# FROM CANALS TO FIBER OPTICS: HOW SOLITON METHODS ENTERED PHYSICS 

SAMUEL J. FERGUSON

M.S./Critical Essay, May 2014

Advisor: Prof. Donald A. Gurnett, Dept. of Physics \& Astronomy, University of Iowa, Iowa City, Iowa


#### Abstract

Solitons and the mathematical machinery for their study now play crucial roles in many subareas of physics. These subareas are diverse, drawn from such scientific pursuits as classical and fluid mechanics, nonlinear optics, and the theories of quantum fields, solids, plasmas, and DNA. How did soliton methods enter physics, and what are some of the tools needed to begin using them? After a reasonably historical introduction to the phenomenon of the solitary wave, an overview is made of key turning points that led to the creation of soliton theory. This overview includes a close examination of some of the papers that particularly influenced the historical development. The inverse scattering method-a tool par excellence in solving soliton equations or so-called completely integrable systems-is then introduced, with emphasis placed on the simplest initial-value problem for the equation describing solitary waves in a canal. Some mathematical sidelights pertaining to this equation of Korteweg-de Vries are included for the interested reader, and these point the way to other well-studied equations for which the soliton perspective is indispensable.


Nonlinear wave phenomena can sometimes exhibit particlelike behaviors. Under certain technical conditions these phenomena are called solitons. A variety of nonlinear physical systems, arising in fluid mechanics, nonlinear optics, plasmas, quantum fields, and several other areas of physical interest, may be understood in terms of solitons-but solitons were once subject to skepticism, and even their existence was doubted. The purpose of this essay is to show how soliton methods succeeded in entering physics, and to introduce some of the ideas and applications that they brought with them.

## 1. Historical Introduction to Solitary Waves

Before solitons were studied, there were observations of solitary waves. Later soliton theory built on the success of equations that modeled these waves, and so it may be helpful to pursue solitary waves before considering solitons in general. Interest in modeling solitary waves may be heightened by considering them in historical context, as many physicistssome well known-examined this problem.

Solitary waves were first given scientific attention by John Scott Russell, who studied waves in order to design better ships. In August of 1834, Russell was sitting on his horse beside the Union Canal of Edinburgh, Scotland, when he made an observation that changed his life.

I was observing the motion of a boat which was rapidly drawn along a narrow channel by a pair of horses, when the boat suddenly stopped-not so the mass of water in the channel which it had put in motion; it accumulated round the prow of the vessel in a state of violent agitation, then suddenly leaving it behind, rolled forward with great velocity, assuming the form of a large solitary elevation, a rounded, smooth and well-defined heap of water, which continued its course along the channel apparently without change of form or diminution of speed. I followed it on horseback, and overtook it still rolling on at a rate of some eight or nine miles an hour, preserving its original figure some thirty feet long and a foot to a foot and a half in height. Its height gradually diminished, and after a chase of one or two miles I lost it in the windings of the channel.

Here is a picture of a solitary wave in a canal:


Figure 1. Solitary wave in an Edinburgh canal.

The water waves usually encountered in a pool or at the beach do not move "without change of form or diminution of speed." In fact, such waves either:

- get wider and shallower (as in a swimming pool), quickly disappearing; or
- "break" (as at the beach), the wave peak racing ahead, outrunning its support, and crashing down.
The wave Russell saw, however, did neither of these things - after a push, it kept its shape and speed for miles. If a ship were to behave like the "great solitary wave" Russell described, it would be undeniably energy-efficient-a better ship.

Russell investigated further by building a 30 foot long wave tank in his backyard garden, eventually using a piston to produce solitary waves and study them experimentally. He
discovered a mathematical relationship between the depth of the water at rest ( $h$ ), the maximum height of the wave $\left(\eta_{0}\right)$, and the wave's propagation speed $(c)$. Let $c_{0}=\sqrt{g h}$ be the speed of long-wavelength linear waves in water of depth $h$, where $g$ is the strength of the gravitational field; see below for details. Then, according to [14], Russell found that

$$
c=c_{0}\left(1+b \eta_{0}\right),
$$

for some positive parameter $b$ depending on $h$. Thus, these nonlinear waves always had a higher speed than the corresponding linear waves. Considering his observations of the solitary wave phenomenon to be of scientific value, he presented his "Report on Waves" [22], the source of the above quotes, to the British Association for the Advancement of Science.

This paper was severely criticized by two of the leading scientists of the day, the wellknown astronomer Sir G.B. Airy and the influential mathematical physicist G.G. Stokes. Aside from the mathematical relationship described above, Russell had made little in the way of a mathematical analysis regarding solitary waves, and in particular made no use of differential equations. In his 1844 paper "Tides and Waves," [1] Airy derives a formula for the speed of shallow water waves that appears to disagree with Russell's formula. He then argues that a propagating wave like Russell's must get steeper in front, and less steep behind-in other words, that solitary waves cannot keep their shape. He writes, of Russell's "great solitary waves," "We are not disposed to recognize as deserving the epithets 'great' or 'primary'." In 1847, Stokes worked on a related problem, writing "On the theory of oscillatory waves" [24]. Here Stokes analyzes waves in nonviscous fluids with a periodic profile, such as superpositions of sine waves like

$$
\eta(x, t)=\eta_{0} \sin (k(x-c t)),
$$

which are required by linear water wave equations to travel at speed

$$
c=\sqrt{\frac{g}{k} \tanh (k h)} .
$$

Note that in the long-wavelength limit, $\lambda \gg 1$, where $k=\frac{2 \pi}{\lambda} \ll 1$, we have $\tanh (k h) \approx k h$. Thus, the long-wavelength propagation speed is $c \approx \sqrt{g h}$, explaining $c_{0}$ 's definition as $\sqrt{g h}$. Stokes presents a formula for the profile of a periodic wave with infinitely many humps, which he claims, "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." Airy and Stokes appear to have proved that solitary waves do not exist! This apparently put an end to Russell's scientific research.

Nevertheless, among ship designers, Russell is remembered for determining the natural traveling speed for a given fluid depth, a result which grew directly out of his research on solitary waves. He is also remembered for his work on what was then the largest moving
manmade object, The Great Eastern. In 1865, the Great Eastern was used to lay 4,200 kilometers of the transatlantic telegraph cable between Ireland and Newfoundland, the first electronic communication system between Europe and America [14]. Russell's obituary in the June 10, 1882 edition of The Times, calling Russell's system of design the "wave system," says:

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.

The conclusions Russell drew from his observations of solitary waves apparently developed into solutions of challenging engineering problems. The scientific problem of solitary waves, however, remained.

## 2. From Solitary Waves to the Korteweg-de Vries Equation

Within Russell's own lifetime the French scientist de Boussinesq proposed a new theory of shallow water waves, different from Airy's, in which Russell's observations were no longer inconsistent. According to [13], Boussinesq's most accessible paper on shallow water waves is [2], whose title is too lengthy to give here. In this paper, he considers long-wavelength fluid motion in a shallow canal with rectangular cross section. The fluid is assumed incompressible, irrotational, and inviscid, with friction being ignored. Since the approaches of other physicists-such as Lord Rayleigh and Korteweg-de Vries-were similar to Boussinesq in their use of power series, it is worth explaining his approach here.

For Boussinesq, dynamics are of primary interest, and so at time $t$ the coordinates of a "fluid particle" are denoted $(x, y)=(x(t), y(t))$. Thus, the velocity at time $t$ is given by $(u, v)=\left(x^{\prime}(t), y^{\prime}(t)\right)$. As is usual for problems of two-dimensional potential flows in fluid mechanics, Boussinesq defines a velocity potential $\phi(x, y)$ satisfying $u=\frac{\partial \phi}{\partial x}$ and $v=\frac{\partial \phi}{\partial y}$, as well as a stream function $\psi(x, y)$ satisfying $u=\frac{\partial \psi}{\partial y}$ and $v=-\frac{\partial \psi}{\partial x}$. For $\phi$ to be well-defined, its mixed partials must be equal, that is, the partial derivatives of $u$ and $v$ must be related in a certain way, $\frac{\partial u}{\partial y}-\frac{\partial v}{\partial x}=0$. This condition is equivalent to the hypothesis that the fluid
be irrotational. Similarly, the stream function $\psi$ is well-defined exactly when the fluid is assumed incompressible. The stream function gets its name from the fact that equations of the form $\psi=c$, with $c$ constant, define streamlines for the flow. By definition, $u=\frac{\partial \phi}{\partial x}=\frac{\partial \psi}{\partial y}$ and $v=\frac{\partial \phi}{\partial y}=-\frac{\partial \psi}{\partial x}$, so $\phi$ and $\psi$ satisfy the Cauchy-Riemann equations of complex analysis. This implies that $\phi$ and $\psi$ are the real and imaginary parts of a complex-valued analytic function $f$ of a complex variable $z=x+i y$. In other words, $f(x+i y)=\phi(x, y)+i \psi(x, y)$.

A typical analytic function has singularities. For example, $f(z)=\frac{1}{1-z}$ has a singularity at $z=1$. This prevents its power series about $z=0, f(z)=1+z+z^{2}+z^{3}+\ldots$, from converging outside of the open unit disk.

