Chapter 2

Developing PDE Intuition

We need to develop some intuition about differential equations, the behavior of their solutions and the structure of their solution sets. As we will see, in the case of linear differential equations there is a great deal that can be said in general about the set of solutions while no such general results about the solution sets of nonlinear equations exists. Furthermore, we will learn a bit about distortion and dispersion, two phenomena that can cause the shape of the initial profile to change and become essentially unrecognizable as time passes.

2.1 The Structure of Linear Equations

There is no doubt that there is a tremendous amount of *structure* associated with linear differential equations. This structure often allows us to solve the equations, to study the equations and to fully understand the range of possible solutions and also play a role in our understanding of real world phenomena that are described by linear equations.

2.1.1 Differential Operators I Linear equations have a nice property which allows us to write them in a compact and consistent format. Bring all of the terms involving the unknown function to the left side of the equal sign and all of the other terms to the other side. Then, instead of using any of the other notations for denoting derivatives, write the symbol ∂_x^n to the left of the name of the function to denote the nth derivative with respect to the variable x. The point is that now the equation looks like it involves products of things like ∂_x with things like the unknown function, but there is exactly one factor of the function in each term on the left and so it can be "factored out".

Applying this idea to the two linear differential equations from Section 1.1, we find

$$f = 2f'''(x) \qquad \text{can be written as} \qquad \left(-2\partial_x^3 + 1\right) f = 0$$

and

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} = x^2 + f \qquad \text{can be written as} \qquad (\partial_x + \partial_y - 1)\, f = x^2.$$

The objects in parentheses here are called *differential operators*, and the key point is that they have an independent existence. We can define $L = -2\partial_x^3 + 1$ and write the first equation as simply $L(f) = 0$, and the second equation is just $M(f) = x^2$ where $M = \partial_x + \partial_y - 1$.

Definition 2.1: A *differential operator* is a polynomial in the symbols ∂_{x_1}, ∂_{x_2}, …, ∂_{x_n} with coefficients that are functions of the variables $x_1, \ldots, x_n$:

$$Q = \sum_{i_1=0}^{m_1} \cdots \sum_{i_n=0}^{m_n} c_{i_1 \cdots i_n}(x_1, \ldots, x_n) \partial_{x_1}^{i_1} \cdots \partial_{x_n}^{i_n}.$$

Such an operator acts on a function in the sense that $\partial_{x_i}^m$ takes the m^{th} derivative of the function with respect to the variable x_i and the coefficients simply multiply the result:

$$Q(f) = \sum_{i_1=0}^{m_1} \cdots \sum_{i_n=0}^{m_n} c_{i_1 \cdots i_n}(x_1, \ldots, x_n) \frac{\partial^{i_1 + \cdots + i_n} f}{\partial_{x_1}^{i_1} \cdots \partial_{x_n}^{i_n}}.$$

It is this action on functions that makes us call them *operators*: they turn functions into functions. As is standard practice for linear operators, we call the set of all functions f for which $L(f) = 0$ the *kernel of L*.

Example 2.2 What is the result of applying the differential operator

$$Q = x^3y^2\partial_x\partial_y^2 + \cos(x)\partial_y^3$$

in the variables x and y to the function $f(x,y) = \cos(xy)$?

Solution

$$\begin{aligned} Qf(x,y) &= x^3y^2f_{xyy}(x,y) + \cos(x)f_{yyy}(x,y) \\ &= y^3\sin(xy)x^5 - 2y^2\cos(xy)x^4 + \cos(x)\sin(xy)x^3. \end{aligned}$$

As you see from our derivation above in which we factored out the unknown function from a linear differential equation, differential operators can be useful in writing such equations. In the language of differential operators, a linear differential equation is any equation

$$L(f) = m$$

for an unknown function f in terms of a fixed differential operator L and fixed function m.

We learn in calculus that differentiation distributes over addition of functions and through multiplication by constants, so that if f and g are functions and a and b are constants, then

$$(af + bg)' = af' + bg'.$$

We learned even earlier, in grade school, that multiplication is also distributive. It follows that differential operators are *linear* operators in that if you have two functions f and g and two constants a and b, then for any differential operator L the equation

$$L(af + bg) = aL(f) + bL(g) \tag{2.1}$$

is satisfied. In other words, *application of the operator distributes over addition and scalar multiplication of functions.*

The notation of differential operators is not just a convenience that allows us to write linear differential equations in a compact way. As it turns out, we will later be interested in learning how to do algebra with differential operators themselves: adding them to each other, multiplying them by each other, and even finding their "inverses" in a certain sense. This will provide us with one important clue as to what makes the soliton equations very special nonlinear equations.

However, until then, you will not be too far off if you think of it merely as a notational simplification that will make the proofs of the next section much simpler.

Caveat: Remember that $c_i(x)\partial_x^i$ applied to a function $f(x)$ is

$$c_i(x)\partial_x^i\,(f(x)) = c_i(x)\frac{d^i f}{dx^i},$$

the product of the coefficient function with the i^{th} derivative of f. The case when $i = 0$ then would involve the coefficient function multiplied by the zeroth derivative of f, which is the function f itself. However, many students get this wrong and simply add in the coefficient function of the term without any visible power of ∂_x, rather than adding in the product of this coefficient with the function to which the operator is being applied.

For example, what is the result of applying the operator $\partial_x - \frac{1}{x}$ to the function x? A student making the error described here would say "The first term is ∂_x applied to x, which is 1, and then the second term is just $-1/x$ so the answer is $1 - 1/x$." But, that is incorrect, in fact,

$$\left(\partial_x - \frac{1}{x}\right)(x) = (x)' - \left(\frac{1}{x}\right)x = 1 - 1 = 0.$$

The function x is in the kernel of the differential operator $\partial_x - 1/x$.

2.1.2 Linear Homogeneous Case: Superposition Principle

Recall that every linear differential equation can be written in the form

$$L(f) = m$$

where L is a differential operator and m is some function. If it happens that $m = 0$, then we say that it is a *homogeneous* differential equation. (It is "homogeneous" in the sense that every essential term in the equation has exactly one factor of some derivative of the unknown function f.) Otherwise, it is *inhomogeneous* (which would mean that there are nonzero terms in the equation which do not involve f). To begin with, we will consider the slightly simpler homogeneous case.

Theorem 2.3 If $f_1, \ldots, f_n$ are all solutions to the linear differential equation

$$L(f) = 0,$$

then the linear combination $F = a_1 f_1 + a_2 f_2 + \cdots + a_n f_n$ is *also* a solution to the same equation for any choice of constants a_i.

Proof To show that F is a solution, we apply the operator L to it and algebraically manipulate it until we obtain the value zero. We see that

$$\begin{aligned} L(F) &= L(a_1 f_1 + a_2 f_2 + \cdots + a_n f_n) \\ &\quad \text{(here we have replaced } F \text{ with its definition)} \\ &= a_1 L(f_1) + a_2 L(f_2) + \cdots + a_n L(f_n) \\ &\quad \text{(by the linearity of the operator } L\text{)} \\ &= a_1 \times 0 + a_2 \times 0 + \cdots + a_n \times 0 \\ &\quad \text{(since each } f_i \text{ is a solution)} \\ &= 0. \end{aligned}$$

□

The significance of this claim is quite stunning. Loosely put, it tells us that if we know any finite number of solutions to the equation $L(f) = 0$, then from them we can build *infinitely many* different solutions by scaling each of the known solutions and adding them together.

Even one nonzero solution f_1 of the equation can be used to produce infinitely many other solutions by simply scaling it without adding it to another solution. However, this gives only a *one-dimensional* space of solutions, while the linear combinations of n linearly independent solutions form an n-dimensional space of solutions.

2.1.3 Flashback: Linear Algebra At this point, you may wish to consult the notes or textbook from your linear algebra course to remember that a vector space is a set with certain properties, including closure under addition and scalar multiplication. In those terms, what we have noted above is merely the fact that the solution set of the homogeneous differential equation $L(f) = 0$ is itself a vector space (with the solutions being the "vectors"). Also, in terms of linear algebra, this fact is not at all surprising since we can also describe the same set of functions as being the *kernel* of the linear operator L, and it is well known that the kernel always has the structure of a vector space.

The *dimension* of a vector space is the size of the smallest subset of vectors with the property that any element of the space can be written as a linear combination of them. In the case of a homogeneous, ordinary, linear differential equation there is a very useful fact[1] relating the order of the operator to the dimension of the solution space:

Theorem 2.4 If $L = \sum_{i=0}^{n} c_i(x)\partial_x^i$ (with $n > 0$ and $c_n \not\equiv 0$), then there exist linearly independent functions $f_1, \ldots, f_n$ such that the set of all solutions to the equation $L(f) = 0$ is the n-dimensional space of functions which they span.

Example 2.5 What differential operator L has as its kernel the solution set of the differential equation $f'' = -f$? Write a formula which gives the *general* solution to the equation and verify that it is indeed in the kernel of L.

