A brief history of solitons and the KdV equation*

Alex Kasman

Department of Mathematics, College of Charleston, Charleston, SC 29424, USA

Soliton theory is an interdisciplinary area at the interface of mathematics and physics. It studies a special class of nonlinear partial differential equations (NLPDEs) having solutions that are waves which behave like particles. Amazingly, unlike most NLPDEs, we can write exact formulas for the solutions to these 'soliton equations'. This article is a review providing the historical context necessary to appreciate these spectacular developments, a brief overview of the early history of the field, and a list of references to consult for additional information.

Keywords: KdV equation, nonlinear partial differential equation, solitons, waves.

THERE are many different phenomena in the real world which we describe as 'waves'. Because of tsunamis, microwave ovens, lasers, musical instruments, acoustic considerations in auditoriums, ship design, collapse of bridges due to vibration, solar energy, etc., this is clearly an important subject to study and understand. Generally, studying waves involves deriving and solving some differential equations. As these involve derivatives of functions, they are a part of the branch of mathematics known to professors as analysis and to students as calculus. But, in general, the differential equations involved are tough to work with, that one needs advanced techniques to even get approximate information about their solutions.

It was therefore a big surprise in the 1960s and 1970s when it was realized for the first time that some of these equations were much easier than they first appeared. These equations that are not as difficult as people might have thought are called 'soliton equations' because among their solutions are some quite interesting ones that we call 'solitons'. The original interest in solitons was because they behaved a lot more like particles than we would have imagined. But shortly after that, it became clear that there was something about these soliton equations that made them not only interesting, but also too easy compared to most other wave equations.

The textbook *Glimpses of Soliton Theory*[1] was written to provide an elementary explanation of the mathematics responsible for these 'miracles'. The present article was adapted from that book (especially chapter 3) specifically for the readers of *Current Science*. Its purpose is to provide a brief overview of the history, scientific significance and mathematical structure of soliton theory. For additional details, please consult the reading resources listed at the end of this article.

The observation

In August 1834, Scottish ship designer John Scott Russell was sitting on his horse beside the Union Canal near Edinburgh and staring at the water when he saw something that would change his life.

> '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.'
>
> –J. S. Russell[2]

In other words, he saw a hump of water created by a boat on the canal and followed it for several miles. Certainly, other people had seen such waves before since the circumstances that created it were not particularly unusual. But, it may be that nobody before gave it such careful thought.

The point is that the wave he saw did not do what one might expect. Similar to waves in a swimming pool or at

*The article is adapted by the author from the book *Glimpses of Soliton Theory: The Algebra and Geometry of Nonlinear PDEs*, with permission of the publisher.

e-mail: kasmana@cofc.edu

the beach, one might expect a moving hump of water to either:

- Get wider and shallower and quickly disappear into tiny ripples like that of a wave generated in a swimming pool or
- 'Break' like the waves at the beach, with the peak becoming pointy, racing ahead of the rest of the wave until it has nothing left to support it and comes crashing down.

It was therefore of great interest to Russell that the wave he was watching did neither of these things, but basically kept its shape and speed as it travelled down the canal unchanged for miles. Being a ship designer, he must have been thinking 'Look at that wave go and go with just one little push. I wish that I could get a boat to do that!'

Terminology and backyard study

Russell used the words 'solitary wave' and 'wave of translation' to describe the phenomenon he observed that day. By 'solitary wave', he was clearly referring to the fact that this wave has only a single hump, unlike the more familiar repeating sine wave pattern that one might first imagine upon hearing the word 'wave'. Although this may not be quite what Russell intended, for our purposes 'translation' refers to the fact that the wave profile, i.e. the shape it had when viewed from the side, stays the same as time passed, as if it was a cardboard cutout that was merely being pulled along.

To study his solitary waves, Russell built a 30-foot long wave tank in his back garden. Among the most interesting things he discovered was that there was a mathematical relationship among the height of the wave, the depth of the water when at rest and the speed at which the wave travels. He believed that this phenomenon would be of great importance and so reported on it to the British Association for the Advancement of Science[2].

A less-than-enthusiastic response

Although we can say with hindsight that he was correct to have had high expectations for the future of the solitary wave, his ideas were not well received by the scientific establishment of his day. In particular, mathematical physicists George Biddell Airy and George Gabriel Stokes each argued that Russell's wave theory was completely inaccurate.

Perhaps Russell's real problem was that although he was clearly a great thinker, he had little expertise in mathematics. Aside from the relationship between wave height and speed reported above, he did not attempt any serious mathematical analysis of the phenomenon. Stokes and Airy, however, were experts in the use of differential equations to model wave phenomena. And, unfortunately, they both mistakenly believed that their analysis had demonstrated that Russell's theory was incorrect.

In his an article[3], Airy derives a different formula for the speed of a wave that he believed was in disagreement with Russell's and wrote: 'We are not disposed to recognize [Russell's Solitary Wave] as deserving the epithets "great" or "primary".'

Stokes wrote an article[4] about waves with a periodic profile (e.g. sine waves) and presented a formula for such a wave with infinitely many humps which he claimed 'is the only form of wave which possesses the property of being propagated with a constant velocity and without change of form – so that a solitary wave cannot be propagated in this manner. Thus the degradation observed by Russell is ... an essential characteristic of the solitary wave'.

Other known wave phenomena

Considering some of the mathematical analysis of wave phenomena that was known at the time provides an insight into why Stokes and Airy would have found Russell's observations difficult to believe.

Linear solitary waves

The equation

$$u_{xx} - u_{tt} = 0 \tag{1}$$

is so fundamental, it is often called 'the wave equation'. It was studied by Jean le Rond d'Alembert in the 18th century as a model of a vibrating string on a musical instrument. For example, it is possible to see the solution $u(x, t) = \sin(x + t) + \sin(x - t)$ to this equation shown in Figure 1 as being such a string tied to the x-axis at the points $x = 0$ and $x = \pi$ and viewing the graphs for different values of t as being like the frames in a movie.