Particularly nice analytic functions, such as $e^{z}$, do not have this problem, and have an everywhere-convergent power series about each point. Thus, if we understood $e^{x}$ for $x$ real and wanted a better idea of the behavior of $f(z)=e^{z}$ for $z$ complex, we could use its Taylor expansion about $x$ :

$$
f(x+i y)=\sum_{n=0}^{\infty} \underbrace{f^{(n)}(x)}_{=e^{x}} \frac{(i y)^{n}}{n!}=e^{x}\left(\sum_{k=0}^{\infty}(-1)^{k} \frac{y^{2 k}}{(2 k)!}+i \sum_{k=0}^{\infty}(-1)^{k} \frac{y^{2 k+1}}{(2 k+1)!}\right)
$$

This would then tell us that $e^{x+i y}=e^{x}(\cos (y)+i \sin (y))$.
Boussinesq apparently felt that in the long-wavelength limit, his unkown analytic function $f(z)$ would be nice enough with respect to $y$-convergence to justify a similar procedure:

$$
f(x+i y)=\sum_{n=0}^{\infty} f^{(n)}(x) \frac{(i y)^{n}}{n!}=\sum_{k=0}^{\infty}(-1)^{k} f^{(2 k)}(x) \frac{y^{2 k}}{(2 k)!}+i \sum_{k=0}^{\infty}(-1)^{k} f^{(2 k+1)}(x) \frac{y^{2 k+1}}{(2 k+1)!}
$$

Later, Korteweg and de Vries explained the $y$-convergence on page 426 of [15] with these words: "For long waves these series are rapidly convergent. Indeed, for such waves the state of motion changes slowly with $x$, and therefore the successive differential-quotients with respect to this variable of all functions referring, as $f$ does, to the state of motion, must rapidly decrease." Boussinesq, who rarely addresses complex-valued functions directly, appears to assume implicitly that $f(x)$ is real-valued for real $x$; this is false for such analytic functions as $f(z)=i z$. However, if we grant him this, then we can follow him by equating the two terms above with the real and imaginary parts of $f$, namely $\phi$ and $\psi$ :

$$
\phi(x, y)=\sum_{k=0}^{\infty}(-1)^{k} f^{(2 k)}(x) \frac{y^{2 k}}{(2 k)!}
$$

and

$$
\psi(x, y)=\sum_{k=0}^{\infty}(-1)^{k} f^{(2 k+1)}(x) \frac{y^{2 k+1}}{(2 k+1)!}
$$

Having come this far, there is nothing to stop Boussinesq from calculating $u=\frac{\partial \phi}{\partial x}$ and $v=\frac{\partial \phi}{\partial y}$ to arrive at similar series expansions for $u$ and $v$.

Boussinesq then required boundary conditions to hold on the surface profile $y=\eta(x, t)$, a kinematic condition

$$
v=\frac{\partial \eta}{\partial t}+u \frac{\partial \eta}{\partial x}
$$

(this comes from taking $\frac{\mathrm{D}(y-\eta)}{\mathrm{D} t}=\frac{\partial}{\partial t}(y-\eta)+(u, v) \cdot \nabla(y-\eta)=0$ ) and a dynamic condition,

$$
\frac{\partial \phi}{\partial t}+\frac{P}{\rho}+\frac{1}{2} \nabla \phi \cdot \nabla \phi+g \eta=F(t)
$$

which may be recognized as a form of Bernoulli's equation. Here, $P$ is a supposed constant value that the pressure $p$ assumes at the surface, and $\rho$ is a supposed constant fluid density. $F(t)$ is an unkown function, not to be confused with $f(z)$. These equations-and the notation we've adopted-may be found on page 207 of the fluid mechanics textbook [4], except that the kinematic condition found there accounts for a $z$-component in the velocity.

The dynamic boundary condition can be written in a more appropriate form by moving $\frac{P}{\rho}$ to the right-hand side and writing out $\nabla \phi$ and $\frac{\partial \phi}{\partial t}$ in terms of known series. The latter may appear worrisome because $\phi$ is a function of $(x, y)$, but recall that $(x, y)=(x(t), y(t))$ and $(u, v)=\left(x^{\prime}(t), y^{\prime}(t)\right)$, so

$$
\frac{\partial \phi}{\partial t}=\frac{\partial \phi}{\partial x} \frac{d x}{d t}+\frac{\partial \phi}{\partial y} \frac{d y}{d t}=u^{2}+v^{2}
$$

and $\nabla \phi \cdot \nabla \phi=u^{2}+v^{2}$. Using the series for $u$ and $v$ gives an equation relating $\eta$ and $f(x)$; Boussinesq removed $P / \rho$ and the time-dependence of $F(t)$ by requiring everything in sight (except a $g h$ term embedded in $F(t)$ to properly account for gravity) to go to 0 as $x \rightarrow \pm \infty$. This gives two equations (from the two boundary conditions) in $\eta$ and $f$. Despite the presence of derivatives on $f$ in the series expansions, Boussinesq proceeds to truncate the series and eliminate $f$ to obtain an equation valid for $\eta$ up to a certain order; however, the writing is not as transparent as might be wished. The reader interested in seeing the tedious details of this purely mechanical approximation process is referred to [2], and may profitably read [21]. The point is that, to a first approximation, Boussinesq obtains the wave equation

$$
\frac{\partial^{2} \eta}{\partial t^{2}}=g h \frac{\partial^{2} \eta}{\partial x^{2}}
$$

To second order, he arrives at what we now refer to as the Boussinesq equation,

$$
\frac{\partial^{2} \eta}{\partial t^{2}}=g h \frac{\partial^{2} \eta}{\partial x^{2}}+g h \frac{\partial^{2} \zeta}{\partial x^{2}}
$$

where we ${ }^{1}$ let $\zeta=\frac{3 \eta^{2}}{2 h}+\frac{h^{2}}{3} \frac{\partial^{2} \eta}{\partial x^{2}}$. This appears, for the most part, to satisfy him.
The theory of de Boussinesq was confirmed by 1876 in [21] due to some investigations by Lord Rayleigh, a former student of Stokes. By supposing the existence of a solitary wave

[^0]and adding a flow to negate its velocity-thus making the wave profile time-independentLord Rayleigh was able to derive an explicit formula for the wave form $\eta$ independently of Boussinesq. To wit, making use of standard techniques (velocity potential, stream function, series, pressure equation) and approximations (especially for square roots), he was eventually able to obtain an equation for the wave profile $\eta(x)$. In modern notation, his equation was
$$
\left(\frac{d \eta}{d x}\right)^{2}+\frac{3}{h^{3}} \eta^{2}\left(\eta-\eta_{0}\right)=0
$$

Here $\eta_{0}$ is the maximum height of the wave above the equilibrium fluid depth $h$. Rayleigh's solution to this equation, satisfying $\eta(0)=\eta_{0}$, was

$$
\eta(x)=\eta_{0} \operatorname{sech}^{2}\left(\sqrt{\frac{3 \eta_{0}}{4 h^{3}}} x\right)
$$



Figure 2. Lord Rayleigh's function $\eta(x)$ with $\eta_{0}=h=3$; the appearance is similar to a gaussian bump. In realistic situations we actually expect $\eta_{0} \ll h$.

This famous $\operatorname{sech}^{2}(\theta)$ profile of Rayleigh reccurs frequently in the literature, e.g., as the $\operatorname{sech}^{2}(\theta)$ potential in quantum mechanics, and has surprising predictive power in the study of canal waves. For example, suppose we want to compare the order of magnitude of the ratio $\delta \equiv \frac{h}{w}$ with $\varepsilon \equiv \frac{\eta_{0}}{h}$, where $w$ is the half-height width of the wave profile. If $w_{0}$ is the
half-height width of $\operatorname{sech}^{2}(x)$, the coefficient of $x$ in Rayleigh's function gives $w=w_{0} / \sqrt{\frac{3 \eta_{0}}{4 h^{3}}}$. Ignoring the constants, this yields $\frac{h}{w} \approx \sqrt{\frac{70}{h}}$, whence $\delta^{2} \approx \varepsilon$, a result that appears to hold true for nonlinear canal waves in general. I encourage the interested reader to derive Rayleigh's solution from its differential equation and initial condition as an exercise. The reader may take heart in the knowledge that, as Lord Rayleigh assumed the presidential chair of the British Association for the Advancement of Science, Lord Kelvin confessed, "Some of the pages of Lord Rayleigh's work have taxed me most severely, but the strain was well repaid."

At this point, a few intuitive comments about the wave equation in one spatial dimension may be in order. We recall that the equation for waves traveling at speed $c$ in one spatial dimension is

$$
\frac{\partial^{2} u}{\partial t^{2}}=c^{2} \frac{\partial^{2} u}{\partial x^{2}}
$$

and that its general solution is $u(x, t)=f(x-c t)+g(x+c t)$. That is, the general solution is the sum of a profile propagating to the right at speed $c$ and another profile traveling to the left at speed $c$. Suppose, however, that we were only interested in waves propagating to the right. In other words, we are only interested in functions of the form $u(x, t)=f(x-c t)$. In that case, a much simpler (meaning: lower order) equation is sufficient to characterize the wave translation. Calculating the partial derivatives with respect to $t$ and $x$, we see in this case that $\frac{\partial u}{\partial t}=(-c) f^{\prime}(x-c t)$ and $\frac{\partial u}{\partial x}=f^{\prime}(x-c t)$. Comparing the two, we find that $\frac{\partial u}{\partial t}+c \frac{\partial u}{\partial x}=0$, or equivalently,

$$
\frac{\partial u}{\partial t}+\frac{\partial(c u)}{\partial x}=0
$$

Moreover, since $\frac{\partial^{2}}{\partial t^{2}}-c^{2} \frac{\partial^{2}}{\partial x^{2}}=\left(\frac{\partial}{\partial t}+c \frac{\partial}{\partial x}\right)\left(\frac{\partial}{\partial t}-c \frac{\partial}{\partial x}\right)$, this equation-in differential operator form-may be said to be "inside" the wave equation, in the sense that it is a multiplicative factor of its differential operator. From this point of view, it is perhaps much easier to understand the suspicion of Airy and Stokes toward the solitary wave. Clearly if, to a good approximation, the solitary wave has a profile of the form $\eta(x-c t)$, then $\eta$ does satisfy $\left(\frac{\partial^{2}}{\partial t^{2}}-c^{2} \frac{\partial^{2}}{\partial x^{2}}\right) \eta=0$. If this wave differential operator is related to $\eta$ 's (presumed) governing equation as a multiplicative factor, then it seems inconceivable that this wave differential operator should be anything other than the one Boussinesq found to be correct to first order,

$$
\frac{\partial^{2} \eta}{\partial t^{2}}=g h \frac{\partial^{2} \eta}{\partial x^{2}} .
$$

In that case, $c=\sqrt{g h}$, and we have the usual long-wavelength linear waves. For the speed of Russell's waves to exceed $\sqrt{g h}$-yet still have a simple governing equation-something rather subtle must be going on.

Everything might work out correctly if, somehow, the speed $c$ were to be hidden inside the machinery of the equation, or did not appear explicitly. Toward this end, we can replace
$c$ with a wave velocity function ${ }^{2} \varpi=\varpi(x, t)$ determined by the Boussinesq equation. The idea is that $\varpi$ just happens to be constant for solitary waves, but for general nonlinear waves it can depend on $x$ and $t$. Having replaced $c$ with $\varpi(x, t)$, we still expect the equation

$$
\frac{\partial u}{\partial t}+\frac{\partial(\varpi u)}{\partial x}=0
$$

to hold for nonlinear waves $u$ propagating to the right; in [2], Boussinesq uses conservation of mass to help justify this. Now, if $\varpi$ can be eliminated by writing it in terms of $\eta$, then we will have derived the governing equation for nonlinear long-wavelength water wave profiles $\eta$ propagating to the right. We will have obtained the famous Korteweg-de Vries equation.

