ with respect to s gives

$$\frac{dx}{dt} = -\frac{u_t}{u_x}$$

and now substitution into 3.5.6 gives

$$\frac{dx}{dt} = f(u(x,t)) = f(c)$$

so the characteristic curves are straight lines, whose slope is $f(c)$. In particular, if we take the straight line with slope $f(u_0(x))$ starting from the point $(x, 0)$, then $u(x,t)$

will have the constant value $u_0(x)$ along this line, i.e.,

$$x(t)|_{u(x,t)=u_0(x)} = x(0) + f(u_0(x))t$$

c.v.d. □

It is now easy to explain steepening and breaking. We assume that the function f is monotonically increasing and that $u_0(x)$ is strictly decreasing on some space interval I. If we follow the part of the wave profile that is initially over the interval I, we see from the theorem 3.5.1 that the higher part (to the left) will move faster than the lower part (to the right), and so gradually overtake it. The result is that the wave "bunches up" and its slope increases (steepening) and eventually there will be a first time T_B when the graph has a vertical tangent (breaking). Clearly the solution cannot be continued past $t = T_B$, since for $t > T_B$ the theorem 3.5.1 would give a multi-valued graph for $u(x,t)$. It is an easy exercise to show that the breaking time is given by $T_B = \frac{1}{|min(u_0'(x))|}$.

Consider now what happens by eliminating the dispersion terms in the KdV equation,

$$u_t + 6uu_x = 0 \tag{3.5.7}$$

It is the NLA with $f(u) = 6u$, also known as the *non-viscous Burgers equation*, whose initial value problem can be solved numerically by a variety of methods. It can also be solved in implicit form analytically, for short times, by the method of characteristics,

$$u(x,t) = u_0(x - 6ut) \tag{3.5.8}$$

so that the velocity of a point of constant displacement u is equal to that displacement. The main disadvantage of the 3.5.8 is that it is not conveniently represented on a fixed numerical grid.

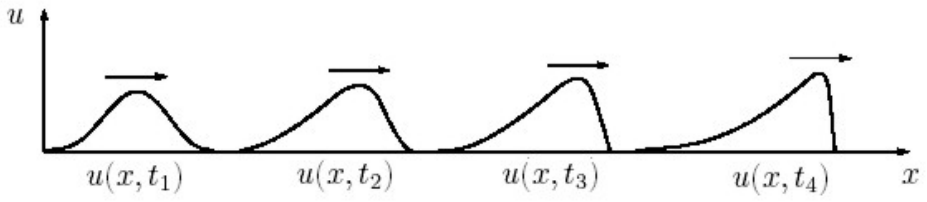


Figure 11: Shock front development in an acoustic wave.

As a result, the wave "breaks"; that is, portions of the wave undergoing greater displacements move faster than, and therefore overtake, those undergoing smaller displacements. *A basic attribute of nonlinear waves is feedback.* As the faster-moving peak overtakes the front of the wave, the more it grows. And the more it grows, the faster it overtakes the front, steepening the wavefront.

In an optical fiber, for example, nonlinearity is due to the so-called *Kerr effect*, where the refractive index depends on the intensity of the optical pulse

$$n = n(\omega, E) = n_0(\omega) + n_2|E|^2$$

In brief, the dispersive and the nonlinear term exhibit two opposing tendencies: nonlinear convection tends to steepen wavecrests, while dispersion tends to flatten wave crests. When both the dispersive and the nonlinear term are present in the equation the two effects can neutralize each other. In other words, if the wave has a special shape the effects are exactly counterbalanced and the wave, with a single wavecrest, propagates without changing shape or losing energy. The soliton shape can be found by direct integration of the KdV equation.

Moreover KdV is the simplest non-dissipative wave-equation with these properties. But real solitary waves undergoes not only dispersion and nonlinearity, but also dissipation and external disturbances can affect them.

3.5.3 Dissipation and perturbations

Solitons are formed as a result of a fine balance between competing effects of dispersion and nonlinearity on the profile of the input wave pulse when both these effects are much stronger than the influence of dissipation which is always present in any real medium. Thus, envelope solitons can only be observed experimentally in the weakly dissipative media.

For example, when Zabusky and Kruskal had used a numerical scheme to simulate the KdV equation, the numerical errors affected the solution.

In the real world solitons are subject to disturbances, and real systems often contain mechanisms that destroy exact soliton behavior and cause the solitons to broaden, such as viscosity, friction, impurities, inhomogeneities, and temperature fluctuations. Experiments nowadays are more sensitive to these external disturbances, allowing for the observation of effects, which could not be detected before. This means that the corresponding theoretical models need to be modified in order to account for these disturbances, so that the broadening can be controlled.

For example, in handling fiber loss, early methods used repeaters to strengthen the signal every several of tens of kilometers. As the width of the available pulses was decreased, the spacing of the repeaters was increased. In the mid 1980's it was proposed that a second wave could be sent along the fiber, which could be used to halt

the dispersion of a soliton through a process known as **Raman scattering**. In 1988 the group at Bell Laboratories used this technique to successfully propagate a soliton over 6000 km without the need for repeaters.

It is well known that a small dissipative perturbation does not change qualitatively the envelope soliton profile. A fully formed envelope soliton simply widens due to dissipation, and its amplitude decreases with time twice as fast as the amplitude of a linear sinusoidal signal in the same medium. On the other hand, a weak dissipation strongly affects the process of soliton formation from an initially linear wave pulse when its amplitude is gradually increased.

Perturbation methods (with the perturbation taking place around the soliton) have been developed to compute the response of the soliton to external forces, damping, etc. The perturbed KdV equation is

$$u_t + cu_x + 6uu_x + u_{xxx} = f(u, x, t)_x \tag{3.5.9}$$

where f is an external force. For example one could consider the time-periodic one-harmonic case $f = f_0 \cos \omega(x - vt)$.

Often the result is that the parameters characterizing the soliton (such as velocity and amplitude) are now time dependent, with the time dependence governed by simple ODEs.

Recently it is observed a significant interest in the *chaotic behaviour* of solutions to PDEs. For example, chaos was found for a perturbed sine-Gordon equation, the NLS equation, and for the KdV-Burgers equation and its generalization on high-order non-linearities, and to the case of Kadomtsev-Petviashvili (KP) equation. For the KdV equation incorporating dissipation and instability it was shown that for the strongly dissipative case the overall evolution of solutions is chaotic with irregular soliton interactions (T.Kawahara and S.Toh, 1989).

3.6 Exact Soliton Solution to the KdV Equation

The aim of the this paragraph is to find the general exact soliton solution to the KdV equation, without any boundary conditions. The KdV is a hyperbolic PDE in the general sense of the hyperbolicity definition. From that it follows that it describes a reversible dynamical process. The main part of the subsequent deduction of an analytical solution is taken from Vvedenskii (1992), who also performed non-linear superposition of two, three and more solutions corresponding to two, three or more soliton waves by using Bäcklund transform.

So, let's rewrite the KdV equation in standard form

$$u_t + 6uu_x + u_{xxx} = 0 \tag{3.6.1}$$

Let us start with the most elementary of all one-dimensional wave equations, the *forward wave equation*,

$$u_t + vu_x = 0 \tag{3.6.2}$$

Any wave function of the form $f(x - vt)$ is a solution to 3.6.2 where v denotes the speed of the wave. For the well known wave equation

$$u_{tt} - v^2 u_{xx} = 0 \tag{3.6.3}$$

the famous d'Alembert solution leads to two wave fronts represented by terms $f(x + vt)$ and $f(x - vt)$. Hence we start here with a trial solution

$$u(x,t) = z(x - vt - x_0) \equiv z(\xi) \tag{3.6.4}$$

with v and x_0 arbitrary constants. Substituting the trial solution 3.6.4 into 3.6.1 we are led to the ODE

$$-v\frac{dz}{d\xi} + 6z\frac{dz}{d\xi} + \frac{d^3z}{d\xi^3} = 0 \tag{3.6.5}$$

Integrating with respect to ξ

$$-vz + 3z^2 + \frac{d^2z}{d\xi^2} = c_1 \tag{3.6.6}$$

Note that 3.6.6 is an equation of motion for a nonlinear oscillator in a cubic potential energy field $V(z) = z^3 - \frac{vz^2}{2}$, with the Hamiltonian $H(z,x) = \frac{z^2}{2} + V(z)$, and c_1 playing role of energy times 2. It generically has two equilibria.

Integrating with respect to z

$$-v\frac{z^2}{2} + z^3 + \frac{1}{2}\Big(\frac{dz}{d\xi}\Big)^2 = c_1 z + c_2 \tag{3.6.7}$$

where c_1 and c_2 are the constants of integration. Now, from the requirements that z and its 1st and 2nd derivatives are 0 for $\xi \to \infty$ it follows that $c_1 = c_2 = 0$. Equation 3.6.7 can be written as

$$\Big(\frac{dz}{d\xi}\Big)^2 = z^2(v - 2z) \tag{3.6.8}$$

By extracting the square root and by separation of variables we obtain

$$\int_0^z \frac{d\zeta}{\zeta\sqrt{v - 2\zeta}} = \int_0^\xi \eta \tag{3.6.9}$$

The first integral can be solved by using a transformation

$$\zeta = \frac{v}{2} sech^2 w \tag{3.6.10}$$

Remember that $sech(w) = \frac{2}{e^w + e^{-w}}$. We have

$$\sqrt{v - 2\zeta} = \sqrt{v} \cdot tanh^2 w$$

$$d\zeta = -v \cdot \frac{sinh(w)}{cosh^3 w} dw$$

$$w = sech^{-1} \sqrt{\frac{2\zeta}{v}}$$

Sobstituting in 3.6.9 and
transforming back to z we obtain

$$u(x,t) = \frac{v}{2} sech^2 \left[\frac{\sqrt{v}}{2} (x - vt) \right] \tag{3.6.11}$$

In order to have a real solution v must be a positive number, i.e. the solitary wave moves to the right. Its amplitude is proportional to the speed v, while its width is inversely proportional to $\sqrt{v}$. *Thus larger amplitude solitary waves move with a higher speed than smaller amplitude waves.* This fact has not correspondences in the domain of the linear systems, where amplitude and propagation speed are quite independent quantities.

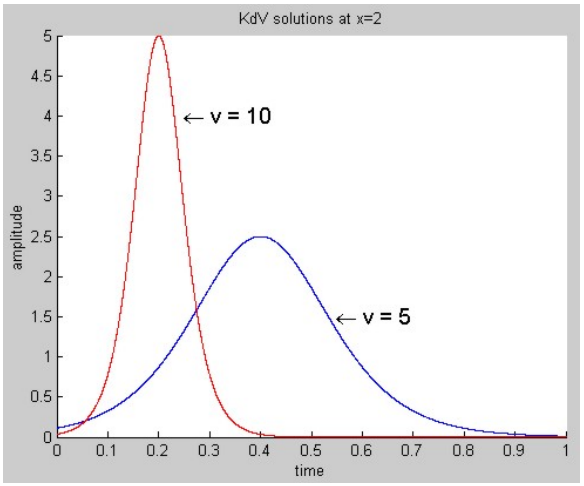


Figure 12: Plots for the KdV solution

It is straightforward to note that a generic coefficient δ in the nonlinear term of the KdV equation would have lead to

$$\int_0^z \frac{d\zeta}{\zeta\sqrt{v-\frac{\delta}{3}\zeta}} = \int_0^\xi \eta \tag{3.6.12}$$

or, using $v' = \frac{6}{\delta}v$,

$$\int_0^z \frac{d\zeta}{\zeta\sqrt{v'-2\zeta}} = \sqrt{\frac{\delta}{6}}\,\xi \tag{3.6.13}$$

and finally we would have found

$$u(x,t) = \frac{v'}{2} sech^2\big[\tfrac{\sqrt{v'}}{2}\sqrt{\frac{\delta}{6}}\,(x-vt)\big] \tag{3.6.14}$$

i.e. a simply re-scaled solution

$$u(x,t) = \frac{3v}{\delta} sech^2\big[\tfrac{\sqrt{v}}{2}(x-vt)\big] \tag{3.6.15}$$

In the same way as we obtained the solution 3.6.11 we could obtain another solution which is:

$$u(x,t) = -\frac{v}{2} csch^2\big[\tfrac{\sqrt{v}}{2}(x-vt)\big] \tag{3.6.16}$$

This is an irregular solution to the KdV Equation, for it has a singularity along the line $t = \frac{x}{v}$.

3.7 Fixed Points and Linearization Analysis

Let us study the phase portrait of the ODE 3.6.6. We can start by transforming this 2^{nd} order equation to coupled 1^{st} order equations. Defining $y_1 \equiv z$ and $y_2 \equiv \frac{dz}{d\xi}$, 3.6.6 gives

$$\begin{cases} y_1' = y_2 \\ y_2' = -3y_1^2 + vy_1 + c_1 \end{cases}$$

An equilibrium point is the set of values of y_1 and y_2 that makes $y_1'(\xi) = y_2'(\xi) = 0$ for all ξ.

In the context of a system of 1^{st} order equations describing the motion of a particle, an equilibrium point corresponds to the position for which the speed and acceleration are zero at that position. Hence the term *fixed point*(if a particle starts out in that

position, it will remain there for all time). But the system we investigate is different in nature, because z is not a position, and ξ is not time. Instead of time as the independent variable, we have $\xi = x - vt$. For our purposes it can be considered as a position shifted by a fixed quantity at some finite time. c_1 is the integration constant which can be chosen equal to zero from the condition that the solution is bounded at infinity. There are two equilibrium points,

$$(y_1, y_2) = (0, 0)$$
$$(y_1, y_2) = (\tfrac{v}{3}, 0)$$

As is common procedure, we now linearize the system about these points. In the context of systems describing the positions of particles in time, linearizing the system tells us about whether the fixed point is stable, that is, if the particle is perturbed, if it will oscillate about the fixed point or run away from it. We don't enjoy this interpretation of the linearization for our systems. Instead, we just learn about the solution in the vicinity of that point. The Jacobian is

$$\mathbf{J} = \begin{bmatrix} 0 & 1 \\ v - 6x_1 & 0 \end{bmatrix} \qquad \mathbf{J}|_{(0,0)} = \begin{bmatrix} 0 & 1 \\ v & 0 \end{bmatrix} \qquad \mathbf{J}|_{(\frac{v}{3},0)} = \begin{bmatrix} 0 & 1 \\ -v & 0 \end{bmatrix}$$

Zeroing the determinant $det(\lambda\mathbf{I} - \mathbf{J})$ we find that for $(0, 0)$ eigenvalues are real and distinct, $\lambda_{1,2} = \pm\sqrt{v}$, hence this equilibrium point *is a saddle*. For $(\frac{v}{3}, 0)$ eigenvalues are $\lambda_{1,2} = \pm j\sqrt{v}$, hence this equilibrium point *is a center*.

Now we know the behavior of the system about these equilibrium points. But where do the soliton solutions show up? Perhaps we can learn more by investigating the phase diagrams numerically.

As expected, we have one fixed point near which the solution is periodic, and one at the origin where the solution is saddle-like.

We have a *sepatrix* which starts at the origin, winds once around the other equilibrium point, and returns to the origin. Inside this sepatrix trajectories are diffeomorphic to ellipses but "sharper" at their left hand ends. These correspond to periodic solutions of our ODE i.e. to periodic traveling wavetrains. These periodic solutions may be written explicitly in terms of the Jacobi elliptic function.

Additionally, we have closed orbits encompassing all the fixed points (another periodic solution type).

What should the trajectories of the soliton look like in the phase plane? Solitons will have some nonzero maximum amplitude for some point in space, at which the first derivative with respect to space will be zero ($y_2 = \frac{dz}{d\xi} = 0$ at $y_1 = z = z_{max}$). This corresponds to a point on the y_1 axis in the (y_1, y_2) phase plane. From the example shown in the figure, we see the soliton peak being $z_{max} = \frac{v}{2}$, exactly what expected from the analytic soliton solution 3.6.11.

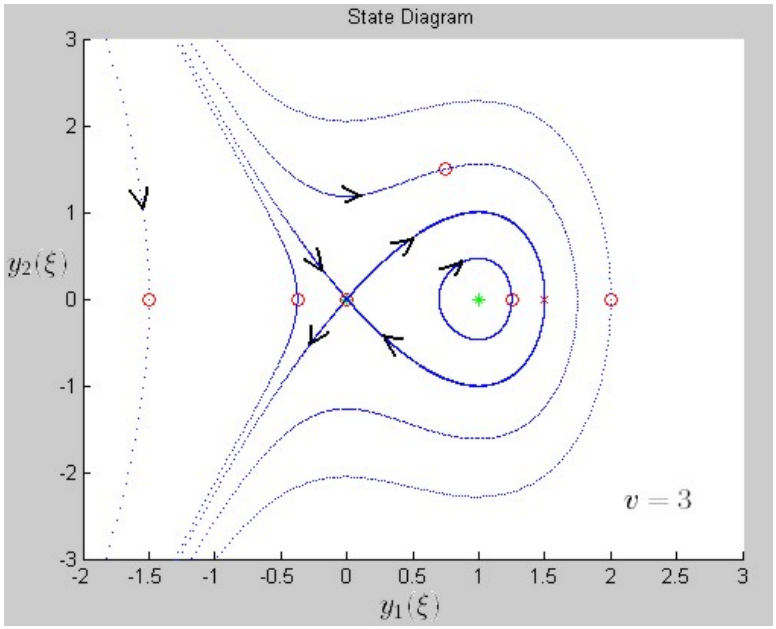


Figure 13: Phase diagrams for the KdV equation with $v = 3$. (*) - equilibrium points; (o) - initial condition for plotting trajectories; (x) - The soliton peak $(\frac{v}{2}, 0)$

Additionally, the soliton amplitude will go to zero sufficiently far away from this point in either direction (and consequently the derivative of the soliton amplitude will also go to zero). This corresponds to the soliton trajectory approaching the origin of the phase plane from two different directions. In math-language, such a trajectory in the phase plane that begins and ends at the same fixed point is a **homoclinic orbit**.

This is very informative; it tells us that *a system hosts a bright soliton if it has the following ingredients: a fixed point the origin, and a homoclinic orbit to this fixed point that crosses the y_1 axis.* Note that these conditions imply that the fixed point at the origin cannot be circular, otherwise the homoclinic orbit could not approach the origin from two different directions.

Another interesting note is with regard to how the soliton trajectory approaches the origin. First, the trajectory approaches the origin, but never reaches the origin. Indeed, it starts from the origin at $\xi \to -\infty$, and it reaches the the origin at $\xi \to +\infty$. Second, the trajectory must approach the origin along the eigenvectors of the Jacobian evaluated at the origin (otherwise the trajectory would not approach the origin). Solving for

$$\mathbf{y}\mathbf{J}|_{(0,0)} = \pm\sqrt{v}\mathbf{y}$$

we find that the stable and unstable eigenvectors are $(1, -\sqrt{v})$ and $(1, \sqrt{v})$, respec-

tively. This determines how fast the soliton wave rises and falls, both in space and time.

As already said, the homoclinic orbit represents a disturbance with a single crest (the only non periodic cnoidal wave) which travels indefinitely without changing shape, so it is a "solitary wave". But saying it is a "soliton" implies much more! To see why, observe that from what we've said so far, we might well expect that our solitary wave is unstable, i.e. that small perturbations should either change it into a long period cnoidal wave, or into a disturbance which blows up. Alternatively, because this solitary wave arises from a governing equation which balances the opposite tendencies of nonlinearity and dispersion, we might well expect solitary waves to be unstable: surely this "balancing" is a delicate affair! But in fact, decades of work (both experimental and theoretical, beginning with Russell's work) show that *solitons are stable against small nonlinear perturbations.* This is, or should be, quite surprising! The remarkable stability of solitons is well illustrated by their very surprising nonlinear interactions. The unusual stability of solitons is explained by the existence of an infinite hierarchy of independent conservation laws for the parent PDE; compare ordinary Hamiltonian systems, which admit only finitely many independent conservation laws, that is, a solution $u(x,t)$ is characterized by an infinite list of quantities which are invariant during the evolution of the solution $u(x,t)$.

3.8 Exact Solution with Bäcklund transform

We are now going to find the same solutions to KdV with the help of Bäcklund transform (BT). This technique is used later to build non-linear superposition of basic solutions. In BT, 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.

We introduce a function w with the property $w_x = u$. It follows from KdV

$$w_{tx} + 6w_x w_{xx} + w_{xxxx} = 0 \tag{3.8.1}$$

that can be rewrite

$$\frac{\partial}{\partial x}\big(w_t + 3w_x^2 + w_{xxx}\big) = 0 \tag{3.8.2}$$

Integration with respect to x yields

$$w_t + 3w_x^2 + w_{xxx} = f(t) \tag{3.8.3}$$

The integration function $f(t)$ can be set equal to 0 without loss of generality. In order to understand this fact let's introduce another variable

$$w^* = w - \int^t f(\tau)d\tau \tag{3.8.4}$$

The solution u to KdV will be obtained from the solution w to 3.8.3 and since $w_x^* = w_x$ there is no need for a distinction between w^* and w. Now our purpose is to solve the PDE

$$w_t + 3w_x^2 + w_{xxx} = 0 \tag{3.8.5}$$

An Auto-Bäcklund transform, that leaves equation 3.8.5 invariant, is given by the following set of equations:

$$\widetilde{w}_x = -w_x + \beta - \frac{1}{2}(w - \widetilde{w})^2 \tag{3.8.6}$$

$$\widetilde{w}_t = -w_t + (w - \widetilde{w})(w_{xx} - \widetilde{w}_{xx}) - 2(w_x^2 + w_x\widetilde{w}_x + \widetilde{w}_x^2) \tag{3.8.7}$$

The parameter β is the so called *Bäcklund parameter*. We now may generate a non trivial solution by applying the transform to the trivial solution $\widetilde{w} \equiv 0$ of 3.8.5. The transform equations result into

$$w_x = \beta - \frac{1}{2}w^2 \tag{3.8.8}$$

$$w_t = ww_{xx} - 2w_x^2 \tag{3.8.9}$$

The first line may easily be integrated and be inserted into the second. Two types of solutions result:

$$w(x,t) = \sqrt{2\beta}\tanh\left[\sqrt{\tfrac{\beta}{2}}(x-2\beta t)\right] \tag{3.8.10}$$

$$\overline{w}(x,t) = \sqrt{2\beta}\coth\left[\sqrt{\tfrac{\beta}{2}}(x-2\beta t)\right] \tag{3.8.11}$$

It is easy to verify that both functions do fulfill 3.8.8 and 3.8.5. The second function has a singularity for $t = \frac{x}{2\beta}$. This irregular solution here and in future is denoted with a bar. By simple derivation of the equations above with respect to x we deduce two solutions (a regular one and an irregular one) to the KdV equation. These solutions are:

$$u(x,t) = w_x(x,t) = \beta sech^2\left[\sqrt{\tfrac{\beta}{2}}(x-2\beta t)\right] \tag{3.8.12}$$

$$\overline{u}(x,t) = \overline{w}_x(x,t) = -\beta csch^2\left[\sqrt{\tfrac{\beta}{2}}(x-2\beta t)\right] \tag{3.8.13}$$

Again we are led to the same solution we already found with 3.6.11 and 3.6.16, with $\beta = \frac{v}{2}$.

3.9 Nonlinear superposition

Be Φ any solution to 3.8.5 and let $\Phi(\beta_1)$ and $\Phi(\beta_2)$ be two other solutions obtained from Φ by applying the Bäcklund transform with the Bäcklund parameters β_1 and β_2 respectively. From 3.8.6 we have (only the first line is used):

$$\Phi_x(\beta_1) = -\Phi_x + \beta_1 - \frac{1}{2}(\Phi - \Phi(\beta_1))^2 \tag{3.9.1}$$

$$\Phi_x(\beta_2) = -\Phi_x + \beta_2 - \frac{1}{2}(\Phi - \Phi(\beta_2))^2 \tag{3.9.2}$$

Let $\Phi(\beta_2, \beta_1)$ be the solution obtained from Φ by successive application of 3.8.6 by first applying parameter β_1 and afterwards parameter β_2. Then we have

$$\Phi_x(\beta_2, \beta_1) = -\Phi_x(\beta_1) + \beta_2 - \frac{1}{2}(\Phi(\beta_1) - \Phi(\beta_2, \beta_1))^2 \tag{3.9.3}$$

In the same way let $\Phi_x(\beta_1, \beta_2)$ be the solution by first applying β_2 and afterwards β_1. We then have

$$\Phi_x(\beta_1, \beta_2) = -\Phi_x(\beta_2) + \beta_1 - \frac{1}{2}(\Phi(\beta_2) - \Phi(\beta_1, \beta_2))^2 \tag{3.9.4}$$

We now demand that

$$\Phi(\beta_1, \beta_2) = \Phi(\beta_2, \beta_1) \equiv \Psi \tag{3.9.5}$$

This is Bianchi's *Theorem of permutability*, and is an example of a nonlinear superposition formula!

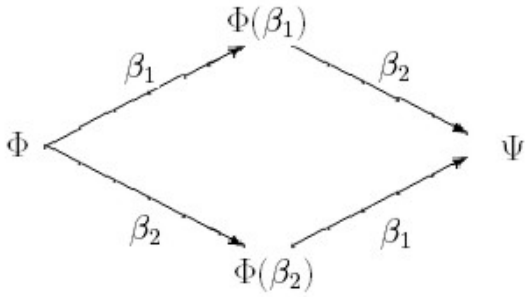


Figure 14: A commutative Bianchi diagram for constructing the 2-soliton solution

Combining the results obtained subtracting 3.9.2 from 3.9.1 and 3.9.4 from 3.9.3 yields:

$$\Psi = \Phi + 2\frac{\beta_1 - \beta_2}{\Phi(\beta_1) - \Phi(\beta_2)} \tag{3.9.6}$$

We can show that the function Ψ is a solution to 3.8.5. The 3.9.6 is he main formula for building up further solutions.