As this is a linear differential equation, one can talk about having a basis of solutions and forming other solutions as linear combinations of them. Usually, trigonometric functions such as those above are being used to write solutions of this equation. Although it is not necessary to work upon with that basis, in order to compare this equation with the one shown in the next section below, let us try to form a solution out of a trigonometric basis which would look like the solitary wave that Russell observed on the canal.

For any value of the parameter k, the function

$$u_k^*(x, t) = \cos(kx + kt)$$

is a solution of eq. (1). Note that the speed with which the wave translates left is independent of the choice of the

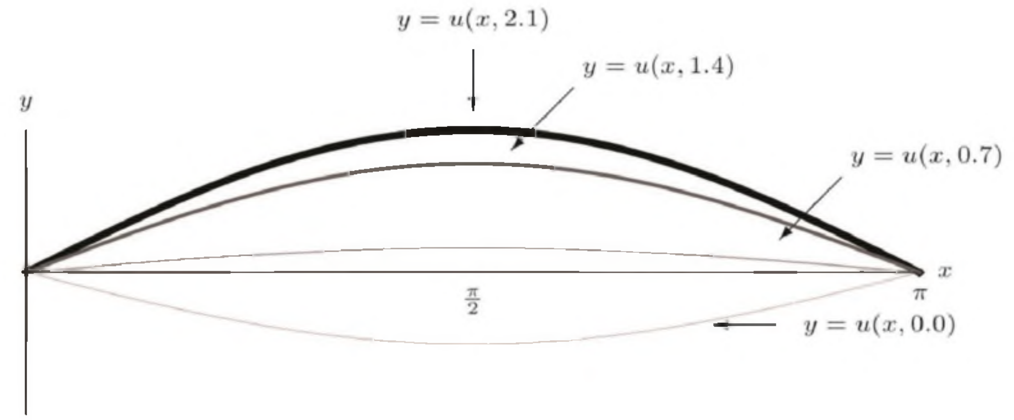

Figure 1. D'Alembert's wave eq. (1) models the dynamics of a vibrating string as a function $u(x, t)$ which gives the height of the string at horizontal position x and time t. By viewing a few different values of t (as shown above) it is possible to see how the string will move. Note that we are assuming $u(0, t) = u(\pi, t) = 0$ so that the string is π units long when at rest and fixed at the ends.

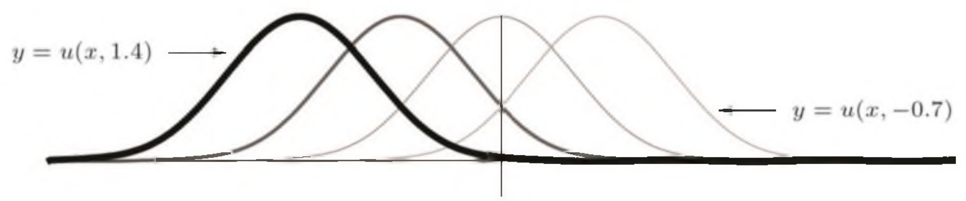

Figure 2. The solution (2) to eq. (1) looks like a single-humped wave translating to the left at constant speed even though it is a linear combination of cosine waves of different frequencies. For this to happen, it is important that the waves of different frequencies all move at the same speed.

constant k which determines the spatial frequency of the wave. The solution $u_1^*(x, t)$ is a wave that has one peak and one trough every 2π units while $u_2^*(x, t)$ has a peak and a trough in only π units, but an animation of either solution would show the solution moving to the left with constant speed one unit of space per unit of time regardless of this frequency.

As D'Alembert's wave equation is linear, any linear combination of these functions will also be a solution. As shown in Figure 2, the solution

$$u(x, t) = 0.25 + 0.352u_1^*(x, t) + 0.242u_2^*(x, t) + 0.130u_3^*(x, t) + 0.054u_4^*(x, t) + 0.018u_5^*(x, t), \quad (2)$$

when viewed on the interval $-3 \leq x \leq 3$ and $-0.7 \leq t \leq 1.4$ looks like a single-humped wave moving to the left at constant speed one. This particular choice of linear combination of cosine waves has the effect of nearly cancelling out to zero to form what appears on the graph to be a long flat stretch on either side of the hump. Because each component function $u_k^*(x, t)$ in the linear combination translates to the left at speed one, this property of cancelling out to form what looks like a single hump is preserved as time passes. It is precisely this interesting feature which will be altered in the example of the next section.

Linear dispersive waves

In contrast to the example of the previous section, consider the simple looking equation

$$u_t = u_{xxx}. \quad (3)$$

One can easily verify that it has solutions of the form

$$u_k^\Delta(x, t) = \cos(kx - k^3 t) = \cos(k(x - k^2 t)).$$

The initial profile of $u_k^\Delta(x, t)$ at time $t = 0$ looks exactly like $u_k^*(x, t)$; a cosine wave with frequency depending on k. However, since it is of the form $f(x - k^2 t)$, it will move to the right with constant speed k^2. The fact that the speed depends on the frequency is quite important, and so there is a technical term that reflects it; we say that eq. (3) is a dispersive equation.

The term 'dispersive' suggests things being spread out or dispersed, and that is exactly what it means here. A linear combination of different frequencies will separate as time passes and hence the coefficients selected to

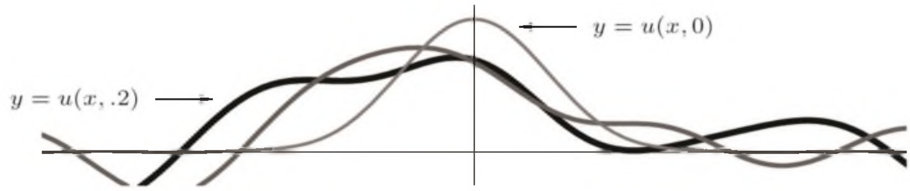

Figure 3. Because the different frequencies translate at different speeds, solution (4) to the dispersive wave eq. (3) looks like a clear single-humped wave at time $t = 0$ but degenerates into a mess by time $t = 0.2$.