Unfortunately, it is unclear where we are going to get our hands on an expression for the wave velocity $\varpi(x, t)$. If Russell's relation $c=c_{0}\left(1+b \eta_{0}\right)$ is our guide, we expect $\varpi$ to take the form $\varpi=\sqrt{g h}(1+(\ldots))$, but " $(\ldots)$ " is a bit of a mystery. We expect it will have something to do with the expression $\zeta=\frac{3 \eta^{2}}{2 h}+\frac{h^{2}}{3} \frac{\partial^{2} \eta}{\partial x^{2}}$ that appears in the Boussinesq equation. The only way I can see to move forward is to attempt to work out the velocity for a particular example - such as Lord Rayleigh's function

$$
\eta(x)=\eta_{0} \operatorname{sech}^{2}\left(\sqrt{\frac{3 \eta_{0}}{4 h^{3}}} x\right)
$$

If $\eta(x-c t)$ is a solution, what is $\zeta=\frac{3 \eta^{2}}{2 h}+\frac{h^{2}}{3} \frac{\partial^{2} \eta}{\partial x^{2}}$ ? In calculating $\frac{\partial^{2} \eta}{\partial x^{2}}$, we can use $\tanh ^{2}(\theta)=$ $1-\operatorname{sech}^{2}(\theta)$ with $\theta=\sqrt{\frac{3 \eta_{0}}{4 h^{3}}}(x-c t)$ to eliminate $\tanh ^{2}$ in favor of $\operatorname{sech}^{2}$ terms, arriving at

$$
\frac{\partial^{2} \eta}{\partial x^{2}}=-2 \eta_{0}\left(\frac{3 \eta_{0}}{4 h^{3}}\right)\left(3 \operatorname{sech}^{4}(\theta)-2 \operatorname{sech}^{2}(\theta)\right)
$$

Therefore,

$$
\frac{h^{2}}{3} \frac{\partial^{2} \eta}{\partial x^{2}}=\eta_{0}\left(\frac{\eta_{0}}{2 h}\right)\left(2 \operatorname{sech}^{2}(\theta)-3 \operatorname{sech}^{4}(\theta)\right)
$$

whence

$$
\zeta=\frac{3 \eta^{2}}{2 h}+\frac{h^{2}}{3} \frac{\partial^{2} \eta}{\partial x^{2}}=\frac{\eta_{0}^{2}}{h} \operatorname{sech}^{2}(\theta) .
$$

In other words, we may conclude ${ }^{3}$ that, for solitary waves,

$$
\zeta=\frac{\eta_{0}}{h} \eta
$$

Substituting $\frac{\eta_{0}}{h} \eta$ for $\zeta$ in the Boussinesq equation yields

$$
\frac{\partial^{2} \eta}{\partial t^{2}}=g h \frac{\partial^{2} \eta}{\partial x^{2}}+g \eta_{0} \frac{\partial^{2} \eta}{\partial x^{2}}
$$

[^1]Now the speed $c$ is in our grasp. The solitary wave $\eta$ traveling at speed $c$ satisfies both $\frac{\partial^{2} \eta}{\partial t^{2}}=c^{2} \frac{\partial^{2} \eta}{\partial x^{2}}$ and $\frac{\partial^{2} \eta}{\partial t^{2}}=g\left(h+\eta_{0}\right) \frac{\partial^{2} \eta}{\partial x^{2}}$. As $\eta$ 's partial derivatives are uniquely determined, we must have $c^{2}=g\left(h+\eta_{0}\right)$. In fact, according to [6], this was a result known to Russell from his solitary wave experiments. Factoring $h$ from the term in parentheses and applying the binomial approximation, we arrive at

$$
c=\sqrt{g h\left(1+\eta_{0} / h\right)} \approx \sqrt{g h}\left(1+\frac{\eta_{0}}{2 h}\right)
$$

when $\eta_{0} / h \ll 1$. Since a relatively low height is observed for solitary waves in practice, the condition $\eta_{0} / h \ll 1$ is satisfied, and we obtain excellent agreement between the implicit Korteweg-de Vries equation and Russell's experiments with solitary waves. (In particular, if $c_{0}=1=h$ for appropriate units, then the solitary wave exceeds the linear wave speed by half its height per unit time, an easily-remembered result.)

Since $\frac{\zeta}{\eta}=\frac{\eta_{0}}{h}$ for solitary waves, we would guess that, in analogy to $c=\sqrt{g h}\left(1+\frac{\eta_{0}}{2 h}\right)$, the wave velocity $\varpi$ is given by

$$
\varpi=\sqrt{g h}\left(1+\frac{\zeta}{2 \eta}\right) .
$$

I have provided motivation for this important equation because, once knowing it, we can derive the Korteweg-de Vries equation from

$$
\frac{\partial \eta}{\partial t}+\frac{\partial(\varpi \eta)}{\partial x}=0
$$

However, Boussinesq provides no apparent motivation for it. Instead, he defines an auxiliary function $\psi$ and offers a sort of mathematical proof that it is 0 . To see how this might be useful, note that to verify our guess for $\varpi$, it suffices to show that $\varpi-\sqrt{g h}=\sqrt{g h} \frac{\zeta}{2 \eta}$, or equivalently, $\eta(\varpi-\sqrt{g h})-\frac{\sqrt{g h}}{2} \zeta=0$. Boussinesq in [2] defines

$$
\psi=\eta(\varpi-\sqrt{g h})-\frac{\sqrt{g h}}{2} \zeta .
$$

If Boussinesq can show that $\psi=0$, then he will have verified our guess for $\varpi$. He does this by comparing $\frac{\partial \psi}{\partial x}$ with $\frac{\partial \psi}{\partial t}$. The calculation of the former is

$$
\frac{\partial \psi}{\partial x}=\frac{\partial(\varpi \eta)}{\partial x}-\sqrt{g h} \frac{\partial \eta}{\partial x}-\frac{\sqrt{g h}}{2} \frac{\partial \zeta}{\partial x}
$$

and, after using $\frac{\partial \eta}{\partial t}+\frac{\partial(\varpi \eta)}{\partial x}=0$, we arrive at

$$
\frac{\partial \psi}{\partial x}=-\frac{\partial \eta}{\partial t}-\sqrt{g h} \frac{\partial \eta}{\partial x}-\frac{\sqrt{g h}}{2} \frac{\partial \zeta}{\partial x} .
$$

To find $\frac{\partial \psi}{\partial t}$, however, we require an expression for $\frac{\partial(\varpi \eta)}{\partial t}$. One may be arrived at by using $\frac{\partial \eta}{\partial t}+\frac{\partial(\varpi \eta)}{\partial x}=0$ to substitute $-\frac{\partial(\varpi \eta)}{\partial x}$ for $\frac{\partial \eta}{\partial t}$ in the Boussinesq equation. We get

$$
\frac{\partial}{\partial t}\left(-\frac{\partial(\varpi \eta)}{\partial x}\right)=g h \frac{\partial^{2} \eta}{\partial x^{2}}+g h \frac{\partial^{2} \zeta}{\partial x^{2}}
$$

or, equivalently,

$$
\frac{\partial}{\partial x}\left(-\frac{\partial(\varpi \eta)}{\partial t}\right)=g h \frac{\partial^{2} \eta}{\partial x^{2}}+g h \frac{\partial^{2} \zeta}{\partial x^{2}}
$$

which may be integrated with respect to $x$. With the constant ignored due to vanishing at infinity, we obtain

$$
\frac{\partial(\varpi \eta)}{\partial t}=-g h \frac{\partial \eta}{\partial x}-g h \frac{\partial \zeta}{\partial x} .
$$

Finally, we may calculate $\frac{\partial \psi}{\partial t}$; recall that $\psi=\varpi \eta-\sqrt{g h} \eta-\frac{\sqrt{g h}}{2} \zeta$. We get

$$
\frac{\partial \psi}{\partial t}=\frac{\partial(\varpi \eta)}{\partial t}-\sqrt{g h} \frac{\partial \eta}{\partial t}-\frac{\sqrt{g h}}{2} \frac{\partial \zeta}{\partial t}
$$

Now Boussinesq says that we may use $\frac{\partial \zeta}{\partial t} \approx-\sqrt{g h} \frac{\partial \zeta}{\partial x}$ to first order since $\eta$-and hence $\zeta$-is propagating to the right; recall that we want $\varpi$ for such $\eta$. Apparently this approximation will not disturb the desired order of approximation elsewhere. Together with our equation for $\frac{\partial(\varpi \eta)}{\partial t}$, we get

$$
\frac{\partial \psi}{\partial t}=-g h \frac{\partial \eta}{\partial x}-g h \frac{\partial \zeta}{\partial x}-\sqrt{g h} \frac{\partial \eta}{\partial t}+\frac{g h}{2} \frac{\partial \zeta}{\partial x}
$$

which may be simplified to

$$
\frac{\partial \psi}{\partial t}=-g h \frac{\partial \eta}{\partial x}-\frac{g h}{2} \frac{\partial \zeta}{\partial x}-\sqrt{g h} \frac{\partial \eta}{\partial t}=\sqrt{g h} \frac{\partial \psi}{\partial x} .
$$

This says that $\psi$ is a function of the form $\psi=\psi(x+\sqrt{g h} t)$, so that $\psi$ propagates to the left! Since, however, $\psi$ is written in terms of various rightward-traveling functions, this is manifestly impossible unless $\psi=0$. According to [13], one of Boussinesq's contemporaries, de Saint Venant, found this proof so unilluminating that he considered it a high priority to provide a more physical demonstration. However, I consider this one proof to provide sufficient illustration of French elegance for our purposes.