Solution To find the differential operator, we pull all of the terms involving f over to the left side of the equation, write the derivatives in the notation[2] involving ∂, and "factor out" the common fs:

$$f'' + f = 0 \Rightarrow \left(\partial^2 + 1\right) f = 0.$$

So, the differential operator we want is $L = \partial^2 + 1$.

Since the operator has *order* 2, we know that its kernel should be two-dimensional. Thus, we only need to find two linearly independent solutions to the equation $f'' = -f$ to obtain a basis for the entire solution space. In fact, we know of two solutions. Since

[1] A proof of this theorem appears in most textbooks on ordinary differential equations. See, for example, Theorem 5.2 in Boyce and DiPrima *[17]*.

[2] When writing an *ordinary* differential operator, one involving differentiation with respect to a single variable only, it is common to leave off the subscript notation and simply write ∂^n. In particular, here it should be understood that "∂" means "∂_x".

$\frac{d^2}{dx^2}\sin(x) = -\sin(x)$ and $\frac{d^2}{dx^2}\cos(x) = -\cos(x)$ each is a solution. They are linearly independent because there is no constant a such that $a\cos(x) = \sin(x)$ (for all x). Consequently, the general solution of the equation is

$$f(x) = a\cos(x) + b\sin(x)$$

where a and b can be chosen to be arbitrary constants. (In other words, $\sin(x)$ and $\cos(x)$ form a *basis* for the entire solution set of this equation.)

Finally, we verify that this formula does indeed give us a function in the kernel of L regardless of the choice of the constants:

$$\begin{aligned} L(f) &= aL(\cos(x)) + bL(\sin(x)) \\ &= a(-\cos(x) + \cos(x)) + b(-\sin(x) + \sin(x)) \\ &= 0. \end{aligned}$$

Caveat: The case of homogeneous, linear, ordinary differential equations is especially nice since then we get a finite-dimensional vector space as the solution set. However, if any of these characterizing adjectives is changed (e.g., to inhomogeneous, nonlinear, or partial), then the situation becomes anywhere from a little to a lot more complicated. In the next section, we will look at the inhomogeneous case which – as you may recall from linear algebra – requires that we "shift" the vector space of solutions. In the section after that, we will see that moving to partial differential equations puts us in a situation that is quite similar, but for which the solution set is *infinite*-dimensional and so a basis of infinitely many functions must be used for the general formula. Finally, when we move to the nonlinear case, we will see that there is generally no reason to expect any sort of nice structure for the solution space whatsoever.

2.1.4 Linear Inhomogeneous Case: Structure Just as in the homogeneous case, there is a theorem which characterizes the structure of the solution set of an inhomogeneous linear equation and which allows one to make new solutions out of old ones:

Theorem 2.6 If f_1 and f_2 are both solutions to the linear equation $L(f) = m$, then $f_3 = \lambda f_1 + (1 - \lambda) f_2$ is also a solution for any value of the constant λ. Or, more generally, if $f_1, \ldots, f_k$ are solutions to $L(f) = m$, then

$$F = \sum_{i=1}^{k} \lambda_i f_i$$

is a solution to the same equation so long as $\sum_{i=1}^{k} \lambda_i = 1$.

(The proof of Theorem 2.6 is left as a homework exercise. See Problem 4 at the end of this chapter.)

Note that this is a bit more restrictive than the homogeneous case in which an arbitrary linear combination works. We can think of it geometrically by thinking of the set of points $(\lambda_1, \ldots, \lambda_k)$ in k-dimensional vector space which lie on a hyperplane. In fact, all of this should sound quite familiar, because you probably learned in your linear algebra course that the set of vectors $\vec{x}$ satisfying the equation $A\vec{x} = \vec{b}$ for a given matrix A and vector $\vec{b}$ is a subspace if $\vec{b} = 0$ and a subspace shifted by a fixed vector away from the origin in general.

2.2 Examples of Linear Equations

2.2.1 Vibrating Strings and D'Alembert's Wave Equation

One of the applications usually seen in a course on partial differential equations is the derivation of an equation that models the dynamics of a vibrating string, like the string on a guitar. This is done by considering a function $u(x, t)$ so that for each fixed value of the time variable t the graph of $y = u(x, t)$ on $0 \leq x \leq \pi$ represents the position of the string at that particular instant. (See Figure 2.2-1.) Then by working out the forces that would be felt by each tiny piece of the string being pulled by the neighboring pieces (and making a few simplifying assumptions and considering x and t to be in the appropriate coordinate for distance and time), it is possible to reach the conclusion that u will satisfy the linear equation

$$u_{xx} - u_{tt} = 0. \tag{2.2}$$

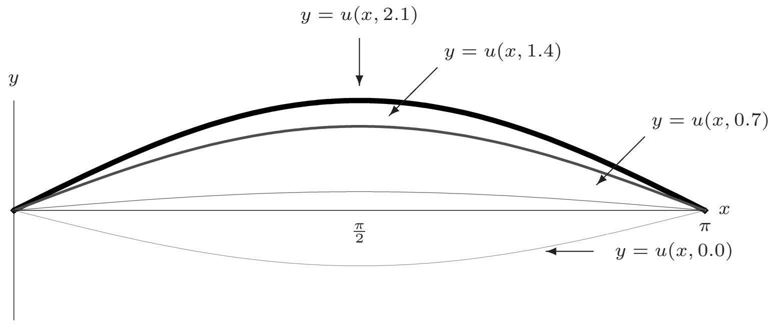

Figure 2.2-1: D'Alembert's Wave Equation (2.2) models the dynamics of a vibrating string as a function $u(x,t)$ which gives the height of the string at horizontal position x and time t. By viewing a few different values of t (as shown above) it is possible to see how the string will move. Note that we 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.

We will not worry about this derivation here, but will focus on the solutions of the equation and how we can combine them to produce new solutions. (As a result of the assumptions made, the solutions will not dampen over time as do the strings on a real guitar. Moreover, it is not an accurate model if the amplitude of the vibration is very large.)

As will be verified in homework Problem 1 at the end of this chapter,

$$u_k(x,t) = \sin(kx + kt) + \sin(kx - kt) \tag{2.3}$$

is a solution to the equation for every value of k and, moreover, when k is a positive integer then $u_k(0,t) = u_k(\pi,t) = 0$ for all values of t. (The latter property corresponds to the fact that the string on a guitar is fastened at the ends and therefore is unable to move.)

So, let us take advantage of *Mathematica* to watch a "movie" of these solutions to the wave equation. Create the definition

```
u[k_,x_,t_]:=Sin[k*x+k*t]+Sin[k*x-k*t]
```

so that we can easily make any one of the solutions for arbitrary values of the parameter k. Using `MyAnimate` to plot the function `u[1,x,t]` in the viewing window $0 \leq x \leq \pi$ and $-3 \leq y \leq 3$ for $0 \leq t \leq 2\pi$ shows a "string" which seems fixed at the ends but which vibrates up and down periodically. Doing the same for `u[2,x,t]`, and `u[3,x,t]` we see a line again fixed at the ends, but now there are additional fixed points in between where we see a vibration similar to the vibration seen for `u[1,x,t]`.

In general, `u[k,x,t]` looks like a simple vibration of a string that is fixed at the ends and at `k` additional *nodes* (points where the string remains stationary) in between. They are the simple harmonic solutions and if a string vibrated in one of these patterns it would produce a very pure and simple sound. But, if you simply pluck a guitar string, something more complicated happens.

We can easily make more complicated solutions by taking *superpositions* (or *linear combinations*) of these functions to create other solutions. Try, for example, animating the solution

```
u[1,x,t] + .4 u[2,x,t] -.2 u[5,x,t]
```

on the same viewing window and time interval as before. This also is something that a guitar string can do. In fact, a plucked guitar string is made up of a combination of simple harmonic vibrations very much like the one in this example.

It is not just stringed instruments which work in this way. If we imagine the function as representing the sound wave produced by the vibrating string rather than as the string itself, then it is a fundamental principle that all musical tones are built up as combinations of these harmonics. Choosing appropriate coefficients on a linear combination of such pure tones makes it possible to reproduce the sound wave corresponding to a trumpet or a clarinet rather than a guitar string.

The same idea can be applied in reverse, breaking a sound wave down into its component frequencies. This is what is known as Fourier Analysis[3], and it is an important mathematical technique in many engineering applications. This decomposition of a solution into its component harmonics can be illustrated in practice with an actual guitar. By lightly touching a plucked string at a point which is a node

[3] *Fourier Analysis* is the theory that studies building functions out of linear combinations of trigonometric functions. It is a fundamentally *linear* theory in that the trigonometric functions form an orthonormal basis for a vector space of solutions, and the coefficients needed to obtain any given function as such a combination can then be obtained by taking a product with individual basis elements. (This is how the mysterious decimal coefficients in (2.4) below are obtained!) It is almost exactly like what you would have seen in a standard linear algebra course, except for the fact that the basis is infinite and that the product involves integration of functions rather than multiplying of vectors.

for one component of the linear combination, it is possible to dampen out all of the other components and hear the pure tone associated with just a single solution `u[k,x,t]`. Most guitarists know this technique and call it "playing a harmonic".