If we choose the trivial solution $\Phi \equiv 0$ we obtain regular and irregular solutions like 3.8.10. Each of these can play the role of $\Phi(\beta_1)$ and $\Phi(\beta_2)$ in equation 3.9.6, in order to produce other solutions. Using two regular solutions (with two different values of β) would lead to an irregular solution by using 3.9.6. This can be seen as follows

$$\begin{aligned}\lim_{x\to\pm\infty}[\Phi(\beta_1) - \Phi(\beta_2)] &= \lim_{x\to\pm\infty}\left[\sqrt{2\beta_1}\tanh[\sqrt{\tfrac{\beta_1}{2}}(x - 2\beta_1 t)] - \sqrt{2\beta_2}\tanh[\sqrt{\tfrac{\beta_2}{2}}(x - 2\beta_2 t)]\right] \\ &= \pm(\sqrt{2\beta_1} - \sqrt{2\beta_2})\end{aligned}$$

Since the function Φ is continuous in x there will always be one value of the argument x for which the denominator in 3.9.6 is zero, leading to a singularity in Ψ. So we use both the regular and the irregular solution from to construct a regular superpositioned function solution:

$$\Psi(x,t) = \frac{2(\beta_1 - \beta_2)}{w(\beta_1, x, t) - \overline{w}(\beta_2, x, t)} \tag{3.9.7}$$

Analogous to what we did here before, the bar again denotes the irregular solution. The two Bäcklund parameters are chosen as $\beta_2 > \beta_1$. This ensures that the denominator in 3.9.7 does not vanish for any values of the arguments.

Since the function Ψ is a solution to 3.8.5 the superpositioned solution to the original KdV equation is given by the first derivative with respect to x:

$$u(x,t) = \frac{d}{dx}\Big(\frac{2(\beta_1 - \beta_2)}{w(\beta_1, x, t) - \overline{w}(\beta_2, x, t)}\Big) \tag{3.9.8}$$

This process of superpositioning a regular and an irregular solution to 3.8.5 can be iterated. We have to apply 3.9.6 again to 3 solutions.

$$\Phi(\beta_1, \beta, x, t) = \frac{2(\beta - \beta_1)}{w(\beta, x, t) - \overline{w}(\beta_1, x, t)} \tag{3.9.9}$$

$$\overline{\Phi}(\beta_2, \beta, x, t) = \frac{2(\beta_2 - \beta)}{w(\beta_2, x, t) - w(\beta, x, t)} \tag{3.9.10}$$

$$\Psi(x,t) = w(\beta, x, t) + \frac{2(\beta_1 - \beta_2)}{\Phi(\beta_1, \beta, x, t) - \overline{\Phi}(\beta_2, \beta, x, t)} \tag{3.9.11}$$

$$\beta_2 > \beta_1 > \beta \tag{3.9.12}$$

Principally by superposition a regular N-Solitons can be constructed from a pair of irregular and regular (N-1)-solutions. Contour plots for the two- and the three-soliton solutions are shown below

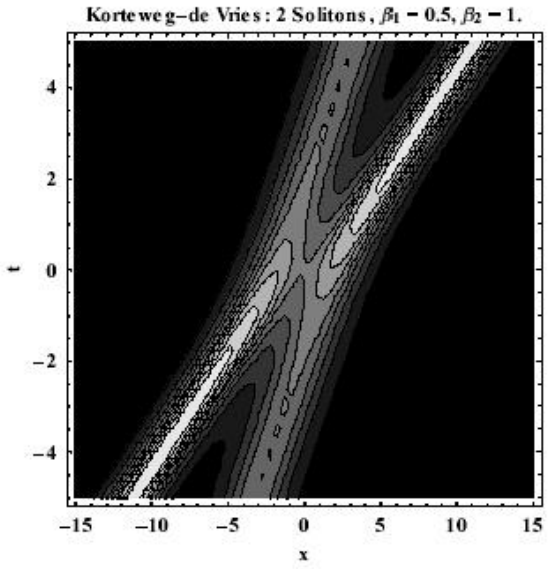


Figure 15: Contour plots for the two-soliton solutions. Note the phase shift introduced by the collision.

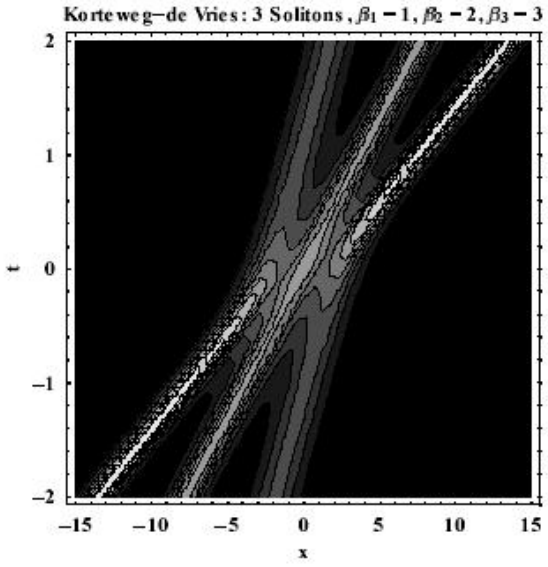


Figure 16: Contour plots for the three-soliton solutions

4 Local Conservation Laws

Be $u(x,t)$ a solution of a given PDE, and be $\mathcal{T}$ and $\mathcal{X}$ two polynomials in x,t,u and its x-derivatives, such that

$$\frac{\partial \mathcal{T}}{\partial t} + \frac{\partial \mathcal{X}}{\partial x} = 0 \tag{4.0.13}$$

4.0.13 is said *Conservation Law*. The quantities $\mathcal{T}$ and $\mathcal{X}$ are respectively called conserved **density** and **flux**. A well known example is the conservation law of the current density, that is a local expression of the conservation of charge

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{J} = 0 \tag{4.0.14}$$

With appropriate boundary conditions equation 4.0.13 leads to a conserved quantity (constant of the motion). Integrating with respect to x we get

$$\frac{\partial}{\partial t} \int_{-\infty}^{\infty} \mathcal{T}\, dx = -\left[\mathcal{X}\right]_{-\infty}^{\infty} \tag{4.0.15}$$

Under periodic boundary conditions in x, with $B - A$ an integer multiple of the period, or with $A = +\infty$, $B = -\infty$ and $\lim_{x \to \pm\infty} u = 0$, the square bracket vanishes and we have the constant of motion

$$\int_{-\infty}^{\infty} \mathcal{T} dx = costant \tag{4.0.16}$$

For instance, in the KdV it can be easily seen that

$$\begin{aligned} 0 &= u_t - 6uu_x + u_{xxx} && (4.0.17)\\ &= \frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(-3u^2 + u_{xx}) && (4.0.18)\\ & && (4.0.19) \end{aligned}$$

Thus, using

$$\begin{cases} \mathcal{T} = u \\ \mathcal{X} = u_{xx} - 3u^2 \end{cases}$$

it follows that

$$\int_{-\infty}^{\infty} u \; dx = constant \tag{4.0.20}$$

This conservation law has some physical interpretation. Remembering that $u(x,t)$ is proportional to the height of the water with respect to the unperturbed level, we deduce that 4.0.20 takes the meaning of a mass conservation law.

Besides, each $\mathcal{T}$ is only determined up to an exact x-derivative, so it defines an equivalence class of conserved densities

$$\mathcal{T} \sim \mathcal{T} + \frac{d\mathcal{S}}{dx} \tag{4.0.21}$$

since this only adds $\frac{d\mathcal{S}}{dt}$ to $\mathcal{X}$ and leaves the value of $\int_{-\infty}^{\infty} \mathcal{T} dx$ unchanged

$$\frac{\partial}{\partial t} \int_{-\infty}^{\infty} (\mathcal{T} + \frac{d\mathcal{S}}{dx}) dx = - \left[\mathcal{X}\right]_{-\infty}^{\infty} = 0$$

$$\frac{\partial}{\partial t} \int_{-\infty}^{\infty} \mathcal{T} dx = - \left[\frac{d\mathcal{S}}{dt} + \mathcal{X}\right]_{-\infty}^{\infty} = 0$$

Therefore two conserved densities differing by an x-derivative determine the same conservation law (they are said to be gauge-equivalent). For the KdV equation the first three non-equivalent conservation laws are

mass conservation law	$\mathcal{T}_0 = u$	$\mathcal{X}_0 = -u_{xx} - 3u^2$
momentum conservation law	$\mathcal{T}_1 = \frac{1}{2}u^2$	$\mathcal{X}_1 = -uu_{xx} + \frac{1}{2}u_x^2 - 2u^3$
energy conservation law	$\mathcal{T}_2 = \frac{1}{2}u_x^2 - u^3$	$\mathcal{X}_2 = -u_x u_{xxx} + \frac{1}{2}u_{xx}^2 + 3u^2 u_{xx} - 6uu_x^2 + \frac{9}{2}u^4$
?? conservation law	...	...

In his initial investigation of local conservation laws valid for the KdV equation, Miura et al. (1968) discovered nine conservation laws by direct calculation and Miura (1968) found a tenth, hinting that an infinite number of conserved quantities might exist (Tabor 1989, p. 288). He wrote:

> "*The author's own introduction to this field of research began here with the tedious work of deriving more explicit conservation laws. Four additional ones were discovered in rapid succession leading to the obvious conjecture that there were infinitely many. However, during the summer of 1966 there was a rumor that only nine polynomial conservation laws existed-exactly the number that had been found. Consequently, the author spent a week's vacation at a beautiful lake near Peterboro, Ontario, Canada, working out the tenth conservation law, which exists. An algorithm had been developed for computing the conserved densities and Donald Stevens of the Courant Institute of Mathematical Sciences devised a computer program for the AEC CDC 6600 computer which successfully computed the eleventh conserved density consisting of terms. This was a significant accomplishment in view of the fact that a program using the FORMAC symbol manipulating language for an IBM 7094 computer written at the Los Alamos Scientific Laboratory for the same purpose only successfully computed up through the fifth conserved density before exceeding the available storage space.*". (R.M.Miura. The Kortewege de Vries equation, a survey of results. SLAM Review, 18 (1976) 412-559.)

4.1 MKdV equation and Miura Transformation

In order to ascertain whether the KdV equation was the only such equation with so many conservation laws Miura investigated the equation

$$u_t - 6u^2u_x - u_{xxx} = 0 \quad x \in \mathbf{R}, \quad t > 0 \tag{4.1.1}$$

that is the modified KdV equation (MKdV). The MKdV equation can be written

$$w_t + 6w^2w_x - w_{xxx} = 0 \quad x \in \mathbf{R}, \quad t > 0 \tag{4.1.2}$$

$$\frac{\partial w}{\partial t} + \frac{\partial}{\partial x}(2w^3 - w_{xx}) = 0 \tag{4.1.3}$$

Thus, the first few conservation laws for MKdV correspond to conserved densities and fluxes

$$\begin{array}{ll} \overline{\mathcal{T}}_{-1} = w & \overline{\mathcal{X}}_{-1} = -w_{xx} + 2w^3 \\ \overline{\mathcal{T}}_0 = w^2 & \overline{\mathcal{X}}_0 = -2ww_{xx} + w_x^2 + 3w^4 \\ \overline{\mathcal{T}}_1 = \frac{1}{2}(w_x^2 + w^4) & \overline{\mathcal{X}}_1 = -w_xw_{xxx} + \frac{1}{2}w_{xx}^2 - 2w^3w_{xx} + 6w^2w_x^2 + 2w^6 \\ \dots & \dots \end{array}$$

Miura then noticed a remarkable fact, by substituting

$$u = -w_x - w^2 \tag{4.1.4}$$

into the conserved densities of the KdV equation they were transformed into those of the MKdV equation (except for $\overline{\mathcal{T}}_{-1}$, which is not included this way). For example

$$\mathcal{T}_0 = u = -w_x - w^2 = -\frac{d}{dx}(w) - \overline{\mathcal{T}}_0 \sim -\overline{\mathcal{T}}_0 \tag{4.1.5}$$

$$\mathcal{T}_1 = \frac{1}{2}u^2 = \frac{1}{2}(w_x^2 + w^4 + 2w^2w_x) = \frac{d}{dx}(\frac{1}{3}w^3) + \overline{\mathcal{T}}_1 \sim \overline{\mathcal{T}}_1 \tag{4.1.6}$$

$$\dots \tag{4.1.7}$$

where $\sim$ refers to the gauge-equivalence relation introduced above. The substitution 4.1.4 is now referred to as the **Miura transformation** (more properly, **Miura map**) Such maps play a very important role in the Hamiltonian theory of soliton equations. A transformation due to Gardner provides an algorithm for computing an infinite number of conserved densities of the KdV equation, which are connected to those of the MKdV equation through the Miura transformation. Furthermore 4.1.4 gives a direct relation between equations KdV and MKdV, since if w satisfies the MKdV equation then u satisfies the KdV equation.

5 Burgers Equation

The Burgers equation is a nonlinear PDE of second order, as follows

$$u_t + uu_x - \lambda u_{xx} = 0 \qquad x \in \mathbf{R}, \quad t > 0 \tag{5.0.8}$$

This equation is used in fluid dynamics teaching and in engineering as a simplified model for turbulence about a point, shock wave formation, and mass transport. Many different closed-form and numerical solutions are known for particular sets of boundary conditions. λ is the damping coefficient. As already done with KdV we start with a change of variable, $\xi = x - vt$. But

$$\frac{du}{d\xi} = \frac{\partial u}{\partial x} = -\frac{1}{v}\frac{\partial u}{\partial t}$$

so that the Burgers PDE becomes an ODE

$$-v\frac{du}{d\xi} + u\frac{du}{d\xi} - \lambda\frac{d^2u}{d\xi^2} = 0 \tag{5.0.9}$$

Integrating with respect to ξ

$$-vu + \frac{1}{2}u^2 - \lambda\frac{du}{d\xi} + c_1 = 0 \tag{5.0.10}$$

Separating variables and integrating we are led to

$$\frac{\xi + c_2}{2\lambda} = \frac{1}{\sqrt{2c_1 - v^2}} \tan^{-1}\left(\frac{u - v}{\sqrt{2c_1 - v^2}}\right) \tag{5.0.11}$$

Following the same procedure, with a change of variable $\xi = x+vt$, and rearranging, we obtain the two travelling-wave general solution to the Burgers equation

$$u(x,t) \equiv u_1(x - vt) = v + \sqrt{2c_1 - v^2}\tan\left[\frac{\sqrt{2c_1 - v^2}}{2\lambda}(x - vt + c_2)\right] \tag{5.0.12}$$

$$u(x,t) \equiv u_2(x + vt) = -v + \sqrt{2c_3 - v^2}\tan\left[\frac{\sqrt{2c_3 - v^2}}{2\lambda}(x + vt + c_4)\right] \tag{5.0.13}$$

$$\tag{5.0.14}$$

where c_1,c_2,c_3,c_4 are the constants of integration (real or complex, all arbitrary, except that complementary pairs cannot be equal unless both are zero). Let's consider for the moment the first equation. Using the fact that

$$\tan(j\alpha) = j\ \tanh(\alpha)$$

we can rewrite it

$$u(x,t) = v + j\sqrt{2c_1 - v^2}\tanh\big[-j\frac{\sqrt{2c_1 - v^2}}{2\lambda}(x - vt + c_2)\big] \tag{5.0.15}$$

c_1 and c_2 are determined by boundary conditions. We can choose meaningful real values for the constants such that we have a purely real expression. For purposes of illustration and interpretation, consider the case wherein $\lim_{x\to-\infty} u = u_0$, $\lim_{x\to\infty} u = 0$, $u(0,0) = \frac{u_0}{2}$. We obtain $c_1 = c_2 = 0$ and

$$u(x,t) = v - v\tanh[\frac{v}{2\lambda}(x - vt)] \tag{5.0.16}$$

with $v = \frac{u_0}{2}$.

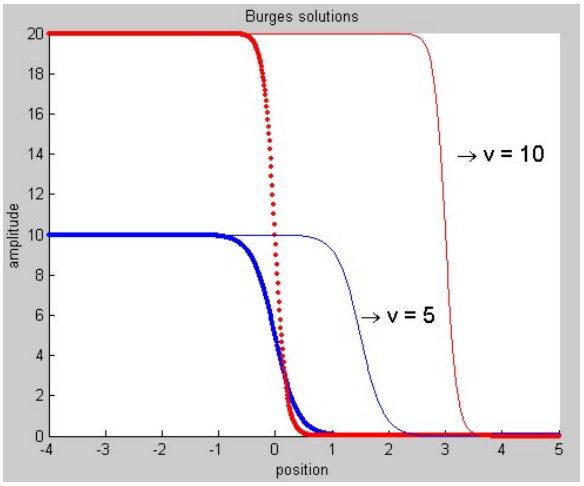


Figure 17: Plots for u_1 solution to the Burgers equation with the BC given in the example. Two curves with $v = 10$ and $v = 5$ are plotted at time instants $t = 0$ ad $t = 0.3$. $\lambda = 1$.

Recall the shape of the tanh() function. These functions could represent a state change from one level to another, with a smooth s-shaped transition, that is propagated, or diffuses, throughout the medium. Parameter λ controls the speed of transition. This could represent an incremental change in temperature, pressure, or concentration that diffuses away from the point of disturbance with a wavefront speed v. Moreover, the first derivatives of these functions represent a finite-amplitude solitons moving away to infinity in either direction from the site of the original disturbance at the same speed.

6 KdV-Burgers Equation

In this case to the Burgers equation the 3rd order partial derivative with respect to x is added, as follows

$$u_t + 2uu_x - \lambda u_{xx} + \beta u_{xxx} = 0 \tag{6.0.17}$$

This equation has various physical applications, for instance, it can serve as a nonlinear wave model of a fluid in an elastic tube, of a liquid with small bubbles and turbulence. The coefficient λ and β are the damping and the dispersion coefficient, respectively.

7 Sine-Gordon Equation (SGE)

This historically important equation has the form

$$u_{tt} - u_{xx} + \sin(u) = 0 \qquad x \in \mathbf{R}, \quad t > 0 \tag{7.0.18}$$

The Sine-Gordon Equation (SGE) is a nonlinear version of the Klein-Gordon equation $u_{tt} - u_{xx} + u = 0$, by which its name derives. It is even older than the KdV equation. It arise 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 (*fluxons*) in long Josephson transmission lines and in Josephson junctions (a junction between two superconductors), the motion of rigid pendula attached to a stretched wire, and the propogation of dislocations in crystal lattices. Further, the SGE has also been proposed as a simplified model for a unified field theory.

Originally, the Sine–Gordon equation was introduced in the field of differential geometry, for the theory of "pseudo-spherical surfaces", i.e. surfaces of constant negative Gaussian curvature immersed in $\mathbf{R}^3$, around 1860. The Gauss-Codazzi equations for such surfaces reduce to the SGE, so that by the "Fundamental Theorem of Surface Theory", there is a bijective correspondence between isometry classes of isometric immersions of the hyperbolic plane into $\mathbf{R}^3$ and solutions to the SGE (Strictly speaking, we relax the immersion condition to admit cusp singularities along curves). Because of this, and the great interest in non-Euclidean geometry during the latter half of the last century, a prodigious amount of effort was devoted to the study of the SGE by the great geometers of that period, resulting in a beautiful body of results. Thus, the equation, as well as several solution techniques, were known in the 19th century, but the equation grew greatly in importance when it was realized that it led to solutions ("kink" and "antikink") with the collisional properties of solitons.

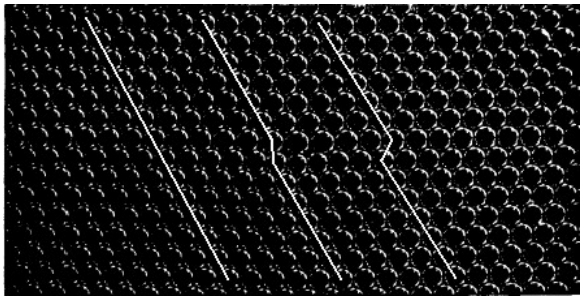


Figure 18: Dislocation in a cristal model. Dislocations are ubiquitous line defects responsible for many properties of crystalline materials. In response to an external load, dislocations move through the crystal producing a permanent change of shape. Generally, materials that yield can avoid a greater danger—catastrophic failure by fracture. Consequently, exactly how dislocations move in response to stress has been a matter of serious concern and intensive study.

An equivalent form for the SGE is

$$u_{\xi\eta} = \sin(u) \qquad \xi, \eta \in \mathbf{R} \tag{7.0.19}$$

which is obtained by variable change $\xi = \frac{1}{2}(x-t)$, $\eta = \frac{1}{2}(x+t)$.

$$u_t = u_\eta \eta_t + u_\xi \xi_t = \frac{1}{2}u_\eta - \frac{1}{2}u_\xi \tag{7.0.20}$$

$$u_x = u_\eta \eta_x + u_\xi \xi_x = \frac{1}{2}u_\eta + \frac{1}{2}u_\xi \tag{7.0.21}$$

$$u_{tt} = \frac{1}{2}(u_{\eta\eta}\eta_t + u_{\xi\eta}\xi_t) - \frac{1}{2}(u_{\xi\xi}\xi_t + u_{\eta\xi}\eta_t) = \frac{1}{4}(u_{\eta\eta} - 2u_{\eta\xi} + u_{\xi\xi}) \tag{7.0.22}$$

$$u_{xx} = \frac{1}{2}(u_{\eta\eta}\eta_x + u_{\xi\eta}\xi_x) + \frac{1}{2}(u_{\xi\xi}\xi_x + u_{\eta\xi}\eta_x) = \frac{1}{4}(u_{\eta\eta} + 2u_{\eta\xi} + u_{\xi\xi}) \tag{7.0.23}$$

Plugging in 7.0.18 gives 7.0.19.

In 1939 Frenkel and Kontorova introduced SGE equation to model the relationship between dislocation dynamics and plastic deformation of a crystal (Frenkel & Kontorova, 1939). In this context $u(x,t)$ describes the atomic displacement in the x-direction and the sin function represents periodicity of the crystal lattice.

Notice that 7.0.18 allows for an infinite number of trivial solutions, namely $u = 0, \pm\pi, \pm 2\pi, \pm 3\pi, \ldots$. Systems with a multitude of such "degenerate ground states" also allow solutions that connect two neighboring ground states. Solutions of this type are often called **kinks**, and for the sine-Gordon equation they are exact solitons; that is, they collide elastically without generation of dispersive radiation.

7.0.18 admits the following two waves as nontrivial solitonic solutions

$$u(x,t) \equiv z(x - x_0 - vt) = 4 \arctan(e^{\pm \frac{x - x_0 - vt}{\sqrt{1-v^2}}}) \tag{7.0.24}$$

with velocity v in the range $(-1, +1)$. As seen for KdV exact solutions of equation 7.0.18 involving arbitrary numbers of dislocation components as 7.1.10 can be generated through a succession of Bäklund transforms, but this was not known to Frenkel and Kontorova.

In 7.1.10 the + sign holds for the so called **Kink**, and the - sign holds for the **Antikink**.

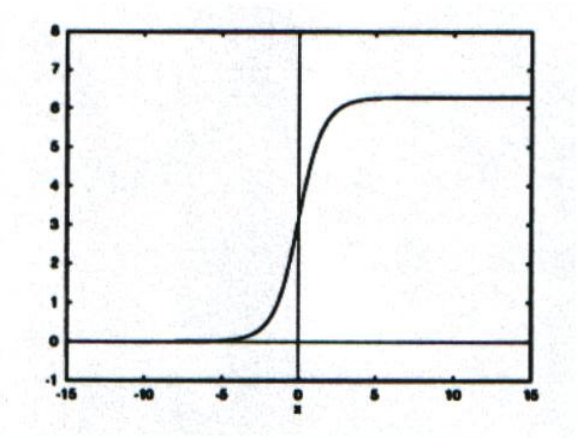


Figure 19: Profile of the Kink for $t = 0$ and $v = 0$.

7.1 Derivation of soliton solutions to the SGE

Soliton solutions of the SGE equation are far richer than those of the KdV and MKdV equations. Even the 1-soliton solution consists of two different cases, "Kink" and "Antikink". A Kink is a solution whose boundary values at the left infinity is 0 and at the right infinity is 2π; the boundary values of an Antikink is 0 and -2π, respectively.