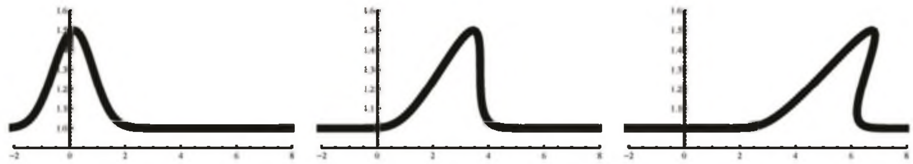

Figure 4. The dynamics of the initial profile $u(x, 0) = 1 + 0.5e^{-x^2}$ under the evolution of the Inviscid Burgers equation illustrates that even with a clear initial shape, problems such as a shock wave (in the centre) and 'multi-valued functions' (at right) can arise.

affect the shape of the graph such as in eq. (2) will not last long.

Observe what happens to the solution

$$u(x, t) = 0.25 + 0.352u_1^{\Delta} + 0.242u_2^{\Delta}$$
$$+0.130u_3^{\Delta} + 0.054u_4^{\Delta} + 0.018u_5^{\Delta}, \quad (4)$$

as time passes (Figure 3). The figure shows that even though it has the same clear single-humped shape at time $t = 0$, it quickly degenerates into a mess. (The figure shows the solution at times $t = 0$, $t = 0.1$ and $t = 0.2$.)

Breaking nonlinear waves

Many of the common features of nonlinear equations can be understood with the Inviscid Burgers' equation

$$u_t + uu_x = 0. \quad (5)$$

One important difference between this equation and those we have seen earlier, is that apart from such solutions whose initial profile is a straight line, we cannot find closed formulas for the solutions $u(x, t)$ to this equation. This is what generally occurs with nonlinear equations, even when they appear as simple as eq. (5). Consequently, various methods have been developed to explain the behaviour and dynamics of solutions to those equations without any explicit solutions.

The 'method of characteristics' is useful for figuring out the behaviour of solutions to some differential equations. The basic idea is that the behaviour is tracked along a curve (or 'characteristic') $x = c(t)$ in the xt-plane. With an appropriate choice of the curve, things work out effectively. In the case of eq. (5), for e.g. the method of characteristics shows that the initial profile of a wave will evolve in time so that its points shift to the right at a speed proportional to their heights. In particular, if the initial profile is the 'bell-curve' $u(x, 0) = 1 + 0.5e^{-x^2}$, problems arise as shown in Figure 4. As the highest point travels to the right at a higher speed than the lower points, it eventually catches up with them. This leads at first to a vertical 'wall' as seen in the middle image, officially known as a shock wave[5]. Continuing further the peak of the wave finally passes the lower points.

This is actually not an unrealistic set of pictures. This equation is a simple model of waves as they approach the beach, and so this 'wave breaking' phenomenon is the one to be recognized. However, despite the fact that these figures could be associated with a familiar physical phenomenon, they are mathematically troubling since the curves in the centre and right graphs of Figure 4 fail to satisfy the 'vertical line test'. In other words, these are not even functions.

Implications for Russell's solitary wave

As seen in Figure 2, we can find solutions to differential equations which take the form of a single-humped wave translating at constant speed. However, these solutions were for linear differential equations. One consequence of this linearity is that the solution can be multiplied by the constant 2 (thereby doubling its height) and it would still be a solution having the same speed. The fact that Russell claimed that the speed of his wave would depend on its height, clearly indicated that a mathematical model of the situation would necessarily be nonlinear, in which case it would be reasonable to expect the sort of distortion seen in Figure 4. Moreover, previous experience would have led Airy and Stokes to expect dispersion to be an important factor in the dynamics of water waves, in which case something like the 'mess' in Figure 3 would also be occurring at the same time. Between the distortion and dispersion, it is difficult to see how a properly shaped, translating, single-humped solution could possibly exist, and this is what Stokes and Russell tried to capture rigorously in their mathematical analysis.

As we will see shortly, such assumption would be at least partly correct. The distortion and dispersion that they would have expected are both present. However, their conclusion that this would eliminate the possibility of a solitary wave was incorrect. In fact, an appropriate combination of the two produces several surprising and unexpected results.

'The Great Eastern'

It is unfortunate that these two mathematicians erroneously rejected Russell's theory. Certainly, it must have disappointed Russell. It may have looked as if his interest in solitary waves was either misplaced or unappreciated. However, among ship designers he was well remembered for determining the natural travelling speed for a given depth (a result which grew directly out of his research on solitary waves) and for his work on what was at the time the largest moving manmade object, *The Great Eastern*. His obituary in the June 10, 1882 edition of *The Times* says:

> The first vessel on the wave system was called the Wave, and was built in 1835; it was followed in 1836 by the Scott Russell, and in 1839 by the Flambeau and Fire King. Mr. Scott Russell was employed at this time as manager of the large shipbuilding establishment at Greenock, now owned by Messrs. Caird and Co. In this capacity he succeeded in having his system employed in the construction of the new fleet of the West India Royal Mail Company, and four of the largest and fastest vessels – viz., was the Teviot, the Tay, the Clyde and the Tweed – were built and designed by himself ... The most important work he ever constructed was the Great Eastern steamship, which he contracted to build for a company of which the late Mr Brunel was the engineer. The Great Eastern, whatever may have been her commercial failings, was undoubtedly a triumph of technical skill. She was built on the wave-line system of shape ... It is not necessary now to refer to this ship in any detail. In spite of the recent advances made in the size of vessels, the Great Eastern, which was built more than a quarter century ago, remains much the largest ship in existence, as also one of the strongest and lightest built in proportion to tonnage.

It is especially interesting to note that in 1865, *the Great Eastern* was used to lay 4200 km of the transatlantic telegraph cable between Ireland and Newfoundland, which was the first electronic communication system between Europe and America.

The KdV equation

By the year 1895, Russell and Airy were both dead and George Gabriel Stokes was essentially in retirement. So, the controversy over Russell's wave was less emotionally potent, if not completely forgotten. It was at that time that a Dutch mathematician, Diederik Korteweg, and his student Gustav de Vries, decided to model water waves on a canal using differential equations. (Perhaps they were inspired by the fact that their home country of the Netherlands has so many canals!)