The point of this example is to illustrate the importance of the structure of the solution space to a homogeneous equation. One can produce infinitely many solutions by taking arbitrary linear combinations of just a few known ones. Furthermore, if we were willing to take *infinite* sums of the functions $u_k(x,t)$ with coefficients getting smaller at a rate that would allow the sum to converge, we would in fact have *all* of the solutions to the equation (2.2) which have this "vibrating string" property of being fixed on the x-axis at $x = 0$ and $x = \pi$. It is in this sense that the solution set forms an *infinite-dimensional* vector space with the functions $u_k(x,t)$ being a basis.

2.2.2 Traveling Waves and D'Alembert's Wave Equation If we relax the condition that the solution needs to be zero at $x = 0$ and $x = \pi$ we can see other solutions to the same equation (2.2). Note, for example, that

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

is also a solution[4]. Since it is of the form $f(x + \alpha t)$ where $f(z) = \cos(kz)$ and $\alpha = 1$, we know from homework question 4 in the last chapter that this solution represents a cosine wave moving to the left at *constant speed one.*

Note that this speed with which the wave translates left is independent of the choice of the 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.

In contrast, in the next section we will see a different equation for which the frequency of the solution does determine the speed of translation. The difference between these two scenarios and how

[4] We will continue to try to write our solutions as combinations of trigonometric functions, even though one can check that any function of the form $f(x \pm t)$ is a solution to this equation. This is because, as explained in the previous footnote, such a decomposition of solutions to linear equations into different "harmonics" is a fundamental technique and better contrasts this example with the one that is to appear in the next section.

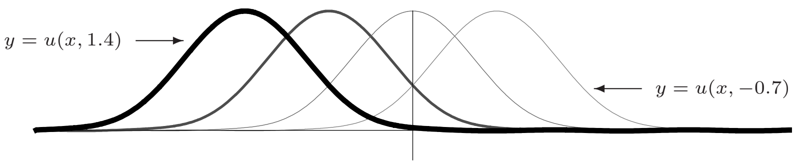

Figure 2.2-2: The solution (2.4) to equation (2.2) looks like a single-humped wave translating to the left at constant speed even though it is a linear combination of cosine waves of different frequencies. For this to happen, it is important that the waves of different frequencies all move at the same speed.

they affect the dynamics of the shape of the graph are important components of the intuition that you should be developing at this point.

As you may know if you have flipped ahead and looked at the pictures, solutions which have a single peak and no trough are going to be very important later in this book. We can make a solution that looks almost like that as a superposition of the functions u_k^* for a few choices of k. (Again, we could do better if we were willing to consider more terms in the sum or even infinite sums, but this approximation is sufficient to demonstrate the main idea.)

Observe in Figure 2.2-2 that the solution

$$\begin{aligned} u(x,t) = .25 + .352u_1^*(x,t) + .242u_2^*(x,t) + .130u_3^*(x,t) \\ +.054u_4^*(x,t) + .018u_5^*(x,t) \end{aligned} \tag{2.4}$$

when viewed on the interval $-3 \leq x \leq 3$ and $-.7 \leq t \leq 1.4$ looks like a single-humped wave moving to the left at constant speed one. I say "looks like" because viewing it on even a slightly larger piece of the x-axis reveals that there are other peaks on the solution that we cannot see. Nevertheless, note that 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. This may surprise you, if you have not previously seen that trigonometric functions can be combined to produce shapes that do not look obviously trigonometric. However, it should *not* surprise you that this shape simply translates to the left at constant speed. Since 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 nice feature which will be altered in the example of the next section.

2.2.3 A Dispersive Wave Equation Imagine that you and your friends stood out in a field and positioned yourselves so that you spelled out a word when viewed from above. If you all moved in the same direction at the same speed, then to an observer in a helicopter, this word would appear to "travel" across the field. On the other hand, if you each moved at a different speed, then the word would only be visible briefly and would quickly degenerate into a "mess" to the observer. This latter situation is what we will see in this section, when we study an equation for which the speed of translation of a trigonometric profile depends upon its frequency.

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

$$u_t = u_{xxx}. \tag{2.5}$$

You can easily verify that it has solutions of the form

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

The initial profile of $u_k^\star(x,t)$ 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 $f(x - k^2 t)$ we know from Problem 4 that 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 equation (2.5) 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, so that a careful choice of coefficients to affect the shape of the graph such as we made in (2.4) will not last long.

Observe what happens to the solution

$$u(x,t) = .25 + .352u_1^\star + .242u_2^\star + .130u_3^\star + .054u_4^\star + .018u_5^\star \tag{2.6}$$

as time passes. Figure 2.2-3 shows that even though it has the same nice single-humped shape at time $t = 0$ it quickly degenerates into a mess. (The figure shows the solution at times $t = 0$, $t = .1$ and $t = .2$.)

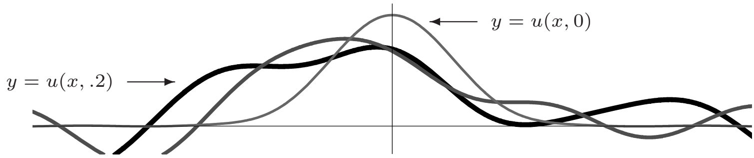

Figure 2.2-3: Because the different frequencies translate at different speeds, solution (2.6) to the dispersive wave equation (2.5) looks like a nice single-humped wave at time $t = 0$ but degenerates into a mess by time $t = .2$.

2.3 Examples of Nonlinear Equations: Not Quite as Nice?

Whether linear equations are "easy" may be subjective, but the results of the preceding sections clearly indicate that there is always a great deal of *structure* to their solution sets and this certainly helps in finding and understanding solutions. In contrast, there are no comparable general results about the solution sets of nonlinear differential equations.

For example, if we tell you that the functions f_1 and f_2 are two solutions to a linear differential equation, then you know that $\frac{1}{4}f_1 + \frac{3}{4}f_2$ is also a solution without having to know anything in particular about the equation or the solutions. But, having two solutions to a nonlinear differential equation does not in general give you any way to produce even one more solution, let alone the infinitely many that we can produce for linear equations.

2.3.1 Waves that Shock We can encounter many of the common features of nonlinear equations with a famous example, the Inviscid Burgers' Equation:

$$u_t + uu_x = 0. \tag{2.7}$$

One important difference between this equation and those we have seen earlier is that aside from the rare solutions whose initial profile is a straight line (see homework Problem 5), we cannot find closed formulas for the solutions $u(x, t)$ to this equation. This is what one generally expects to occur with nonlinear equations, even if their formula looks as simple as (2.7). Consequently, a wide variety of methods have been developed to say *something* about the behavior and dynamics of solutions to equations even in the absence of explicit solutions. Since (2.7) is an "evolution equation", we could make some simple predictions using our `SimpleEvolver[]` program if we know the initial profile $u(x,0) = f(x)$. However, in this case, it is possible to provide a more specific and more accurate description of the dynamics of a solution with any given initial profile.

The "Method of Characteristics" is useful for figuring out the behavior of solutions to some differential equations. The basic idea is that you track the behavior along a curve (or "characteristic") $x = c(t)$ in the xt-plane. By an appropriate choice of the curve, things can work out nicely. Perhaps you've seen it in another class. If not, here is just an example.

Rather than looking at all values of x and all values of t, let us focus at just one x value for each time. Then the function becomes $u(c(t),t)$, which is just a function of one variable. What is its derivative?

$$\frac{d}{dt}u(c(t),t) = u_x(c(t),t)c'(t) + u_t(c(t),t).$$

Now, we can pick a point where we want the curve to start at time $t = 0$: (x_0, t). We know the height of the wave there is $f(x_0)$. This is where we have to be clever in picking the function $c(t)$. What happens if I just define the curve to be $x = c(t) = f(x_0)t + x_0$? Then the derivative becomes

$$u_x(f(x_0)t + x_0, t)f(x_0) + u_t(f(x_0)t + x_0, t).$$

Plugging in $t = 0$ gives

$$u_x(x_0,0)f(x_0) + u_t(x_0,0) = u_x(x_0,0)u(x_0,0) + u_t(x_0,0) = 0$$

where the last equality is true precisely because $u(x,t)$ is assumed to be a solution of the equation.