Several classes of solutions can be found by making the assumption that the solution is of of the form

$$u(x,t) = 4 \arctan[\frac{\phi(x)}{\psi(t)}] \tag{7.1.1}$$

This can be physically motivated on the grounds that the identity

$$\arctan(y) = \begin{cases} -\frac{\pi}{2} - \arctan(\frac{1}{y}), & for\ y < 0 \\ \frac{\pi}{2} - \arctan(\frac{1}{y}), & for\ y > 0 \end{cases}$$

means that interchanging space and time variables preserves the solution, as required by the symmetry of the sine-Gordon equation 7.0.18. (Although the reason for the factor of 4 is not entirely clear)

Plugging 7.1.1 into the SGE 7.0.18 then gives

$$\frac{\psi^2}{\phi}\phi_{xx} + \frac{\phi^2}{\psi}\psi_{tt} = (\psi^2 + 2\psi_t^2 - \psi\psi_{tt}) + (-\phi^2 + 2\phi_x^2 - \phi\phi_{xx}) \tag{7.1.2}$$

Since the right-hand side contains two terms, one dependent only on t and one only on x, it can be eliminated by differentiating both side with respect to both x and t. Doing so and dividing the result by $2\phi\phi_x\psi\psi_t$ gives

$$\left(\frac{\phi_{xxx}}{\phi^2\phi_x} - \frac{\phi_{xx}}{\phi^3}\right) + \left(\frac{\psi_{ttt}}{\psi^2\psi_t} - \frac{\psi_{tt}}{\psi^3}\right) = 0 \tag{7.1.3}$$

or

$$\frac{(\phi_{xx}/\phi)_x}{\phi\phi_x} + \frac{(\psi_{tt}/\psi)_t}{\psi\psi_t} = 0 \tag{7.1.4}$$

Since the left term depends on x only and the right term depends on t only, separation of variables can be used to write

$$\frac{(\phi_{xx}/\phi)_x}{\phi\phi_x} = -\frac{(\psi_{tt}/\psi)_t}{\psi\psi_t} = -6k^2 \tag{7.1.5}$$

Rewriting these two equations then gives

$$\frac{d}{dx}\left(\frac{\phi_{xx}}{\phi}\right) = -6k^2\phi\phi_x \qquad \frac{d}{dt}\left(\frac{\psi_{tt}}{\psi}\right) = 6k^2\psi\psi_t \tag{7.1.6}$$

These can be integrated directly to give

$$\frac{\phi_{xx}}{\phi} = -3k^2\phi^2 + A^2 \qquad \frac{\psi_{tt}}{\psi} = 3k^2\psi^2 + B^2 \tag{7.1.7}$$

or

$$\phi_{xx} = -3k^2\phi^3 + A^2\phi \qquad \psi_{tt} = 3k^2\psi^2 + B^2\psi \tag{7.1.8}$$

where A^2 and B^2 are constants of integration. These equations can be solved in general in terms of elliptic integrals, but interesting classes of solution can be investigated by picking particularly simple values of the integration constants. A simple class of solution is obtained by taking $k = 0$, in which case the equations have solutions

$$\phi(x) = a_1 e^{Ax} + a_2 e^{-Ax} \qquad \psi(t) = b_1 e^{Bt} + b_2 e^{-Bt} \tag{7.1.9}$$

where a_1, a_2, b_1, b_2, A, B are complex numbers. A single-soliton solution is obtained choosing $a_1 = b_1 \neq 0$, $a_2 = b_2 = 0$ $A = \pm(1 - v^2)^{-1/2}$ and $B = -Av$, with $v \in (-1, 1)$

$$u(x,t) = 4 \arctan(e^{\pm \frac{x-vt}{\sqrt{1-v^2}}}) \tag{7.1.10}$$

The positive solution is a soliton also known as the "kink solution," while the negative solution is an antisoliton also known as the "antikink solution"

A two-soliton solution exists for $a_1 = a_2 = vb_1 = -vb_2$ and $A = B = v(1 - v^2)^{-1/2}$,

$$u(x,t) = 4 \arctan \left[\frac{v \sinh(\frac{vx}{\sqrt{1-v^2}})}{cosh(\frac{vt}{\sqrt{1-v^2}})} \right] \tag{7.1.11}$$

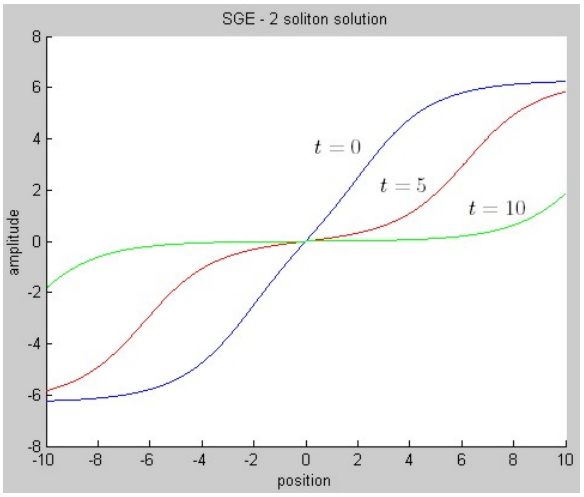


Figure 20: Time evolution of a two-soliton solution to SGE. $v = 0.5$

In a similar way a two-kink solution is given by

$$u(x,t) = 4 \arctan \left[\frac{\frac{1}{\sqrt{1-v^2}} \sinh(\frac{x}{\sqrt{1-v^2}})}{cosh(\frac{vt}{\sqrt{1-v^2}})} \right] \tag{7.1.12}$$

Other solutions exist for the Sine-Gordon equation, they describe the Kink-Kink and Kink-Antikink (**Breather**) interaction processes.

The Breather can be interpreted as a bound state soliton-antisoliton. With $m^2 < 1$ it has the form

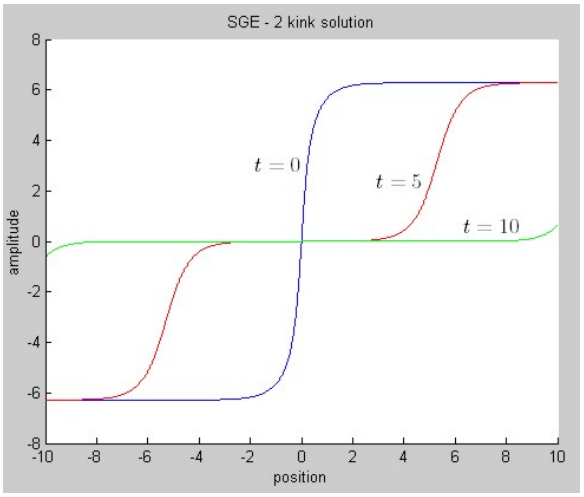


Figure 21: Time evolution of a two-kink solution to SGE. $v = 0.5$

$$u(x,t) = -4\arctan\left[\ \frac{m}{\sqrt{1-m^2}}\frac{\sin(t\sqrt{1-m^2})}{\cosh(mx)}\ \right] \tag{7.1.13}$$

For a fixed x, this is a periodic function of t with frequency $\frac{2\pi}{\sqrt{1-m^2}}$

Other PDE equations with degenerate ground states also have kink and antikink solutions, but they are not exact solitons like those of the SGE.

7.2 Bäcklund transform (BT)

Starting from any solution of the SGE, Bäcklund transformation (BT) creates a two-parameter family of new solutions. One slight complication is that the construction of the new solutions requires solving a certain ODE. However the so-called "Bianchi Permutability Formula" allows us to easily compose Bäcklund transformations. That is, once we have found this first set of new solutions, we can apply another Bäcklund transformation to any one of them to get still more solutions of SGE, and this second family of new solutions can be written down explicitly as algebraic functions of the first set, without solving any more ODEs. Moreover, we can continue inductively in this manner, getting an infinite sequence of families of more and more complex solutions (and related pseudospherical surfaces). If we take as our starting solution the identically zero (or "vacuum") solution to the SGE, at the first stage we get the so-called Kink (or one-soliton) solutions, and the corresponding family of pseudospherical surfaces is the Dini family (including the well-known pseudosphere). Using the Bianchi Formula once gives rise to the two-soliton solutions and the corresponding Küen Surface, and

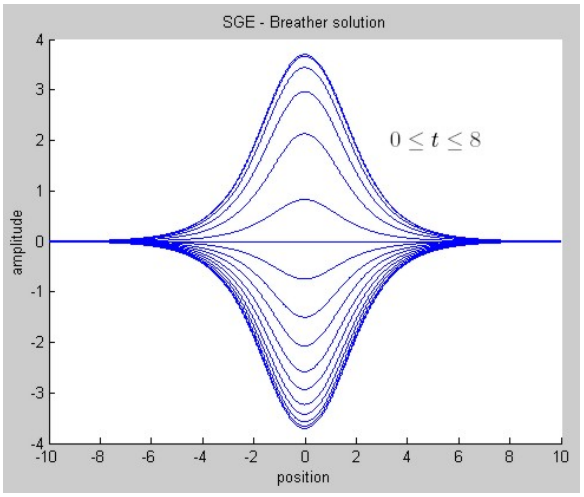


Figure 22: Time evolution of breather. $m = 0.8$

repeated application leads in principle to all the higher soliton solutions. As soon as it was realized that SGE was a soliton equation, it was natural to ask whether KdV also had an analogous theory of Bäcklund transformations. It was quickly discovered that this was in fact so. Only special nonlinear PDEs are found to have BTs.

to do

7.3 La teoria di De-Broglie della doppia soluzione

De Broglie initially regarded particles and his waves as distinct entities. However, in the later part of his career, he came to view particles as singularities in the waves, adopting Schrödinger's basic idea of particles as wave-based phenomena. He observed:

> "*One is therefore led to consider every particle as an extended wave phenomenon involving a very small region of high concentration of the wave field, a kind of singularity region, which would be the particle in the strict sense of the word. It seems probable that the wave field would, in that very small region, obey non-linear wave equations, while throughout the exterior region of the wave field would be almost entirely governed by the linear equations of propagation of the type now classical in wave mechanics.*".

L'idea di fondo può così essere sintetizzata: lo stato di una particella elementare viene rappresentato da due funzioni

$$\varphi(\mathbf{r}, t) = \Phi(\mathbf{r}, t) e^{j\bar{\theta}(\mathbf{r},t)} \tag{7.3.1}$$

$$\psi(\mathbf{r},t) = \Psi(\mathbf{r},t)e^{j\theta(\mathbf{r},t)} \tag{7.3.2}$$

dove φ soddisfa una certa equazione di evoluzione non-lineare e ψ soddisfa l'equazione di Schrödinger.

$$j\hbar\psi_t + \frac{\hbar^2}{2m}\nabla^2\psi - V(\mathbf{r})\psi = 0 \tag{7.3.3}$$

Inoltre De Broglie suppone che la relazione $\theta = \overline{\theta}$ tra le fasi delle due onde sia valida in ogni punto dello spazio tranne che in una piccola regione di esso localizzabile in un'intorno sferico racchiudente la particella. Ovviamente l'equazione non.lineare a cui deve ubbidire φ deve essere tale che nella regione esterna al raggio della particella essa debba coincidere con l'equazione di Schroedinger. De Broglie riteneva che l'onda φ dovesse racchiudere le informazioni concernenti la struttura e le proprietà della particella in esame. Però egli non fu in grado di proporre l'equazione nonlineare a cui doveva ubbidire la φ.Queste idee hanno spinto molti studiosi a cercare delle equazioni nonlineari le quali rispondano alle richieste formulate dalla teoria di De-Broglie.Esempi in tale direzione sono stati proposti sulla base dell'equazione di Sine-Gordon. Possiamo riassumere dicendo che in molti concordano sul riconoscere il Breather, soluzione della S.G., come un esempio dell'onda φ di De-Broglie. Tuttavia i contenuti di tali modelli sono stati elaborati soltanto nel caso di una sola dimensione spaziale e perciò non sono direttamente applicabili. Comunque allo stato attuale tali problemi sono ancora aperti anche se non mancano, da diversi punti di vista, dei tentativi nella direzione di una migliore comprensione di questi aspetti legati alle equazioni nonlineari.

8 Toda's Equation

By the first work of Zabusky and Kruskal, who explained the discrete FPU problem, it was evident that local solutions of both continous PDEs and difference differential equations could exhibit solitonic properties. Another development of the 1960s was Toda's discovery of exact two-soliton interactions on a discrete 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

$$\frac{a}{b}(e^{-bu_j} - 1) + au_j \tag{8.0.4}$$

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 $a \to +\infty$ and $b \to 0$ with ab finite, this reduces to the quadratic potential of a linear spring (elastic constant $\frac{1}{2}ab$). In the limit $a \to 0$ and $b \to +\infty$ with ab finite, it describes the interaction between hard spheres.

9 Nonlinear Schröedinger Equation (NSE)

The Nonlinear Schröedinger Equation (NSE) has an interesting pre-history. It was discovered "in disguise" (and then re-discovered at least three times) in the early part of this century. In 1906, Da Rios wrote a master's thesis [Da Rios, Rend. Circ. Mat. Palermo 22 (1906), 117-135] under the direction of Levi-Civita, in which he modeled the free evolution of a thin vortex-filament in a viscous liquid by a time-dependent curve $\overrightarrow{\gamma}(x,t)$ in $\mathbf{R}^3$ satisfying the equation $\overrightarrow{\gamma}_t = \overrightarrow{\gamma}_x \times \overrightarrow{\gamma}_{xx}$. Now by the Frenet equations, $\overrightarrow{\gamma}_x \times \overrightarrow{\gamma}_{xx} = \kappa \overrightarrow{B}$ where $\kappa(x,t)$ is the curvature and $\overrightarrow{B}$ the binormal, so the filament evolves by moving in the direction of its binormal with a speed equal to its curvature. This is now often called the *vortex-filament equation* (VFE) or the *smoke-ring equation*. In 1971, Hasimoto noticed a remarkable gauge transformation that transforms the VFE to the NSE. In fact, if $\tau(x,t)$ denotes the torsion of the curve $\overrightarrow{\gamma}(x,t)$, then the complex scalar quantity $q(x,t) = \kappa(x,t)\ exp(j \int \tau(\xi,t) d\xi)$ satisfies NLS if and only if $\overrightarrow{\gamma}$ satisfies the VFE.

But it is as an "envelope equation" that NLS has recently come into its own. If a one-dimensional, amplitude modulated, high-frequency wave is moving in a highly dispersive and non-linear medium, then to a good approximation the evolution of the wave envelope (i.e., the modulating signal) in a coordinate system moving at the group velocity of the wave will satisfy NLS. Without going into detail about what these hypotheses mean they do in fact apply to the light pulses travelling along optical fibers that are rapidly becoming the preferred means of communicating information at high bit-rates over long distances. Soliton solutions of NLS seem destined to play a very important role in keeping the Internet and the *World Wide Web* from being ruined by success.

The scalar nonlinear Schrödinger equation (NSE, or NLS) is defined by

$$ju_t + j\gamma u + u_{xx} + k\,|u|^2\,u = 0 \qquad x \in \mathbf{R},\ \ t > 0 \tag{9.0.5}$$

Really, 9.0.5 is a system of 2 equations, since u is complex-valued, in contrast to KdV, SGE and the Toda Lattice.

9.0.5 is a nonlinear generalization of a linear equation $ju_t + u_{xx} + u = 0$, 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 was natural to call equation 9.0.5 the nonlinear Schrödinger equation. When the NSE is used to model wave packets in such fields as hydrodynamics, nonlinear optics, nonlinear acoustics, plasma waves and biomolecular dynamics, however, its derivation and solution has nothing to do with quantum mechanics.

In an optical fiber, for example, the nonlinear term $k\,|u|^2\,u$ is due to the refractive index dependency on the electric field of the lightwave $n = n(\omega, E) = n_0(\omega) + n_2|E|^2$,

so that the dispersion relation is $k(\omega, E) = \frac{\omega}{c} n(\omega, E)$.

Sometimes NSE comes with an extra additive term $\propto j u_x$, but a simple change of variables reduce it to 9.0.5.

The term $j\gamma u$ is the coefficient characterizing dissipation. It is assumed a small perturbation compared to all other terms, and often it is neglected (lossless medium approximation).

Sometimes a system of N coupled NSE equations is given, a system that we refer to as **vector NSE** (or VNSE)

$$j\mathbf{u}_t + \mathbf{u}_{xx} + k \|\mathbf{u}\|^2 \mathbf{u} = 0 \tag{9.0.6}$$

where we omitted dissipation. Both the NSE and the VNSE equations admit integrable discretizations. A natural discretization of 9.0.5 is

$$j\frac{d}{dt}u_n + j\gamma u_n + \frac{1}{\Delta x^2}(u_{n+1} - 2u_n + u_{n-1}) + \frac{k}{2}|u_n|^2 (u_{n+1} - u_{n-1}) = 0 \tag{9.0.7}$$

where the index n ranges over the finite or infinite 1D lattice. 9.0.7 is a $O(\Delta x^2)$ finite-difference approximation that describes a simple model for a lattice of coupled anharmonic oscillators, where k plays the role of a anharmonic parameter. This equation is also integrable and has soliton solutions.

NSE appears in many physical problems,like plasma interactions, and many sorts of quantum feedback loops. This kind of equation is very important in mathematics and physics, and there are many researches on its well-posedness of solution and the existence of solution.

The solitonic solution derived below exist under the condition of the same signs of the parameters of second-order linear dispersion u_{xx} (coefficient $\delta = 1$) and cubic nonlinearity $|u|^2 u$ (coefficient k), when the nonlinear compression of the field is compensated by the dispersion spreading. This constraint implies $k > 0$ and states the so called *Lighthill criterion*. In fact the case $k < 0$ doesn't admit soliton solutions vanishing at infinity, but a class of solitons with nontrivial background intensity (*dark solitons*).

It results that NSE has an infinite number of conserved quantity.

It is interesting to note that the NSE equation is in a way more universal than the KdV equation and the SGE since an almost monochromatic, small-amplitude solution of the KdV equation (or SGE) will evolve according to the NLS equation.

9.1 Exact Solution to the NSE in a lossless medium

At a first approximation, consider a lossless medium ($\gamma = 0$). Let's consider $u = \Phi(x,t)e^{j\theta(x,t)}$ with real θ and Φ. Substituting into 9.0.5 and equating to zero both real

and imaginary parts we arrive to the system

$$\Phi\theta_{xx} + 2\Phi_x\theta_x + \Phi_t = 0 \tag{9.1.1}$$
$$\Phi_{xx} - \Phi\theta_x^2 - \Phi\theta_t + k\Phi^3 = 0 \tag{9.1.2}$$

Let's introduce v, the speed of the envelope function Φ, and w, the speed of the phase function θ. With change of variables $\Phi(x,t) = \Phi(x - vt) \equiv \Phi(\xi)$ and $\theta(x,t) = \theta(x - wt) \equiv \theta(\eta)$ we obtain

$$\Phi\theta_{\eta\eta} + 2\Phi_\xi\theta_\eta - v\Phi_\xi = 0 \tag{9.1.3}$$
$$\Phi_{\xi\xi} - \Phi\theta_\eta^2 + w\Phi\theta_\eta + k\Phi^3 = 0 \tag{9.1.4}$$

Multiplying (9.1.3) by Φ, then using $\frac{\partial}{\partial\xi} = \frac{\partial}{\partial\eta} = \frac{\partial}{\partial x}$, and finally integrating with respect to ξ we obtain

$$\Phi^2(2\theta_\eta - v) = C \tag{9.1.5}$$

Choosing the integration constant $C = 0$ we have $\theta_\eta = \frac{v}{2}$. Substituting into (9.1.4)

$$\Phi_{\xi\xi} - \frac{v^2}{4}\Phi + \frac{wv}{2}\Phi + k\Phi^3 = 0 \tag{9.1.6}$$

multiplying by $2\Phi_\xi$

$$2\Phi_\xi\Phi_{\xi\xi} - \frac{v^2}{2}\Phi\Phi_\xi + wv\Phi\Phi_\xi + 2k\Phi^3\Phi_\xi = 0 \tag{9.1.7}$$

and integrating with respect to ξ, we obtain

$$\Phi_\xi^2 = -\frac{k}{2}\Phi^4 + \frac{1}{4}(v^2 - 2vw)\Phi^2 + D \tag{9.1.8}$$

Separating variables and integrating once again

$$\int \frac{d\Phi}{\sqrt{-\frac{k}{2}\Phi^4 + \frac{1}{4}(v^2 - 2vw)\Phi^2 + D}} = x - vt \tag{9.1.9}$$

$D = 0$ gives the simplest and well known soliton solution:

$$\Phi(x,t) = \sqrt{\frac{2a}{k}}\ sech(\sqrt{a}\ \xi)\ e^{j\frac{v}{2}\eta} \tag{9.1.10}$$

where $a = \frac{1}{4}(v^2 - 2vw)$. This soliton exists only for $v > 2w$. The maximum amplitude of the wave packet is given by $\sqrt{\frac{2a}{k}}$, while the width depends on the ratio

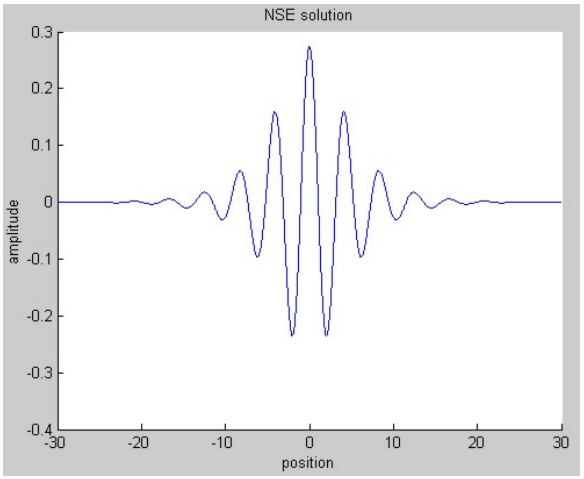


Figure 23: Real part of the NSE envelope soliton solution; $a = 0.075$, $k = 2$, $v = 3$.

$\frac{v}{w} > 2$. As $\frac{v}{w}$ approaches to 2, it spreads out in space. Solitons like 9.1.10 are known as *envelope solitons*.

Any initial excitation for the NLS equation will decompose into solitons and/or dispersive radiation. Propagating envelope soliton is characterized by its phase, which is changing during propagation.

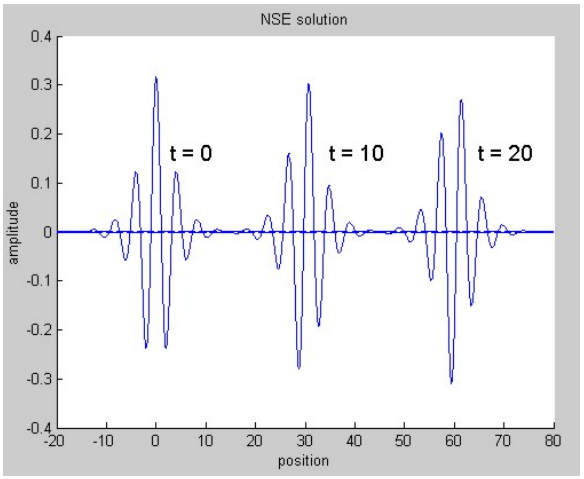


Figure 24: Phase changing along propagation.

Other solitonic waves exist that satisfy NSE. Zakharov and Shabat have shown that this soliton solution is a unique stably localized solution of 9.0.5: any localized pulse in the frame of the 9.0.5 tends to a system of the solitons 9.1.10 having different

amplitudes. Interaction of the solitons in the frame of the 9.0.5 is elastic.

10 The Inverse Scattering Transform (IST)

In 1967, in what would prove to be one of the most cited mathematical papers in history, [C.S.Gardner, J.M.Greene, M.D.Kruskal, and R.M.Miura, *Method for solving the Korteweg de Vries equation*, Phys. Rev Lett. 19 (1967), 1095-1097], Clifford Gardner, John Greene, Martin Kruskal, and Robert Miura (shortly, GGKM) introduced an ingenious method, called the *Inverse Scattering Transform* (IST), for solving the KdV equation.

In the years that followed, the IST changed applied mathematics like no other tool since the Fourier Transform (to which it is closely related), and it soon became clear that it was the key to understanding the remarkable properties of soliton equations.

The remarkable discovery of GGKM was followed by the work of Zakharov and Shabat (1972) which showed that another well-known nonlinear wave equation, the nonlinear Schrödinger equation, was also integrable by an inverse scattering transform. Their demonstration that the integrability of the KdV equation was not an isolated result, was followed closely by analogous results for the modified KdV equation (Wadati, 1972) and the Sine-Gordon equation (Ablowitz et al, 1973). In 1974, Ablowitz, Kaup, Newell and Segur provided a generalisation and unification of these results in the AKNS scheme. From this point there has been an explosive and rapid development of soliton theory in many directions (see, for instance, Ablowitz and Segur 1981, Dodd et al 1982, Newell 1985 and Drazin and Johnson 1989)

In this paragraph, the Korteweg-de Vries equation (KdV) is considered, and it is derived by using the Lax method. Next, we present an elementary outline of the inverse scattering transform (IST) to solve the initial-value problem for the KdV. Soliton solutions to the KdV are also derived.

10.0.1 IST applied to the KdV

In 1967, Gardner, Greene, Kruskal, and Miura presented a method, known as the **Inverse Scattering Transform** (IST), to solve the Cauchy problem (initial-value problem) for the KdV

$$u_t + 6uu_x + u_{xxx} = 0 \qquad x \in \mathbf{R}, \quad t > 0 \tag{10.0.11}$$

$$u(x,t) \equiv u(x,0) \qquad \text{for } t = 0 \tag{10.0.12}$$

assuming that $u(x,0)$ approaches a constant sufficiently rapidly as $x \to \pm\infty$ (there is no loss of generality in choosing that constant as zero). They showed that $u(x,t)$ can be obtained from $u(x,0)$ with the help of the solution to the inverse scattering problem for the linear Schrödinger equation

$$\Psi_{xx}(k,x) = -k^2\Psi(k,x) + V(x)\Psi(k,x) \tag{10.0.13}$$

In appropriate units, it describes the quantum mechanical behavior of a particle of total energy k^2 under the influence of the potential V. For our purpose V have to be replaced by $u(x,t)$, and Ψ have to depend on time

$$\Psi_{xx}(k,x,t) = -k^2\Psi(k,x,t) + u(x,t)\Psi(k,x,t) \tag{10.0.14}$$

This is the so-called *time-evolved Schrödinger equation*, where $t > 0$ is just a parameter that is usually interpreted as time. They also explained that soliton solutions to the KdV corresponded to the case of zero reflection coefficient in the scattering data. They observed from various numerical studies of the KdV that, for large t, $u(x,t)$ in general consists of a finite train of solitons traveling in the positive x direction and an oscillatory train spreading in the opposite direction.

10.0.2 IST applied to the NSE

During the summer of 1972 Newell and his colleagues organized the first soliton research workshop (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 IST for the NSE (Zakharov & Shabat, 1972). Let's re-write this nonlinear PDE,

$$ju_t + u_{xx} + 2\left|u\right|^2 u = 0 \quad x \in \mathbf{R}, \quad t > 0 \tag{10.0.15}$$

The initial-value problem for 10.0.15 can be solved with the help of the solution to the inverse scattering problem for the 1st-order linear system

$$\begin{cases} \xi_x = -j\lambda\xi + V(x)\eta \\ \eta_x = j\lambda\eta - V^*(x)\xi \end{cases} \tag{10.0.16}$$

where V is the potential, λ is the spectral parameter, and "*" denotes complex conjugation. The system 10.0.16 is known as the *Zakharov-Shabat system*.