Beginning with the accurate but unwieldy Navier–Stokes equations, they made some simplifying assumptions including a sufficiently narrow body of water so that the wave could be described with only one spatial variable and constant, shallow depth as one would find in a canal. Putting all of this together, they settled on the equation[6]

$$u_t = \frac{3}{2} u u_x + \frac{1}{4} u_{xxx}. \tag{6}$$

Due to their initials, this important equation is now known as the 'KdV equation' (note 1).

It may be that the mathematical progress on understanding Russell's solitary wave was delayed until the appropriate mathematical techniques were available. The study of elliptic curves in the decades after Russell's original observation did not have any application in the study of water waves. However, it was by making use of results from this area of 'pure mathematics', that Korteweg and de Vries were able to derive a large set of solutions to eq. (6) which could translate and maintain their shape. Among these solutions were the functions

$$u_{\mathrm{sol}(k)}(x,\ t) = \frac{8k^2}{(e^{kx+k^3t} + e^{-kx-k^3t})^2} = 2k^2\ \mathrm{sech}^2(kx + k^3t), \tag{7}$$

which satisfy the KdV equation for any value of the constant k. This formula gives a translating solitary wave, like Russell's, that travels at speed k^2 and has height $2k^2$. For instance, in Figure 5 solutions $u_{\mathrm{sol}(1)}(x, t)$ and $u_{\mathrm{sol}(2)}(x, t)$ are compared side-by-side. Note that in each case the height of the wave is twice its speed.

Two things here should be surprising to those who created prejudices on differential equations, viz. they found an exact formula for various solutions to a nonlinear partial differential equation (NLPDE), and the

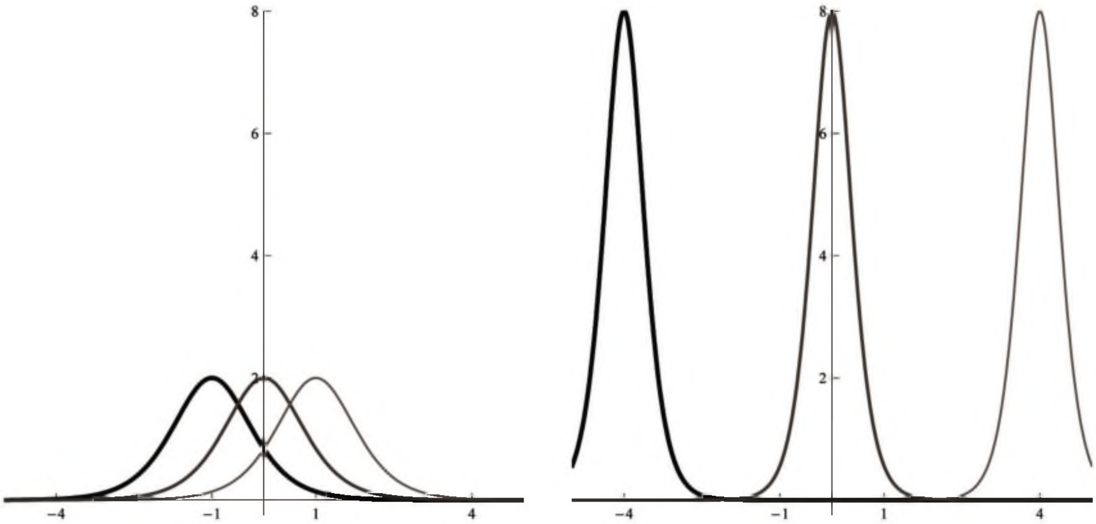

Figure 5. Two solitary wave solutions of the form of eq. (7) to the KdV eq. (6). The figure on the left shows solution with $k = 1$ and the right is $k = 2$. In each case, the figure illustrates the solution at times $t = -1$, $t = 0$ and $t = 1$. Note that the speed with which the wave translates is k^2 and that the height is twice the speed.

solutions could avoid distortion and dispersion despite Stokes' intuition to the contrary. Consider, for instance, that eq. (6) is an evolution equation which looks like a combination of two equations that we saw previously. The u_{xxx} term in the evolution eq. (3) resulted in the separation of different frequency components of a 'single-humped' initial profile, leading to its dissipation. On the other hand, the uu_x term in the Inviscid Burgers' eq. (5), for which we could not find explicit solutions, induced a nonlinear distortion in its solutions that destroyed any 'single-humped' initial profile. However, strangely, the combination of these two terms seems to avoid both of these problems.

It would be easy to dismiss these surprises as being mere coincidences, not worthy of further study, and this is likely the way that anyone interested in the solitary wave controversy might have reacted at the time. Specifically, the fact that solutions could be written explicitly was a consequence of the coincidence that the KdV equation bears some similarity to an equation related to elliptic curves. And, one might say that it is a coincidence here that the effects of distortion (from the uu_x term) and dispersion (from u_{xxx}) are perfectly balanced so they cancel out. However, it would be a long time before anyone realized that these were not mere coincidences. In fact, many more solutions to the KdV equation can be written exactly and have geometric origins, and the 'perfect balance' that allows the existence of a solitary wave solution to a nonlinear equation is not so rare as one might think.

Early 20th century

Researchers in the early 20th century showed little interest in the KdV equation or Russell's solitary wave. Thus, nothing directly related to this story occurred during this time. However, two tangentially related developments are worth mentioning.

The theory of physics underwent a major revolution during that period in the form of quantum mechanics. At the risk of oversimplifying a very complicated theory, let me say that quantum mechanics comes from two basic assumptions: that particles themselves are waves and that quantities that were previously thought of as numbers (such as 'speed') are actually operators, like the differential operators mentioned below. (See for e.g., an article by Terence Tao on the Schrödinger operator; http://www.math.ucla.edu/~tao/preprints/schrodinger.pdf.)

In the context of this article what matters is that there is lot of interest in waves that behave like particles and/or particles that behave like waves as this seems to be what the world is made of. In that sense, Russell's observation of an isolated wave that maintains its shape and speed – just as a hypothetical particle would do under its own inertia – could have been of interest to the scientists who created quantum physics, but that did not happen.