What does it mean to say that the derivative is equal to zero? It means that (infinitesimally) the value of u is neither going up nor down as we move along the characteristic. In other words, the height stays the same: If you were on a boat riding along a wave shaped like $u(x,t)$ and you are at the point $x = c(t)$ at time t, then you stay at the same height. Now, we can think of a bunch of different people doing this at different points along the wave. Each one moves along a straight line in the xt-plane, but does not go up or down. The slopes of the lines are determined by their initial height. In fact, they move to the right at a speed exactly equal to their initial height (or move left if they begin at a negative height).

This idea of keeping track of the different points on the graph of the initial profile may sound familiar, as we used a similar idea in Section 1.6.2. However, the method of characteristics is far more reliable, and so in this case rather than simply seeing a "still picture" of one later time step, it is possible to animate the dynamics with relatively high accuracy.

The *Mathematica* command:

```
AnimBurgers[f_, {x,a_,b_}] := Module[{data,i,j,n,min,max},
  data[0] = Table[{i, N[f /. x -> i]}, {i, a - (b - a),
     b + (b - a), (b - a)/100}];
  min = Min[Table[N[f /. x -> i], {i,a,b,(b-a)/100}]];
  max = Max[Table[N[f /. x -> i], {i,a,b,(b-a)/100}]];
  For[i = 1, i <= 100, i = i + 1,
   data[i] =
    Table[{data[i - 1][[j, 1]] + .1 data[i - 1][[j, 2]],
      data[i - 1][[j, 2]]}, {j, 1, Length[data[0]]}]];
  ListAnimate[
   Table[Show[
     Graphics[{AbsoluteThickness[2], RGBColor[0, 0, 1],
       Line[data[i]]}], Axes -> True, AspectRatio -> .5,
     PlotRange -> {{a, b}, {min - .2 (max - min),
        max + .2 (max - min)}}], {i, 1, 99, 1}]]]
```

will take a function of x and an interval as arguments and will output an animation of the dynamics as determined by the Inviscid Burgers' Equation using this method of shifting points to the right by a distance proportional to their height.

For instance, suppose we want to see what would happen under the dynamics induced by equation (2.7) to the bell-curve shaped initial profile $u(x,0) = 1 + .5e^{-x^2}$. We just type

```
AnimBurgers[1+.5 E^(-x^2),{x,-2,8}]
```

into *Mathematica* (after defining `AnimBurgers[]` as above, of course) and we see an animation which begins with the nice smooth "hump" shape seen at the left in Figure 2.3-4. However, it is clear that problems will arise. Since the highest point will be traveling to the right at a higher speed than the lower points, it will eventually catch up with them. This leads at first to a vertical "wall" as seen in the middle image, known officially as a *shock wave [114]*. Continuing further we see that the peak of the wave has actually passed the lower points.

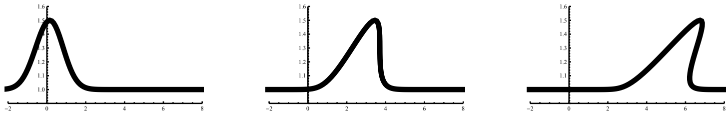

Figure 2.3-4: The dynamics of the initial profile $u(x,0) = 1 + .5e^{-x^2}$ under the evolution of the Inviscid Burgers' Equation illustrates that even with a nice initial shape problems such as a shock wave (in the center) and "multi-valued functions" (at right) can arise.

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 one you should recognize. However, despite the fact that we can associate these figures with a familiar physical phenomenon, they are mathematically troubling since the curves in the center and right graphs of Figure 2.3-4 fail to satisfy the "vertical line test". In other words, these are not even *functions*.

2.3.2 The Navier-Stokes Equations Our next example is also a very famous collection of nonlinear partial differential equations, but it is one that remains the subject of active research at the cutting edge of mathematics. The Navier-Stokes Equations very accurately describe the dynamics of water, other fluids, and even gases. They are used by scientists, engineers and even the special effects crews who animate water in movies.

The unknown functions involved are $u_i(x_1, \ldots, x_n, t)$, $f_i(x_1, \ldots, x_n, t)$ and $p(x_1, \ldots, x_n, t)$ where $1 \leq i \leq n$. So the solution here would not be a single function but a collection of $2n + 1$ functions of $n + 1$ variables that together satisfy the equations. Moreover, there are $n + 1$ equations:

$$\frac{\partial}{\partial t} u_i + \sum_{j=1}^{n} u_j \frac{\partial u_i}{\partial x_j} = \nu \sum_{j=1}^{n} \frac{\partial^2 u_i}{\partial x_j^2} - \frac{\partial p}{\partial x_i} + f_i \qquad 1 \leq i \leq n, \tag{2.8}$$

$$\sum_{i=1}^{n} \frac{\partial u_i}{\partial x_i} = 0. \tag{2.9}$$

The notation of these equations may at first appear daunting. Our goal in introducing them here is so that the reader can see how difficult some differential equations can be. It will only be necessary for us to learn to *read* the equations, to see how it is related to the Inviscid Burgers' Equation (2.7), and to appreciate famous open problems involving these equations.

Let us consider the different players in the Navier-Stokes Equations separately so that we can understand their significance. First, note that n measures the number of dimensions of space we are considering. In realistic situations, we generally consider $n = 3$ so that the fluid can move in a 3-dimensional space, but considering $n < 3$ is sometimes useful (e.g., if the fluid is just a film on a surface) and $n > 3$ may also have some theoretical significance.

Now, the variables x_i are the spatial variables that identify a point in space and t is the temporal parameter. The function u_i measures the velocity of the fluid in the positive x_i-direction at each point in space and at each time. (Thus, $\vec{F} = \langle u_1, \ldots, u_n \rangle$ is the velocity field of the flow and equation (2.9) merely states that $\nabla \cdot \vec{F} = 0$, so the vector field is divergence free.)

The function p measures the *pressure* at each point and at each time, and the functions f_i determine an external force being applied. (Thus, with all of the $f_i \equiv 0$ one would see the motion of the fluid in the absence of any external force. However, if gravity or wind are influencing it, then this would be apparent in f_i.) Finally, the number ν is the *viscosity* of the fluid. As stated above, a solution to these equations behaves *remarkably* like an incompressible fluid or gas in realistic situations, with the n equations (2.8) being essentially just Newton's formula $F = ma$ applied to a liquid.

By choosing n, f_i, p and ν correctly, it is possible to reproduce the Inviscid Burgers' Equation as a special case of (2.8). In particular, if we take $\nu = 0$ so that the second derivative terms vanish, and if we additionally let $n = 1$, $f_1 = 0$ and $p = 0$, then (2.8) becomes exactly (2.7) for $u = u_1$. The significance term "inviscid" in the name of equation (2.7) now becomes clear. It refers to the fact that the viscosity ν is assumed to have the value 0.

It was relatively simple for us to show in Section 2.3.1 that there are solutions to the Inviscid Burgers' Equation that have a nice smooth shape at time $t = 0$ but cease to be a function (due to the formation of a shock wave) in a finite amount of time. Thus, it is *amazing* that a similar question for the Navier-Stokes Equations is one of the most famous open problems in applied mathematics. The question of whether a realistic, smooth initial condition $\vec{F}(x_1, x_2, x_3, 0)$ for the velocity field will necessarily *stay* smooth and divergence free for $t > 0$ in the case $n = 3$ for the Navier-Stokes Equations is one of the five

Millennium Problems for which the Clay Mathematics Institute has offered a large cash prize *[30]*.

Clearly, this question must be *incredibly* hard to answer, otherwise the importance of the problem and the $1 million prize would have enticed someone to answer it by now. Therefore, this is a good illustration of both the value and difficulty of working with nonlinear partial differential equations.

2.3.3 Reviewing Our Prejudices Based on the few examples we have considered, we may be tempted to state the following generalizations:

- Linear equations are easier to work with than nonlinear equations because of the superposition principle that allows us to produce new solutions from known solutions, and to understand the structure of the solution set.
- Not only do we lack structure to the solution set in the case of nonlinear equations, but we generally also lack the ability to write down an explicit formula for the general solution.
- Although we saw an example of a single-humped wave that keeps its shape in Figure 2.2-2, the presence of nonlinearity or dispersion apparently prevents such a solution from existing.

Of course, it is dangerous to generalize from just a few examples. That is why these have been called "prejudices" rather than "facts". Nevertheless, many people who work with differential equations on a regular basis share these prejudices since they are, in many cases, accurate. As we will see, however, for the special class of nonlinear partial differential equations that are studied in soliton theory, these prejudices are misleading. In particular, we will find that there is a structure to the solution set analogous to the vector space structure for linear equations, that the solutions can be written explicitly, and that a traveling wave solution that is localized into a single hump is not only possible in the presence of nonlinearity and dispersion but is the very symbol of the theory.

Chapter 2: Problems

1. Here we will work out a few of the details that were left out of Section 2.2.1.

(a) For what *two* values of c is the function $\sin(kx+ckt)$ a solution to the equation (2.2)? (Note: Suppose that $k \neq 0$ is some unknown constant.)