10.0.3 IST applied to the MKdV equation

Soon afterwards, again in 1972 Wadati showed that the modified KdV equation (MKdV)

$$u_t - 6u^2u_x + u_{xxx} = 0 \quad x \in \mathbf{R}, \quad t > 0 \tag{10.0.17}$$

can be solved with the help of the inverse scattering problem for the linear system

$$\begin{cases} \xi_x = -j\lambda\xi + V(x)\eta \\ \eta_x = j\lambda\eta + V(x)\xi \end{cases} \tag{10.0.18}$$

10.0.4 IST applied to the SGE

Next, in 1973 Ablowitz, Kaup, Newell, and Segur showed that also the SGE

$$u_{xt} = \sin(u) \quad x \in \mathbf{R}, \quad t > 0 \tag{10.0.19}$$

can be solved in the same way by exploiting the inverse scattering problem associated with the linear system

$$\begin{cases} \xi_x = -j\lambda\xi - \frac{1}{2}V_x(x)\eta \\ \eta_x = j\lambda\eta + \frac{1}{2}V_x(x)\xi \end{cases} \tag{10.0.20}$$

Since then, many other nonlinear PDEs have been discovered to be solvable by the method of IST (in 1974 IST formulations had been constructed also for the Toda's equation).

The IST can be viewed as a nonlinear extension of Fourier transforms, and numerous physically significant nonlinear equations have been solved by generalizations of this technique which include the following: PDEs in one spatial and one temporal dimensions, differential-difference equations, ODEs, ordinary difference equations, singular integro-differential equations, PDEs in two spatial and one temporal dimensions.

An evolution equation, e.g. a partial differential equation, that is solvable by some form of the IST is often referred to as being **completely integrable**. Such integrable evolution equations have been shown to usually possess several remarkable properties including:

- the 'elastic' interaction of solitary waves, i.e. multi-soliton solutions
- Bäcklund transformations
- an infinite number of independent conservation laws
- a complete set of action-angle variables
- an underlying Hamiltonian formulation
- a Lax representation
- a bilinear representation and soliton solutions via a finite series of exponential solutions

- an associated linear eigenvalue problem whose eigenvalues are constants of the motion

However, the precise interrelationships between all these (and other) properties has yet to be rigorously formulated. Then, important questions arise:

1. What is it that really characterises completely integrable equations?
2. Can we characterize the set of nonlinear PDEs solvable by an IST? In other words, can we find a set of necessary and sufficient conditions that guarantee an initial-value problem for a nonlinear PDE to be solvable via an IST related to a linear problem? (this is the so-called *characterization problem*)
3. given a nonlinear PDE that is known to be solvable by an IST, can we determine the corresponding linear problem?

There does not yet seem to be a completely satisfactory answers to these open questions.

10.1 The Lax Method

In 1968 Peter Lax presented a method to show that the KdV can be derived as a *compatibility condition* related to the time evolution of solutions to 10.2.1. A similar result holds for other nonlinear PDEs solvable by the IST method.

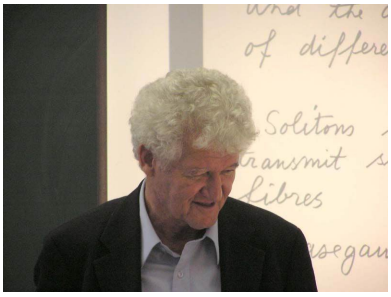


Figure 25: Peter D.Lax

We first outline the general idea behind the Lax method and next demonstrate its application on the KdV.

The basic idea behind the Lax method is the following. Given a spectral problem $\mathcal{L}\Psi = \lambda\Psi$ find a linear differential operator $\mathcal{A}$ (the operators $\mathcal{L}$ and $\mathcal{A}$ are said to form a **Lax pair**) such that:

1. the spectral parameter λ does not change in time
2. the quantity $\Psi_t - \mathcal{A}\Psi$ remains a solution to $\mathcal{L}\Psi = \lambda\Psi$
3. the quantity $\mathcal{L}_t + \mathcal{L}\mathcal{A} - \mathcal{A}\mathcal{L}$ is a multiplication operator, i.e. it is not a differential operator

It can be easily shown that these assumptions imply the so-called **Lax relation**

$$\mathcal{L}_t + \mathcal{L}\mathcal{A} - \mathcal{A}\mathcal{L} = 0 \tag{10.1.1}$$

which is interpreted as an integrable PDE and in general is a nonlinear evolution equation containing a 1st-order time derivative. From this implication we can derive KdV and other nonlinear PDEs solvable by the IST by appropriate choice of $\mathcal{L}$, assuming that $\mathcal{A}$ has the form

$$\alpha_3\partial_x^3 + \alpha_2\partial_x^2 + \alpha_1\partial_x + \alpha_0 \tag{10.1.2}$$

where $\partial_x \equiv \frac{\partial}{\partial x}$ and α_j may depend on x and t, but not on the spectral parameter λ.

There are other methods to derive such nonlinear PDEs. One such method (**AKNS** method) is due to Ablowitz, Kaup, Newell, and Segur, who first used it to derive the SGE.

10.1.1 The Lax Method applied to the KdV equation

Let us illustrate the Lax method to derive the KdV from the 1-D Schrödinger equation. For this purpose, we re-write 10.2.1 in the form $\mathcal{L}\Psi = \lambda\Psi$ with $\lambda = k^2$, where

$$\mathcal{L} = -\partial_x^2 + u(x,t) \tag{10.1.3}$$

Replacing 10.1.2 and 10.1.3 into the necessary implication 10.1.1, we can find an equation in the form

$$()\partial_x^5 + ()\partial_x^4 + ()\partial_x^3 + ()\partial_x^2 + ()\partial_x + () = 0 \tag{10.1.4}$$

The coefficient of ∂_x^5 automatically vanishes. Setting to zero the four subsequent coefficients, we obtain

$$\mathcal{A} = c_1\partial_x^3 + c_2\partial_x^2 + (c_3 - \frac{3}{2}c_1 u)\partial_x + c_4 - \frac{3}{4}c_1 u_x - c_2 u \tag{10.1.5}$$

with c_1, c_2, c_3, c_4 denoting arbitrary constants. Choosing $c_1 = -4$ and $c_3 = 0$ in the last coefficient in 10.1.4 and setting that coefficient to zero, we get the KdV equation

$$u_t - 6uu_x + u_{xxx} = 0 \qquad x \in \mathbf{R}, \quad t > 0 \tag{10.1.6}$$
$$\tag{10.1.7}$$

Moreover, by letting $c_2 = c_4 = 0$, we obtain the operator $\mathcal{A}$ as

$$\mathcal{A} = -4\partial_x^3 + 6u\partial_x + 3u_x \tag{10.1.8}$$

10.1.2 The Lax Method applied to the NSE

As for the Zakharov-Shabat system 10.0.16 we can write its time-evolved version as $\mathcal{L}\Psi = \lambda\Psi$, where the linear operator L is given by

$$\mathcal{L} = j\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\partial_x - j\begin{bmatrix} 0 & u \\ u^* & 0 \end{bmatrix}$$

and the operator $\mathcal{A}$ is obtained as

$$\mathcal{A} = 2j \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \partial_x^2 - 2j \begin{bmatrix} 0 & u \\ u^* & 0 \end{bmatrix} \partial_x - j \begin{bmatrix} -|u|^2 & u_x \\ u_x^* & |u|^2 \end{bmatrix}$$

It can be verified that the compatibility condition 10.1.1 is equivalent to the nonlinear Schrödinger equation 10.0.15

Peter David Lax (born May 1, 1926) is a highly-respected mathematician working in the areas of pure and applied mathematics. He has made important contributions to integrable systems, fluid dynamics and shock waves, solitonic physics, hyperbolic conservation laws, and mathematical and scientific computing, among other fields. Lax was born in Budapest, Hungary and moved with his parents to the United States in 1941. Lax holds a faculty position in the Department of Mathematics, Courant Institute of Mathematical Sciences, New York University. He is a member of the National Academy of Sciences, USA. He was awarded the National Medal of Science in 1986, the Wolf Prize in 1987 and the Abel Prize in 2005. He is an alumnus of New York University, where he received both his bachelor's degree in 1947 and his PhD in 1949 with thesis advisor Richard Courant.

10.2 The Schrödinger Equation and the Scattering Data

In this paragraph we present an elementary review of the IST for the KdV. We consider the time-evolved Schrödinger equation

$$\Psi_{xx}(k,x,t) = -k^2\Psi(k,x,t) + u(x,t)\Psi(k,x,t) \tag{10.2.1}$$

Here $u(x,t)$ plays the role of a time-varying potential. The initial value for $\Psi(k,x,t)$ is $\Psi(k,x,0) \equiv \Psi(k,x)$ obtained as solution to the time-independent equation

$$\Psi_{xx}(k,x) = -k^2\Psi(k,x) + u(x,0)\Psi(k,x) \tag{10.2.2}$$

Thus, we view $V(x)$ as the initial value $u(x,0)$ of the time-varying potential $u(x,t)$.

The scattering coefficients $T(k,t)$, $R(k,t)$, and $L(k,t)$ associated with 10.2.1 are viewed as evolving from the corresponding coefficients $T(k)$, $R(k)$ and $L(k)$ of 10.2.2, respectively, from $t = 0$; Thus, our notation is such that

$$V(x) = u(x,0) \tag{10.2.3}$$

$$\Psi(k,x) = \Psi(k,x,0) \tag{10.2.4}$$

$$T(k) = T(k,0); R(k) = R(k,0); L(k) = L(k,0) \tag{10.2.5}$$

10.3 The direct scattering problem

Now, go back to the linear Schrödinger equation

$$\Psi_{xx}(k,x) = -k^2\Psi(k,x) + V(x)\Psi(k,x) \qquad x, k^2 \in \mathbf{R} \tag{10.3.1}$$

The determination of the scattering matrix

$$\mathbf{S}(k) = \begin{bmatrix} T(k) & R(k) \\ L(k) & T(k) \end{bmatrix}$$

from the potential $V(x)$ is known as the **direct scattering problem**. Solving the **inverse scattering problem** for 10.3.1 consists of the determination of $V(x)$ from an appropriate set of scattering data. In the 10.3.1 the potential is assumed to belong to the Faddeev class, i.e. $V(x)$ is real valued, measurable, and

$$\int_{-\infty}^{+\infty} (1+|x|)|V(x)|dx < +\infty$$

There are two types of solutions to 10.3.1, namely, scattering solutions and bound-state solutions. The scattering solutions are those that consist of linear combinations of e^{jkx} and e^{-jkx} as $x \to \pm\infty$; they occur for $k \in \mathbf{R} \setminus \{0\}$.

Real k values correspond to positive energies of the particle (k^2), and a particle of positive energy can be visualized as capable of escaping to $\pm\infty$ as a result of scattering by V. Heuristically, since $V(x)$ vanishes at $\pm\infty$, the particle will still have some kinetic energy at infinity and hence is allowed to be at infinity.

Among the scattering solutions to 10.3.1 are the *Jost solution from the left*, f_l, and the *Jost solution from the right*, f_r, satisfying the respective boundary asymptotic conditions

$$e^{-jkx}f_l(k,x) = 1 + o(1); \quad e^{-jkx}\frac{df_l}{dx} = jk + o(1) \quad \text{for } x \to +\infty \tag{10.3.2}$$

$$e^{jkx}f_r(k,x) = 1 + o(1); \quad e^{jkx}\frac{df_r}{dx} = -jk + o(1) \quad \text{for } x \to -\infty \tag{10.3.3}$$

$$f_l(k,x) = \frac{e^{jkx}}{T(k)} + \frac{L(k)e^{-jkx}}{T(k)} + o(1) \quad \text{for } x \to -\infty \tag{10.3.4}$$

$$f_r(k,x) = \frac{e^{-jkx}}{T(k)} + \frac{R(k)e^{jkx}}{T(k)} + o(1) \quad \text{for } x \to +\infty \tag{10.3.5}$$

On the other hand, a bound state of 10.3.1 is a solution that belongs to $L^2(\mathbf{R})$ in the x variable. Each bound state corresponds to a negative total energy of the particle ($k^2 < 0$), and as a result the particle is bound by the potential and does not have sufficient kinetic energy to escape to infinity. For V belonging to the Faddeev class,

these solutions decay exponentially as $x \to \pm\infty$, occur only at certain distinct k-values on the imaginary axis, are finite in number (N) and are linearly independent solution to 10.3.1. Thus, let's suppose that the bound states occur at $k = j\kappa_i$ $(i = 1 \dots N)$ with the ordering $0 < \kappa_1 < \kappa_2 < \cdots < \kappa_N$.

It can be shown that the normalized bound-state solution $\varphi_i(x)$ at $k = j\kappa_i$ is defined as

$$\varphi_i(x) \equiv f_l(j\kappa_i, x)\big(\int_{-\infty}^{+\infty} f_l(j\kappa_i, x)^2 dx\big)^{-\frac{1}{2}}, \qquad i = 1 \dots N$$

To obtain the scattering coefficients, namely, the **transmission coefficient** T, the **reflection coefficient from the left** L and the **reflection coefficient from the right** R, is it possible to use the formulas

$$T(k) = \frac{2jk}{f_r(k,x)f_l'(k,x) - f_r'(k,x)f_l(k,x)} \tag{10.3.6}$$

$$L(k) = \frac{f_l(k,x)f_r'(-k,x) - f_l'(k,x)f_r(-k,x)}{f_r(k,x)f_l'(k,x) - f_r'(k,x)f_l(k,x)} \tag{10.3.7}$$

$$R(k) = \frac{f_l(-k,x)f_r'(k,x) - f_l'(-k,x)f_r(k,x)}{f_r(k,x)f_l'(k,x) - f_r'(k,x)f_l(k,x)} \tag{10.3.8}$$

where we used the notation $f' = \frac{df}{dx}$. While L and R exists only for $k \in \mathbf{R} \setminus \{0\}$, T has a meromorphic extension in the upper-half complex plane $\{\Im(k) > 0\}$. For this extension a pole exists for each bound state and vice versa. In other words, at $k = j\kappa_i$ $(i = 1 \dots N)$ the meromorphic extension of $T(k)$ has a pole.

The scattering matrix $\mathbf{S}$ can be constructed in terms of the bound-state energies and either one of the reflection coefficients R and L: For example, given $R(k)$ for $k \in \mathbf{R}$ and the bound-state poles, one can construct T and L as

$$T(k) = \big(\prod_{i=1}^{N} \frac{k + j\kappa_i}{k - j\kappa_i}\big) \exp\big(\frac{1}{2\pi j}\int_{-\infty}^{+\infty} \frac{log(1 - |R(s)|^2)}{s - k - j0^+}\, ds\big), \quad k \in \{\Im(k) \geq 0\} \tag{10.3.9}$$

$$L(k) = -\frac{R(k)^* T(k)}{T(k)^*}, \qquad k \in \mathbf{R} \tag{10.3.10}$$

Similarly, given $L(k)$ for $k \in \mathbf{R}$ and the bound-state poles, one can construct T and R as

$$T(k) = \big(\prod_{i=1}^{N} \frac{k + j\kappa_i}{k - j\kappa_i}\big) \exp\big(\frac{1}{2\pi j} \int_{-\infty}^{+\infty} \frac{log(1 - |L(s)|^2)}{s - k - j0^+}\ ds\big), \quad k \in \{\Im(k) \geq 0\} \tag{10.3.11}$$

$$R(k) = -\frac{L(k)^* T(k)}{T(k)^*}, \qquad k \in \mathbf{R} \tag{10.3.12}$$

These equations solve the direct scattering problem for the 10.3.1.

10.4 The inverse scattering problem

The inverse scattering problem for 10.3.1 consists of the determination of the potential V from an appropriate set of scattering data. If there are no bound states, then either one of the reflection coefficients R and L uniquely determines the corresponding potential in the Faddeev class. However, when there are bound states, for the unique determination of V, in addition to a reflection coefficient and the bound-state energies $\{\kappa_i\}_{i=1}^N$, one also needs to specify for each bound-state a *norming constant* $\{c_{ri}\}_{i=1}^N$ (or $\{c_{li}\}_{i=1}^N$), or equivalently a *dependency constant* $\{\gamma_i\}_{i=1}^N$. The left and right bound-state norming constants are defined as

$$c_{li} = \big[\int_{-\infty}^{+\infty} f_l(j\kappa_i, x)^2 dx\big]^{-\frac{1}{2}} \qquad c_{ri} = \big[\int_{-\infty}^{+\infty} f_r(j\kappa_i, x)^2 dx\big]^{-\frac{1}{2}} \tag{10.4.1}$$

while the the dependency constants are defined as

$$\gamma_i = \frac{f_l(j\kappa_i, x)}{f_r(j\kappa_i, x)}, \qquad i = 1 \dots N \tag{10.4.2}$$

These coefficients are related to each other via the residues of T as

$$Res(T, j\kappa_i) := \lim_{k \to j\kappa_i} T(k)(k - j\kappa_i) = j c_{li}^2 \gamma_i = j\frac{c_{ri}^2}{\gamma_i}, \qquad i = 1 \dots N \tag{10.4.3}$$

Since the sign of γ_i is the same as that of $(-1)^{N-i}$, and $c_{li}, c_{ri} > 0$, we have

$$c_{ri} = (-1)^{N-i} \gamma_i \; c_{li}, \qquad i = 1 \dots N \tag{10.4.4}$$

Hence, the normalized bound-state solutions can be expressed with

$$\varphi_i(x) := c_{li} f_l(j\kappa_i, x) = (-1)^{N-i} c_{ri} f_r(j\kappa_i, x), \qquad i = 1 \dots N$$

To recover the potential V uniquely, we can use various sets of scattering data such as the left scattering data $\{R(k); \{\kappa_i\}_{i=1}^N; \{c_{li}\}_{i=1}^N\}$; the right scattering data $\{L(k); \{\kappa_i\}_{i=1}^N; \{c_{ri}\}_{i=1}^N\}$; or $\{\mathbf{S}(k); \{\gamma_i\}_{i=1}^N\}$; where $\mathbf{S}$ is the scattering matrix.

There are various methods to recover V from an appropriate set of scattering data. One method consists in the use of the *Faddeev-Marchenko equation*. For example, suppose to know the left scattering data, the first step is to calculate

$$M_l(y) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} R(k) e^{jky} dk \; + \; \sum_{i=1}^{N} c_{li}^2 e^{-\kappa_i y} \tag{10.4.5}$$

then solve the so-called **left Faddeev-Marchenko equation**

$$K_l(x,y) + M_l(2x+y) + \int_0^{+\infty} M_l(2x+y+z)K_l(x,z)dz = 0, \qquad y > 0 \qquad (10.4.6)$$

an finally construct the potential as

$$V(x) = -2\frac{dK_l(x,0^+)}{dx} \qquad (10.4.7)$$

Similarly, using the right scattering data as the input one obtains the **right Marchenko integral equation**

$$K_r(x,y) + M_r(-2x+y) + \int_0^{+\infty} M_r(-2x+y+z)K_r(x,z)dz = 0, \qquad y > 0 \quad (10.4.8)$$

where

$$M_r(y) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} L(k)e^{jky}dk \; + \; \sum_{i=1}^{N} c_{ri}^2 e^{-\kappa_i y} \qquad (10.4.9)$$

Following this way one can finally construct the potential as

$$V(x) = 2\frac{dK_r(x,0^+)}{dx} \qquad (10.4.10)$$

10.5 Time evolution of the scattering data

We are interested in determining the time evolution of $\Psi(k,x,t)$ appearing in the time-evolved Schrödinger equation 10.2.1 from its initial state $\Psi(k,x)$ appearing in 10.2.2. The scattering data associated with 10.2.1 are envisioned as evolving from the corresponding quantities associated with 10.2.2, from $t=0$, in such a way that

$$u(x,0)=V(x), \quad \Psi(k,x,0)=\Psi(k,x), \tag{10.5.1}$$

$$T(k,0)=T(k), \quad R(k,0)=R(k), \quad L(k,0)=L(k), \tag{10.5.2}$$

$$c_{li}(0)=c_{li}, \quad c_{ri}(0)=c_{ri}, \quad \gamma_i(0)=\gamma_i \tag{10.5.3}$$

Notice that the spectral parameter k and the bound-state energies κ_i do not change in time. As the initial potential $u(x,0)$ evolves to $u(x,t)$, condition 2 of the Lax method described above allows us to determine the time evolution of any solution to the SchrÄodinger equation 10.2.1. For example, let us find the time evolution of $f_l(x,kt)$, the Jost solution from the left. From condition 2 of the Lax method $\partial_t f_l - \mathcal{A} f_l$ remains a solution to 10.2.1 and hence we can write it as a linear combination of the two linearly independent Jost solutions as

$$\partial_t f_l(k,x,t) - \mathcal{A} f_l(k,x,t) = p(k,t) f_l(k,x,t) + q(k,t) f_r(k,x,t) \tag{10.5.4}$$

where the operator $\mathcal{A}$ is given by 10.1.8, and the coefficients $p(k,t)$ and $q(k,t)$ are yet to be determined. The asymptotic conditions for the time-evolved Jost solutions are analogous to those valid for the time independent ones, i.e.

$$f_l(k,x,t)=e^{jkx}[1+o(1)]; \quad \frac{df_l}{dx}=e^{jkx}[jk+o(1)] \quad \text{for } x\to+\infty \tag{10.5.5}$$

$$f_r(k,x,t)=e^{-jkx}[1+o(1)]; \quad \frac{df_r}{dx}=-e^{-jkx}[jk+o(1)] \quad \text{for } x\to-\infty \tag{10.5.6}$$

$$f_l(k,x,t)=\frac{e^{jkx}}{T(k,t)}+\frac{L(k,t)e^{-jkx}}{T(k,t)}+o(1) \quad \text{for } x\to-\infty \tag{10.5.7}$$

$$f_r(k,x,t)=\frac{e^{-jkx}}{T(k,t)}+\frac{R(k,t)e^{jkx}}{T(k,t)}+o(1) \quad \text{for } x\to+\infty \tag{10.5.8}$$

For each fixed t, we can use in equation 10.5.4 the asymptotic conditions valid for $x\to+\infty$. More, we can assume $u(x,t)=o(1)$, $u_x(x,t)=o(1)$ as $x\to+\infty$, obtaining

$$p(k,t)=-4jk^3; \qquad q(k,t)=0$$

Thus, $f_l(k,x,t)$ evolves in time by obeying the linear 3rd-order PDE

$$\partial_t f_l(k,x,t) - \mathcal{A} f_l(k,x,t) = -4jk^3 f_l(k,x,t) \quad x \in \mathbf{R}, \quad t > 0 \tag{10.5.9}$$

Proceeding in a similar manner, we find that $f_r(k,x,t)$ evolves in time according to

$$\partial_t f_r(k,x,t) - \mathcal{A} f_r(k,x,t) = 4jk^3 f_r(k,x,t) \quad x \in \mathbf{R}, \quad t > 0 \tag{10.5.10}$$

with $\mathcal{A} = -4\partial_x^3 + 6u\partial_x + 3u_x$. Notice that the time evolution of each Jost solution is fairly complicated. We will see, however, that the time evolution of the scattering data is very simple. Letting $x \to -\infty$ in 10.5.9, using asymptotic conditions valid for $x \to -\infty$ and assuming $u(x,t) = o(1)$, $u_x(x,t) = o(1)$ as $x \to -\infty$, by comparing the coefficients of e^{jkx} and e^{-jkx} on both sides, we get

$$\partial_t T(k,t) = 0; \quad \partial_t L(k,t) = -8jk^3 L(k,t)$$

which give us

$$T(k,t) = T(k,0) = T(k); \quad L(k,t) = L(k,0)e^{-8jk^3 t} = L(k)e^{-8jk^3 t} \tag{10.5.11}$$

In a similar way, from 10.5.10 as $x \to +\infty$ we obtain

$$R(k,t) = R(k,0)e^{8jk^3 t} = R(k)e^{8jk^3 t} \tag{10.5.12}$$

Thus, the transmission coefficient remains unchanged and only the phase of each reflection coefficient changes as time progresses.

Let us also evaluate the time evolution of the dependency constants $\gamma_i(t)$, which are defined analogous to 10.4.2 as

$$\gamma_i(t) = \frac{f_l(j\kappa_i,x,t)}{f_r(j\kappa_i,x,t)}, \qquad i = 1 \ldots N \tag{10.5.13}$$

Evaluating 10.5.9 at $k = j\kappa_i$ and replacing $f_l(j\kappa_i,x,t) = \gamma_i(t) f_r(j\kappa_i,x,t)$ we get

$$f_r(j\kappa_i,x,t)\partial_t\gamma_i(t) \; + \; \gamma_i(t)\partial_t f_r(j\kappa_i,x,t) \; - \; \gamma_i(t)\mathcal{A} f_r(j\kappa_i,x,t) = -4j\kappa_i^3 f_r(j\kappa_i,x,t) \tag{10.5.14}$$

On the other hand, evaluating 10.5.10 at $k = j\kappa_i$, we obtain

$$\gamma_i(t)\partial_t f_r(j\kappa_i,x,t) \; - \; \gamma_i(t)\mathcal{A} f_r(j\kappa_i,x,t) = 4j\kappa_i^3 f_r(j\kappa_i,x,t) \tag{10.5.15}$$

Comparing 10.5.14 and 10.5.15 we see that $\partial_t\gamma_i(t) = -8j\kappa_i^3\gamma_i(t)$, or equivalently

$$\gamma_i(t) = \gamma_i(0)e^{-8j\kappa_i^3 t} = \gamma_i e^{-8j\kappa_i^3 t} \tag{10.5.16}$$

In the light of the first identity in 10.5.11, from the time-evolved version of 10.4.3, we get

$$c_{li}(t) = c_{li}e^{4j\kappa_i^3 t} \qquad c_{ri}(t) = c_{ri}e^{-4j\kappa_i^3 t} \tag{10.5.17}$$

10.6 Solution to KdV and Solitons

Let us use $\mathcal{D}(t)$ to denote the scattering data for the time-evolved Schrödinger equation 10.2.1 with the time-evolved potential $u(x,t)$

$$\mathcal{D}(t) \equiv \{R(k,t); L(k,t); T(k,t); \{c_{lj}(t)\}; \{c_{rj}(t)\}; \{\gamma_j(t)\}\}$$

From the previous section we have

$$\begin{cases} R(k,t) = R(k)e^{8jk^3t}; \quad L(k,t) = L(k)e^{-8jk^3t}; \quad T(k,t) = T(k) \\ c_{li}(t) = c_{li}e^{4j\kappa_i^3 t}; \quad c_{ri}(t) = c_{ri}e^{-4j\kappa_i^3 t}; \quad \gamma_i(t) = \gamma_i e^{-8j\kappa_i^3 t} \end{cases}$$

Note that the time-evolution of the scattering data is really simple. On the other hand, the time evolution of the potential $u(x,0) \mapsto u(x,t)$ will be much more complicated.