Another important base to the story of solitons is that the mathematical physicists treat differential operators like the 'Schrödinger Operator' $L = \partial^2 + u(x)$ as having a physical reality and not merely as formal mathematical notations. Among the other things done with them is to theoretically 'scatter' a wave off of them.

Also in the early 20th century, the British mathematicians Burchnall and Chaundy were doing their own research in which the numbers of the usual theories were replaced by differential operators. However, rather than doing concrete physics, they were working in one of the most 'pure' areas of math research: algebraic geometry[7].

As it turned out, the algebraic geometry of differential operators and scattering of waves off of $\partial^2 + u(x)$ became important parts of the theory of solitons in the second half of the 20th century.

Numerical discovery of solitons

Just as the first big mathematical advance towards understanding Russell's solitary wave had to wait until the theoretical machinery of the theory of elliptic functions was in place, the next big step required some actual machinery: the digital computer. In the 1950s, computers were not the user-friendly machines of today but were considered tools for mathematicians.

Among those doing 'numerical experiments' on these early computers were physicist Enrico Fermi and mathematicians John Pasta and Stanislaw Ulam at the Los Alamos National Laboratory, USA. Together with Mary Tsingou, they programmed a Los Alamos computer to obtain approximate solutions to nonlinear equations with the prescient intention of developing better intuition about nonlinearity. It was their assumption that if a nonlinear system was to start with a nice, ordered initial state, it would not take long before it was distorted and destroyed beyond recognition; but they wanted to see it happen in experiments on the computer. What they found surprised them. Just as Stokes and Airy were mistaken in their assumption that a nonlinear wave equation would necessarily destroy a nice single-humped initial state, the Los Alamos investigators were surprised to see that their intuitions were not confirmed[8]; or, as Ulam described it:

> Fermi expressed often a belief that future fundamental theories in physics may involve nonlinear operators and equations, and that it would be useful to attempt practice in the mathematics needed for the understanding of nonlinear systems ... . The results of the calculations (performed on the old MANIAC machine) were interesting and quite surprising to Fermi. He expressed to me the opinion that they really constituted a little discovery in providing intimations that the prevalent beliefs in the universality of 'mixing and thermalization' in nonlinear systems may not be always justified[9].

This mystery, that nonlinearity was seemingly better than expected, was known as the Fermi–Pasta–Ulam Problem and was described in an article published at Los Alamos. Because Los Alamos is the site of much classified work on nuclear weapons, the article was not officially distributed until the 1960s.

It was then that mathematicians Martin Kruskal at Princeton University and Norman Zabusky at Bell Labs, USA, conducted their own computer experiments[10]. Rather than considering a discrete system of connected vibrating masses as in the Fermi–Pasta–Ulam experiments, they wanted to consider a nonlinear wave equation. Taking the Fermi–Pasta–Ulam model and considering its continuum limit gave them such a nonlinear partial differential equation for a function $u(x, t)$. However, it was not a new equation; they had rediscovered the KdV eq. (6).

At this point, the existence of solutions in the form of eq. (7) was known. However, there was no reason to expect that any additional solutions could be written in an exact form. So, Kruskal and Zabusky conducted numerical experiments using computers. There were two interesting results from this study:

- If the initial profile was positive and 'localized' (if it was equal to zero everywhere except on one finite interval where it took positive values), then their experiments showed the solution breaking apart into a finite number of hums, each behaving like one of Russell's solitary waves, along with some 'radiation' which travelled away from them in the other direction. This would suggest that the solutions of eq. (7) play a fundamental role in describing a general localized positive solution to the KdV equation, similar to the way in which basic vibrating modes form a basis for solutions to D'Alembert's wave eq. (1). (Of course, they cannot actually form a basis for solutions as the equation is nonlinear and its solution set does not have the structure of a vector space!)
- Something interesting also happens when one views solutions that just appear to combine two different solitary waves (without 'radiation'). For these solutions (Figure 6), there are two humps each moving to the left with speed equal to half their height. However, it is not just a case of a sum of two of the solitary wave solutions found by Korteweg and de Vries; if the taller of the two humps is on the left, then they simply move apart. Here, it would be intriguing if we consider a situation in which a taller hump is to the right of a shorter one. As it is moving to the left at a higher speed it will eventually catch up. Intuition about nonlinear differential equations would have made an expert at the time realize that though the KdV equation has this unique property of having solitary wave solutions, when two solitary waves come together like this, the result would be a mess. One would expect that whatever coincidence allows them to exist in isolation would be destroyed by overlap and that the future dynamics of the solution would not resemble solitary waves at all. However, the numerical experiments of Kruskal and Zabusky showed hump shapes surviving the 'collision' and seemingly separating again into two separate solitary waves translating left at speeds equal to half their heights! Moreover, the same phenomenon occurred when three or more separate peaks were combined to form an initial profile where the peaks moved at appropriate speeds, briefly 'collide' and separate again.

The name 'solitary wave' coined by Russell more than one hundred years ago was intended to reflect the fact that these waves, unlike the periodic sine wave solutions generally considered at the time, had only a single peak.

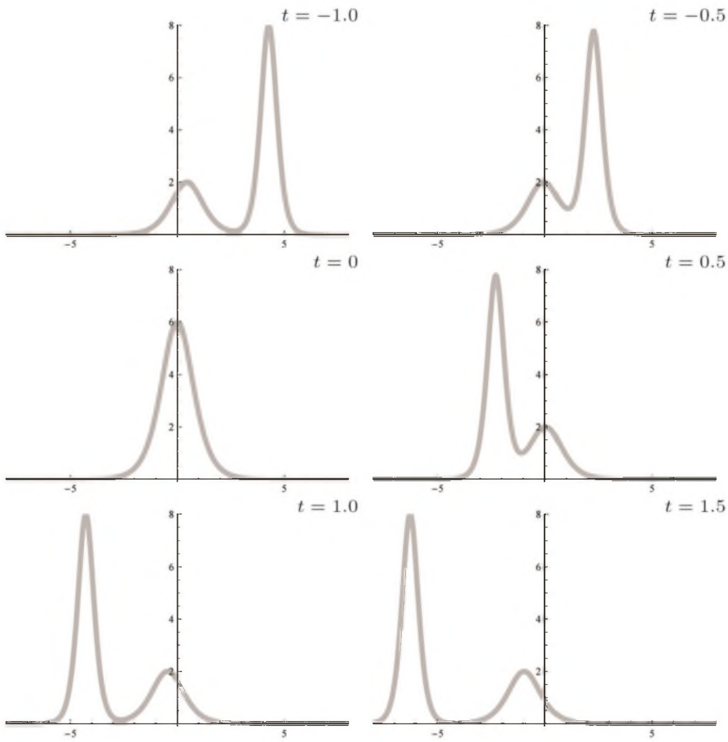

Figure 6. A solution to the KdV equation as it would have appeared to Kruskal and Zabusky in their numerical experiments. Note that two humps, each looking like a solitary wave, come together and then separate.