(b) Let $u_k(x,t)$ be the sum of the two solutions you found in part (a). How do we *know* that it is also a solution?

(c) Show that $u_k(0,t) = 0$. (The value is zero at $x = 0$ for all time.) Hint: Use the fact that $\sin(-x) = -\sin(x)$.

(d) Let $u_k(x,t)$ be as in (b). Show that if k is an *integer*, then $u_k(\pi,t) = 0$ for all t. (You'll need some trickier trig identities here.)

2. In this question, we will note an important difference between the linear equation (2.2) and the nonlinear equation (2.7).

(a) Using the notation $u_k^*(x,t) = \cos(kx+kt)$ from Section 2.2.2, note that $u_1^*(x,t)$ and $2u_1^*(x,t)$ are both solutions to (2.2). How do their dynamics compare? Specifically, make an animation that shows both of these solutions and note how their heights and/or speeds differ (or do not differ).

(b) Now, use the `AnimBurger` command to watch the dynamics of the solutions to the Inviscid Burgers' Equation (2.7) which have initial profiles $u(x,0) = e^{-x^2}$ and $u(x,0) = 2e^{-x^2}$ on appropriate intervals of the x-axis. Again, the initial profile is stretched out to twice the height, but how do they compare in time? Specifically, compare how long it takes until the wave breaks for the two different initial profiles.

(c) **General scaling phenomenon for each equation:** Let λ be any nonzero constant and define

$$\tilde{u}(x,t) = \lambda u(x,\gamma t).$$

If $u(x,t)$ is a solution to (2.2), what value can you pick for γ (possibly depending on λ) to be certain that $\tilde{u}(x,t)$ is also a solution? Alternatively, if $u(x,t)$ is a solution to (2.7), what value of γ would guarantee that $\tilde{u}(x,t)$ is also a solution? Describe in words the ways that the scaled solutions $\tilde{u}(x,t)$ differ from the unscaled solutions in each case. (Note: Your answer to this question should include a *proof* that $\tilde{u}(x,t)$ is a solution when γ has the value you specify.)

3. Which *two* of the functions

$$f(x) = x, \qquad f(x) = x^2, \qquad f(x) = e^x, \qquad f(x) = e^{2x},$$

are in the kernel of the operator

$$Q = \partial^2 + \frac{4x}{1-2x}\partial - \frac{4}{1-2x}?$$

Using them, name *another* function in the kernel of Q (other than the function $f(x) = 0$).

4. Prove Theorem 2.6. In particular, show that if $f_1, \ldots, f_k$ are solutions to the (inhomogeneous) linear differential equation $L(f) = m$, then

$$F = \sum_{i=1}^{k} \lambda_i f_i$$

is a solution to the same equation so long as $\sum_{i=1}^{k} \lambda_i = 1$.

5. Although most solutions to the Inviscid Burgers' Equation (2.7) are impossible to write in terms of formulas involving functions we know, it is possible to write some exact solutions to this equation in the form

$$u(x,t) = c_1(t)x + c_2(t).$$

Note that for each fixed value of t in the domain of c_1 and c_2, the graph is a straight line. Can you find an example of non-constant functions c_1 and c_2 for which this gives an exact solution of the Inviscid Burgers' Equation? Can you find the *general* formula for the solution? (Show your work and explain your reasoning.)

6. Suppose I have a solution $u(x,t)$ of the Inviscid Burgers' Equation in the form (2.7) and I want to produce a solution of the (equivalent) equation

$$2U_t + 9UU_x = 0$$

by choosing nonzero constants λ and γ so that $U(x,t) = u(\lambda x, \gamma t)$ is a solution of this new equation. What choices of λ and γ will be sure to work?

7. The usual rule for identifying what is a linear differential equation and what is a nonlinear differential equation, which has been adopted in this book, may not be the best one. The purpose of this question is merely to emphasize that there are subtler points one might want to consider if this dichotomy is to be taken seriously.

   (a) Show that the set of solutions to D'Alembert's Wave Equation (2.2) is exactly the same as the set of solutions to the *nonlinear* equation:

$$u_{xx}^2 + u_{tt}^2 = 2u_{xx}u_{tt}.$$

Should we really consider this to be a nonlinear equation, or is it a linear equation written in an unnecessarily messy way?

(b) Show that whether the set of functions satisfying the nonlinear equation

$$(u_t - u_{xxx})(1 + u_x^2) = 0$$

is the same as the set of solutions satisfying (2.5) depends on the number field in which the function $u(x,t)$ is presumed to take values. In particular, show that the answer is "yes" if one considers real-valued functions, but that the solution set is not closed under taking linear combinations when one considers complex valued solutions.

Chapter 2: Suggested Reading

Consider consulting the following sources for more information about the material in this chapter.

- Both "An Introduction to the Mathematical Theory of Waves" by Knobel *[64]* and Drazin and Johnson's "Solitons: An Introduction" *[28]* are books that build on the ideas introduced in this chapter which are accessible to undergraduate readers.
- Two books which are on the shelves of many researchers in this field are Ince's book on linear differential equations *[53]* and "Linear and Nonlinear Waves" by Whitham *[114]*. Both should be consulted by anyone wishing to delve deeply into this topic.
- An article by McAdams, Osher and Teran *[77]* gives an elementary introduction to the Navier-Stokes Equations and their use in computer animation.

Chapter 3

The Story of Solitons

The tale of John Scott Russell and his observation of an interesting wave on a canal in Scotland in 1834 is repeated so often in the literature on soliton theory that it has taken on almost mythological importance. This section will repeat the myth with just enough details to be understood and to motivate our further investigations. Throughout the rest of the book, we will add some important missing details of both mathematical and historical significance.

3.1 The Observation

John Scott Russell was born in Scotland in 1808. His father, a clergyman, intended his son to also have a career in the church. However, due to his interest and abilities in engineering and science, his father allowed him to attend university, where he earned his degree at the age of 16. By the time he was 24 years old, Russell was awarded a temporary position as a professor at the University of Edinburgh. It was at this time that he began to study waves with the goal of designing better ships through a greater understanding of the way that the ship's hull and the water interact.

So, in August of 1834, J.S. 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 sud-*

denly 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, "Report on Waves" (14th Meeting of the British Association for the Advancement of Science, 1844)

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 you might expect. From our experiences with waves in a bathtub or at the beach, you might expect a moving hump of water to either:

- get wider and shallower and quickly disappear into tiny ripples as we see with a wave that you might generate with your hand 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. He must have been thinking "Wow, they just gave that wave a little push and off it goes – if only I could figure out how to get a boat to do the same!"

3.2 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". As for "wave of translation", it may be that he was referring to the question of whether the individual molecules of water were moving along with the hump or merely moving up and down, but that is not how the term is generally used in soliton theory today nor how we will use it in this book. To us, "translation" refers to the fact that the profile of the wave – the shape it has when viewed from the side – stays the same as time passes, as if it was a cardboard cutout that was merely being pulled along rather than something whose shape could vary moment to moment.

To study his solitary waves, Russell built a 30 foot long wave tank in his back garden. He found that he could reliably produce them in his tank and study them experimentally. Among the most interesting things he discovered was that there was a mathematical relationship between the height of the wave (k), the depth of the water when at rest (h), and the speed at which the wave travels (c). He believed that this phenomenon would be of great importance and so reported on it to the British Association for the Advancement of Science *[94]*.

3.3 A Less-than-enthusiastic Response

Although we can say with hindsight that he was correct to have had great expectations for the future of the solitary wave, his ideas were not well received by the scientific establishment of his day. In particular, the great mathematical physicists George Biddell Airy and George Gabriel Stokes each argued that Russell's wave theory was completely inaccurate. You probably recognize the name "Stokes" from math and physics classes that you have taken (e.g., "Stokes' Theorem" or "Stokes Phenomenon"). Airy's name similarly gets mentioned today far more frequently than Russell's (though not generally in undergraduate courses) for his many important contributions to math and science.

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 1845 "Tides and Waves" *[2]*, 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 a paper in 1847 called "On the theory of oscillatory waves" *[104]* about waves with a periodic profile (e.g., sine waves) and presents 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."*

It is easy to see why they would have found Russell's observations difficult to believe. As we saw in Figure 2.2-2, one can find solutions to differential equations which take the form of a single-humped wave translating at constant speed. However, that was a solution to a *linear* differential equation. One consequence of this linearity is that one can multiply the solution by the constant 2 (thereby doubling its height) and it would still be a solution, and still have 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 *non*linear, in which case it would be reasonable to expect the sort of *distortion* that we saw in Figure 2.3-4. Moreover, previous experience would have led them to expect *dispersion* to be an important factor in the dynamics of water waves, in which case something like the "mess" in Figure 2.2-3 would also be occurring at the same time. Between the distortion and the dispersion, it is difficult to see how a nicely shaped, translating, single-humped solution could possibly exist, and this is what they tried to capture rigorously in their mathematical analysis.