In 1885, de Saint Venant was able to offer mathematical explanations for Stokes and Airy's mistakes. It seems that these mistakes erroneously suggested that a wave profile cannot be maintained unless nonlinear effects are negligible, in which case dispersion would be an issue-in solitary waves, dispersion and nonlinearity somehow balance. More subtle mistakes involved confusion regarding the differences between linear and nonlinear phenomena. For example, the solitary waves Russell observed always manifested as a bump ( $\eta>0$ ), and never as a depression $(-\eta)$; these are "gravity waves," so it perhaps makes sense that up and down must be treated differently. On the other hand, if a given function $\eta$ solves a linear equation, any multiple of it-including $-\eta$-solves it too. Thus, for example, if Stokes
is somehow convinced he has shown there can be no solitary-wave-like depression, then (assuming he treats the phenomenon with linear methods) he will be erroneously convinced of the impossibility of the existence of solitary wave bumps.

Due to the ad hoc nature of Rayleigh's approximations, doubts still remained about the solitary wave's ability to persist without dispersing beyond short-time regimes. Finally, in 1895, the Dutch mathematician Korteweg and his student de Vries derived an explicit nonlinear differential equation modeling rightward-traveling water waves on a canal (contained in Boussinesq's paper [2] only implicitly). Their equation, now called the "KdV" equation after their initials (thus explaining the name of the "KdV Institute" in Amsterdam), is obtained from the Euler equation for a nonviscous, incompressible fluid with irrotational flow; a derivation may be found ${ }^{4}$ in Appendix A of [5]. Taking this equation into a weakly nonlinear regime, "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," according to [14]. Importantly, the stability of their equation-in the absence of friction-provided a clear, unified framework that removed the lingering suspicions regarding the existence of solitary waves. They point up these suspicions in closing the first paragraph of their famous paper [15]:

In such excellent treatises on hydrodynamics as those of Lamb and Basset, we find that even when friction is neglected long waves in a rectangular canal must necessarily change their form as they advance, becoming steeper in front and less steep behind. Yet since the investigations of de Boussinesq, Lord Rayleigh, and St. Venant on the solitary wave, there has been some cause to doubt the truth of this assertion. Indeed, if the reasons adduced were really decisive, it is difficult to see why the solitary wave should make an exception; but even Lord Rayleigh and McCowan, who have successfully and thoroughly treated the theory of this wave, do not directly contradict the statement in question. They are, as it seems to us, inclined to the opinion that the solitary wave is only stationary to a certain approximation.

The equation of Korteweg and de Vries for the height $\eta(x, t)$ of a nonlinear surface wave ${ }^{5}$ above its equilibrium level, where $h$ is the fluid depth, is given by

$$
\frac{1}{c_{0}} \frac{\partial \eta}{\partial t}+\frac{\partial \eta}{\partial x}+\frac{3}{2 h} \eta \frac{\partial \eta}{\partial x}+\frac{h^{2}}{6} \frac{\partial^{3} \eta}{\partial x^{3}}=0
$$

[^2]where $c_{0}=\sqrt{g h}$ is the propagation speed of the linear waves in the long-wavelength limit. The reader is welcome to plug our above expression for $\varpi$ into $\frac{\partial \eta}{\partial t}+\frac{\partial(\varpi \eta)}{\partial x}=0$ and divide by $c_{0}$ to obtain precisely this equation, though Boussinesq did not ${ }^{6}$ do this. The reader may also wish to verify directly that if $\eta>0$ and $-\eta$ are both solutions of this equation then $\eta$ must be constant. Indeed, after multiplying the equation for $\eta$ by -1 and subtracting this from the equation for $-\eta$, we will have $\eta \frac{\partial \eta}{\partial x}=0$, whence $\frac{\partial \eta}{\partial x}=0$; plugging this into the Korteweg-de Vries equation then yields $\frac{\partial \eta}{\partial t}=0$ as well, so that $\eta$ is a time-independent constant. This is a good illustration of how intuitions regarding linear equations need not extend to the nonlinear.

Following [5] and Korteweg-de Vries, we can attempt to isolate the nonlinear effects by changing to a frame moving at speed $c_{0}$. Introducing the new variable $X=x-c_{0} t$ eliminates the second term, leaving

$$
\frac{1}{c_{0}} \frac{\partial \eta}{\partial t}+\frac{3}{2 h} \eta \frac{\partial \eta}{\partial X}+\frac{h^{2}}{6} \frac{\partial^{3} \eta}{\partial X^{3}}=0
$$

Having this equation in hand, we can ask whether it confirms Rayleigh's calculation of the solitary wave profile. Assume that we have a solution $\eta$ of this equation propagating at constant speed, so we may write $\eta(X, t)=\eta(X-c t)$. Plugging this into the equation, we get

$$
\frac{1}{c_{0}}(-c) \eta^{\prime}+\frac{3}{2 h} \eta \eta^{\prime}+\frac{h^{2}}{6} \eta^{\prime \prime \prime}=0 .
$$

Rearranging, we have

$$
\eta^{\prime \prime \prime}+\frac{9}{h^{3}} \eta \eta^{\prime}-\frac{6 c}{h^{2} c_{0}} \eta^{\prime}=0 .
$$

Now $\eta \eta^{\prime}=\left(\eta^{\prime}\right)^{2} / 2$, so integrating once yields, for some constant $A$,

$$
\eta^{\prime \prime}+\frac{9}{2 h^{3}} \eta^{2}-\frac{6 c}{h^{2} c_{0}} \eta=A
$$

If we multiply by $\eta^{\prime}$ and observe that $\eta^{\prime \prime} \eta^{\prime}=\left(\left(\eta^{\prime}\right)^{2} / 2\right)^{\prime}$, we can integrate again to obtain

$$
\frac{\left(\eta^{\prime}\right)^{2}}{2}+\frac{3}{2 h^{3}} \eta^{3}-\frac{6 c}{2 h^{2} c_{0}} \eta^{2}=A \eta+B .
$$

If we require $\eta, \eta^{\prime}, \eta^{\prime \prime} \rightarrow 0$ as $\xi=X-c t \rightarrow \pm \infty$, then the equation for $\eta^{\prime \prime}$ makes $A=0$, and the equation for $\eta^{\prime}$ makes $B=0$. (If we do not take $A$ and $B$ to be zero, then we must rely on Jacobi elliptic functions to find the solution. This is more sophistication than is needed for finding the solitary wave profile.) We arrive at the equation

$$
\left(\eta^{\prime}\right)^{2}=\eta^{2}\left(\frac{6 c}{h^{2} c_{0}}-\frac{3}{h^{3}} \eta\right)=\frac{6 c}{h^{2} c_{0}} \eta^{2}\left(1-\frac{c_{0}}{2 c h} \eta\right)
$$

[^3]We take the square root of both sides:

$$
\eta^{\prime}= \pm \sqrt{\frac{6 c}{h^{2} c_{0}}} \cdot \eta \sqrt{1-\frac{c_{0}}{2 c h} \eta} .
$$

This renders the equation separable:

$$
\int \frac{d \eta}{\eta \sqrt{1-\frac{c_{0}}{2 c h} \eta}}=\int \pm \sqrt{\frac{6 c}{h^{2} c_{0}}} d \xi
$$

To evaluate the integral on the left we let $\eta=\frac{2 c h}{c_{0}} \operatorname{sech}^{2}(\theta)$, recalling that $1-\operatorname{sech}^{2}(\theta)=$ $\tanh ^{2}(\theta)$ in order to eliminate the square root.

$$
\int \frac{d \eta}{\eta \sqrt{1-\frac{c_{0}}{2 c h} \eta}}=\int \frac{\frac{2 c h}{6} 2 \operatorname{sech}(\theta)(-\operatorname{sech}(\theta) \tanh (\theta)) d \theta}{\frac{2 c h}{6} \operatorname{sech}^{2}(\theta) \tanh (\theta)}
$$

Since $\left(\operatorname{sech}^{2}\right)^{\prime}=2 \operatorname{sech}(-\operatorname{sech} \cdot \tanh )$, all functions of $\theta$ cancel. Thus, $\eta(\xi)=\frac{2 c h}{c_{0}} \operatorname{sech}^{2}(\theta)$, where (taking 0 for the constant of integration and ignoring signs since $\operatorname{sech}^{2}(\theta)$ is an even function of $\theta$ ) the equation

$$
2 \theta=\sqrt{\frac{6 c}{h^{2} c_{0}}} \xi
$$

determines $\theta$. That is,

$$
\eta(\xi)=\frac{2 c h}{c_{0}} \operatorname{sech}^{2}\left(\frac{1}{2} \sqrt{\frac{6 c}{h^{2} c_{0}}} \xi\right) .
$$

To compare this profile with Lord Rayleigh's, let $\eta_{0}=\frac{2 c h}{c_{0}}$. Then $c=c_{0} \frac{\eta_{0}}{2 h}$. In terms of $\eta_{0}$,

$$
\eta(\xi)=\eta_{0} \operatorname{sech}^{2}\left(\frac{1}{2} \sqrt{\frac{3 \eta_{0}}{h^{3}}} \xi\right)=\eta_{0} \operatorname{sech}^{2}\left(\sqrt{\frac{3 \eta_{0}}{4 h^{3}}} \xi\right)
$$

in perfect agreement with Lord Rayleigh's function.
Introducing appropriate dimensionless variables reduces the Korteweg-de Vries equation to its "standard form." Unfortunately, the choice of "standard form" is not uniform across disciplines, authors, or even chapters within the same book. Fortunately, they differ only by linear changes of dependent and independent variables. For example, while Kasman, in [14], chooses an equation of the form

$$
\frac{\partial u}{\partial t}=\frac{3}{2} u \frac{\partial u}{\partial x}+\frac{1}{4} \frac{\partial^{3} u}{\partial x^{3}}
$$

as standard form, Dauxois and Peyrard, in [5] p. 165, take

$$
\frac{\partial u}{\partial t}-6 u \frac{\partial u}{\partial x}+\frac{\partial^{3} u}{\partial x^{3}}=0
$$

as theirs. To reduce the equation in $t$ and $X$ to the latter (note the minus sign!), let $u=-\eta / h$, $\xi=X / \frac{2 h}{\sqrt{6}}$, and $\tau=t / \frac{8 h}{c_{0} \sqrt{6}}$ (I found these substitutions by working backward), renaming $\xi$ and $\tau$ as $x$ and $t$ after sufficient use of the old $x$ and $t$ through the chain rule; finding
the appropriate substitutions for the former (Kasman) equation are left to the interested reader. (Note that these new variables are really dimensionless: the units for $\frac{8 h}{c_{0} \sqrt{6}}=\frac{8}{\sqrt{6}} \sqrt{\frac{h}{g}}$ are $\sqrt{\frac{\mathrm{m}}{\mathrm{m} / \mathrm{s}^{2}}}=\mathrm{s}$, so $\tau$ has units of seconds/second, i.e., is dimensionless.)