However, now seeing how gregarious they are, the name no longer seems appropriate. The term 'soliton' was used by Kruskal and Zabusky to describe these solutions, by combining the beginning of the word 'solitary' with an ending meant to suggest the concept of a fundamental particle in physics like a 'proton' or 'electron'.

More specifically, we now refer to the solitary wave solutions as 1-soliton solutions of the KdV equation. In general, an n-soliton solution of the KdV equation has n separate peaks (at most times). One can loosely refer to each of the separate peaks as being 'a soliton', even though they are part of the graph of the same function, similar to a local maximum in the graph of a polynomial.

For instance, Figure 6 illustrates a 2-soliton solution of the KdV equation in which a taller soliton travelling at speed 4 catches up to a shorter one with speed 1. Briefly, at time $t = 0$, we cannot see two separate peaks, but later again they separate so that we can clearly see a soliton of height 2 and another of height 8. However, this should not be mistaken to be the same as two 1-solitons viewed together. The next section will explore the ways in which the two solitons 'noticed' and affected each other as they met.

Hints of nonlinearity

As the KdV equation is nonlinear, there is no reason to think that the sum of two solutions would be a solution. In particular, the function $u^{\times}(x, t) = u_{\mathrm{sol}(1)}(x, t) + u_{\mathrm{sol}(2)}(x, t)$ shown in Figure 7 is not a solution to the KdV equation. However, if one were to watch an animation that shows its dynamics, one would have to look very closely to see how different it is than $u_{\mathrm{sol}(1,2)}(x, t)$, shown in Figure 6. These differences, though subtle, are quite important.

First, consider the graphs of $u^{\times}(x, 0)$ and $u_{\mathrm{sol}(1,2)}(x, 0)$. In both cases, only a single hump is seen in the graph of the function at that time. However, the height of the hump is different. Since $u^{\times}(x, 0)$ is the sum of peaks of heights 2 and 8, it has a peak of height 10. In contrast, Figure 6 clearly shows that $u_{\mathrm{sol}(1,2)}(x, 0)$ has a peak of height 6. This is one clear difference between the 2-soliton solution and the sum of two 1-soliton solutions.

More subtle is the fact that there is something slightly different about the positions of the peaks in the 2-soliton solution. Note that the shorter soliton is nearly centered on the y-axis at time $t = -0.5$. At time $t = 0$ one cannot see two separate peaks, but then at time $t = 0.5$ when the peaks have separated again, one still sees the smaller soliton nearly centered on the y-axis. In contrast, as the smaller peak in $u^{\times}(x, t)$ always moves to the left at constant speed 1 (note 2), it will have moved one unit to the left at the time interval $-0.5 \leq t \leq 0.5$.

Clearly there is some sort of nonlinear interaction going on in the 2-soliton solution. If solitons are considered as particles, they would have not simply passed through each other without any effect, but have actually 'collided' and in some sense the KdV equation incorporates their 'bounce'.

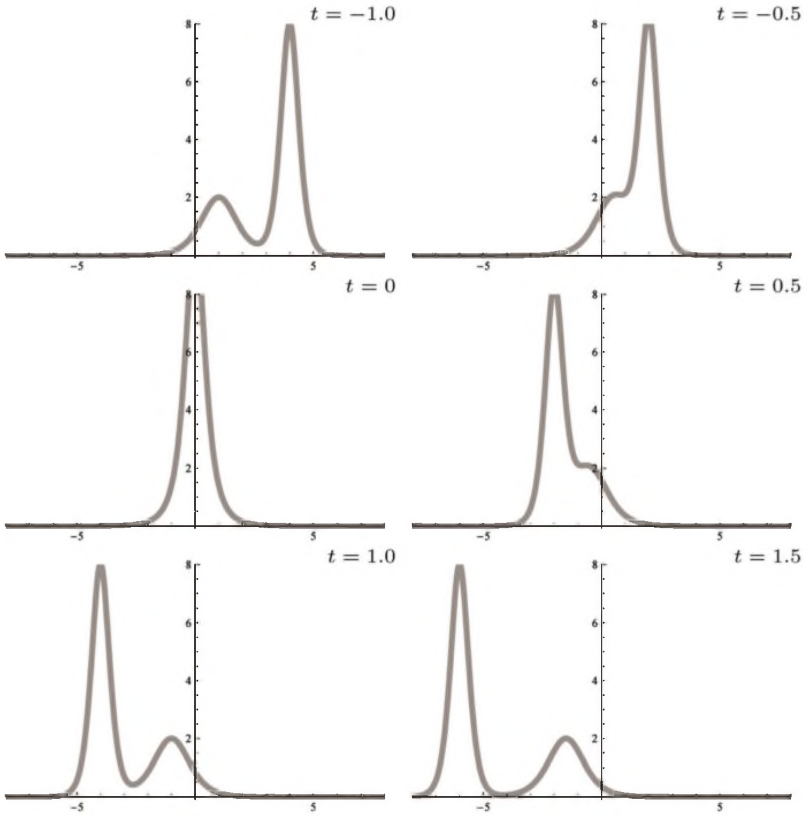

Figure 7. This is not a solution to the KdV equation. This is a sum of the one-soliton solutions $u_{sol(1)}(x, t)$ and $u_{sol(2)}(x, t)$. Compare this with Figure 6, which is a KdV solution, to see subtle differences despite the fact that each shows a hump moving to the left at speeds 1 and 4 respectively, at most times and a single hump centred on the x-axis at time $t = 0$.