As we will see shortly, such intuition 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, the appropriate combination of the two produces a number of surprising and unexpected results.

3.4 The Great Eastern

It is unfortunate that these two great mathematicians erroneously rejected Russell's theory. Certainly, it must have been a source of unhappiness for Russell. It may have looked as if his interest in solitary waves was either misplaced or unappreciated. However, among ship designers he was remembered for determining the natural traveling 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 man-made 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 4,200 kilometers of the transatlantic telegraph cable between Ireland and Newfoundland, which was the first electronic communication system between Europe and America.

3.5 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 famous 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 extremely accurate but unwieldy Navier-Stokes Equations (2.8) and (2.9), they made some simplifying assumptions including a sufficiently narrow body of water so that the wave can 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 *[65]*

$$u_t = \frac{3}{2}uu_x + \frac{1}{4}u_{xxx}. \tag{3.1}$$

Due to their initials, this famous equation is now known as the "KdV Equation[1]".

It may be that 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 would not have seemed to have any applications 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 family of solutions to (3.1) which translate and maintain their shape. (This will be the subject of Section 4.3). Among these solutions were the functions

$$u_{sol(1,k)}(x,t) = \frac{8k^2}{(e^{kx+k^3t} + e^{-kx-k^3t})^2} \tag{3.2}$$

which satisfy the KdV Equation for *any* value of the constant k. This formula gives a translating solitary wave, like Russell's, that travels

[1] In fact, the equation they wrote was not exactly in the form (3.1), but was equivalent up to the conventions of this book. In particular, their equation had explicit parameters for various physical constants. However, as we will be more interested here in the theoretical significance of the equation than in using it as an actual physical model, this particular form of the equation will be most convenient. 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 is 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 *[16, 24, 86]*.

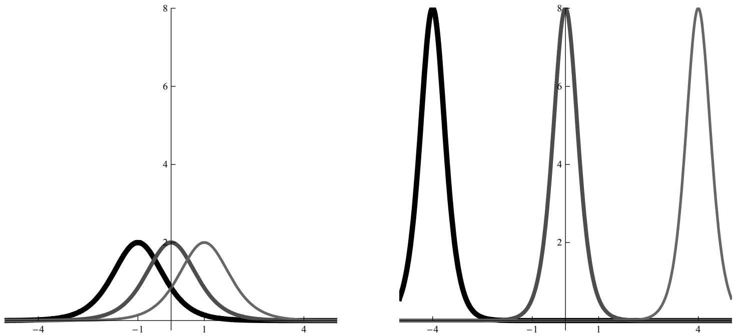

Figure 3.5-1: Two solitary wave solutions of the form (3.2) to the KdV Equation (3.1). The figure on the left shows the solution with $k = 1$ and the right is $k = 2$. In each case, the figure illustrates the solution at times $t = -1$, $t = 0$ and $t = 1$. Note that the speed with which the wave translates is k^2 and that the height is twice the speed.

at speed k^2 and has height $2k^2$. See, for instance, Figure 3.5-1 where the solutions $u_{sol(1,1)}(x,t)$ and $u_{sol(1,2)}(x,t)$ 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 have developed our prejudices on differential equations in the previous chapter: they found an exact formula for many solutions to a nonlinear PDE, and the solution seems to be able to avoid distortion and dispersion despite Stokes' intuition to the contrary. Consider, for instance, that equation (3.1) is an evolution equation which looks like a combination of two equations that we saw previously. The u_{xxx} term which we saw previously in the evolution equation (2.5) resulted there in separation of the different frequency components of a "single-humped" initial profile, resulting in its dissipation. More dramatically, the uu_x term appeared also in the Inviscid Burgers' Equation (2.7), for which we could not find explicit solutions and induced a nonlinear distortion in its solutions that soon destroyed any "single-humped" initial profile. However, somehow, 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 investigation, and this is likely the way that anyone interested in the solitary wave controversy might have reacted at the time. Specifically, the fact that the 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 (as we will see in the next chapter). And, one might say that it is a coincidence here that the effects of the 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.

3.6 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 stunning revolution in the early 20th century in the form of *quantum mechanics*. (The other stunning revolution of the time, that of relativity, can also be related to soliton theory in a way, but that is beyond the scope of this book and will not be discussed.) 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 we previously thought of as numbers (such as "speed") are actually operators like the differential operators from Section 2.1.1. Although this sounds strange and perhaps nonsensical, this is currently our most accurate description of the behavior of tiny particles, producing verifiable predictions about experiments that have been tested many times and to a very high degree of accuracy.

There are entire books on the subject of quantum mechanics, and it is mentioned here only in passing, so you may be left with the feeling that you do not understand what this actually means. Be assured that the readers of those books and even the writers also do not yet seem to understand what this means[2]. In any case, the significance to the subject of this book is that there is great interest in waves which behave like particles and/or particles that behave like waves because that seems to be what the world is made of. With that

[2] Lest anyone think that this is an unfair slander of modern physics, let me quote the great physicist Richard Feynman who said "I think I can safely say that nobody understands quantum mechanics" *[32]*.

in mind, 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 they do not seem to have taken any notice of it.

Moreover, it will soon become important to the story of solitons that 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 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 *[19]*.

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

3.7 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, Stanislaw Ulam, and Mary Tsingou at the Los Alamos National Laboratory. In a now famous experiment, they programmed a Los Alamos computer to give them approximate solutions to nonlinear equations with the prescient intention of developing better intuition about nonlinearity. They presumed that even 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. However, they wanted to see this happen in experiments on the computer. What they found surprised them. Just as Stokes and Airy were mistaken in their assumption that the evolution induced by 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 *[31]*; 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 *[111]*.

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

It was then that mathematicians Martin Kruskal at Princeton University and Norman Zabusky at Bell Labs conducted their own computer experiments *[120]*. Rather than considering a discrete system of connected vibrating masses as in the Fermi-Pasta-Ulam-Tsingou experiments, they wanted to consider a nonlinear wave equation. Taking the Fermi-Pasta-Ulam-Tsingou 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 Equation (3.1).

At this point, of course, the existence of solutions in the form (3.2) were 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 computer programs similar to our `AnimBurgers` program from page 40. However, being far more accurate than this very simplistic program, theirs were able to produce animations showing the dynamics of solutions to (3.1) over a longer period of time. There were two amazing results that came from this investigation:

- 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 the animation showed the solution breaking apart into a finite number of humps, each behaving like one of Russell's solitary waves, along with some "radiation" which travels away from them in the other direction. This seems to suggest that the solutions of the form (3.2) play a fundamental role in describing the general localized positive solution to the KdV Equation, similar to the way in which the basic vibrating modes (2.3) form a basis for solutions to D'Alembert's Wave Equation (2.2). (Of course, they cannot *actually* form a basis for the solutions since 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 (see Figure 3.7-2), there are two humps each moving to the left with speed equal to half their height. As we will see, it is *not* the case that this is simply 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. The amazing thing, however, it to consider the situation in which a taller hump is to the right of a shorter one. Since it is moving to the left at a greater speed it will eventually catch up. Intuition about nonlinear differential equations would have suggested to any expert at the time that even though the KdV Equation has this remarkable 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 the 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 the 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 could be seen to occur when three or more separate peaks were combined to form an initial profile: the peaks would move at appropriate speeds, briefly "collide" and separate again.

The name "solitary wave" coined by Russell more than one hundred years earlier 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. However, now seeing how *gregarious* they are, the name no longer seems appropriate. The term "*soliton*"

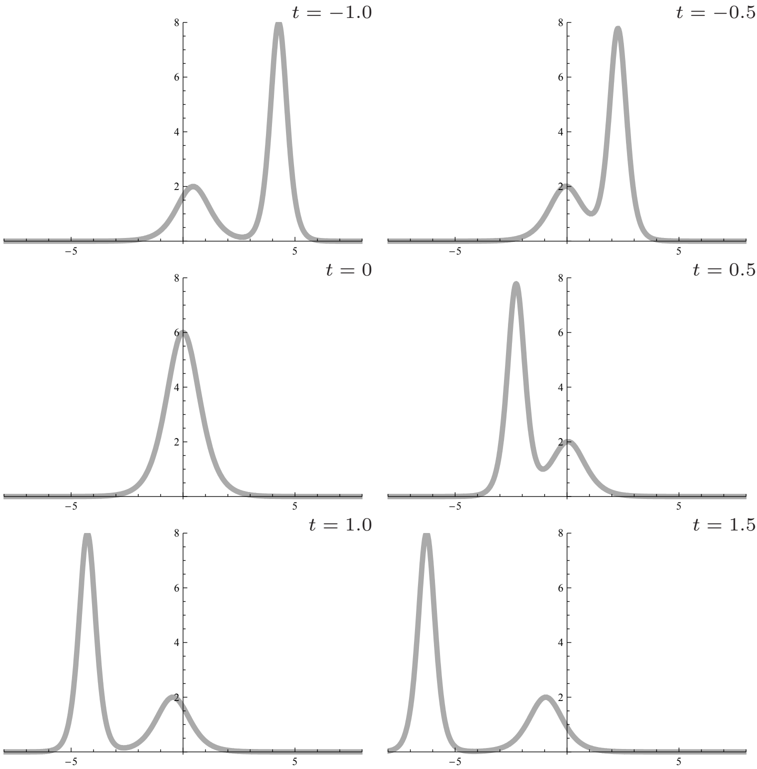

Figure 3.7-2: A solution to the KdV Equation as it would have appeared to Kruskal and Zabusky in their numerical experiments. Note that two humps, each looking like a solitary wave, come together and then separate.

was used by Kruskal and Zabusky to describe these solutions, 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, much as one could refer to a local maximum in the graph of a polynomial.