Perhaps surprisingly, Korteweg and de Vries in [15] were able to use cutting-edge techniques from pure mathematics to find further solutions to their nonlinear partial differential equation. These solutions were in terms of Jacobian elliptic functions such as $\operatorname{sn}(x)$ and $\mathrm{cn}(x)$, which describe what are now called, in honor of $\mathrm{cn}(x)$, the "cnoidal waves." The ability to find such solutions may have been dismissed as a coincidence (the supposed coincidence being between equations pertaining to KdV solutions and certain equations from the theory of elliptic functions). Either way, little further work appears to have been done on solitary waves or the Korteweg-de Vries equation until the 1960's, when a paper by Fermi, Pasta, and Ulam written at Los Alamos, [9], became declassified.

## 3. Turning Points in the Development of Soliton Theory

Just as Airy and Stokes mistakenly assumed that a nonlinear wave equation would destroy any initial wave profile, so the Los Alamos investigators believed that any nice, ordered initial data set would quickly be distorted beyond recognition under evolution by a nonlinear equation-but decided to test their intuition with a computer. Ulam, in [25] later, said:

Fermi expressed often a belief that future fundamental theories in physics may involve non-linear operators and equations, and that it would be useful to attempt practice in the mathematics needed for the understanding of non-linear 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 non-linear systems may not always be justified.

Fermi, Pasta, and Ulam wanted to study how a crystal-like structure evolves toward thermal equilibrium with a computer. They did this in [9] by simulating a chain of $N=64$ particles of unit mass linked by a qadratic interaction potential and a weakly nonlinear cubic interaction. The system they considered may be described by the one-dimensional Hamiltonian ${ }^{7}$

$$
H=\sum_{j=0}^{N-1} \frac{1}{2} p_{j}^{2}+\sum_{j=0}^{N-1} \frac{1}{2} K\left(u_{j+1}-u_{j}\right)^{2}+\frac{K \alpha}{3} \sum_{j=0}^{N-1}\left(u_{j+1}-u_{j}\right)^{3} .
$$

[^4]Here $u_{j}$ is the displacement along the chain of particle or atom $j$ with respect to its equilibrium position, and $p_{j}$ is its momentum. The small coefficient $\alpha \ll 1$ measures the strength of the cubic contribution to the interaction potential (which gives rise to a quadratic force term). The two ends of the chain were kept fixed by analogy with a vibrating string, so $u_{0}=0=u_{N}$. In their paper [9], they explain, "The purpose of our computations was to see how, due to nonlinear forces perturbing the periodic linear solution, the string would assume more and more complicated shapes, and, for $t$ tending to infinity, would get into states where all the Fourier modes acquire increasing importance." For the discrete system under consideration, a typical approach would be to think in terms of "normal modes," which play the same role as the Fourier modes in the continuum limit. After appropriate normalization ${ }^{8}$ these are related to the displacements through $A_{k}=\sqrt{\frac{2}{N}} \sum_{j=0}^{N-1} u_{j} \sin (j k \pi / N)$ and have the frequency values $\omega_{k}^{2}=4 K \sin ^{2}(k \pi / 2 N)$. Thinking of these modes as a basis, we can rewrite the Hamiltonian $H$ in terms of them as

$$
H=\frac{1}{2} \sum_{k}\left(\left(\dot{A}_{k}\right)^{2}+\omega_{k}^{2} A_{k}^{2}\right)+\alpha \sum_{k, \ell, m} c_{k, \ell, m} A_{k} A_{\ell} A_{m}
$$

The $c_{k, \ell, m}$ witness a coupling between modes; if $\alpha=0$, then the equations of motion become linear-and decoupled. The problem in that case is treated in books on statistical mechanics such as [23].

Due to the coupling between modes, Fermi, Pasta, and Ulam thought that energy introduced in a single mode ( $k=1$ in their simulation) would drift into the other modes, until the energy would be shared between all modes. This would be an example of equipartition of energy, essentially a hypothesis that lies at the foundation of statistical physics. In the beginning their calculations suggested this would be true. As their paper states, "Starting in one problem with a quadratic force and a pure sine wave as the initial position of the string, we indeed observe initially a gradual increase of energy in the higher modes as predicted."

One day, according to Metropolis ${ }^{9}$ in [19], they accidentally left the program ${ }^{10}$ running long after the steady state had apparently been reached. When they realized they had forgotten to kill the program and returned to the computer room, they noticed the equipartition of energy had ceased. "For example, mode 2 decides, as it were, to increase rather rapidly at the cost of all other modes and becomes predominant. At one time, it has more energy than all the others put together!" To their great surprise, after 157 of mode 1's periods, almost all of the energy was back in the lowest mode, as if the system had reset itself. "Finally, at a later

[^5]time mode 1 comes back to within one percent of its initial value so that the system seems to be almost periodic." Contrary to the authors' expectations, the drift of mode 1's energy to a steady, energy-sharing state does not occur. "In other words, the systems certainly do not show mixing."

The results of these numerical experiments challenged readers in the sixties to conduct their own numerical investigations to resolve this "FPU paradox," that nonlinearity does not guarantee equipartition of energy. To understand it, researchers were forced by necessity to stop thinking in terms of the normal modes from the linear theory. They had to consider the full nonlinearity intrinsically, not in Fourier space, but in real space. According to [5], Kruskal at Princeton University (among those in the boat depicted in Figure 1) and Zabusky at Bell Labs found an explanation by taking a suitable limit of the discrete system of connected vibrating masses that Fermi, Pasta, Ulam, and Tsingou had used in the original experiment. To see what they were doing, begin with the equations of motion derived from the Fermi-Pasta-Ulam Hamiltonian $H=\sum_{j=0}^{N-1} \frac{1}{2} p_{j}^{2}+\sum_{j=0}^{N-1} \frac{1}{2} K\left(u_{j+1}-u_{j}\right)^{2}+\frac{K \alpha}{3} \sum_{j=0}^{N-1}\left(u_{j+1}-u_{j}\right)^{3}$, namely

$$
\ddot{u}_{j}=K\left(u_{j+1}+u_{j-1}-2 u_{j}\right)+K \alpha\left[\left(u_{j+1}-u_{j}\right)^{2}-\left(u_{j}-u_{j-1}\right)^{2}\right] .
$$

It is helpful to highlight solutions having a small amplitude with respect to the lattice spacing of the particles at rest, which we denote by $a$. This is done by substituting $u=\epsilon a v$ with $\epsilon \ll 1$. Dividing through by $\epsilon a$ and perhaps thinking of $c \equiv a \sqrt{K}$ as the speed of sound, we obtain

$$
\ddot{v}_{j}=\frac{c^{2}}{a^{2}}\left(v_{j+1}+v_{j-1}-2 v_{j}\right)+\frac{\epsilon c^{2} \alpha}{a}\left[\left(v_{j+1}-v_{j}\right)^{2}-\left(v_{j}-v_{j-1}\right)^{2}\right] .
$$

Recall that Fermi, Pasta, and Ulam began by exciting the normal mode with the lowest wave number, $k=1$. It turns out that the system began to experience recurrence or near-periodicity before any of the large-wavenumber modes became excited. Because of this, Zabusky and Kruskal would have been justified in restricting their attention to low-wavenumber modes, which amounts to investigating long-wavelength behavior in the continuum limit. To more readily estimate orders of magnitude, the dimensionless variables $X=\epsilon x / a$ and $\theta=c t / a$ are introduced. The presence of $\epsilon$ in the definition of $X$ expresses the long-wavelength idea that $v$ varies slowly with respect to spatial displacement. Substituting $v_{j \pm 1}=v\left(x_{j} \pm a\right)=v(X \pm \epsilon)$ in Taylor-expansion form into the $(X, \theta)$-equation for $\ddot{v}_{j}$, we get

$$
\frac{c^{2}}{a^{2}} \frac{\partial^{2} v}{\partial \theta^{2}}=\frac{c^{2}}{a^{2}}\left(\epsilon^{2} \frac{\partial^{2} v}{\partial X^{2}}+\frac{\epsilon^{4}}{12} \frac{\partial^{4} v}{\partial X^{4}}\right)+\frac{\epsilon c^{2} \alpha}{a}\left[(\ldots)_{+}^{2}-(\ldots)_{-}^{2}\right]
$$

where

$$
(\ldots)_{+}=\left(\epsilon \frac{\partial v}{\partial X}+\frac{\epsilon^{2}}{2} \frac{\partial^{2} v}{\partial X^{2}}+\ldots\right)
$$

and

$$
(\ldots)_{-}=\left(\epsilon \frac{\partial v}{\partial X}-\frac{\epsilon^{2}}{2} \frac{\partial^{2} v}{\partial X^{2}}+\ldots\right)
$$

Only keeping terms up to order $\epsilon^{4}$ (and noticing the $\epsilon$ preceding the term in brackets) yields

$$
\frac{\partial^{2} v}{\partial \theta^{2}}=\epsilon^{2} \frac{\partial^{2} v}{\partial X^{2}}+\frac{\epsilon^{4}}{12} \frac{\partial^{4} v}{\partial X^{4}}+2 \epsilon^{4} a \alpha \frac{\partial v}{\partial X} \frac{\partial^{2} v}{\partial X^{2}} .
$$

If we kept terms just up to order $\epsilon^{2}$, we would have the wave equation for profiles moving at speed $\epsilon$. By changing to the moving-frame space variable $\xi=X-\epsilon \theta$ we can isolate any remaining behavior. Due to the $\frac{\partial^{4} v}{\partial X^{4}}$ term and nonlinear $\frac{\partial v}{\partial X} \frac{\partial^{2} v}{\partial X^{2}}$ term, the solution will still evolve in the frame translating at speed $\epsilon$, but because the coefficient $\epsilon^{4}$ is so small it will change slowly. We will be able to see the physics more clearly if we "speed up time," or equivalently, pick a new timescale appropriate for slow-time change. To bring the physics to the surface, then, we let $\theta=\epsilon^{-3} \tau$; when $\tau$ increases, $\theta$ increases a lot, so time is sped up. Equivalently, we let $\tau=\epsilon^{3} \theta$; thus, $\tau$ is the parameter for evolving slow-time change. The choice of exponent for $\epsilon$ is forced on us by the need to pick out a scale where the terms containing spatial derivatives and those containing time derivatives interact at the same order in $\epsilon$.