Explicit formulas for *n*-soliton solutions

An interesting fact revealed in a later study by Gardner *et al.*[11], was that these n-soliton solutions of the KdV equation did not have to be studied in numerical simulations because it was possible to write exact formulas for them. For example

$$u_{sol(1,2)}(x, t) = \frac{24(e^{2x+2t} + 6e^{6x+18t} + 4e^{4x+16t} + 4e^{8x+20t} + e^{10x+34t})}{(1+3e^{2x+2t} + e^{6x+18t} + 3e^{4x+16t})^2}, \quad (8)$$

is an exact solution of eq. (6) and is shown in Figure 6.

This is quite surprising as it means that we have explicit formulas for a large and interesting family of solutions to this NLPDE. It is quite intriguing to note that the technique which was used to find those solutions is based on quantum mechanics (the theory in which particles have a wave-like nature). In that theory, some of the quantities which were numbers in previous theories of quantum physics were replaced by differential operators. To study the one-dimensional scattering problem of how an incoming wave $\psi(x)$ will 'bounce off' of another wave $u(x)$ (thought of as an obstacle), one is led to work with the differential operator $\partial^2 + u(x)$. Eventually, this shows that the n-soliton solutions $u(x, t)$ to the KdV equation have the property of being reflectionless for this scattering problem (for any value of t and any n-soliton solution). An additional property that turns out to be important is that they depend isospectrally on the variable t (i.e. the eigenvalues of the operator do not change in time). Pursuing this line of reasoning, Gardner *et al.*[11] were able to use a technique called inverse scattering to write exact formulas for the n-soliton solutions.

Soliton theory and applications

The KdV equation is quite interesting; despite being nonlinear and dispersive, it has solutions which avoid the damaging effects of dispersion and nonlinear distortion and maintain their clear, localized shapes indefinitely. These solutions have a certain 'particle-like' nature, which is contrary to our intuition of how waves ought to behave but might prove useful in understanding the behaviour of both waves and particles. Interestingly, the n-soliton solutions look almost like linear combinations of n 1-soliton solutions, suggesting that there might be some nonlinear analogue of the superposition principle for linear equations at work. Finally, unlike most nonlinear equations whose solutions can be studied only numerically or qualitatively, we can write explicit formulas for exact KdV solutions.

Soliton theory is a branch of mathematics which was developed to understand this phenomenon. Some of the big questions it addresses are: (i) Why is it that we can write several exact solutions to the KdV equation when we cannot do so for most nonlinear equations?. (ii) The relationship between the n-soliton solutions and n different 1-soliton solutions that it suggests that there is a way in which the KdV equation solutions can be combined. We know that they are not actually linear combinations and do not form a vector space. What is the method in which solutions are combined and can we give them a geometric structure analogous to the vector space structure for solutions to linear equations?. (iii) How can we identify other equations – either known already to researchers or yet to be discovered – that have these same interesting features?. (iv) What can we do with this new information?

The briefest possible answer to these questions is to note that the KdV equation has a hidden underlying algebraic structure that generic NLPDEs do not share and by understanding this structure we can find many different equations that share all of these features and thus deserve the name 'soliton equations'.

These equations often have physical significance as they model those phenomena which we encounter in the real world such as waves on a 2-dimensional surface like the ocean, light in optical fibre, electrons in a thin wire, transcription bubble in DNA, or energy transfer in proteins. In this sense, solitons have become tools of scientists and engineers for understanding the universe.

Soliton theory is also useful in mathematics. As Fermi predicted, it gives us a window into the world of nonlinearity. Previously, it was difficult to predict about a nonlinear situation. Now, we have a large set of nonlinear equations whose solutions can be studied explicitly (note 3). So, in some senses, the algebro-geometric structure of soliton equations allows us to use our knowledge of algebra and geometry to understand nonlinear differential equations better than we did before. However, soliton theory is also surprisingly useful in the other direction as well. That is, there are questions in algebraic geometry which have been answered using soliton theory.

Mathematics is sometimes seen as being divided into 'pure' and 'applied' subjects. The analysis of NLPDEs and especially the dynamics of waves, generally fall in the 'applied' side of this division while algebraic geometry is a part of the 'pure'. To some of us, it is endearing that each of these can inform us of the other in the intersection that is the soliton theory.

My textbook *Glimpses of Soliton Theory*[1] attempts to elaborate further on the answers to big questions (i)–(iii) at a level which would be accessible to any reader who has taken traditional undergraduate courses in linear algebra and multivariable calculus. If I achieved my goal, by the end of the book a reader should have a sense of satisfaction, much as one feels a certain thrill upon learning how a magician performed a particularly surprising trick. Of course, many other authors have written about this topic as well. After the 'epilogue', I have suggested some additional literature about solitons that can be consulted.

Epilogue

It was not only other researchers who were uninterested in the article by Korteweg and de Vries in the early 20th century, even Korteweg and de Vries themselves failed to show much interest in it. At the time, it must have seemed like a relatively minor result, not noticeable among the other important discoveries of Korteweg, and not important enough to interest de Vries who stopped doing research and became a teacher.

Both Korteweg and de Vries would be very surprised to learn what became of their one collaboration. I was inspired to look at their article on its 100th anniversary, and so in 1995 I found my way to a rarely used corner of the MIT library where the old journals were kept. There were shelves and shelves of journals from the late 19th century, all covered in dust. One volume stood out as its binding was clean, and when I took it off the shelf it fell open to the KdV article. Clearly, this article which attracted little attention when it was first published was of great interest one hundred years later.

Korteweg and de Vries are honoured in other ways that they probably would never have imagined. The mathematics institute in Amsterdam is called the 'KdV Institute', and one of the headings in the mathematics subject classification scheme is 'KdV-like Equations'.