So, for instance, we can say that Figure 3.7-2 illustrates a 2-soliton solution of the KdV Equation in which a taller soliton traveling 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, you should not mistakenly think that this is 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.

3.8 Hints of Nonlinearity

In Chapter 5 we will investigate the soliton solutions of the KdV Equation in greater detail. At this point, however, we will simply look at Figure 3.7-2 to note the way in which it differs from the corresponding graphs of the linear combination $u^{\times}(x,t) = u_{sol(1,1)}(x,t) + u_{sol(1,2)}(x,t)$ of two different 1-soliton solutions shown in Figure 3.8-3.

For homework you will check to see whether $u^{\times}(x,t)$ is a solution to the KdV Equation. (At the risk of spoiling your surprise, we use the notation $u^{\times}(x,t)$ so as to indicate that it is *not* a solution to the KdV Equation, using the $\times$ like the "x" that your teacher uses to indicate that something is wrong.) However, if you were to watch an animation that shows its dynamics you would have to look very closely to see how it is different than the actual KdV 2-soliton solution $u_{2sol}(x,t)$ shown in Figure 3.7-2. These differences, though subtle, are quite important.

First, consider the graphs of $u^{\times}(x,0)$ and $u_{2sol}(x,0)$. In both cases, one sees only a single hump 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 *ten*. In contrast, Figure 3.7-2 clearly shows that $u_{2sol}(x,0)$ has a peak of height *six*. 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 $t = -.5$. At time $t = 0$ one cannot see two separate peaks, but then at time $t = .5$ when the peaks have separated again, one *still* sees the smaller soliton nearly centered on the y-axis! In contrast, since the smaller peak in $u^{\times}(x,t)$ always moves to the left at constant[3] speed 1, it will

[3]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. How-

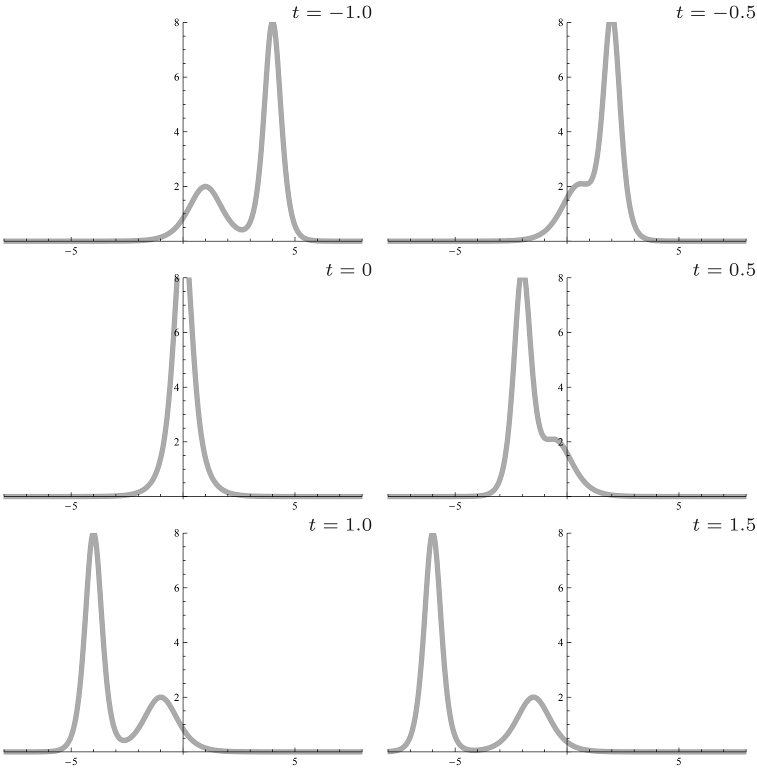

Figure 3.8-3: This is *not* a solution to the KdV Equation! This is a sum of the one soliton solutions $u_{sol(1,1)}(x,t)$ and $u_{sol(1,2)}(x,t)$. Compare to Figure 3.7-2, which is a KdV solution, to see the subtle differences despite the fact that each shows a hump moving to the left at speeds 1 and 4, respectively, at most times and a single hump centered on the x-axis at time $t = 0$.

have moved one unit to the left during the time interval $-.5 \leq t \leq .5$.

Later we will consider exactly what these differences imply. For now, however, it is enough to think of it as an indication that there is some sort of nonlinear *interaction* going on in the 2-soliton solution. If we think of the solitons as particles, then they have not simply passed through each other without any effect, but have actually "collided"

---

ever, 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.

and in some sense the KdV Equation incorporates their "bounce".

It is important to note that – like Russell's solitary waves – these solutions exist not only as solutions to the KdV Equation but also as physical phenomena which can be observed as surface waves on a canal. It is possible to generate waves in a water tank and see dynamics exactly like those described in this section. This has been carefully verified in the years since the Kruskal-Zabusky numerical experiments, but such experiments were *also* carried out by J.S. Russell over one hundred years earlier. Since the scientific establishment doubted his claims that solitary waves existed, they completely ignored his investigations into how pairs of them would interact. So the results of these experiments were forgotten and ignored until the phenomenon they described was rediscovered in these computer experiments.

3.9 Explicit Formulas for n-soliton Solutions

A separate surprise, which was revealed in a later paper by Gardner, Greene, Kruskal and Miura *[36]*, was that these n-soliton solutions of the KdV Equation did not have to be studied in numerical simulations because it is possible to write exact formulas for them. For example,

$$u_{2sol}(x,t) = \frac{24\left(e^{2x+2t} + 6e^{6x+18t} + 4e^{4x+16t} + 4e^{8x+20t} + e^{10x+34t}\right)}{\left(1 + 3e^{2x+2t} + e^{6x+18t} + 3e^{4x+16t}\right)^2} \tag{3.3}$$

is an exact solution of (3.1) and it is the one that is illustrated in Figure 3.7-2.

This is quite surprising since it means that we have explicit formulas for a very large and interesting family of solutions to this nonlinear PDE. The method used to find these formulas is also quite interesting, and is not a subject that this book will be considering aside from a few remarks here[4]. It is quite intriguing to note that the method is based on the techniques of *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 have been replaced by differential operators, such as those in Chapter 2. To study the one-dimensional *scattering problem* of how an incoming wave $\psi(x)$ will "bounce off" of another wave $u(x)$ (thought

[4] Interested students with a background in analysis or quantum physics are encouraged to read about the method of "inverse scattering" in the sources suggested at the end of the chapter.

of as an obstacle), one is led to work with the differential operator $\partial^2 + u(x)$. Amazingly, it turns out 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 will turn out to be important is that they depend *isospectrally*[5] on the variable t. Pursuing this line of reasoning, Gardner, Greene, Kruskal and Miura were able to use a technique called *inverse scattering* to write exact formulas for the n-soliton solutions.

This is a very powerful and useful technique in soliton theory. However, it depends on advanced mathematics and physics which goes beyond what we intend to cover in this book. Instead, we will be looking at other methods – more algebro-geometric in nature – which also allow us to write exact solutions.

3.10 Soliton Theory and Applications

There is something very interesting about the KdV Equation. Despite being nonlinear and dispersive, it has solutions which seem to be able to avoid the damaging effects of these "forces" and maintain their nice, localized shapes indefinitely. These solutions have a certain "particle-like" nature, which runs contrary to our intuition of how waves ought to behave but might prove useful in understanding the behavior 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 we can only study numerically or qualitatively, we can write down explicit formulas for many exact KdV solutions.

Soliton theory is the branch of mathematics which was developed to understand this phenomenon. Among the big questions it asks are:

Big Question I: Why is it that we can write so many exact solutions to the KdV Equation when we cannot do so for most nonlinear equations?

Big Question II: The relationship between the n-soliton solutions and the n different 1-soliton solutions that it resembles suggests there is some way in which solutions of the KdV Equation can be combined.

[5] This term will be explained in Section 7.1.1.

We know that the multi-solitons are not actually linear combinations of solitary wave solutions 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?