Changing from $x$ and $\theta$ to $\xi=x-\epsilon \theta$ and $\tau=\epsilon^{3} \theta$, we have

$$
\frac{\partial}{\partial x}=\frac{\partial \xi}{\partial x} \frac{\partial}{\partial \xi}+\frac{\partial \tau}{\partial x} \frac{\partial}{\partial \tau}=\frac{\partial}{\partial \xi}
$$

and

$$
\frac{\partial}{\partial \theta}=\frac{\partial \xi}{\partial \theta} \frac{\partial}{\partial \xi}+\frac{\partial \tau}{\partial \theta} \frac{\partial}{\partial \tau}=(-\epsilon) \frac{\partial}{\partial \xi}+\epsilon^{3} \frac{\partial}{\partial \tau} .
$$

We have $\frac{\partial^{2} v}{\partial \theta^{2}}$ on the left side of the equation, so we calculate

$$
\frac{\partial^{2}}{\partial \theta^{2}}=\epsilon^{2} \frac{\partial^{2}}{\partial \xi^{2}}-2 \epsilon^{4} \frac{\partial^{2}}{\partial \xi \partial \tau}+\epsilon^{6} \frac{\partial^{2}}{\partial \tau^{2}} .
$$

Up to order $\epsilon^{4}$, our equation is now

$$
\epsilon^{2} \frac{\partial^{2} v}{\partial \xi^{2}}-2 \epsilon^{4} \frac{\partial^{2} v}{\partial \xi \partial \tau}=\epsilon^{2} \frac{\partial^{2} v}{\partial \xi^{2}}+\frac{\epsilon^{4}}{12} \frac{\partial^{4} v}{\partial \xi^{4}}+2 \epsilon^{4} a \alpha \frac{\partial v}{\partial \xi} \frac{\partial^{2} v}{\partial \xi^{2}}
$$

Deleting the common term $\epsilon^{2} \frac{\partial^{2} v}{\partial \xi^{2}}$ from both sides and transposing the term with mixed partials, we conclude that

$$
2 \epsilon^{4} \frac{\partial^{2} v}{\partial \xi \partial \tau}+\frac{\epsilon^{4}}{12} \frac{\partial^{4} v}{\partial \xi^{4}}+2 \epsilon^{4} a \alpha \frac{\partial v}{\partial \xi} \frac{\partial^{2} v}{\partial \xi^{2}}=0
$$

Everything is to the same order in $\epsilon$, and all derivatives include $\frac{\partial}{\partial \xi}$, so we can clean up by defining $w=\frac{\partial v}{\partial \xi}$ and dividing by the coefficient of the time-derivative term:

$$
\frac{\partial w}{\partial \tau}+\frac{1}{24} \frac{\partial^{3} w}{\partial \xi^{3}}+a \alpha w \frac{\partial w}{\partial \xi}=0
$$

Remarkably, we've arrived at a form of the Korteweg-de Vries equation, the key to Zabusky and Kruskal's explanation of the FPU paradox. In fact, their 4-page paper [16] is focused
exclusively on observations of numerical solutions to the Korteweg-de Vries equation. A "phenomenological description" of recurrence in the Fermi-Pasta-Ulam problem is considered an application of their paper, and is only touched upon in the first paragraph. I have included this derivation because I was curious to see how the connection to the Korteweg-de Vries equation could be made explicit.

In their numerical ${ }^{11}$ experiments [16], Zabusky and Kruskal made a couple of striking observations:

- Positive initial profiles supported on intervals of finite extent tended to evolve under the KdV equation into a finite number of humps, each behaving like one of Russell's solitary waves; and


Figure 3. The "collision" of two solitary waves, that is, a 2-soliton.

- if two such humps approach each other, they "interact," e.g., the sum of the heights of the wave profiles during an interaction often decreases and they may decelerate briefly as the continuum or lattice is compressed. After "collision," two humps emerge which are hardly distinguishable in shape or speed, if at all, from the originals.

The first observation suggests that solitary waves in real space may play the role of normal modes for solutions of the KdV equation, similar to how the motion of a simplified vibrating string may be described with fundamental sine wave modes in Fourier analysis. The second observation, where waves are colliding and coming apart like fundamental particles (or even billiard balls), calls into question whether these should be called "solitary waves" at all: they aren't always solitary, and don't always behave like waves. Looking for a name reminiscent of fundamental particles like the term "electron," Kruskal and Zabusky settled on "solitons," with the prefix "solit-" deriving from "solitary." More precisely, a solution (even if it depends

[^6]on more than one spatial variable or solves a different "soliton equation" than the KdV equation) which typically has $n$ humps (outside of interactions) is called an $n$-soliton.


Figure 4. A picture of a plot from the Zabusky-Kruskal paper [16]. It depicts the evolution of the Fermi-Pasta-Ulam initial condition in real space, at three different times. The dotted initial graph is of a sine wave, corresponding to wavenumber $k=1$. The dashed later graph depicts a sharp fall around distance 0.50 that looks like it could turn into a discontinuity. The solid final graph demonstrates the linear relation between the amplitude of a soliton and the excess of its speed over the sound velocity.

What happens when a fundamental sine wave provides the initial condition ${ }^{12}$ for the Korteweg-de Vries equation-do we see what happens when the first normal mode is excited in the Fermi-Pasta-Ulam paper? At first, the nonlinear term in the Korteweg-de Vries equation tends to make sharp fronts in the data, as if a discontinuity is about to form. This steepening requires higher and higher Fourier modes to describe it, similarly to how Fermi, Pasta, and Ulam observed higher normal frequencies being employed when their system appeared to be mixing. Then this steepness is jerked into a series of pulses-the third derivative $\frac{\partial^{3} w}{\partial \xi^{3}}$ in the Korteweg-de Vries equation prevents a shock from being realized.

[^7]The pulses become solitons; like solitary waves, their speed is roughly proportional to their maximum height in the frame moving at the long-wavelength linear speed. This explains why a family of solitons which stay equidistant should have their peaks lie on a moving line, an observation illustrated in [16]. Under periodic boundary conditions-where the system has finite spatial extent-the solitons must occasionally return to positions near their initial locations, almost restoring initial conditions. Our expectation of this occurring in an actual physical system is due to the ergodic hypothesis-that in general, each particle, or soliton, will travel within any prescribed closeness to a given point, arbitrarily often-the same principle at the root of the equipartition theorem. This provides a kind of satisfaction with underlying principles ${ }^{13}$ that may make up for dissatisfaction with the consequences.

Investigations into soliton equations might have continued to develop qualitatively, growing alongside numerical simulations of partial differential equations. It could have gradually extended into the analysis of two-spatial-dimension generalizations of the KdV equation, such as the Kadomtsev-Petviashvili (KP) equation (cf. Figures 5 and 6). Instead, interest in this subject experienced explosive growth with the discovery of the "inverse scattering transform," introduced in the 1967 paper, [10], of Gardner, Greene, Kruskal, and Miura. They demonstrated that $n$-soliton solutions of the KdV equation could be obtained more-or-less explicitly (often in closed form) from an initial profile using analytical techniques. This provided a precise analytical counterpart to the pictures of initial profiles breaking into a family of advancing solitary waves. Furthermore, it spurred a variety of further developments in physics and mathematics in which "soliton" or "KdV-like" equations were used in the physical descriptions of, e.g., plasma physics, solid state physics, and molecular biology, with the inverse scattering transform being generalized to solve these equations. What is valuable about modeling with nonlinear soliton equations rather than "linearizing" problems is that significant features of the physics may be preserved that would otherwise be obscured by linearization. Where nonlinearity is unavoidable, usage of soliton equations and solution by inverse scattering transform may be essential to a researcher's toolkit, and may be best used in conjunction with linearization-as in "linearization around a soliton solution," discussed in [5]. However, because of the mathematical sophistication required to use the techniques, we will merely illustrate inverse scattering for the KdV equation, emphasizing how the process works for the simplest example.

[^8]

Figure 5. A 2-soliton collision dependent on two spatial variables.


Figure 6. A 2-soliton collision on the surface of the ocean.

## 4. Korteweg-de Vries Solutions and Potentials

Given a potential energy function $V(x)$-in either classical or quantum mechanics-we can use it to glean a variety of information about the corresponding physical system. Sometimes we don't even need to solve any differential equations to exploit this. For example, recall that, along the way to deriving Rayleigh's function $\eta(x)$ from the Korteweg-de Vries equation, we came upon the equation

$$
\eta^{\prime \prime}+\frac{9}{2 h^{3}} \eta^{2}-\frac{6 c}{h^{2} c_{0}} \eta=A,
$$

where $A=0$ due to the boundary conditions. This can be written in the form $F=m a$ from classical mechanics with $m=1$, that is, $F=\eta^{\prime \prime}$. We then have $F(\eta) \equiv \frac{6 c}{h^{2} c_{0}} \eta-\frac{9}{2 h^{3}} \eta^{2}$. Now imagining this to describe a one-dimensional physical system, we can write down the potential $V(\eta)$ from $F(\eta)=-V^{\prime}(\eta)$. Taking the integration constant to be 0 , we get

$$
V(\eta)=\frac{3}{2 h^{3}} \eta^{3}-\frac{6 c}{2 h^{2} c_{0}} \eta^{2} .
$$

Remarkably, knowledge of $V(\eta)$ explains why solitary canal waves always manifest as bumps, and never as depressions. Here I draw a graph of $V$ against $\eta$; see Figure 7 for the picture.


Figure 7. A picture of the author's graph of $V$ against $\eta$.

Imagine that a solitary wave profile is observed from left to right. At $x=-\infty$ it is at height $\eta=0$. If nontrivial, however, it will eventually deviate from this. If, at any point, $\eta$ is positive, then by conservation of energy $\eta(x)$ will increase up to $\eta_{0}$; on the graph this corresponds to traveling right, down the potential well and up to the second zero of $V(\eta)$. As $\eta=0$ at $x=+\infty$, the height $\eta$ will then decrease; this corresponds to traveling left, back to the origin of the potential diagram. Actually, one could imagine an $\eta(x)$ with multiple positive bumps, corresponding to multiple trips about the potential well. However, we never have $x_{0}$ with $\eta\left(x_{0}\right)<0$. If this ever happened then the slippery slope of the potential well for $\eta<0$ would cause $\eta$ to continue to decrease as lower and lower potential energies are reached, eventually diverging. Thus, only nonnegative solitary wave profiles occur in practice.