One of the applications of soliton theory has also provided an ironic epilogue to the story of J. S. Russell and his interest in solitary waves. As they did in the 19th century, people have once again laid cables for communication between North America and Europe under the ocean. This time, of course, Russell's boat is not being used. However, Russell's work is still central to this newer effort at trans-Atlantic communication. The cables this time are not electronic but optical. The interesting point is that the information in the optical fibre is carried in the form of solitons – solitary waves of light. One can see why the property that Russell's wave on the canal 'kept on going' would be a useful feature for communication over such long distances. As the *Fiber Optic Reference Guide* puts it, 'The ability of soliton pulses to travel on the fiber and maintain its launch wave shape makes solitons an attractive choice for very long distance, high data rate fiber optic transmission systems[12]'.

Suggested reading

Consider consulting the following resources for additional information on this topic:

- Filippov's *The Versatile Soliton*[13] covers many historical facts which were left out of this brief summary.
- Bullough and Caudrey's historical analysis[14] appears in the proceedings of a conference honouring the 100th anniversary of the article by Korteweg and de Vries.
- The article on symmetries of solitons by Palais in *Bulletin of the AMS*[15] begins with a history of solitons before moving onto a more rigorous mathematical discussion.
- Fields Medalist, Sergei Novikov, wrote an article in Russian which was translated into English and provides a glimpse of the history of solitons from a Soviet perspective[16].
- Please consult the book by Remoissenet[17] for a more physical approach to this subject, including many laboratory experiments. This book also contains discussions of solitons in optical fibre and electrical circuits.
- A brief survey of the applications of the KdV equation, emphasizing those which had been confirmed by experiments as of 1995, can be found in the review article by Crighton[18].
- Of course, if you enjoyed this article please also take a look at the book *Glimpses of Soliton Theory*[1] from which it was excerpted.

Notes

1. In fact, the equation they wrote was not exactly as in the form of eq. (6). In particular, their equation had explicit parameters for various physical constants which have been eliminated here for convenience by a change of variables. Moreover, it should be noted that the history of mathematics is rarely as simple as it is portrayed in textbooks, and many would argue that this equation was not accurately named, as the equation and its connection to Russell's solitary wave were studied in earlier publications by another mathematician, Joseph Valentin Boussinesq[19,20].
2. Admittedly, the peaks in $u^{\times}(x, t)$ are not necessarily located in exactly the same places as the corresponding peaks in the two solitary wave solutions. However, if one takes this into account, then the apparent shifting of the expected locations of the peaks in the 2-soliton solution is actually worse, not better, so we will simply ignore it.
3. However, it should be noted that these are rather special nonlinear equations and so we should be careful not to over-generalize. Much more is possible in 'the nonlinear world' than we see through the window of soliton theory. Chaos theory, another important development of 20th century mathematics, provides a 'window' that looks at nonlinearity from the other side that gives a different view.

1. Kasman, A., *Glimpses of Soliton Theory: The Algebra and Geometry of Nonlinear PDEs*, Student Mathematical Library, American Mathematical Society, US, 2010, vol. 54.
2. Russell, J. S., Report on waves. Report of the Fourteenth Meeting of the British Association for the Advancement of Science, September 1844, London, York, 1845, pp. 311–390.
3. Airy, G. B., Tides and waves. *Encyc. Metrop.*, 1845, **192**, 241–396.
4. Stokes G. G., On the theory of oscillatory waves. *Trans. Cambridge Philos. Soc.*, 1847, **8**, 441–455.
5. Whitham, G. B., *Linear and Nonlinear Waves*, Pure and Applied Mathematics, A Wiley Interscience Series of Texts, Monographs and Tracts, John Wiley, 1973.
6. Korteweg, D. J. and de Vries, G., On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves. *Philos. Mag. Ser. 5*, 1895, **39**, 422–443.
7. Burchnall, J. L. and Chaundy, T. W., Commutative ordinary differential operators. *Proc. R. Soc. London A*, 1928, **118**, 557–583.
8. Fermi, E., Pasta, J. and Ulam, S., Studies of nonlinear problems. I. In *Nonlinear Wave Motion, Lectures in Applied Mathematics*, American Mathematical Society, US, 1974, vol. 15, pp. 143–145.
9. Ulam, S., *Collected Papers of Enrico Fermi*, University of Chicago Press, US, 1965.
10. Zabusky, N. J. and Kruskal, M. D., Interaction of solitons in a collisionless plasma and the recurrence of initial states. *Phys. Rev. Lett.*, 1965, **15**, 240–243.
11. Gardner, C. S., Greene, J. M., Kruskal, M. D. and Miura, R. M., Method for solving the Korteweg–de Vries equation. *Phys. Rev. Lett.*, 1967, **19**, 1095–1097.
12. Goff, D. R. and Hansen, K. S., *Fibre Optic Reference Guide: A Practical Guide to Communications Technology*, CRC Press, 2002, p. 260.
13. Filippov, A. V., *The Versatile Soliton*, Birkhäuser, Basel, 2000.
14. Bullough, R. K. and Caudrey, P. J., Solitons and the Korteweg-de Vries equation. *Acta Appl. Math.*, 1995, **39**(1–3), 193–228.
15. Palais, R. S., The symmetries of solitons. *Bull. Amer. Math. Soc.*, 1997, **34**(4), 339–403.
16. Novikov, S. P., Integrability in mathematics and theoretical physics: solitons. *Math. Intell.*, 1992, **4**, 13–21.
17. Remoissenet, M., *Waves called Solitons. Concepts and Experiments*, Springer-Verlag, Berlin, 3rd edn, 1999.
18. Crighton, D. G., Applications of KdV. *Acta Appl. Math.*, 1995, **39**(1–3), 39–67.
19. Boussinesq, J. V., Théorie de l'´ecoulement tourbillonnant et tumultueux des liquides dans les lits rectilignes a grande section, Gauthier-Villars et fils, 1897.
20. Pego, R., Letter to the editor. *Not. Math. Soc.*, 1998, **45**, 358.

ACKNOWLEDGEMENTS. I am grateful to Phoolan Prasad for suggesting the idea for this review article and for his editorial advice. In addition, I thank the American Mathematical Society (especially Erin M. Buck) for their permission to adapt this article from the copyright protected material in my textbook.

Received 5 September 2018; accepted 5 September 2018

doi: 10.18520/cs/v115/i8/1486-1496