Note that $\mathcal{D}(0)$ corresponds to the initial scattering data associated with the potential $u(x,0)$. The initial-value problem for the KdV consists of finding $u(x,t)$ when $u(x,0)$ is known. Its solution is obtained in three steps as indicated in the following diagram.

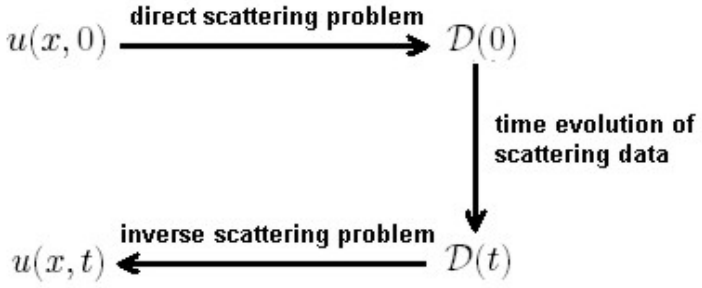


Figure 26: The inverse scattering transform

1. From $u(x,0)$ obtain the corresponding scattering data $\mathcal{D}(0)$. The direct scattering problem $V(x) \mapsto \mathcal{D}(0)$ is equivalent to solving 10.3.1 and obtaining a Jost solutions $f_l(k,x)$ and $f_r(k,x)$, from which $\mathcal{D}(0)$ can be constructed.

2. Let the scattering data evolve in time $\mathcal{D}(0) \mapsto \mathcal{D}(t)$

3. With $\mathcal{D}(t)$ as input, solve the inverse scattering problem $\mathcal{D}(t) \mapsto u(x,t)$ for 10.2.1. This problem is known to be uniquely solvable when the initial potential $V(x) \equiv u(x,t)$ belongs to the Faddeev class. this is the most difficult step, but for the KdV equation can be reduced to solution of a linear integral equation.

Any method applicable to solve the inverse scattering problem $\mathcal{D}(0) \mapsto u(x,0)$ can also be used to solve the inverse scattering problem $\mathcal{D}(t) \mapsto u(x,t)$. For example, one can solve the time-evolved left Faddeev-Marchenko equation: the first step is to calculate

$$M_l(y,t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} R(k)e^{8jk^3t+jky}dk \; + \; \sum_{i=1}^{N} c_{li}^2 e^{8\kappa_i^3 t-\kappa_i y} \tag{10.6.1}$$

then solve

$$K_l(x,y,t) + M_l(2x+y,t) + \int_0^{+\infty} M_l(2x+y+z,t)K_l(x,z,t)dz = 0, \qquad y>0 \tag{10.6.2}$$

an finally and recover $u(x,t)$ by using

$$u(x,t) = -2\frac{\partial K_l(x,0^+,t)}{\partial x} \tag{10.6.3}$$

Equivalently, one can solve the time-evolved right Faddeev-Marchenko equation:

$$M_r(y,t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} L(k)e^{-8jk^3t+jky}dk \; + \; \sum_{i=1}^{N} c_{ri}^2 e^{-8\kappa_i^3 t-\kappa_i y} \tag{10.6.4}$$

$$K_r(x,y,t) + M_r(-2x+y,t) + \int_0^{+\infty} M_r(-2x+y+z,t)K_r(x,z,t)dz = 0, \qquad y>0 \tag{10.6.5}$$

$$u(x,t) = 2\frac{\partial K_r(x,0^+,t)}{\partial x} \tag{10.6.6}$$

Let us now turn our attention to solitons. **The soliton solutions to the KdV are obtained when the reflection coefficients are zero** ($L(k)\equiv 0, R(k)\equiv 0$). It is possible to obtain the solutions to the Marchenko equations in an algebraic manner, without really solving the integral equations themselves. When $R\equiv 0$, from 10.3.9,

$$T(k) = \prod_{i=1}^{N}\frac{k+j\kappa_i}{k-j\kappa_i}, \quad k\in\{\Im(k)>0\} \tag{10.6.7}$$

In that case, the left scattering data consists of the sets of bound state energies $\{-\kappa_i^2\}_{i=1}^N$ and norming constants $\{c_{li}\}_{i=1}^N$. Corresponding to this data we obtain the N-soliton solution to the KdV that can be found, e.g., with the help of 10.6.1-10.6.3.

$$u(x,t) = -2\frac{\partial}{\partial x}\left(\frac{det\Lambda}{det\Gamma}\right) \tag{10.6.8}$$

where $det\Lambda$ $det\Gamma$ denote the determinants of the respective matrices

$$\Lambda \equiv \begin{bmatrix} 0 & \varepsilon_1 & \varepsilon_2 & \dots & \varepsilon_N \\ 1 & 1+\frac{\varepsilon_1}{\kappa_1+\kappa_1} & \frac{\varepsilon_2}{\kappa_1+\kappa_2} & \dots & \frac{\varepsilon_N}{\kappa_1+\kappa_N} \\ 1 & \frac{\varepsilon_1}{\kappa_2+\kappa_1} & 1+\frac{\varepsilon_2}{\kappa_2+\kappa_2} & \dots & \frac{\varepsilon_N}{\kappa_2+\kappa_N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \frac{\varepsilon_1}{\kappa_N+\kappa_1} & \frac{\varepsilon_2}{\kappa_N+\kappa_2} & \dots & 1+\frac{\varepsilon_N}{\kappa_N+\kappa_N} \end{bmatrix}$$

$$\Gamma \equiv \begin{bmatrix} 1+\frac{\varepsilon_1}{\kappa_1+\kappa_1} & \frac{\varepsilon_2}{\kappa_1+\kappa_2} & \dots & \frac{\varepsilon_N}{\kappa_1+\kappa_N} \\ \frac{\varepsilon_1}{\kappa_2+\kappa_1} & 1+\frac{\varepsilon_2}{\kappa_2+\kappa_2} & \dots & \frac{\varepsilon_N}{\kappa_2+\kappa_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\varepsilon_1}{\kappa_N+\kappa_1} & \frac{\varepsilon_2}{\kappa_N+\kappa_2} & \dots & 1+\frac{\varepsilon_N}{\kappa_N+\kappa_N} \end{bmatrix}$$

where we have defined

$$\varepsilon_i = c_{li}^2 e^{-2\kappa_i x + 8\kappa_i^3 t}, \quad i = 1, \dots, N \tag{10.6.9}$$

It can be shown that $det\Lambda = \partial_x det\Gamma$, so that

$$u(x,t) = -2\frac{\partial}{\partial x}\Big(\frac{1}{det\Gamma}\frac{\partial(det\Gamma)}{\partial x}\Big) \tag{10.6.10}$$

For example, when $N = 1$, we get the single soliton solution

$$u(x,t) = -2\kappa_1^2 sech^2\big[\kappa_1 x - 4\kappa_1^3 t + \sqrt{ln\ \gamma_1}\big] \tag{10.6.11}$$

In this case, we have $\gamma_1 = \frac{2\kappa_1}{c_{l1}^2}$. This wave moves in the positive x direction with speed $v \equiv 4\kappa_1^2$ and the dependency constant γ_1 plays a role in the initial location of the soliton, x_0. The width of the soliton is inversely proportional to κ_1; this can be seen by using one of the facts that for one-soliton solution we have

$$\int_{-\infty}^{+\infty} u(x,t)dx = -4\kappa_1; \qquad \int_{-\infty}^{+\infty} u(x,t)^2\ dx = \frac{16\kappa_1^3}{3}$$

These relations are the first two local conservation laws for the KdV. In the case of the N-soliton solution $u(x,t)$ it results

$$\int_{-\infty}^{+\infty} u(x,t)dx = -4\sum_{i=1}^{N}\kappa_i; \qquad \int_{-\infty}^{+\infty} u(x,t)^2\ dx = \frac{16}{3}\sum_{i=1}^{N}\kappa_i^3$$

which, as noted by P. C. Sabatier, also follow from the fact that as $t \to +\infty$ the N-soliton solution to the KdV consists of a train of N separate solitons each traveling at speed its own speed $v_i \equiv 4\kappa_i^2$.

Remark 2. Let us note that a potential in the Faddeev class need not even be continuous. On the other hand, from KdV equation we see that classical solutions are three times differentiable with respect to x. Informally speaking, the discontinuities that may be present in the initial value $u(x,0)$ disappear and $u(x,t)$ becomes smoother for $t > 0$.

If the reflection coefficients are different from zero, as $t \to +\infty$ the solution $u(x,t)$ evolves into N rank-ordered solitons propagating to the right (i.e. $x > 0$), and some decaying radiation propagating to the left ($x < 0$)

$$u(x,t) \sim -2\sum_{i=1}^{N} \kappa_i^2 sech^2\left[\kappa_i x - 4\kappa_i^3 t + x_i\right] \; + \; \text{radiation} \tag{10.6.12}$$

The decaying radiation decays at each fixed $x < 0$ as $t^{-1/3}$.

11 Solitary waves in a variable environment

Consider again the KdV equation

$$u_t + cu_x + \mu u u_x + \lambda u_{xxx} = 0 \qquad x \in \mathbf{R}, \quad t > 0 \tag{11.0.13}$$

In a variable environment, the coefficients c, μ and λ are functions of x. This problem can arise in many situations. For instance, it describes the propagation of shallow-water solitary waves moving over variable depth (Johnson, 1973), internal solitary waves in regions of variable topography and background stratification (Grimshaw, 1981), inertial solitary waves in rotating fluids (Leibovich and Randall, 1971).

After transforming to new variables, t with $\tau = \int^x \frac{dx}{c} - t$, u with $U = \sqrt{\sigma}\, u$, where $\sigma(x)$ is a magnification factor, the governing equation which replaces 11.0.13 is the *variable coefficient KdV equation*

$$U_x + \alpha(x) U U_\tau + \beta(x) U_{xxx} = 0 \qquad x \in \mathbf{R}, \quad \tau > 0 \tag{11.0.14}$$

Here $\alpha = \frac{\mu}{c\sqrt{\sigma}}$, $\beta = \frac{\lambda}{c^3}$.

In general, this is not an integrable equation, and is usually solved numerically. However, there are two distinct limiting situations in which some analytical progress can be made.

12 Other Physically-Interesting, Soliton-type PDEs

There are by now dozens of "soliton equations", but not only were the three examples from the preceding section the first to be discovered, they are also the best known, and in many ways still the most interesting and important. In fact, in addition to their simplicity and their Hamiltonian nature, each has certain special properties that give them a "universal" character.

Equations admitting solitonic solutions arise as mathematical models in fluid and plasma dynamics. They falls into two distinct categories: relativistically invariant (like Sine-Gordon) and non relativistically invariant (like KdV and Burgers). In some physical situations, solitary waves propagate through a variable environment, which means that the coefficients appearing in the equation are functions of x.

1. The KdV equation for 1-dimensional, unidirectional water waves in a shallow channel,

$$u_t + 6uu_x + u_{xxx} = 0 \tag{12.0.15}$$

2. The modified KdV equation weakly non-linear lattices

$$u_t - u^2u_x - u_{xxx} = 0 \tag{12.0.16}$$

   It has a cubic nonlinearity and it is integrable.

3. A higher-order KdV equation

$$u_t - 45u^2u_x - 15(u_xu_{xx} + uu_{xxx}) - u_{xxxxx} = 0 \tag{12.0.17}$$

4. The Gardner equation, also known as extended Korteweg-de Vries (eKdV) equation

$$u_t + 6uu_x + \beta u^2u_x + u_{xxx} = 0 \tag{12.0.18}$$

   The solutions of the eKdV equation describe various kinds of nonlinear waves (solitons of positive and negative polarities, "thick" solitary waves, breathers, dissipationless shock waves) for various signs of the coefficients of the quadratic and cubic nonlinear terms. Further, although it is integrable, like the KdV, the number of explicitly solved problems is quite small. In particular, soliton interactions have been investigated by Slunyaev & Pelinovsky (1999) and Slunyaev (2001) for both signs of the cubic nonlinear term. The Gardner equation is often used as a model for strongly nonlinear internal waves in the ocean (Holloway et

al, 2001). The coefficient β can be either positive or negative depending on the oceanic (or atmospheric) stratification.

The analytical expression for the soliton shape can be obtained for either sign of β as

$$u(x,t) = \frac{v}{1 + \sqrt{1 + \frac{\beta v}{6}}\cosh(\sqrt{v}(x - vt))} \tag{12.0.19}$$

5. The Benjamin-Bona-Mahony (BBM) equation

$$u_t + 6uu_x - u_{xxt} = 0 \tag{12.0.20}$$

It has the same asymptotic validity as the KdV equation, and since it has rather better high wavenumber properties, is somewhat easier to solve numerically. However it is not integrable, and consequently has not attracted the same interest as the KdV equation.

6. The KP equation (Kadomtsev and Petviashvili, 1970)

$$(u_t + 6uu_x + u_{xxx})_x \pm u_{yy} = 0 \tag{12.0.21}$$

It is a spatial two-dimensional version of the KdV equation, i.e (2+1)D. Taking soliton theory out of the purely (1+1)D regime was a very important step. This equation includes the effects of weak diffraction in the y-direction. When the "+" sign holds in, this is the KPII equation, for which the solitary wave is stable to transverse disturbances. On the other hand if the "-" sign holds, this is the KPI equation for which the solitary wave is unstable. Both KPI and KPII are integrable equations.

7. The sine-Gordon equation for self-induced transparency in nonlinear optical materials,

$$u_{tt} - u_{xx} + \sin(u) = 0 \tag{12.0.22}$$

8. The Liouville equation Interesting problems in statistical mechanics, and conformal field theory

$$u_{xy} - e^{-2u} = 0 \tag{12.0.23}$$

9. The Tzitzeca-Dodd-Bullough equation problems varying from fluid flow to quantum field theory

$$u_{xy} - e^{-u} - e^{-2u} = 0 \tag{12.0.24}$$

10. The classical Heisenberg ferromagnet equation

$$\overrightarrow{S}_t = \overrightarrow{S} \times \overrightarrow{S}_{xx} \tag{12.0.25}$$

11. An equation from ϕ^4 field theory

$$\phi_{xt} - \phi\sqrt{1 - |\phi_t|^2} = 0 \tag{12.0.26}$$

12. The Boussinesq equation for 2-directional, 1-dimensional water floow in channels

$$u_{tt} - u_{xx} - 3(u^2)_{xx} - u_{xxxx} = 0 \tag{12.0.27}$$

which possesses the solitary wave solution

$$u(x,t) = 2\kappa^2 sech^2[\kappa(x - ct - x_0)], \qquad c^2 = 1 + 4\kappa^2 \tag{12.0.28}$$

13. The Kadomtsev-Petviashvilli (KP) equation, for 2-dimensional water flow

$$(u_t + uu_x + u_{xxx})_x + u_{yy} = 0 \tag{12.0.29}$$

14. (2-dimensional) Toda-lattice equations for lattice structures, and conformal Quantum Field Theory

$$u^i_{xy} + e^{A^i j u^j} = 0 \qquad i = 1, \cdots, n \tag{12.0.30}$$

13 Nonlinear optical fibers

One of the most important applications of soliton theory is in the study of nonlinear fibre optics. For over a hundred years, analogue signals travelling over copper wires provided the main medium for point-to-point communication. Early implementations of this medium (twisted pair) were limited in bandwidth to about 100 Kbit/s. By going over to digital signalling instead of analogue, one can get up to ∼1 Mbit/s, and using coaxial cable one can squeeze out another several orders of magnitude. Until recently this seemed sufficient. A bandwidth of about 1 Gbit/s is enough to satisfy the needs of the POTS (plain old telephone system) network that handles voice communication for the entire US, and that could be handled with coaxial cable and primitive fiber optic technology for the trunk lines between central exchanges, and twisted pairs for the low bandwidth "last mile" from the exchange to a user's home. And as we all know, a coaxial cable has enough bandwidth to provide us with several hundred channels of television coming into our homes.

But suddenly all this has changed. As more and more users are demanding very high data-rate services from the global Internet, the capacities of the communication providers have been stretched to and beyond their limits. The problem is particularly critical in the transoceanic links joining North America to Asia and Europe. Fortunately, a lot of fiber optic cables have been laid down in the past decade, and even more fortunately these cables are being operated at bandwidths that are very far below their theoretical limits of about 100 Gbit/s. To understand the problems involved in using these resources more efficiently, it is necessary to understand how a bit is transmitted along an optical fiber. In principle it is very simple. In so-called RZ (return-to-zero) coding, a pulse of high-frequency laser-light is sent to indicate a one, or not sent to indicate a zero. The inverse of the pulse-width in seconds determines the maximum bandwidth of the channel. A practical lower bound for the pulse-width is about 10^{-12} sec, giving an upper bound of about 1000 Gbit/s for the bandwidth. But of course there are further practical difficulties that limit data-rates to well below that figure (e.g., the pulses should be well-separated, and redundancy must be added for error correction), but actual data transmission rates over optical fibers in the 100 Gbit/s range seem to be a reasonable goal (using wavelength-division-multiplexing). A major obstacle to attaining such rates is the tendency of these very short picosecond pulses to disperse. For example, if an approximate square-wave pulse is sent, then dispersion will cause very high error rates after only several hundreds of miles. However if the pulses are carefully shaped to that of an appropriate NLS soliton, then the built-in stability of the soliton against dispersion will preserve the pulse shape over very long distances, and theoretical studies show that error-free propogation at 10 Gbit/s across the Pacific is feasible with current technology, even without multiplexing.

Following the discovery of the IST for the KdV equation, in 1971 Zakharov and Shabat developed the method of solution for the NSE 9.0.5.

In 1973, Akira Hasegawa of AT&T Bell Labs was the first to suggest that solitons could exist in optical fibers. He also proposed the idea of a soliton-based transmission system to increase performance of optical telecommunications. In 1973, Hasegawa and his colleague Tappert discussed the relevance of the NSE Equation 9.0.5 in optical fibers and their associated solitary wave solutions. They showed that optical fibers could sustain envelope solitons, both bright with anomalous (positive) dispersion and dark with normal (negative) dispersion. These solitons propagate in the longitudinal dimension having a single mode guided in the direction perpendicular to the propagation direction.

Although the Zakharov and Shabat paper was published in 1972, Hasegawa and Tappert were unaware of it until they had completed their studies.

In the early 1980's, Linn Mollenauer, Stolen and Gordon showed that solitons could be produced in laboratory experiments. They observed the narrowing of the light wave pulse as the input power was increased, thereby verifying the soliton phenomenon in the fibre. In 1988, Mollenauer and his team transmitted soliton pulses over 4,000 km using a phenomenon called the **Raman effect**, named for the Indian scientist Sir C. V. Raman who first described it in the 1920s, to provide optical gain in the fiber.

However, there was a serious problem, namely that the amplitude decreased by a factor of 10 over 100 km. In 1987 were developed *Erbium–Doped Fibre Amplifiers* (EDFA), spliced-in segments of optical fiber containing the rare earth element erbium, which counteracted the dissipation. Thus the idea of an all optical transmission system became realistic. In 1991, a Bell Labs research team transmitted solitons error-free at 2.5 Gbit/s over more than 14,000 km, using EDFA amplifiers. Pump lasers, coupled to the optical amplifiers, activate the erbium, which energizes the light pulses.

Unfortunately these solitons have some serious technological defects, one of them due to amplifier noise. Effective long distance soliton transmission system utilizing only optical amplifiers was shown to be problematic (Gordon and Haus), indeed amplifier noise would induce soliton velocity variations which in turn creates arrival time jitter, an effect known as the **Gordon-Haus effect**. In the 1990's, several techniques were developed to overcome this problem. Frequency filters and sliding frequency filters which continuously shift the central frequency of the filters were proposed as mechanisms to alleviate Gordon-Haus jitter.

The technological challenge of transmitting large amounts of information via solitons is further complicated by the need to send many solitons at the same time (namely, multi-soliton transmission) in the same fiber, but having different carrier frequencies. This technique is referred to as *wavelength division multiplexing* (WDM). Invariably this means that solitons undergo many collisions as they travel long distances, and these produce jitter and resonant growth of waves in neighbouring frequency channels via what is usually referred to as **four wave mixing** (FWM).

It is has been shown that the timing jitter in WDM systems is greatly reduced by employing filtering and a method referred to as **dispersion management**, a tech-

nique in which fibers with different dispersion characteristics are merged together. For instance, the *strong periodic dispersion management* employs merged fibres with large changes in dispersion characteristics and with alternating signs; i.e. concatenated fibres consisting of both positive (anomalous) and negative (normal) subportions. Such configurations can substantially reduce Gordon-Haus's as well as collision induced timing jitters.

In 1998, Thierry Georges and his team at France Télécom R&D Center, demonstrated a WDM data transmission of 1 Tbit/s.

In 2001, the practical use of solitons became a reality when Algety Telecom deployed submarine telecommunications equipment in Europe carrying real traffic using solitons.

Although the original motivation to create solitons in optical fibers was for use in communications, the nonlinear phenomena which arise in optical fibers have also been used to develop new laser sources, measuring devices, sensors and switches.

14 Soliton Theory of DNA Transcription

It has been suggested by several authors that nonlinear excitations, in particular solitary waves, could play a fundamental functional role in the process of DNA transcription, effecting the opening of the double chain needed for RNA Polymerase to be able to copy the genetic code. Some models have been proposed to model the relevant DNA dynamics in terms of a reduced number of effective degrees of freedom.

The first step in genome expression is DNA transcription from the original DNA template contained in the cell to a copy – the RNA messenger (mRNA) which will then be used as a 'master copy' in determining protein sequences in accordance with the genetic information. The transcription process is carried out by a specialized enzyme, the **RNA Polymerase** (RNAP). The RNAP opens a 'transcription bubble' of a size of about 15–20 bp, and then travels along the DNA chain keeping the size of the open region more and less constant, i.e. providing at the same time to open the chain in front of it and to close back the one behind. In the active phase of the process, the RNAP proceeds along the DNA chain at a speed of $\sim 10 - 100\ bp/s$ (different authors give different estimates). Since each base pair is linked by two or three hydrogen bonds, the energy involved in such a process, even considering only the one to open (and close) the DNA chain, is of the order of at least $\sim 100 - 1000$ H bonds per second. Obviously, as biological systems live at a temperature of the order of 300 Kelvin, thermal energy is widely available, but the problem is *how is this energy focused at the right place and with the right timing* to operate the process.

It has been suggested by several authors that nonlinear dynamics can have a word to say concerning this, and that energy transport could happen by means of nonlinear excitations travelling along biological chains, such as long proteins or indeed DNA. Roughly speaking, the idea is that there are nonlinear excitations travelling along the DNA chains, causing a local opening of the double chain. The RNAP could then travel along with these, and use the opening of the chain to read the DNA sequence without having to focus the energy needed to open the double chain. These hypothetic nonlinear excitations are not necessarily solitons in strict mathematical sense, but rather solitary waves; however, it is by now common to call them solitons, and I will conform to this usage.

The most widely known theory in this context is probably that of **Davydov's soliton**, which is of essentially quantum nature.

The DNA molecule, like any other one, obeys quantum mechanics; moreover, we are interested in its behaviour in living conditions, i.e. in particular at a quite precise temperature (around 300 K, or a bit more for humans). At this temperature most of the degrees of freedom of the atoms constituting the DNA molecule will be essentially frozen due to quantum mechanical considerations (their excitation energy being much higher), and thus many bonds between atoms can be considered as rigid; on the other side, there are degrees of freedom whose excitation energy is much lower than the

thermal scale, so that they can be considered to behave classically. The consequence of this on modeling is clear: we can safely consider the degrees of freedom of the first kind as non-existent, and concentrate on the remaining – effective – degrees of freedom. These should still be treated quantum mechanically, but for some of them – those of the second kind – we can also use a classical model.

In the Yakushevich model, we consider rotations of the bases in a plane orthogonal to the double helix axis (which of course is not completely true), and any other movement is not considered.

to do

15 Nonlinear Transmission Line

15.1 Introduction

With the recent advances in the performance of high speed transistors, device operating bandwidths have increased far beyond the available measuring instrument bandwidth. This limits the ability to understand the high frequency physics of the devices and also makes it difficult to design circuits that can take advantage of their available bandwidth. To overcome this measurement limitation, sampling circuits with bandwidths exceeding those of the fastest transistors are necessary. The first necessary component of a high frequency sampling circuit is a strobe signal for the sampling bridge that has a pulse width much shorter than the time response of the signal that is being sampled. The technology that enables the electrical generation of subpicosecond pulses is the **nonlinear transmission line** (NLTL).

A NLTL is a GaAs integrated circuit consisting of sections of high impedance transmission line (coplanar waveguide) periodically loaded with reverse-biased Schottky diodes, or other nonlinear capacitive devices, such as *Heterostructure Barrier Varactors* (HBVs). These diodes appear as voltage-variable capacitors (varactors) and cause the propagation delay through the NLTL to depend on the wave amplitude. Nonlinearity arises from the voltage-dependent propagation characteristics of the NLTL. Dispersion arises from the periodicity of the loading diodes. Since the capacitance changes with applied voltage, the propagation characteristics depend on the wave amplitude. The variable capacitance of the diodes results in a voltage dependent velocity for waves traveling on the NLTL, and this variable velocity causes the falling edge of a signal to be compressed. The step function generated by the NLTL can then be differentiated to form a pulse, and this pulse is used to strobe the sampling circuits. With the advent of GaAs based NLTLs in 1987, the growth of this technology has been rapid, with commercial instruments operating to 50 GHz and prototype instruments demonstrated at 250 GHz. Future improvements make feasible to fabricate 2-3 THz sampling circuits.

Figure 27: Photo of an actual Nonlinear transmission line.

The NLTL has three fundamental and quantifiable characteristics just as any non-ideal transmission line. These are nonlinearity, dispersion, and dissipation. Along with

some other characteristics (e.g. impedance, length, etc.), they define a transmission line's behavior with arbitrary stimulation. What distinguishes one class of line from another is the degree to which these characteristics occur and interact. For example, optical fiber has very small nonlinearity and dissipation but moderate dispersion; a small amplitude impulse will spread on propagation due to the dispersion while a large amplitude impulse may become compressed due to the nonlinearity.