What about the potential $V(x)$ in a nonrelativistic quantum mechanical system? In that case we can use $V(x)$ in the Schrödinger equation to determine the allowed energy levels for bound states of the system. For our purposes, we normalize the Schrödinger equation to

$$
-\psi_{x x}+\underset{23}{V(x)} \psi=\lambda \psi,
$$

where $\psi_{x x} \equiv \frac{\partial^{2} \psi}{\partial x^{2}}$, with eigenvalues $\lambda$ being allowed energies. We also write the equation as

$$
\psi_{x x}+(\lambda-V(x)) \psi=0 .
$$

If we are more ambitious, we can also consider the continuous spectrum associated with the equation and find its corresponding "scattering data."

Let's be a bit more precise about the distinction between the discrete spectrum (allowed energies) and the continuous. Since we generally assume $V(x) \rightarrow 0$ as $x \rightarrow \pm \infty$, when $\lambda<0$ we get $\psi_{x x} \sim-\lambda \psi$, exponential asymptotics for $\psi$, as $x \rightarrow \pm \infty$. If we suppose $\psi$ is proportional, asymptotically, to $e^{\sqrt{|\lambda|} x}$ for $x \rightarrow-\infty$, then in general $\psi$ will be asymptotic to a linear combination of exponentially growing and decaying terms as $x \rightarrow+\infty$. If we can find a $\psi$ with exponential decay at both $+\infty$ and $-\infty$ ends, then it is square-integrable, and we have a bound state. In that case, $\lambda$ is part of the discrete spectrum for the equation. For typical $V(x)$ the discrete spectrum will be finite, consisting of $\lambda_{1}, \ldots, \lambda_{N}$, and we often define $\kappa_{i}=\sqrt{\left|\lambda_{i}\right|}$ for notational convenience. When $\lambda>0$ we get $\psi_{x x} \sim-\lambda \psi$, so solutions $\psi$ are, asymptotically, linear combinations of $e^{ \pm i \sqrt{\lambda} x}$. These functions are not square-integrable, and so don't form bound states, but they are relevant to the overall description of the system. Indeed, given $k=\sqrt{\lambda}$ we can generally define a solution $\hat{\psi}$ which is a special linear combination of the possible oscillatory behaviors at infinity, namely $\hat{\psi} \sim e^{-i k x}+b e^{i k x}$ as $x \rightarrow+\infty$ and $\hat{\psi} \sim a e^{-i k x}$ as $x \rightarrow-\infty$. It turns out that the constants $a$ and $b$, or rather the functions $a(k)$ and $b(k)$, are uniquely determined on their domain for a given potential $V(x)$, and that $|a|^{2}+|b|^{2}=1$. This gives rise to the image of an incident wave $e^{-i k x}$ traveling leftward from $+\infty ;|a|^{2}$ is the probability the wave is transmitted past the potential to $-\infty$ (think of $a e^{-i k x}$ ), and $|b|^{2}$ is the probability that the wave is reflected back to $+\infty$ (think of $b e^{i k x}$ ). If $b \equiv 0$, then we say that the potential $V(x)$ is "reflectionless." Remarkably, turning a $\operatorname{sech}^{2}(x)$ wave profile upside down and rescaling can sometimes yield a reflectionless potential; e.g., $V(x)=-2 \operatorname{sech}^{2}(x)$ is reflectionless.

We work out an example of scattering data collection for the potential

$$
V(x) \equiv-2 \delta(x)
$$

where $\delta(x)$ is the Dirac delta. This example is simple except for the complication that solutions to $\psi_{x x}+(\lambda-V(x)) \psi=0$ need not have a continuous derivative at $x=0$. Nipping this in the bud, we integrate $\psi_{x x}+(2 \delta(x)+\lambda) \psi=0$ from $x=-\varepsilon$ to $\varepsilon$ and then take the limit as $\varepsilon$ approaches 0 . In suggestive notation, this yields $\psi_{x}\left(0^{+}\right)-\psi_{x}\left(0^{-}\right)+2 \psi(0)=0$. In other words, the derivative $\psi_{x}(x)$ of the function $\psi$ jumps discontinuously by $-2 \psi(0)$ as $x$ passes from negative to positive; $\psi$ itself is continuous. Now let's look for the bound states-nontrivial square-integrable solutions $\psi$. For this case we must look at $\lambda<0$, writing $\kappa=\sqrt{|\lambda|}$. For $x \neq 0$, our differential equation is $\psi_{x x}-\kappa^{2} \psi=0$, whose solutions are linear
combinations of $e^{ \pm \kappa x}$. If we must look among such functions, the only way $\psi$ can be squareintegrable is if $\psi(x)=\alpha e^{-\kappa x}$ for $x>0$ and $\psi(x)=\beta e^{\kappa x}$ for $x<0$, for some $\alpha$ and $\beta$. For $\psi$ to be continuous at 0 , we must have $\beta=\alpha$. Then, normalizing so that $\int_{-\infty}^{\infty}|\psi(x)|^{2} d x=1$, we find that $|\alpha|=\sqrt{\kappa}$; for simplicity we take $\alpha=\sqrt{\kappa}$. For $\psi$ to be a solution, however, we also require that $\psi_{x}\left(0^{+}\right)-\psi_{x}\left(0^{-}\right)=-2 \psi(0)$. Calculation reveals that $\psi_{x}\left(0^{+}\right)=\sqrt{\kappa}(-\kappa)$, $\psi_{x}\left(0^{-}\right)=\sqrt{\kappa}(\kappa)$, and $\psi(0)=\sqrt{\kappa}$, so division by $-2 \psi(0)$ gives the requirement that $\kappa=1$. Thus, $\lambda_{1}=-\kappa_{1}^{2}=-1$ is the only element of the discrete spectrum, and the one bound state can be written as $\psi_{1}(x)=e^{-|x|}$. What about the continuous spectrum? If $\lambda>0$, we let $k=\sqrt{\lambda}$ and look for a solution $\hat{\psi}$ satisfying $\hat{\psi} \sim e^{-i k x}+b e^{i k x}$ as $x \rightarrow+\infty$ and $\hat{\psi} \sim a e^{-i k x}$ as $x \rightarrow-\infty$. Clearly this will be accomplished by replacing $\sim$ with $=$ in the former for $x>0$ and $\sim$ with $=$ in the latter for $x<0$, provided $\hat{\psi}$ is required to be continuous and satisfy the jump condition. Continuity at $x=0$ implies $a=1+b$. We calculate $\psi_{x}\left(0^{+}\right)=-i k(1-b)$ and $\psi_{x}\left(0^{-}\right)=-i k a=-i k(1+b)$; subtracting gives $2 i k b$, so the jump condition becomes $2 i k b=-2 \psi(0)=-2(1+b)$. Thus, $(2 i k+2) b=-2$, or equivalently,

$$
b(k)=\frac{-1}{1+i k}
$$

We may also write $a(k)=1+b(k)=\frac{i k}{1+i k}$; the probability of transmission is $|a(k)|^{2}=\frac{k^{2}}{1+k^{2}}$ for incident waves of wavenumber $k$. All of the scattering data is in hand, and it consists of: $a(k), b(k)$, and the discrete spectrum ( $\lambda_{1}=-1$ ). Actually, it is supposed to include one more thing, the number $c_{1}$ such that the eigenfunction $\psi_{1}$ associated with $\lambda_{1}$ (so here $\psi_{1}(x)=e^{-|x|}$ ) satisfies $\psi_{1}(x) \sim c_{1} e^{-\kappa_{1} x}$ as $x \rightarrow+\infty$ (so here $c_{1}=1$ ).

Given a potential, one can hope to find the discrete spectrum $\Lambda=\left\{\lambda_{1}, \ldots, \lambda_{N}\right\}$ (and associated $c_{1}, \ldots, c_{N}$ ) and the scattering data $a(k), b(k)$ from the continuous spectrumwhat about doing the reverse? Given all of that information, can we "inverse scatter" and obtain the potential? Yes, actually we can. It turns out that inverse scattering amounts to solving something called the Marchenko integral equation-a product of the Russian scholars Gelfand, Levitan, and Marchenko circa 1951—and sometimes this can be done explicitly. Neither of these processes-scattering and inverse scattering-appear to have anything to do with the KdV equation. However, what Gardner, Greene, Kruskal, and Miura discovered in their two-page 1967 paper [10] is the following. If

- for each $t \geq 0$ we have a potential $V_{t}(x)$ with, say, $V_{t}(x) \rightarrow 0$ as $|x| \rightarrow \infty$, the associated discrete spectrum and scattering data being denoted by by $\Lambda_{t}, c_{1}(t), \ldots, c_{N_{t}}(t)$, $a_{t}(k), b_{t}(k)$, respectively, and if
- the function $u(x, t) \equiv V_{t}(x)$ solves the KdV equation in the standard form of [6], then:
- we have a time-independent discrete spectrum $\Lambda_{t}=\Lambda_{0}$ with a fixed number of elements $N$ and
- the scattering data evolve in time according to simple rules, namely $c_{n}(t)=c_{n}(0) e^{4 \kappa_{n}^{3} t}$ for $1 \leq n \leq N, b_{t}(k)=b_{0}(k) e^{8 i k^{3} t}$, and $a_{t}(k)=a_{0}(k)$.

This leads to a method for solving initial value problems for the KdV equation. Here's how it works:

- Given the initial data $u(x, 0)$, define $V_{0}(x) \equiv u(x, 0)$ and obtain the scattering data for $\psi_{x x}+\left(\lambda-V_{0}(x)\right) \psi=0$. Although hard work in general, the basic ideas and many commonly-occuring examples have been understood since around 1850.
- Having obtained the scattering data $\lambda_{1}, \ldots, \lambda_{N}, c_{1}(0), \ldots, c_{N}(0), a_{0}(k), b_{0}(k)$, we evolve it in time according to the simple rules above. This gives scattering data at time $t$ for a unique potential $V_{t}(x)$, where $u(x, t) \equiv V_{t}(x)$ turns out to be the sought-after solution to the Korteweg-de Vries equation.
- We calculate our answer $u(x, t)=V_{t}(x)$ by solving the Marchenko integral equation, which makes use of the scattering data at time $t$ collected from the previous step.