Big Question III: How can we identify *other* equations – either known already to researchers or yet to be discovered – that have these same interesting features?

Big Question 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 algebro-geometric structure that generic nonlinear PDEs do not share, but that by understanding this structure we can find infinitely many other equations that share all of these features and so also deserve the name "soliton equations".

The KdV Equation and these other equations often have physical significance as they model phenomena we encounter in the real world – waves on a 2-dimensional surface like the ocean, light in optical fibers, electrons in a thin wire, the transcription bubble in DNA, or energy transfer in proteins. In this sense, solitons have become tools of scientists and engineers for understanding the universe and building tools.

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 say *what* was possible in a nonlinear situation. Now, at least, we have a large set of nonlinear equations whose solutions can be studied explicitly[6]. 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. In fact, there are questions of interest 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 nonlinear partial differential equations and especially the dynamics of waves generally fall squarely in

[6]However, it should be admitted 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, and the view looks very different there.

the "applied" side of this division while algebraic geometry is perceived as being among the purest of the "pure". To some of us, it is beautiful and surprising that each of these can provide information about the other in the intersection that is soliton theory.

The remainder of this book will seek to elaborate further on the answers to Big Questions I, II and III. Even at the end of the book, they will remain only *partial* answers to these deep questions since there are in fact *many* ways to understand what makes the KdV Equation so much nicer than an arbitrary nonlinear PDE, and we will only be focusing on those answers which are algebro-geometric and can conveniently be incorporated into an undergraduate textbook. In the end, the reader should have enough of an understanding to feel a sense of satisfaction even if that understanding is incomplete, much as one feels a certain thrill upon learning how a magician performed a particularly surprising trick.

3.11 Epilogue

It was not only *other* researchers who were uninterested in the paper 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 discoveries of the important mathematician Korteweg, and not important enough to bring any attention to 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 paper 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 paper. Clearly, this paper which attracted little attention when it was first published was of great interest one hundred years later.

Korteweg and de Vries are honored 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". So, both the university department from which he earned his degree and a subdiscipline of mathematics are named after Gustav de Vries.

(Not bad for someone who published only one research paper!)

One of the applications of soliton theory provides a very poetic epilogue to the story of J.S. Russell and his interest in solitary waves. In the 19th century, the huge ship that Russell designed based on his wave theory was used to lay the first telegraph cables across the ocean, giving him an important place in the history of communication technology. But, it turns out that Russell's legacy was to be connected to transatlantic communication in another way also. In the 21st century, some information is transmitted through fiber optic cables that guide light waves rather than through copper cables carrying an electric current. As it turns out, the differential equations which describe the motion of light through certain types of cables are soliton equations and so it is possible to transmit information in the optical fiber in the form of solitons – solitary waves of *light [78]*. It is not difficult to see why the properties of solitons would be well suited for application to communication networks. 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 *[42]*."

In other words, interest in optical solitons for use in communication comes from the property that originally made the concept of a solitary wave controversial, the very thing that Russell first noticed in that wave he witnessed on the canal, that it just "kept on going".

Chapter 3: Problems

1. For what values of the constants c_1 and c_2 is the function

$$u(x,t) = \frac{c_1}{(x+c_2)^2}$$

   a solution to the KdV Equation?

2. For what value(s) of the constant c is the function

$$u(x,t) = \frac{cx}{t+1}$$

   a solution to the KdV Equation? Describe the dynamics: "The initial profile of this function at any fixed time looks like .... and as time passes ..."

3. Consider a transformation which takes a function $u(x,t)$ and turns it into
$$\hat{u}(x,t) = a^2 u(ax, a^n t)$$
where a and n are two fixed numbers. There is a unique value of n with the property that $\hat{u}(x,t)$ is a solution to the KdV Equation for every constant a and every KdV solution $u(x,t)$. (It is known as the *scaling symmetry* of the KdV Equation.)
   (a) What is the value of n for which the transformation $u \mapsto \hat{u}$ is a symmetry of the KdV Equation? (In other words, for what number n will $\hat{u}(x,t)$ be a solution of the KdV Equation for every choice of a and every KdV solution $u(x,t)$?)
   (b) Imagine you are watching a movie animating the dynamics of the KdV solution $u(x,t)$. How would an animation of the KdV solution $\hat{u}(x,t)$ be different if you used $a = -1$ as the constant in the scaling symmetry? And how would it be different if you used $a = 2$?

4. Many mathematicians and physicists post their research papers on the internet at "preprint servers" like `arxiv.org`. Go to that website (or another similar site) now and search for papers which have the word "soliton" in their title. At the moment, you can achieve this by going to:

   `https://arxiv.org/search/?query=soliton&searchtype=title`

   How many articles with the word "soliton" in the title have appeared in the past two months? (Note that the papers are listed in reverse chronological order and the number of the article indicates the year and month. For example, "`arXiv:2203.1875`" is the 1875th article posted during the month of March in 2022.) Look at the titles and maybe even the abstracts of some of the articles. Do you see anything you understand (or anything you *want* to understand)?

5. The KdV Equation and its soliton solutions are quite famous and appear in many different books and articles. In one book I see it written as
$$U_t - 6UU_x + U_{xxx} = 0.$$
Of course, this is simply another way to write the same equation. Assuming the change of variables is simply of the form $u(x,t) = \alpha U(x, \beta t)$, how is this form of the KdV Equation related to our

chosen form (3.1)? In what seemingly significant ways would an animation of the 1-soliton solution to this alternative form of the equation be different than the ones shown in Figure 3.5-1?

6. What is the consequence of changing the sign of the parameter k in a KdV 1-soliton? (Compare the solutions $u_{sol(1,k)}(x,t)$ and $u_{sol(1,-k)}(x,t)$.)
7. The *Mathematica* command
`DAlembert[u_]:=Simplify[D[u,{t,2}]-D[u,{x,2}]]`
will take a function of x and t as an argument and subtract its second x-derivative from its second t-derivative and simplify the result. If the function is a solution to D'Alembert's Wave Equation (2.2), then the result will be *zero*. In this sense, you can use this command to check whether something is or is not a solution. Write a similar command `KdV[u]` which will output zero if the argument is a solution to the KdV Equation (3.1) and a nonzero expression[7] if it is not. Test it by checking that $u_{sol(1,k)}(x,t)$ is a solution for all values of k and that $u_{2sol}(x,t)$ is also a solution, but that $u^{\times}(x,t)$ is not.
8. I claim that if $u(x,t)$ is any solution to the KdV Equation and α is any constant, then

$$\hat{u}(x,t) = u(x + \gamma t, t) + \alpha$$

is also a KdV solution if you choose γ correctly. Derive a formula for γ as a function of α so that this will be true.

9. Suppose u is any KdV solution that has the form $u(x,t) = w(x + ct)$. Using your answer to the previous problem, find the numbers γ and α so that $\hat{u}(x,t) = u(x + \gamma t, t) + \alpha$ is a solution of the KdV Equation that satisfies $\frac{\partial}{\partial t}\hat{u} = 0$. Describe in words how an animation of the dynamics of $\hat{u}$ differs from an animation of the solution u.

[7] Having had much experience with *Mathematica*, I can tell you that sometimes the simplify command does not recognize something as being equal to zero even when a human can tell that it is. So, if the output is not zero, you may still have to look at it and think a bit to make sure that this is not the case before concluding that the input was not a solution.

Chapter 3: Suggested Reading

Consider consulting the following sources for more information about the material in this chapter.

- Filippov's *The Versatile Soliton [33]* covers many historical facts which were left out of this brief summary.
- Bullough and Caudrey's historical analysis *[18]* appears in the proceedings of a conference honoring the 100th anniversary of the paper by Korteweg and de Vries.
- The article on symmetries of solitons by Palais in the *Bulletin of the AMS [85]* 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 *[81]*.
- The footnote on page 63 mentions that soliton theory is but *one* window into the mysterious world of nonlinear dynamics. A view from the opposite direction is provided by *chaos theory*. For a brief but beautiful introduction to this concept, read Robert May's essay "The Best Possible Time to Be Alive" *[74]*. A more detailed introduction can be found in the textbook by Devaney *[25]*.
- Although the origins of soliton theory are hydrodynamic, as explained in this chapter, the physical interpretation of the solutions of soliton equations as waves in water will not be of great importance in the remainder of this textbook. Please consult the book by Remoissenet *[91]* for a more physical approach to this subject, including many laboratory experiments. This same book also contains discussions of solitons in optical fiber 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 *[23]*.
- Several of the ideas introduced in this chapter can be further explored using the suggestions in Appendix C. Problems 3 and 8 are examples of *symmetries* of the KdV Equation. Their algebraic implications can be further explored in Project I. The Fermi-Pasta-Ulam-Tsingou and Zabusky-Kruskal computer experiments can be replicated by following the suggestions in Projects VI and VIII. And you can learn more about the use of optical solitons in communications by working on Project III.