NLTLs provide nonlinearity due to the voltage dependent capacitance, dispersion due to the periodicity, and dissipation due to the finite conductivity of the coplanar waveguide conductors and series resistance of the diodes. An approximate equivalent circuit consisting of series inductors and shunt capacitors is much easier to analyze than the transmission line circuit.

15.2 Circuit Modeling

A nonlinear transmission line can be modelled as a linear periodic structure formed by cells, which consists of a section of transmission line of length l and a shunt variable capacitor to ground. The capacitance C is a function of the voltage. A two-conductor transmission line can be described by lumped parameters that are distributed throughout its length. Therefore, by utilizing a first order approximation, this varactor loaded high impedance interconnecting transmission line can be approximated by L-C sections, as shown below

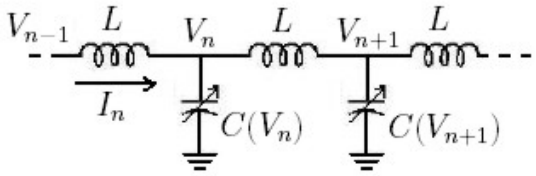


Figure 28: A sigle cell for a NLTL lumped model.

We no longer have the familiar linear relation $Q_n = CV_n$, but rather the nonlinear differential relationship

$$C(V_n) = \frac{dQ_n}{dV_n} \tag{15.2.1}$$

To solve for the voltage as a function of time, we use the relations

$$V_n - V_{n+1} = \frac{d\Phi_{n+1}}{dt} = L\frac{dI_{n+1}}{dt} \tag{15.2.2}$$

$$I_n - I_{n+1} = \frac{dQ_n}{dt} = C(V_n)\frac{dV_n}{dt} \tag{15.2.3}$$

The last equation becomes

$$\frac{dI_n}{dt} - \frac{dI_{n+1}}{dt} = \frac{d}{dt}\big[C(V_n)\frac{dV_n}{dt}\big] \tag{15.2.4}$$

Eliminating I_n and I_{n+1}, we find the following second order differential equation:

$$\frac{d}{dt}\big[C(V_n)\frac{dV_n}{dt}\big] = \frac{1}{L}\big(V_{n+1} + V_{n-1} - 2V_n\big) \tag{15.2.5}$$

The right-hand side of can be approximated with partial derivatives with respect to distance, x, from the beginning of the line, assuming that the spacing between two adjacent sections is δ (i.e., $x_n = n\delta$.)

$$V_{n\pm1} = V(x \pm \delta)_{x=n\delta} = V(x) \pm \delta\frac{\partial V}{\partial x} + \frac{\delta^2}{2!}\frac{\partial^2 V}{\partial x^2} \pm \frac{\delta^3}{3!}\frac{\partial^3 V}{\partial x^3} + \frac{\delta^4}{4!}\frac{\partial^4 V}{\partial x^4} + \ldots \tag{15.2.6}$$

We can assume a small δ and ignore the high order terms for the Taylor expansions of $V(x-\delta)$ and $V(x+\delta)$ obtaining the following approximate continuous partial differential equation

$$\frac{\partial}{\partial t}\big[C(V)\frac{\partial V}{\partial t}\big] = \frac{1}{L}\big(\frac{\partial^2 V}{\partial x^2} + \frac{\delta^2}{12}\frac{\partial^4 V}{\partial x^4}\big) \tag{15.2.7}$$

where C and L are the capacitance and inductance per unit length, respectively. For a continuous transmission line ($\delta \to 0$), it reduces to

$$\frac{\partial}{\partial t}\big[C(V)\frac{\partial V}{\partial t}\big] = \frac{1}{L}\frac{\partial^2 V}{\partial x^2} \tag{15.2.8}$$

Diodes present two sources of nonlinearity: conductive and reactive. The conductive nonlinearity is evident in the I-V curves and the reactive nonlinearity is evident in the C-V curves. The nature of the diode's C-V characteristics depends wholly on the epitaxial structure of the diode. For a given doping profile the approximate capacitance can be determined by

$$\phi - V = \int_0^{x_d(V)} \frac{qx}{\epsilon_s} N(x)dx \tag{15.2.9}$$

then computing the capacitance as $C(V) = \frac{\epsilon_s A}{x_d(V)}$ where ϕ is the barrier potential, V is the applied voltage, x_d is the depletion depth, q is the electron charge, ϵ_s is the dielectric constant, and $N = N_D - N_A$ is the N-type doping concentration as a function of depletion depth. For arbitrary doping profiles, solutions to equation 15.2.9 often require numerical methods and result in ordered pairs of $C(V)$ data. This data can then be fitted to any desired functional relationship. The choice of model depends on the application. Both Ikezi and Hirota model NLTLs, but Ikezi deals exclusively with ferroelectric material loading parallel plate waveguide, while Hirota deals with LC lattices. Ikezi [H. Ikezi, S.S.Wojtowicz, R.E.Waltz, and D.R.Baker, *Temporal Contraction of Solitons in a Nonuniform Transmission Line*, *Journal of Applied Physics*, vol. 64, no. 12, pp. 6836-6838, December 15, 1988] assumes either

$$C(V) = C_{J0}(1 - 2bV) \tag{15.2.10}$$

or

$$C(V) = C_{J0}(1 - 3bV^2) \tag{15.2.11}$$

depending on his approach. Hirota [R.Hirota and K.Suzuki, *Theoretical and Experimental Studies of Lattice Solitons in Nonlinear Lumped Networks*, *Proceedings of the IEEE*, vol. 61, no. 10, pp. 1483-1491, October 1973] uses the model

$$C(V) = \frac{C_{J0}}{(1 - \frac{V}{\phi})} \tag{15.2.12}$$

where C_{J0} is the zero-bias junction capacitance. For most diodes, Hirota's $C(V)$ characteristic fits more accurately than Ikezi's.

The most common function applied to diode $C(V)$ curves is

$$C(V) = \frac{C_{J0}}{(1 - \frac{V}{\phi})^M} \tag{15.2.13}$$

which adds the parameter M to the Hirota's model. M is the grading coefficient ($M = 0.5$ in the case of a uniformly doped diode). 15.2.13 is found in most circuit simulators.

In subsequent analysis we will assume that the voltage is small, and we can approximate the capacitor's voltage dependence by the linear relationship 15.2.10.

Substituting 15.2.10 into 15.2.8, we find

$$\frac{\partial^2 V}{\partial t^2} - \frac{1}{LC_{J0}}\frac{\partial^2 V}{\partial x^2} = \frac{\delta^2}{12}\frac{1}{LC_{J0}}\frac{\partial^4 V}{\partial x^4} + b\frac{\partial(V^2)}{\partial t^2} \tag{15.2.14}$$

where the left-hand side is the classic wave equation and the terms on the right-hand side represent dispersion and non-linearity, respectively.

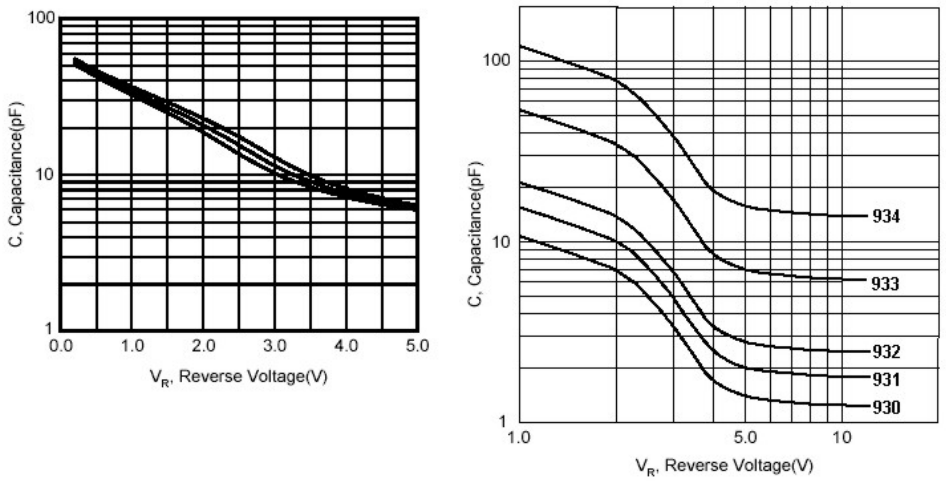


Figure 29: Capacitance-Voltage profile for two discrete Si varactor diodes. Toko KV1471 (left); Zetex ZV931 (right)

If the effect of the dispersive and non-linear terms in the last equation are on the same order of magnitude and opposite signs it is possible to have a single pulse solution with a profile that does not change as it propagates with velocity, v.

A propagating mode solution can be obtained by converting the partial differential equation (PDE) to an ordinary differential equation (ODE) by a simple change of variable: $u = x - vt$. This solution is:

$$V(x,t) = \frac{3(v^2 - v_0^2)}{bv^2} sech^2\Big[\frac{\sqrt{3(v^2 - v_0^2)}}{v_0}\ \frac{x - vt}{\delta}\Big] \tag{15.2.15}$$

where v is the propagation velocity of the pulse and energy, while $v_0 = \frac{1}{\sqrt{LC_{J0}}}$.

This solution is shown in Fig. 2 for three different values of L and C_{J0}, and hence different δ.

Note that this solution is not a function of the input waveform, and thus any arbitrary input will eventually turns into 15.2.15 going through a line which is long enough, if it has enough energy.

As can be seen from 15.2.15, the peak amplitude is a function of the velocity. Defining an effective capacitance, C_{eff}, so that $v = \frac{1}{\sqrt{LC_{eff}}}$, the pulse height is given by

$$V_{max} = \frac{3}{b}\big(1 - \frac{C_{eff}}{C_{J0}}\big) \tag{15.2.16}$$

Using 15.2.10 we can relate C_{eff} to an effective voltage V_{eff}. It is easy to show that

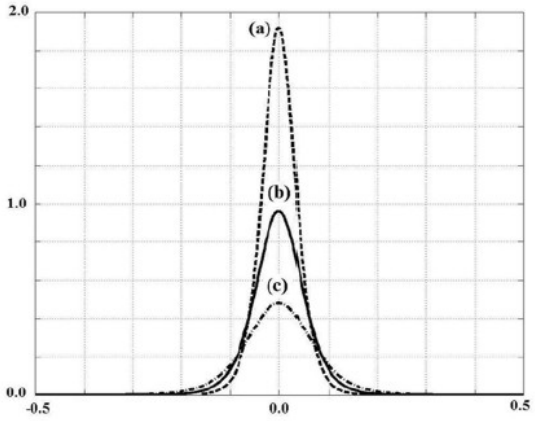


Figure 30: Three normalized soliton shapes for different values of L and C (a) L=1nH, C=1nF (b) L=2nH, C=2nF (c) L=4nH, C=4nF

$$V_{eff} = \frac{V_{max}}{3} \tag{15.2.17}$$

So it is the capacitance at one-third the peak amplitude that determines the effective propagation velocity. Using the relations above we can easily calculate the half-height width of the pulse to be

$$W \approx \frac{\delta}{v} \frac{v_0}{\sqrt{v^2 - v_0^2}} \tag{15.2.18}$$

As can be seen, in a weakly dispersive and non-linear transmission line, the non-linearity can counteract the normally present dispersive properties of the line maintaining solitary waves that propagate without dispersion. This behavior can be explained using the following intuitive argument: The instantaneous propagation velocity at any given point in time and space is given by $\frac{1}{\sqrt{LC}}$. In the presence of a non-linear capacitor with a characteristic given by 15.2.10, the instantaneous capacitance is smaller for higher voltages. Therefore, the points closer to the crest of the voltage waveform experience a faster propagation velocity and produce a shock-wave front due to the nonlinearity. On the other hand, dispersion of the line causes the waveform to spread out. For a proper non-linearity these two effects can cancel each other out.

Sign of solution depends on sign of non-linearity factor, b; For a capacitor with a positive voltage dependence (e.g., an nMOS varactor in accumulation mode) we have

$$C(V) = C_{J0}(1 + 2bV) \tag{15.2.19}$$

resulting in upside down pulses.

Based on these results, to achieve large-amplitude narrow pulses, inductance and capacitance of the NLTL must be as small as possible and non-linearity factor, b, should be large enough to compensate the dispersion of the line. It is also important to know the characteristic impedance of these lines (for impedance matching, etc.) As in a NLTL the capacitance is a function of voltage, we can only define an effective semi-empirical value for the characteristic impedance. The characteristic impedance of a infinite ladder network with longitudinal and transverse impedances Z_1 and Z_2 is

$$Z_0 = \frac{Z_1}{2} + \sqrt{\frac{Z_1^2}{4} + Z_1 Z_2} \tag{15.2.20}$$

For a purely reactive line

$$Z_0 = \sqrt{\frac{L}{C} - \frac{\omega^2 L^2}{4}} \tag{15.2.21}$$

Simulations results indicate that one can approximate Z_{eff} using the capacitance at V_{eff}, i.e.,

$$Z_{eff} \approx \sqrt{\frac{L}{C(V_{max}/3)}} \tag{15.2.22}$$

Another approach is to consider the average capacitance that the wave sees over its entire voltage swing. Rodwell determined that the effective loading capacitance of the diode over the pulse's voltage swing is a constant, as is the effective NLTL impedance. This results in the so called large-signal capacitance

$$C_{LS} = \frac{1}{V_{max} - V_{min}} \int_{V_{min}}^{V_{max}} C(V) dV \tag{15.2.23}$$

This leads naturally to the large signal characteristic impedance of the line, which is the effective loaded impedance of the NLTL under large signal drive:

$$Z_{LS} = \sqrt{\frac{L}{C_0 + C_{LS}}} \tag{15.2.24}$$

where C_0 is the additional capacitance due to the line (not explicitly considered in the model above). If $Z_{load} = Z_{LS}$ then the shock wave is passed to the load without reflection or distortion; and similarly, if $Z_{gen} = Z_{LS}$ then the signal from the generator is launched on the NLTL with no reflection.

15.3 Gradually Scaled NLTL

One problem in pulse narrowing NLTLs is that if the input pulse is wider than W (as in 15.2.18), it is incapable of concentrating all that energy into one pulse and instead the input pulse degenerates into multiple soliton pulses. This is an undesirable effect that cannot be avoided in a standard line. This problem can be solved by using gradually scaled NLTLs. We notice that the characteristic pulse width of the line is controlled by the node spacing, δ, and the propagation velocity, v, which is in turn controlled by, L and C.

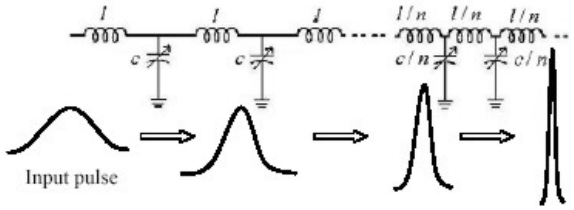


Figure 31: Schematic of the gradually scaled non-linear transmission line

Thus, we use a gradual line consisting of several sections that are gradually scaled to have smaller characteristic pulse width, as shown in the figure. The first few sections have the widest characteristic pulse W, meaning that their output is wider and has smaller amplitude. As a result, the input pulse will cause just one pulse at the output of these sections. The following sections have a narrower response and the last section has the narrowest one. This will guarantee the gradual narrowing of the pulses and avoids degeneration. Each section has to be long enough so that the pulse can reach the section's steady-state response before entering next section.

15.4 Tunnel Diode as a nonlinear element

NLTLs mostly employ Varactor diode as a nonlinear shunt element. Diverting from this approach we would investigate a different nonlinear element, the **tunnel diode**. A tunnel diode is a semiconductor with a negative resistance region that results in very fast switching speeds, up to $5GHz$. The operation depends upon a quantum mechanic principle known as "tunneling" wherein the intrinsic voltage barrier ($0.3V$ for Ge junctions) is reduced due to doping levels which enhance tunneling.

The static current/voltage characteristic of the device, $i = h(v)$ contains a region of negative dynamic resistance. The transitions from positive to negative resistance and back again are smooth - there are no discontinuities in slope and the device does not exhibit hysteresis. The device is only operated in the vicinity of the negative resistance region; typically a span of one or two volts.

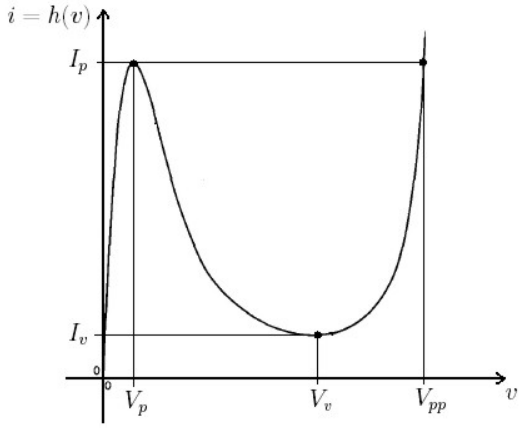


Figure 32: Typical Tunnel Diode I-V characteristic, with all the three relevant points.

Main characteristics of a tunnel diode current-voltage curve are **peak voltage** and **peak current** (V_p, I_p), **valley voltage** and **valley current** (V_v, I_v), **forward peak voltage**, (or **projected peak voltage**), V_{pp}. Typical parameter values for a specific Tunnel device family are shown in the Table below.

To solve for the voltage as a function of time, we have the relations

$$V_n - V_{n+1} = \frac{d\Phi_{n+1}}{dt} = L\frac{dI_{n+1}}{dt} \tag{15.4.1}$$

$$I_n - I_{n+1} = C\frac{dV_n}{dt} + h(V_n) \tag{15.4.2}$$

Following the same way as in the previous paragraph, we arrive at the following approximate continuous PDE

Part Number	I_p mA	I_v mA	C pF	V_p mV	V_v mV	V_{pp} mV	R_s Ω
1N3712	1.0 ± 10%	0.18	10	65	350	500	4.0
1N3713	1.0 ± 2.5%	0.14	5	65	350	510	4.0
1N3714	2.2 ± 10%	0.48	25	65	350	500	3.0
1N3715	2.2 ± 2.5%	0.31	10	65	350	510	3.0
1N3716	4.7 ± 10%	1.04	50	65	350	500	2.0
1N3717	4.7 ± 2.5%	0.60	25	65	350	510	2.0
1N3718	10.0 ± 10%	2.20	90	65	350	500	1.5
1N3720	22.0 ± 10%	4.80	150	65	350	500	1.0
TD-261	2.2 ± 10%	0.31	3.0	70	390	500-700	5.0
TD-261A	2.2 ± 10%	0.31	1.0	80	390	500-700	7.0
TD-262	4.7 ± 10%	0.60	6.0	80	390	500-700	3.5
TD-262A	4.7 ± 10%	0.60	1.0	90	400	500-700	1.7
TD-263	10.0 ± 10%	1.40	9.0	75	400	500-700	1.7
TD-263A	10.0 ± 10%	1.40	5.0	80	410	520-700	2.0
TD-263B	10.0 ± 10%	1.40	2.0	90	420	550-700	2.5
TD-264	22.0 ± 10%	3.80	18.0	90	425	600 Typ.	1.8
TD-264A	22.0 ± 10%	3.80	4.0	100	425	550-700	2.0
TD-265	50.0 ± 10%	8.50	25.0	110	425	625 Typ.	1.4
TD-265A	50.0 ± 10%	8.50	5.0	130	425	640 Typ.	1.5
TD-266	100 ± 10%	17.50	35.0	150	450	650 Typ.	1.1
TD-266A	100 ± 10%	17.50	6.0	180	450	650 Typ.	1.2

Figure 33: Typical parameter values for a specific Tunnel device family. R_s is the series parasitic resistance, C the parallel capacitance.

$$\frac{\partial}{\partial t}\big[C\frac{\partial V}{\partial t} + h(V)\big] = \frac{1}{L}\big(\frac{\partial^2 V}{\partial x^2} + \frac{\delta^2}{12}\frac{\partial^4 V}{\partial x^4}\big) \tag{15.4.3}$$

where C and L are the capacitance and inductance per unit length, respectively. For a continuous transmission line ($\delta \to 0$), it reduces to

$$\frac{\partial}{\partial t}\big[C\frac{\partial V}{\partial t} + h(V)\big] = \frac{1}{L}\frac{\partial^2 V}{\partial x^2} \tag{15.4.4}$$

Modelling a Tunnel Diode can be easily done using polynomials. Assuming to work in the valley of the voltage-current curve, we can approximate $h(V)$ with the following quadratic relationship

$$h(V) = \alpha(V - V_v)^2 + I_v \tag{15.4.5}$$

where $\alpha := \frac{I_p - I_v}{(V_p - V_v)^2} > 0$. Substituting in 15.4.3, and introducing $\mathcal{V} := V - V_v$ we get

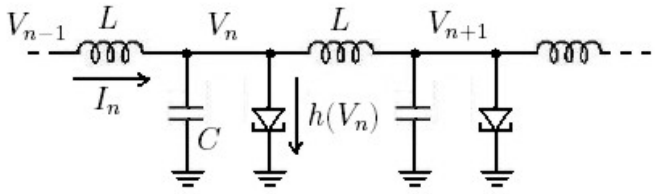


Figure 34: Nonlinear transmission line based on tunnel diodes as nonlinear element.

$$\frac{\partial^2 \mathcal{V}}{\partial t^2} - \frac{1}{LC}\frac{\partial^2 \mathcal{V}}{\partial x^2} = \frac{\delta^2}{12}\frac{1}{LC}\frac{\partial^4 \mathcal{V}}{\partial x^4} - \frac{\alpha}{C}\frac{\partial(\mathcal{V}^2)}{\partial t} \tag{15.4.6}$$

This equation is quite similar to 15.2.14, thus rearranging 15.2.15 we obtain the solitonic solution

$$V(x,t) = V_v - \frac{3C(v^2 - v_0^2)}{\alpha v^2} sech^2\big[\frac{\sqrt{3(v^2 - v_0^2)}}{v_0}\ \frac{x - vt}{\delta}\big] \tag{15.4.7}$$

resulting in negative pulses. Now, assuming to work near the peak of the voltage-current curve, we can approximate $h(V)$ with

$$h(V) = -\alpha(V - V_p)^2 + I_p \tag{15.4.8}$$

obtaining the solitonic solution

$$V(x,t) = V_p + \frac{3C(v^2 - v_0^2)}{\alpha v^2} sech^2\big[\frac{\sqrt{3(v^2 - v_0^2)}}{v_0}\ \frac{x - vt}{\delta}\big] \tag{15.4.9}$$

resulting in upside down pulses. Let's calculate α for a specific device with the following I-V curve parameters

$$V_p = 50mV; \quad I_p = 4.2mA; \quad V_v = 370mV; \quad I_v = 370\mu A; \quad V_{pp} = 525mV$$

we obtain

$$\alpha = \frac{I_p - I_v}{(V_p - V_v)^2} = 0.0374$$

Here, $\frac{\alpha}{C}$ plays the same role of b parameter in the NLTL discussed in the previous paragraph. To achieve large-amplitude narrow pulses, inductance and capacitance of the NLTL must be as small as possible and non-linearity factor, $\frac{\alpha}{C}$, should be large enough to compensate the dispersion of the line.

16 Practical Demonstration of Solitons on a Discrete NLTL

A discrete NLTL could be simulated and built using discrete SMD inductors and varactor diodes (or tunnel diodes) in order to detect solitonic waves and investigate their properties. Inductors and diodes should be characterized using an impedance analyzer, and their equivalent models introduced in a PSPICE simulation. Measurement on the physical circuit should be compared to simulation and the discrepancies explained.

16.1 Unipolar NLTL, varactor based

Let's consider the PSPICE Model of a $N = 33$ node NLTL shown in the picture below.

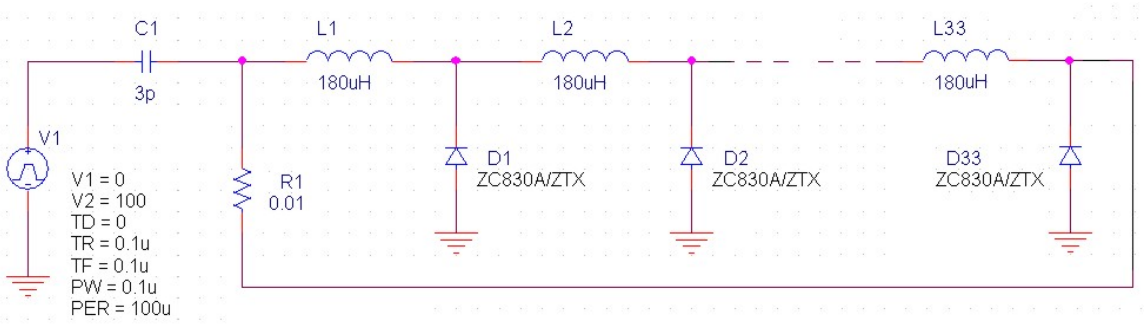


Figure 35: Schematics of the NLTL model.

Resistor $R1$, with a negligible value, only closes the loop. This geometric arrangement of the NLTL will permit us more effectively to put in evidence the space-time evolution of solitons. The transient voltage excitation (source $V1$) is a short rectangular pulse that injects an energy enough to induce a single soliton on the line. Really, due to having closed the line, this single localized excitation will produce two identical solitons propagating in opposite direction, allowing a natural base to observe scattering of solitons having opposite velocities.

Capacitor $C1$ decouples the line from the DC component of the voltage source, letting the AC components to pass through. The diode effect of $D1$ will cut the negative components, so that only the rising edge of the voltage excitation will take an effective part to the soliton generation. The decoupling capacitor introduces just a little extra capacitance in parallel to diode $D33$. The varicap diode is the Zetex ZC830A.

Its capacitance spans linearly from $\approx 19pF$ to $\approx 2pF$ when the reverse voltage varies from $V_R = 0V$ to $V_R = 20V$.

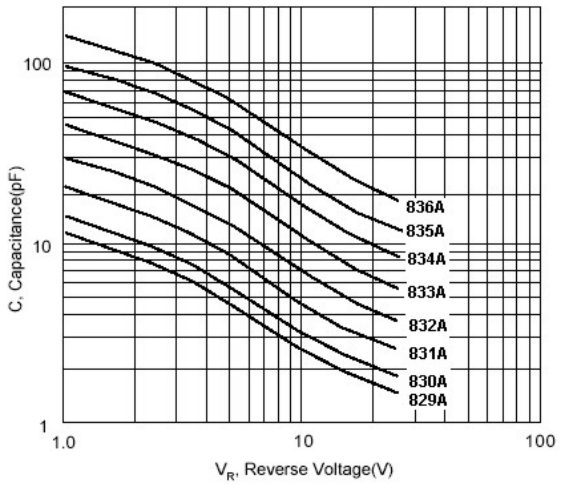


Figure 36: Capacitance-Voltage profile for Si 25V hyperabrupt varactor diodes ZC83xA.

16.1.1 Does the SPICE model match with the diode C-V curve?

PSPICE has built-in models for the semiconductor devices, and the user need specify only the pertinent model parameter values. The general form for a **device model** is

```
.MODEL MNAME TYPE PNAME1=PVAL1 PNAME2=PVAL2 ...
```

Most simple circuit elements typically require only a few parameter values. However, some devices (semiconductor devices in particular) require many parameter values. For ZC830A, the following parameters hold

```
.MODEL ZC830A D IS=5.355E-15 N=1.08 RS=0.1161 XTI=3 EG=1.11
CJO=19.15E-12 M=0.9001 VJ=2.164 FC=0.5 BV=45.1 IBV=51.74E-3
TT=129.8E-9 ISR=1.043E-12 NR=2.01
```

where:

RS	series resistance
IS	saturation current
N	emission coefficient
XTI	temperature coefficient
EG	band-gap voltage
CJO	zero-bias junction capacitance
VJ	bulk junction potential
M	junction grading coefficient
FC	onset of forward-bias depletion capacitance coefficient
BV	reverse breakdown voltage
IBV	reverse breakdown current
ISR	recombination current saturation value
NR	the recombination current emission coefficient
TT	transit time
KF	flicker-noise coefficient
AF	flicker-noise exponent

(For the latest version of the Zetex Semiconductors PSPICE library goto the applications section of the Zetex web site at http://www.zetex.com/).

The last two parameters in the list above are not included in our model, thus they are assigned the default values ($KF = 0$; $AF = 1$). The DC characteristics of the diode are determined by IS and N. An ohmic resistance, RS, is included. Charge storage effects are modelled by a transit time, TT. The temperature dependence of the saturation current is defined by the parameters EG, the energy and XTI, the saturation current temperature exponent. Reverse breakdown is modeled by an exponential increase in the reverse diode current and is determined by the parameters BV and IBV (both of which are positive numbers).

What is important for our discussion are the parameters CJO, VJ, and M. Indeed, those parameters determine the nonlinear depletion layer capacitance model according to the formula

$$C(V_R) = \frac{CJO}{(1 + \frac{V_R}{VJ})^M} \tag{16.1.1}$$

where V_R is the reverse applied voltage. In order to verify the 16.1.1 and the goodness of this model in the Capacitance-Voltage relation we can perform the AC analysis for the resonant LC circuit visible in the inset of the picture below. This kind of PSPICE analysis computes the AC output variables as a function of frequency. The program first computes the DC operating point of the circuit and determines linearized, small-signal models for all of the nonlinear devices in the circuit. The resultant linear circuit is then analyzed over a user-specified range of frequencies. Our desired output is a transfer function (voltage gain). It is convenient to set the input to unity and zero

phase, so that the output variable (the voltage across the diode) has the same value as the transfer function of the output variable with respect to the input. Our analysis is also performed over a set of different DC biases V_{DC}, and the picture represents all the voltage transfer functions.

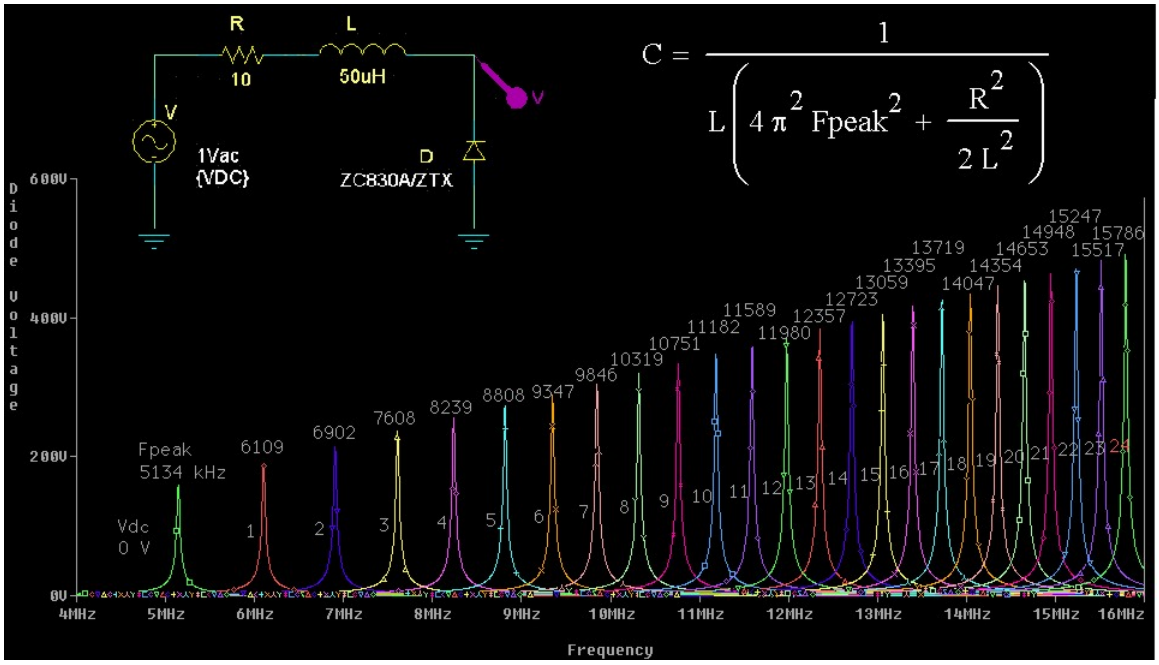


Figure 37:

At the analytical level, zeroing the first derivative of the magnitude of the transfer function give us the capacitance of the varactor diode as a function of the resonant frequency $Fpeak$, as well as a function of the other circuit parameters. This relation is detailed in the up-right inset. Finally, $C(V_{dc})$ is plotted, obtaining the red C-V curve showed in the figure below. As expected, it exactly overlaps with the curve described by the modelling equation , and it is in a reasonable agreement with the Zetex ZC830A curve given in the component datasheet.

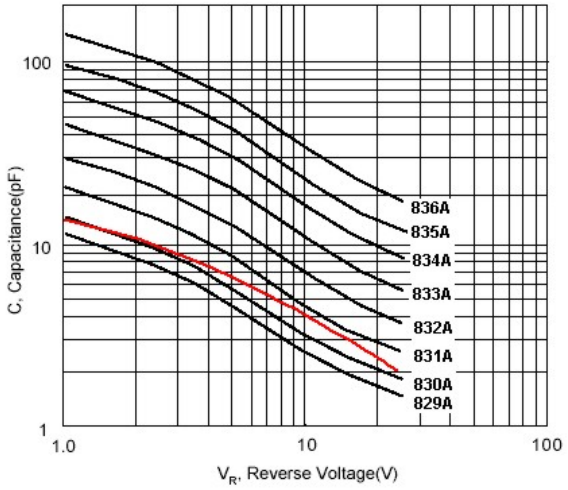


Figure 38:

16.1.2 Simulation

In our simulation we used $L = 180uH$. A lesser value reduces soliton's amplitude and increase its speed. An increase of speed cames also from a reduction in the shunting loss capacitance C (part of this capacitance is embedded in the diode's model, part is due to the transmission lines between consecutive cells). But the amplitude variation is opposite. The inductive element is ideal, so that the only dissipative element is embedded in the varactor's model (RS).

The next picture shows a transient simulation in the time window $t \in [0, 5\mu s]$.

The colored curves represent voltages waveforms at nodes 8, 16, 24 and 33. Let's indicate node voltages as $V(x,t) = V(k,nT) = V_k(nT)$, being $k = 1, \cdots, N$ the node index, and $T = 16.4ns$ the time step period. Instead to see a specific node voltage as a function of time t, an alternative and more effective approach would be to plot the voltage distribution along the entire line at a specific time. To this aim, a PSPICE simulation (OrCAD PSPICE 9.1) was performed for each node voltage, over a time window $t \in [0, 15\mu s]$. The whole data obtained (33*915=30195 real values in floating point) was imported in MATHCAD ver.12, organized as a rectangular matrix for subsequent visualization and manipulation.

In the next picture the spatial voltage distribution is shown as a function of the node index k. The node corresponding to $k = 1$ is that where the impulsive excitation was applied.

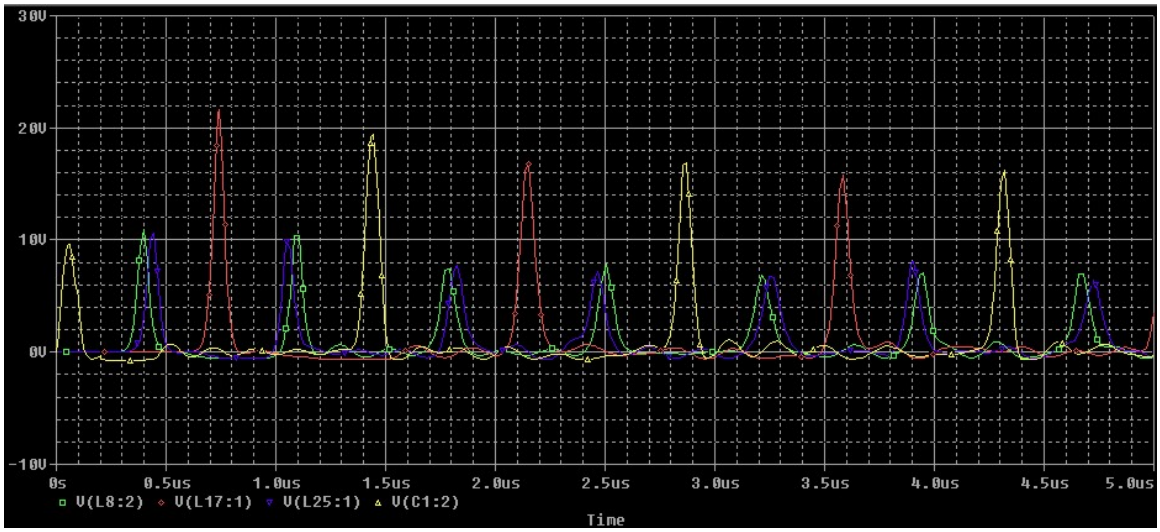


Figure 39: PSPICE Transient simulation of the circuit. Voltage at nodes 8, 16, 24 and 33.

As a consequence of stimulation two $sech()^2$-shaped solitons arise and propagate in opposite direction with the same amplitude and speed. After some instant they collide. We can observe an increase in waveform amplitude during the fusion phase, instead of a decreasing as discussed in a previous paragraph. This is just because the two solitons propagates in opposite direction. Another difference is that velocity before collision is always less than velocity after collision, i.e. $v_{before} < v_{after}$.

A different stimulation can be used to see the collision among solitons moving in the same direction. To this aim we produce a first small amplitude pulse, followed by a large amplitude one.

In this case there is a decrease in waveform amplitude during the fusion phase, and pheraps also $v_{before} = v_{after}$ holds (elastic scattering) (to be verified).

16.2 Bipolar NLTL, varactor based

The varactor-based NLTL discussed in the previous paragraph can be made suitable for supporting solitons having positive or negative polarity. This can be obtained by simply adding a bias voltage to the varicap element, as shown in the picture below.

The effective reverse voltage applied to D is now $V_{POL} - V_n$, so that the junction capacitance approximation becomes

$$C(V_n) = CJ0[1 - 2b(V_{POL} - V_n)] \tag{16.2.1}$$

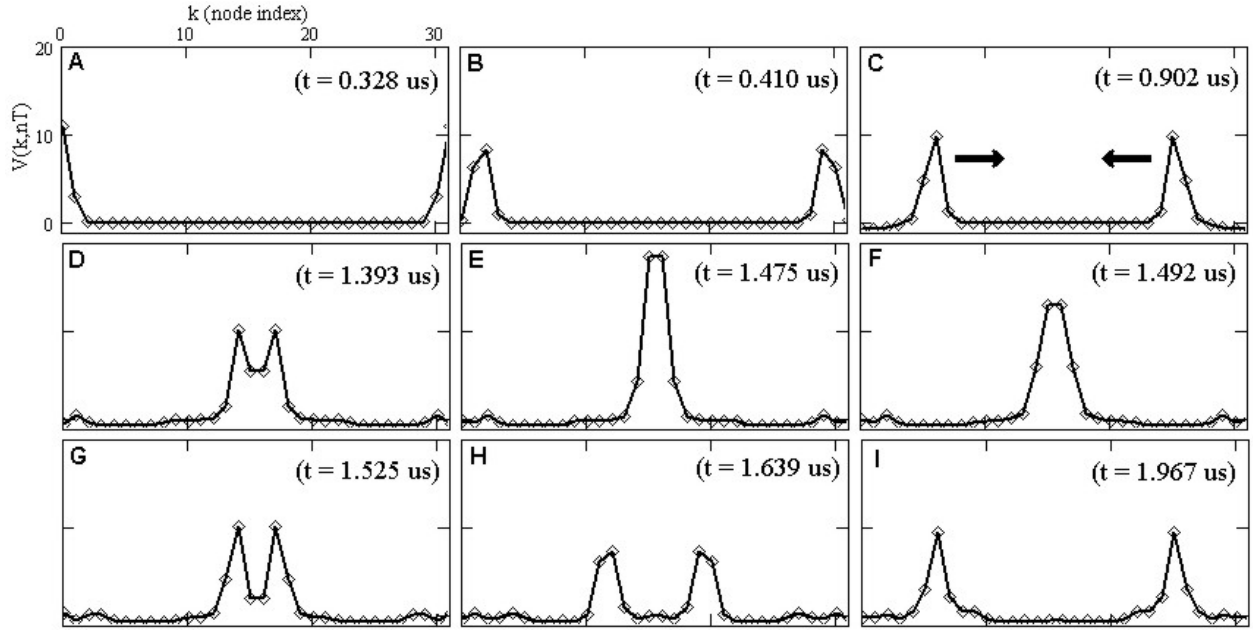


Figure 40: Spatial voltage distribution. Sequence of a solitonic collision.

i.e. the usual capacitor's voltage dependence holds also for $V_n < 0$, provided that $|V_n| < V_{POL}$. In order to allow integral propagation of negative solitonic waves ($V_n < 0$), at the soliton's peak the diode's reverse voltage should not reach the breakdown reverse voltage ($V_M + V_{POL} < BV$). In our simulation we will use $V_M \simeq 5V$ and $V_{POL} = 15V$.

To make the simulation more manageable it is convenient to embed all the cell's components into a single PSPICE component, or a **subcircuit**. A subcircuit that consists of PSPICE elements can be defined and referenced in a fashion similar to device models. The subcircuit is defined in the input file by a grouping of element lines; the program then automatically inserts the group of elements wherever the subcircuit is referenced. There is no limit on the size or complexity of subcircuits, and subcircuits may contain other subcircuits.

For our case the following subcircuit model holds

```
.SUBCKT NLTL   GND IN OUT PPOL R1   IN   2 50E-2 L1   2   OUT
15E-6 C1   OUT   GND 3E-12 D1   OUT   PPOL ZC830A/ZTX .MODEL
ZC830A/ZTX D IS=5.355E-15 N=1.08 RS=0.1161 XTI=3 + EG=1.11
CJO=19.15E-12 M=0.9001 VJ=2.164 FC=0.5 + BV=45.1 IBV=51.74E-3
TT=129.8E-9 + ISR=1.043E-12 NR=2.01 .ENDS
```

This time the NLTL is composed of $N = 48$ cells, not closed in a loop as in the case

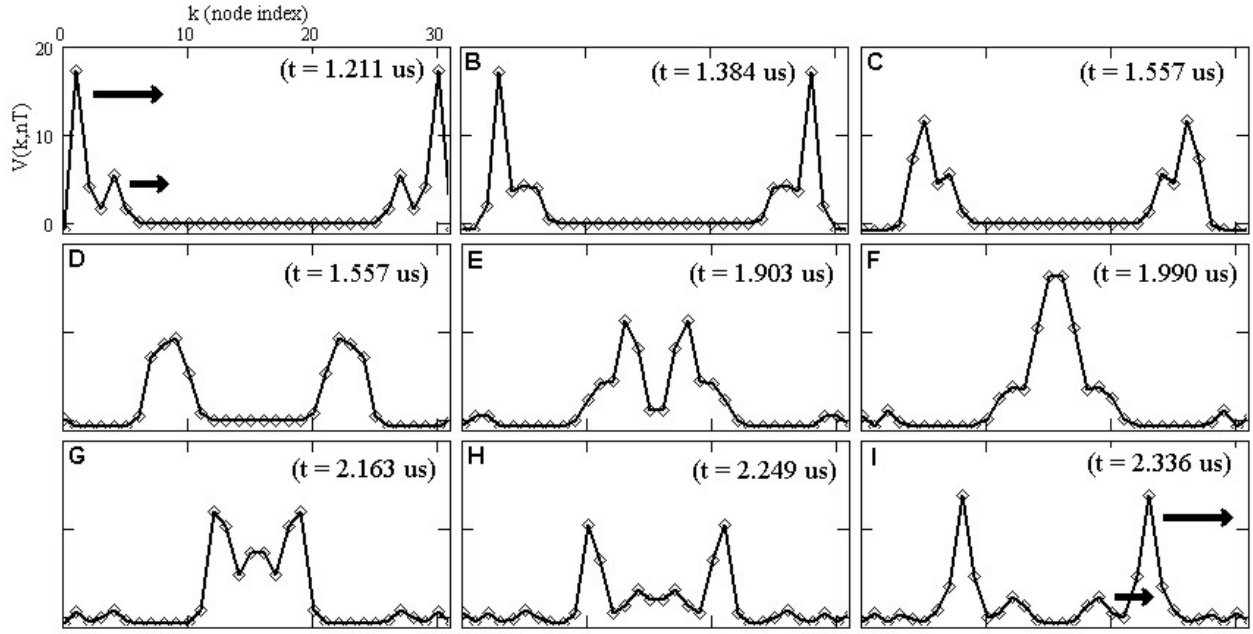


Figure 41: Collision among solitons moving with different speeds in the same direction (B,C,D,E), and in opposite direction (F,G,H). Finally, all solitons exit maintaining their original shape and speed (I).

discussed above. The excitations are represented by voltage source located at both the ends of the line. Termination resistors $R1$ and $R2$ are chosen in such a way energy reflections are minimized. The simulation results are illustrated in the next picture, showing how solitons with equal and opposite polarity interact. Collision is not elastic. Pheraps this is because the two solitons propagates in opposite direction.

Simulations had put in evidence the phase shift predicted in the KdV framework as a consequence of solitonic collisions. Velocity before collision is about $v_{before} \simeq 1.31 \times 10^7$ nodes per second. Velocity after collision is about $v_{after} \simeq 2.62 \times 10^7$ nodes per second, i.e. twice v_{before}. Another observed phenomenon is that solitons with greater amplitudes propagates more slowly if they have positive polarity, while solitons with greater amplitudes propagates more quickly if they have negative polarity. This is not a surprise because an increase in V_n means a decrease in $C(V_n)$. We would have had the opposite effect if we would have put D with anode at $-V_{POL}$.

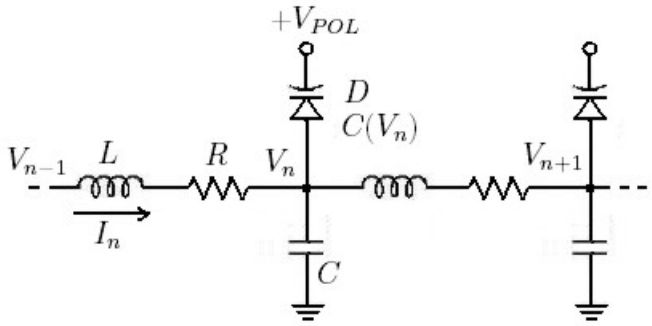


Figure 42:

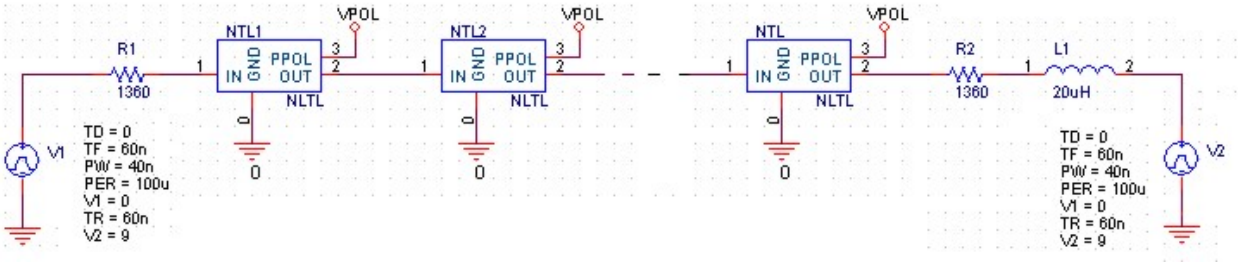


Figure 43: Schematics of the NLTL model.

17 Solitonic Machines (SM)

John von Neumann, concluded in 1945 that

> "*really efficient high-speed computing devices may, in the field of non-linear partial differential equations as well as in many other fields which are now difficult or entirely denied of access, provide us with those heuristic hints which are needed in all parts of mathematics for genuine progress.*" [Collected works of John v. Neumann, vol. V, 1963, p.1–32].

The present paragraph is devoted to the question of whether effective computation can be performed by the interaction of solitons in a bulk medium, linear, planar, or three-dimensional. Various media are possible, including optical fiber, Josephson junctions and electrical transmission lines. Such a structure is "gateless," that is, all computations are determined by an input stream of solitons. We should emphasize that using optical solitons in this way is quite different from what is commonly termed "optical computing", which uses optical solitons to construct gates that replace electronic gates, but which remains within the paradigm of laying out gates and wires.

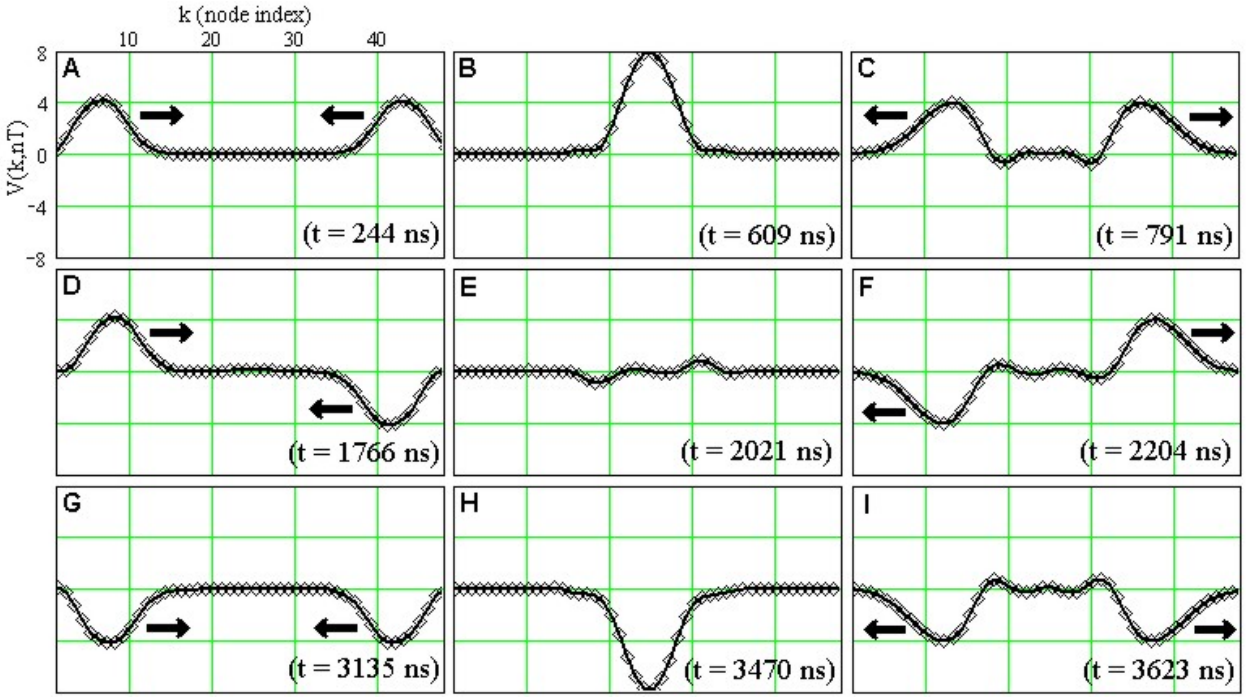


Figure 44: Collision among solitons moving in opposite direction with equal positive polarity (A,B,C), opposite polarity (D,E,F), and equal negative polarity (G,H,I).

The idea here uses a completely homogeneous medium for computation: the entire computation is determined by an input stream of particles.

This idea goes back at least to Steiglitz, Ken, Irfan Kamal, and Arthur H. Watson, (Embedding computation in one-dimensional automata by phase coding solitons, IEEE Transactions on Computers 37 - 1988), where solitons in a **Cellular Automaton** (CA) are used to build a carry-ripple adder.

The figure shows the hierarchy of computational systems in the world of CA. **Particle Machines** (PM) are CA designed to model particle-supporting physical media. **Soliton Machines** (SM) are restricted PMs that bring the abstraction a step closer to physical reality by modeling systems governed by certain well known PDEs, including the Klein-Gordon equation and log-NLSE. Both PMs and SMs are 1D CA that model motion and collision of particles in a uniform medium. **Oblivious Soliton Machines** (OSM) are SMs further restricted to model a class of integrable soliton systems, such as the KdV equation, cubic-NLSE, and SGE. Later we will show that a certain class of integrable PDE's can do only limited computation using SM; has been conjectured that this is true of all integrable equations. The simple behavior of integrable soliton systems makes them unlikely candidates for useful computing media.

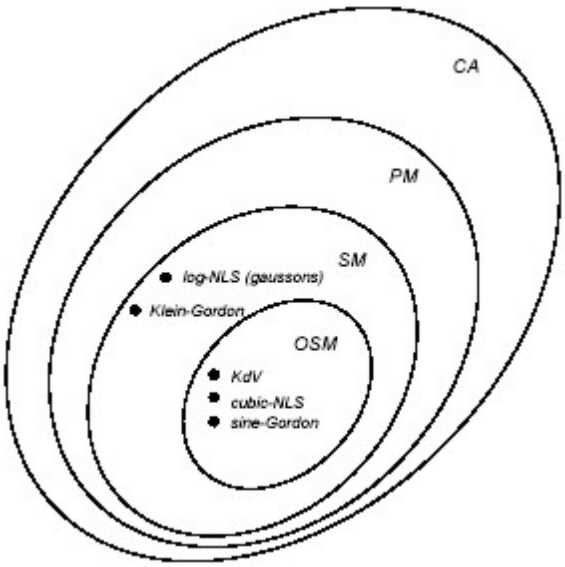


Figure 45: Hierarchy of computational systems in the domain of cellular automata (CA).

17.1 Particle Machines (PM)

The particle machine (PM) model of computation, introduced and shown to be universal by Squier, Richard K. and Ken Steiglitz (1994), is an abstract framework for computing with particles. The PM is a general model, not based on any specific physical system, but which tries to capture the properties of physical particles and particle-like phenomena. Briefly, a PM is a CA with a next-state rule designed to support a set of particles propagating with constant velocities in an infinite 1D medium. Two or more particles may collide; a set of collision rules specifies which particles are created, which are destroyed, and which are unaffected in collisions. A PM begins with a infinite initial configuration of particles and evolves in discrete time steps.

17.2 Oblivious Soliton Machine (OSM)

PM's general model exhibits much more general behavior than that exhibited by integrable systems. For example, the integrable soliton systems like KdV do not support the creation of new solitons or the destruction of existing solitons, and soliton state changes due to collisions are very limited. Thus, we adopt a restricted model of a PM called an Oblivious Soliton Machine (OSM). An OSM is a PM in which each particle has a constant **identity** and a variable **state** that are both vectors of real numbers. A particle's velocity is part of its identity. A typical state may consist of a phase and

a position relative to a *Galilean frame of reference*, whereas a typical identity may include an amplitude in addition to a velocity. No particles can be created or destroyed in collisions, and the identities of particles are preserved. A function of the identities (not states) of the colliding particles determines particle state changes. Immediately after a collision, particles are displaced. Displacement amounts are functions of the identities of the colliding particles, and particles must be displaced into distinct cells. In addition, we require that once two particles collide, the same particles can never collide again.

We refer to OSMs as **oblivious** because the state changes in an OSM do not depend on the variable states of colliding particles, but only on their constant identities. Oblivious collisions in the OSM model correspond to elastic collisions in the integrable PDEs discussed here; however, it is an open question whether or not all elastic soliton collisions in all integrable systems are oblivious. OSMs are not computation-universal, because it can be shown that the maximum time that an OSM can spend performing useful computation is cubic in the size of the input. This implies that computational OSM-based systems governed by the KdV and sine-Gordon equations are not universal, given that positions are used as state, and OSM-based systems governed by the cubic-NLSE are not universal, given that positions and phases are used as state. A conjecture exists, that *all* integrable systems using any choice of state are non-universal using the OSM model.

17.3 Soliton Machine (SM)

Intuitively, OSMs cannot compute universally because particles in an OSM do not transfer enough state information during collisions. We can make a simple modification to the OSM model so that universal computation becomes possible: we make the results of collisions depend on both the identities and states of colliding particles. In addition, we allow particle identities to change. We call the resulting model a soliton machine (SM).

Like an OSM, an SM is also a CA and a PM. The only difference between an SM and a PM is that no particles can be created or destroyed in an SM.

SMs with a quiescent background have at least the computational power of TMs with finite tapes. The question of whether such SMs are universal is open, however. Still, these SMs are more powerful than any OSM, since OSMs can only do computation that requires at most cubic time, while problems exist that require more than cubic time on bounded tape TMs. SMs with a periodic background can simulate an arbitrary TM, and are thus universal.

The class of algorithms that a finite-tape TM can implement depends on the specific function that bounds the size of the TMs tape; for instance, TMs with tapes of length polynomial in the input size can do any problem in PSPACE. Although not universal,

such TMs can do almost any problem of practical significance.

17.4 Nonintegrable Soliton Systems. The Log-NSE

The discussion suggest that we should look to nonintegrable systems for solitons that may support universal computation. It is an open question whether or not there exists such a soliton system.

Certain nonintegrable PDEs support soliton-like waves with behavior more complex than that of integrable solitons. Examples include PDEs such as the Klein-Gordon and logarithmically NSE (**log-NSE**). The solitons in these systems can change their velocities, as well as their phases, upon collisions, and new solitons may be created after collisions. Soliton collisions in nonintegrable systems may be inelastic or near-elastic; that is, colliding solitons can dissipate their energy by producing varying amounts of *radiation*, which erodes other solitons and may eventually lead to complete decay of useful information in the system. To our knowledge, it is an open question whether or not there exists a nonintegrable system with perfectly elastic, or non-radiating, collisions. It also appears to be an open question whether or not perfect elasticity implies obliviousness in any system. A system with collisions that are both perfectly elastic and nonoblivious would offer promise for realizing the SM model using solitons. The system we describe next, the log-NLS equation, has very near-elastic, non-oblivious collisions, and may support perfectly elastic, non-oblivious collisions as well.

The log-NSE, which supports solitons called **gaussons**, was proposed as a nonlinear model of wave mechanics. Gaussons are wave packets with gaussian-shaped envelopes and sinusoidal carrier waves. They are analogous to the wavefunctions of linear wave (quantum) mechanics; that is, the square of the amplitude of a gausson at a given point x can be interpreted as the probability that the particle described by the gausson is at x. Our numerical simulations of gausson collisions verify a published report that they range from deeply inelastic to near-elastic, and perhaps perfectly elastic, depending on the velocities of the colliding gaussons. An approximate range of velocities (the resonance region) has been identified for which collisions are apparently inelastic; outside this region, collisions are reportedly elastic. Three distinct velocity regions exist in which gaussons behave very differently: Depending on the region and gausson phases, gausson collisions can result in amplitude and velocity changes, radiation, or no apparent interaction. Gaussons with low velocities (region 1) offer the most promise for realizing useful computation, since their collisions appear both elastic and non-oblivious. Collisions of gaussons with higher velocities (regions 2 and 3) appear in general to be either oblivious or radiating, though for some combinations of velocities and phases, these collisions are non-oblivious and very near-elastic.

Even if we were to show that the log-NLS equation can be used for universal computation, we would still be left with a gap: *wWe know of no physical realization of this*

equation. But other nonintegrable nonlinear PDEs also offer possibilities for implementing SMs, and many of these do correspond to real physical systems. For example, the Klein-Gordon equation, the NLS equation with additional terms to model optical fiber loss and dispersion, and the coupled NLS equation for birefringent optical fibers all support soliton collisions with complex behavior potentially useful for encoding SM's. Optical solitons that arise from these more complicated equations exhibit gausson-like behavior, and are easily realizable in physical .bers; thus, such optical solitons may be particularly useful as practical means of computing using SMs.

The inelastic and near-elastic soliton collisions we observed in regions 1 and 2 are non-oblivious, thus leaving open the possibility of using them for computation in SMs. One problem with such a method is the potential connection between soliton stability and collision elasticity. We observed that inelastic collisions often resulted in radiation ripples emanating from collisions and eventual disintegration of gaussons in a cylindrical 1D system. In region 2, these ripples and the resulting instability may make the system unsuitable for sustained computation. The more inelastic the collisions, the more quickly the system decayed. However, we do not know if stability and elasticity are necessarily correlated in general, nor do we know if elasticity and obliviousness (and thus lack of computation universality) are related. In fact, collisions of region-1 solitons in the log-NLS equation appear to be both elastic and strongly non-oblivious.

We will discuss the computational power of the ideal machines with which we model physical systems. Being able to simulate a Turing machine (TM), or another universal model, is neither necessary nor sufficient for being able to perform useful computation. For example, certain PMs can perform some very practical regular numerical computations, such as convolution, quite efficiently, and yet these PMs are not necessarily universal. Conversely, simulating a TM is a very cumbersome and inefficient way to compute, and any practical application of physical phenomena to computing would require a more flexible computational environment. Nevertheless, universality serves as a guide to the inherent power of a particular machine model.

Intuitively, the key requirement for accomplishing this is that soliton collisions be nonoblivious; that is, solitons should transfer state information during collisions. All the well-known systems described by integrable PDEs such as the KdV, SGE, cubic NSE, and perhaps all integrable systems, are **oblivious** when displacement or phase is used as state. A cellular automaton (CA) model is presented, the **oblivious soliton machine** (OSM), that captures the interaction of solitons in systems described by such integrable PDEs. We then prove that OSMs with either quiescent or periodic backgrounds can only do computation that requires time at most cubic in the input size; and thus, are far from being computation-universal. Next, a more general class of CA is defined, **soliton machines** (SMs), which describe systems with more complex interactions. It is shown that an SM with a quiescent background can have at least the computational power of a finite-tape TM, whereas an SM with a periodic background can be universal. The search for useful nonintegrable (and nonoblivious) systems is

challenging: We must rely on numerical solution, collisions may be at best only near-elastic, and collision elasticity and nonobliviousness may be antagonistic qualities. As a step in this direction, it is shown that the logarithmically nonlinear Schrödinger equation (log-NLS) supports quasi-solitons (*gaussons*) whose collisions are, in fact, very near-elastic and strongly nonoblivious. It is an open question whether there is a physical system that realizes a computation-universal soliton machine.

18 Additional info

18.1 Spatial Solitons In Microcavities

Solitons are stable "lumps of light" for which the normal spreading due to diffraction and/or dispersion is balanced by an intensity-dependent (i.e. nonlinear) self-guiding effect. These "lumps" can be trapped between a pair of mirrors forming an optical cavity. In a thin, broad area, "microcavity", such solitons are ideal "bits" to represent images or optical information. This project will use a variety of models to study spatial solitons and related phenomena in semiconductor microcavities, and their possible application to information processing. Such microcavities are under intensive development both as commercial sources and as tests of quantum theories of light-matter interaction, and the group has substantial research funding related to solitons in such structures.

18.2 Principles of Dynamics

In this section we will review Classical Mechanics, in both the Lagrangian and Hamiltonian formulations.

The key idea is that we want to set up the equations of motion so that we obtain Newton's laws in an inertial frame; but the structure of the equations should not depend on the choice of coordinates used. The most powerful method for ensuring this is to write the equations as a variational principle. Then we are lead to the elegant and more general formulations of the theory due to Lagrange and Hamilton.

Let's consider the case of a simple particle of mass m moving in a unidimensional *configuration space* $\mathcal{C}$ with a kinetic energy $T(x) = \frac{1}{2}m|x|^2$, subjected to a potential field $U(x)$.

Lagrange's idea was that Newton's laws for this dynamics could also be written in variational form, as the condition for an extremum of the **action functional** (or *action integral*)

$$A(x) = \int_{t_0}^{t_1} \mathcal{L}(x, \dot{x}, t)\ dt \tag{18.2.1}$$

being $\mathcal{L} = T - U$ the **Lagrangian**. Indeed, a smooth curve $x(t) : [t_0, t_1] \to \mathcal{C}$ is a critical point of 18.2.1 if and only if it satisfies the **Euler-Lagrange equation**

$$\frac{\partial}{\partial x}\mathcal{L} - \frac{d}{dt}\frac{\partial}{\partial \dot{x}}\mathcal{L} = 0 \tag{18.2.2}$$

that reads Newton's 2nd law equation

$$m\ddot{x} = -\frac{dU}{dx} \tag{18.2.3}$$

In general 18.2.2 will be a second-order ODE for the path $x(t)$.

This approach can be extended to systems of many particles, perhaps with additional constraints. For instance a pendulum consists of a particle moving in the plane, constrained in such a way that its distance from a fixed point is constant. A 'rigid body' is a collection of many particles, subject to the constraints that the separation between each pair of particles is constant.

The extension to more than one dependent variable $\mathbf{x} = (x_1, \ldots, x_N)$ is straightforward. The action functional becomes

$$A(\mathbf{x}) = \int_{t_0}^{t_1} \mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}, t)\ dt \tag{18.2.4}$$

that leads to a N separate Euler-lagrange equations

$$\frac{\partial}{\partial x_i}\mathcal{L} - \frac{d}{dt}\frac{\partial}{\partial \dot{x}_i}\mathcal{L} = 0 \qquad i = 1 \ldots, N \tag{18.2.5}$$

The advantage of the Lagrangian approach is that the Euler-Lagrange equations always have the same form, so we are no longer restricted to Cartesian coordinates in inertial frames, but we are free to transform coordinates arbitrarily.

We can use the **Legendre transformation** to rewrite Lagrange's equations in the Hamiltonian's equations.

This approach can be extended to systems of many particles, perhaps with additional constraints.

Applications include the approximate description of motion near equilibria, and classification of equilibria according to their stability.

A dynamic system having $D = 1$ degree of freedom is said to be Hamiltonian iff it can be written in the form

$$\begin{cases} \dot{p} = -\frac{\partial H(p,q)}{\partial q} \\ \dot{q} = \frac{\partial H(p,q)}{\partial p} \end{cases}$$

where the *Hamiltonian* $H(p, q)$ is a $\mathbf{C}^2$ time-independent function. For such a system H itself is constant of the motion. It can be easily verified,

$$\dot{H}(p, q) = \frac{\partial H}{\partial p}\dot{p} + \frac{\partial H}{\partial q}\dot{q} = 0$$

This is equivalent to say that the system is conservative, i.e. the relative integral flow Φ^t is conservative (conserves the area).

sistemi hamiltoniani ad un grado di libertà sono integrabili nel senso che la soluzione si può esprimere mediante un algoritmo che include quadrature e funzioni implicite. Tutte le curve di livello regolari, semplici e chiuse della funzione hamiltoniana corrispondono ad orbite periodiche

spazio delle configurazioni: dove varia la variabile indipendente; spazio della fasi: dove si assegna la condizione iniziale.

I sistemi conservativi ad un grado di libertà sono integrabili, e forniscono i principali esempi di sistemi dinamici nonlineari per i quali si può descrivere in modo esplicito la soluzione, mediante formule finite, anche se contenenti quadrature e funzioni implicite o inverse. La descrizione delle proprietà qualitative delle soluzioni è non solo possibile, ma in molti casi facile; in particolare si possono discutere orbite periodiche, insiemi limite, separatrici.

sistemi hamiltoniani ad un grado di libertà sono integrabili nel senso che la soluzione si può esprimere mediante un algoritmo che include quadrature e funzioni implicite. Tutte le curve di livello regolari, semplici e chiuse della funzione hamiltoniana corrispondono ad orbite periodiche.

Le equazioni tipiche della meccanica - e di molti altri modelli matematici di problemi fisici - non si presentano nella forma di equazioni di Hamilton, ma sono espresse direttamente in termini di derivate prime e seconde delle coordinate. Tuttavia, se le equazioni possono essere ricavate dalla definizione di una energia cinetica ed una energia potenziale, esse possono sempre essere espresse nella forma detta di Lagrange. Le equazioni di Lagrange possono a loro volta, sotto certe condizioni, essere equivalenti ad equazioni di Hamilton.

Sommario Molti esempi di hamiltoniane che non sono del tipo semplice $H(p, q) = T(p) + V(q)$ si ottengono per trasformata di Legendre a partire dalle equazioni di Lagrange. I casi più notevoli sono quelli dei moti vincolati, sia di corpi puntiformi, sia di corpi dotati di momento d'inerzia non nullo (purché lo spazio delle configurazioni resti ad una dimensione).

18.3 Simmetries

A symmetry is a transformation which leaves the equation unchanged and transforms one solution into another solution.

Given a dynamical system, a **constant of motion** (or **integral of motion**) is a function that is constant time by time along any path which solves the Euler-Lagrange equations.

Given a Lagrangian (or Hamiltonian) system, for any constant of motion a symmetry exists of the Lagrangian equations. Due to Noether's theorem, each of the conserved quantities is related to a certain symmetry of the equation. Moreover, this result gives an explicit formula for an integral of motion given a symmetry of the Lagrangian.

For example, conservation of power is related to a symmetry relative to the phase shift. The conservation of momentum is related to a symmetry relative to a shift along a spatial coordinate axis. The conservation of the Hamiltonian (energy) is related to the symmetry of shifts along the time axis.

The evolution equations that govern solitons were found to be Hamiltonian and have infinitely many conserved quantities, pointing to the existence of many non-obvious symmetries.

18.4 Integrability

The notion of integrability is rigidly defined for Hamiltonian systems.

If a Hamiltonian system of D degree of freedom has D independent and mutually involutive integrals, i.e., if the system has D independent conserved quantities, and a symmetry corresponding to each, then the system of ODEs is integrable in the sense in which the system can be linearized and solved in terms of successive canonical transformations (or by quadratures), i.e. evaluating and inverting integrals. This is the main result in the Liouville-Arnold theory. For PDEs, there is no rigid definition determined yet. However there are candidates for integrability conditions of those systems. From studies on soliton equations, the following properties are now accepted as definitions of integrability for PDEs

1. Solvability by IST
2. Existence of N-soliton solution
3. Existence of infinite number of conserved quantities or symmetries
4. Existence of Lax pair
5. Existence of bilinear form

Generally it is not easy to obtain explicit solutions and conserved quantities for a given nonlinear equation. So we want to detect whether an equation is integrable or not beforehand. Thus the following integrability criteria have been proposed

- The Painlev´e test for ODE
- The Weiss-Tabor-Carnevale (WTC) method for PDE
- The singularity confinement test for discrete equation
- The algebraic entropy test for discrete equation

Those criteria are also used for deciding the values of parameters of an equation that has a possibility of integrability.

18.5 Applications

As well as having a rich mathematical structure, solitary waves (including solitons) have been obtained as solutions to nonlinear equations modelling a variety of circumstances. For example solitons have been derived in the following physical applications:

1. Water waves in channels, shallow water and the ocean (Scott Russell was a naval architect who designed the Great Eastern, a sailing ship; the design of which was partly based on his knowledge of the dynamics of solitons)
2. Lattice dynamics (e.g. the FPU model); waves in rods and strings
3. Electrical transmission lines
4. General relativity
5. Josephson junctions and superconductors
6. Liquid crystals
7. Optical fibres and telecommunications; nonlinear optics
8. Plasma physics
9. Protein dynamics and DNA
10. Quantum field theory
11. Stratified fluids
12. Rossby waves
13. Statistical mechanics

18.6 Discreteness

Discreteness appears in systems of interacting 'objects', which are separated in space. In crystals these 'objects' are atoms with the distance between them about $7\mathring{A}$, in photonic crystals the 'objects' are waveguides which are equally placed in space with the distance between them about $10\mu m$, in a network of Josephson junctions each junction is an 'object' and the characteristic linear scale between them is $50\mu m$, in a mechanical system of pendulums each of them can be considered as an 'object' and a characteristic distance could be order of 10 cm. The discreteness causes a cut-off of frequency (energy) spectrum. If, in addition to discreteness, the system is nonlinear there is a possibility to excite time periodic and spatially localized states - discrete breathers (DBs). They have been observed numerically and experimentally in a variety of systems.

18.7 Shock waves

Shock waves appear in many phenomena in everyday life. Most easily explained are shock waves coming from airplanes moving at supersonic speeds, or from explosions, but shocks also appear in phenomena involving much smaller velocities. Of particular interest is the flow of hydrocarbons in a porous medium, or to put it more concretely, the flow of oil in an oil reservoir. It is well-known that oil and water do not mix, and the interface between regions with oil and regions with water form what is mathematically defined as a shock. The dynamics of the shocks are vital in the exploitation of hydrocarbons from petroleum reservoirs. Even in everyday phenomena like traffic jams on heavily congested roads, we experience shock waves when there is an accumulation of cars. The shocks do not come from collisions of cars, but rather from a rapid change in the density of cars.

18.8 Differential Equations

Differential equations are nothing but *equations that involve derivatives*. It turns out that all the fundamental laws of nature can be expressed as differential equations, as the following list displays

- Gravitation (Newton's law)
- Quantum mechanics (The Schrödinger equation)
- Electromagnetism (Maxwell's equations)
- Relativity (Einstein's equations)
- The motion of gases and fluids (The Navier–Stokes's equations)

The motion of planets, computers, electric light, the working of GPS (Global Positioning System), and the changing weather can all be described by differential equations. For nonlinear equations the solutions to the equations do not in general form a vector space and cannot (in general) be superposed (added together) to produce new solutions. This further complicates matters.

In a sense, linearity means that the system is the sum of its parts. This allows us to make certain mathematical assumptions and approximations. In nonlinear systems these assumptions cannot be made, the system is more than the sum of its parts, which causes nonlinear systems to be extremely hard (or impossible) to model. In nonlinear systems one encounters such phenomena as chaos effects, strange attractors, and freak waves. Whilst some nonlinear systems and equations of general interest have been extensively studied, the vast majority are poorly understood if at all. Nonlinear systems are probably easiest understood as "everything except the relatively few systems which prove to be linear".

Consider the heat in the room where you are sitting. At each point (x, y, z) in space and time t we let $T = T(x, y, z, t)$ denote the temperature. By assuming that heat flows from hot to cold areas proportional to the temperature difference, that heat does not disappear (which means that the room is completely isolated from the surroundings), and that there are no heat sources, one can derive that the temperature distribution is determined by the so-called **heat equation**, which reads

$$T_t = T_{xx} + T_{yy} + T_{zz} \tag{18.8.1}$$

Here T_t means the derivative of the temperature with respect to the variable t , while T_{xx} denotes the derivative of the derivative, both with respect to the space variable x, and similarly for the remaining terms. Even simple problems give rise to difficult differential equations! Assuming that we know the initial temperature distribution, that is, we know $T = T(x, y, z, t)$ for $t = 0$, our intuition tells us that the temperature should be determined at all later times. This is called an **initial-value problem**. Ideally, when you are given a differential equation, you want the problem to be *well-posed* in the sense that

1. the problem should have at least one solution (existence)

2. the problemshould have no more than one solution (uniqueness)

3. the solution should be stable with respect to perturbations (stability)

The third condition states that a small change in the initial data should give a small change in the solution.

Unfortunately, differential equations normally do not possess solutions that are given by formulas, and so we must add to our "wish list" that we should be able to find a way to compute the solution. The problems are often very complex and require high speed computers to determine an approximate, or numerical, solution. Solutions of differential equations may be very complicated, and there is no unified mathematical theory that covers all, or most, differential equations. Most of the interesting differential equations are nonlinear, where the sum of two solutions is not a solution, which further complicates matters.

References

[1] K. Brauer. *The Korteweg-de Vries Equation:History, exact Solutions, and graphical Representation*, 2004. University of Osnabrück.

[2] E.Afshari and A.Hajimiri. *Non-Linear Transmission Lines for Pulse Shaping in Silicon.* California Institute of Technology (Caltech) Pasadena

[3] *Solitons on a Nonlinear Transmission Line.* 2001, Physics 303 - Physics Dept, UIUC.

[4] *Korteweg-de Vries Equation.* Encyclopedia on Nonlinear Science, edited by A.C. Scott.

[5] F.Banti. *Solitoni ed applicazioni nei plasmi.* Giugno 2002 - Univ. di Pisa.

[6] C.Foell and J.Bessette *Investigation of Solitonic Systems.* December 2004

[7] Clyde M. Davenport *The General Analytical Solution for the Burgers Equation. march 2005*

[8] Andrei N. Slavin Thresholds of Envelope Soliton Formation in a Weakly Dissipative Medium. 1996, Physical Review Letters, vol.77 - N.22

[9] Alwyn Scott Solitons, a brief history of.

[10] Peter S. Lomdahl What Is a Soliton?, in "Solitons in Biology". Spring 1984 - LOS ALAMOS SCIENCE

[11] Ben Kristoffen A New View of the Universe. Neuroscience Research Institute University of California, Santa Barbara

[12] Tuncay Aktosun Solitons and Inverse Scattering Transform. Contemporary Mathematics - Volume 00, XXXX American Mathematical Society

[13] Tuncay Aktosun Inverse Scattering Transform, KdV, and Solitons. Mathematics Subject Classification (2000): 35Q53, 35Q51, 81U40

[14] A.P.Fordy A Historical Introduction to Solitons and Backlund Transformations.

[15] Eric W.Weisstein Sine-Gordon Equation. From MathWorld - A Wolfram Web Resource.

[16] Mariusz H. Jakubowski, Ken Steiglitz Relative Computational Power of Integrable and Nonintegrable Soliton Systems. PhysComp96 Extended abstract. Draft, August 31, 2001

[17] R.Grimshaw, E.Pelinovsky Interaction of a solitary wave with an external force in the extended Korteweg-de Vries equation . November 27, 2001

[18] K.B.BLYUSS CHAOTIC BEHAVIOUR OF SOLUTIONS TO A PERTURBED KORTEWEG-DE VRIES EQUATION. REPORTS ON MATHEMATICAL PHYSICS Vol.49 (2002), p.29-38

[19] HELGE HOLDEN PETER D. LAX, ELEMENTS FROM HIS CONTRIBUTIONS TO MATHEMATICS. Norwegian Academy of Science and Letters. March 17, 2005.

[20] E.Fermi, J.Pasta, and S.Ulam, Studies of non linear problems, Document LA-1940, Los Alamos National Laboratory, May 1955.

[21] Koichi Kondo, Studies on Integrability for Nonlinear Dynamical Systems and its Applications, Master Thesis, Division of Mathematical Science of Osaka University, 2001.