We recapitulate many of these remarks as we work through one of the simplest examples of the inverse scattering method in the first mathematical subsection below.

## 5. The Inverse Scattering Method: Mathematical Questions

5.1. Can I see an explicit example of the inverse scattering method? One begins with initial data, say a wave profile at a given initial time (hopefully physically reasonable). A good example is Rayleigh's function

$$
\eta(x, 0) \equiv \eta(x)=\eta_{0} \operatorname{sech}^{2}\left(\sqrt{\frac{3 \eta_{0}}{4 h^{3}}} x\right)
$$

with its roughly gaussian profile. Recall that we reduced the KdV equation to the standard form

$$
\frac{\partial u}{\partial t}-6 u \frac{\partial u}{\partial x}+\frac{\partial^{3} u}{\partial x^{3}}=0
$$

by the making the substitutions $u=-\eta / h, \xi=X / \frac{2 h}{\sqrt{6}}$, and $\tau=t / \frac{8 h}{c_{0} \sqrt{6}}$, afterwards renaming $\xi$ and $\tau$ as $x$ and $t$, respectively. In terms of $u(x, 0)$, where $x$ relabels $\xi$, we get

$$
u(x, 0)=-2 \operatorname{sech}^{2} x
$$

which we write as $V_{0}(x) \equiv u(x, 0)=-2 \operatorname{sech}^{2} x$. The simple form of $u$ at $t=0$ is the reason for our choice of standard form for the KdV equation.

We recall that the idea of the subsequent technique is from quantum mechanics. We use $V_{0}(x)=u(x, 0)$ as a potential in a normalized Schrödinger equation problem to obtain scattering data. Then, we evolve this data through time $t$ and "inverse scatter" it to obtain $u(x, t)$.

We consider the differential equation $-\psi_{x x}+V_{0}(x) \psi=\lambda \psi$ for an unknown function $\psi$, where subscripts denote differentiation. We rewrite it as $\psi_{x x}+\left(2 \operatorname{sech}^{2} x+\lambda\right) \psi=0$. Similarly, given $V_{t}(x) \equiv u(x, t)$ as a function of $x$ at time $t$, we can associate the equation $-\psi_{x x}+V_{t}(x) \psi=\lambda \psi$. In each case we can ask about the eigenvalues $\lambda$ of bound states, and reflection coefficients $b$, as in quantum mechanics.

For the current example's equation, $\psi_{x x}+\left(2 \operatorname{sech}^{2} x+\lambda\right) \psi=0$, the transformation $T=$ $\tanh x$ given in [6] p. 46 transforms the derivative as $\frac{d}{d x}=\operatorname{sech}^{2} x \frac{d}{d T}=\left(1-T^{2}\right) \frac{d}{d T}$. This yields an associated Legendre equation,

$$
\left(1-T^{2}\right) \frac{d}{d T}\left(\left(1-T^{2}\right) \frac{d \psi}{d T}\right)+\left(\lambda+2\left(1-T^{2}\right)\right) \psi=0
$$

Since our emphasis here is on the process rather than the individual steps, we rely on our source [6] for the scattering data. From knowledge of this Legendre equation one can deduce that here there is only one bound state, typically written as the "Goldstone mode" $\psi_{1}(x)=$ $\frac{1}{\sqrt{2}} \operatorname{sech} x$ with $\lambda_{1}=-1$ (so $\kappa_{1}=1$ ). This state's asymptotic behavior is $\psi_{1}(x) \sim \sqrt{2} e^{-x}$ as $x \rightarrow+\infty$. We denote the exponential's coefficient by $c_{1}(0)=\sqrt{2}$. It is significant that at $t=0$ our potential is reflectionless, that is, $b_{0}(k)=0$ identically; the simple form assumed by the auxiliary function $F$ in the Marchenko equation below is a consequence.

As mentioned above, it turns out (proofs are also in [5] pp. 169-171) that the bound state eigenvalues $\lambda$ for the potentials $V_{t}(x)$ are time-independent, and that the coefficients $c(t)$ and $b$ evolve in a simple way when $u(x, t)$ solves the KdV equation. In fact, here we have $c_{1}(t)=c_{1}(0) e^{4 t}=\sqrt{2} e^{4 t}$ and $b_{t}(k)=0$. Having evolved our scattering data, we then solve an integral equation to inverse scatter. We merely write down the equations and prescriptions, treating time $t$ as a parameter, not always written explicitly. The reader should not be alarmed if the equations and prescriptions appear to come out of nowhere. The interested reader will find a derivation of the Marchenko equation in the next subsection. Our solution $u(x, t)=u(x)$ is given by $u(x)=-2 \frac{d K(x, x)}{d x}$, where, after defining the auxiliary function $F(x)=c_{1}^{2}(t) e^{-x}$ to account for our nontrivial scattering data, $K(x, z)$ is the solution to the Marchenko equation

$$
K(x, z)+F(x+z)+\int_{x}^{\infty} K(x, y) F(y+z) d y=0
$$

Here $F(x+z)=c_{1}^{2}(t) e^{-(x+z)}=2 e^{8 t-x-z}$, so $F(x+z)$ is separable. This means it can be written as a product of the form $X(x) Z(z)$. We take $X(x)=2 e^{8 t-x}$ and $Z(z)=e^{-z}$ for definiteness. Substituting this separated form for $F(x+z)=X(x) Z(z)$ and $F(y+z)=X(y) Z(z)$, we see $K(x, z)$ can be written as
$K(x, z)=-X(x) Z(z)-Z(z) \int_{x}^{\infty} K(x, y) X(y) d y=\left(-X(x)-\int_{x}^{\infty} K(x, y) X(y) d y\right) Z(z)$.

We have shown that $K(x, z)$ is also separable, so $K(x, z)=L(x) Z(z)$ for some $L(x)$. Substituting $L(x) Z(z)$ into our equation for $K(x, z)$ gives

$$
L(x) Z(z)+X(x) Z(z)+Z(z) L(x) \int_{x}^{\infty} Z(y) X(y) d y=0
$$

This is a linear algebraic equation in $L(x)$. Solving it for $L(x)$ yields

$$
L(x)=\frac{-X(x)}{1+\int_{x}^{\infty} Z(y) X(y) d y} .
$$

Recalling that $Z(z)=e^{-z}$, the solution is $K(x, z)=L(x) e^{-z}$ where, in detail,

$$
L(x)=\frac{-2 e^{8 t-x}}{1+\int_{x}^{\infty} 2 e^{8 t-y} e^{-y} d y}=\frac{-2 e^{8 t-x}}{1+e^{8 t-2 x}}
$$

Thus, we have

$$
K(x, z)=\frac{-2 e^{8 t-x}}{1+e^{8 t-2 x}} e^{-z}
$$

so that $K(x, x)=-2 \frac{e^{8 t-2 x}}{1+e^{8 t-2 x}}$. Finally, we obtain $u(x, t)=-2 \frac{d K(x, x)}{d x}=-2 \operatorname{sech}^{2}(x-4 t)$. For more general initial data $u(x, 0)$ we may obtain $N$ bound states, which lead to an $N$-soliton solution requiring somewhat more effort to compute (when $b=0$ one must solve $N$ algebraic equations in $N$ unknowns). When $b \neq 0$ there is also non-solitonic "radiation," which often behaves approximately like a solution to the linear equation obtained from the KdV equation by removing the nonlinear term. In tasking ourselves with finding $K(x, z)$ from $F(x)$ when $b \neq 0$, much depends on how easy it is to work with the inverse Fourier transform of $b$.

Having arrived at our solution, we can return to our original (non-dimensionless) variables to see the physical solution. After some algebra, we find that, in the lab frame (not the frame moving at speed $c_{0}-$ recall $X=x-c_{0} t$, so we go from new $x$, to $X$, to old $x$ ),

$$
\eta(x, t)=\eta_{0} \operatorname{sech}^{2}\left[\sqrt{\frac{3 \eta_{0}}{4 h^{3}}}\left(x-c_{0}\left[1+\frac{\eta_{0}}{2 h}\right] t\right)\right] .
$$

Thus, the shape and size of the profile, under the idealized assumptions of the KdV equation, remains exactly the same, a solitary wave traveling to the right with speed (coefficient of $t$ inside the parentheses)

$$
c=c_{0}\left(1+\frac{\eta_{0}}{2 h}\right)
$$

Seeing this formula emerge once more is beautifully consistent with the results of Russell, Boussinesq, Rayleigh, Zabusky-Kruskal, and Korteweg-de Vries.
5.2. Can I see where the Marchenko integral equation comes from? How can we sketch a quick derivation of the Marchenko integral equation? We need to come up with a scheme for scattering that will allow us-somehow-to back-solve for our potential when the scattering data $b(k)$, etc., is known. We note that our data $c_{n}$ for the asymptotics of
the eigenfunctions $\psi_{n}$ associated with discrete spectrum elements $\lambda_{n}=-\kappa_{n}^{2}$ is of the form " $\psi_{n} \sim c_{n} e^{-\kappa_{n} x}$ as $x \rightarrow+\infty$," so we will likely need to work with conditions at $+\infty$ as opposed to $-\infty$.

Let $\lambda=k^{2}>0$ be given; we want the solution $\hat{\psi}$ to $\psi_{x x}+\left(k^{2}-V\right) \psi=0$ such that $\hat{\psi} \dot{\sim} e^{-i k x}+b(k) e^{i k x}$ as $x \rightarrow+\infty$. A plan for doing this is to first find a solution $\psi_{+}$to $\psi_{x x}+\left(k^{2}-V\right) \psi=0$ which satisfies $\psi_{+} \sim e^{i k x}$ as $x \rightarrow+\infty$, and then let $\hat{\psi}=\psi_{+}^{*}+b(k) \psi_{+}$, where $\psi_{+}^{*}$ is the complex conjugate of $\psi_{+}$. This will give us the right asymptotics at $+\infty$ and we trust that the rest will work out, provided we have $\psi_{+}$. (I will mention a wee detail now to avoid future confusion. It is the following: at the end we're going to deal with Fourier transforms, and we want the $\pm$ signs of the complex exponentials $e^{ \pm i k x}$ in the transforms to come out right. We can arrange for this to happen for $\psi_{+}^{*}$, but it will mean that when we deal with Fourier transforms pertaining to $\psi_{+}$, all of the - signs should be flipped to + in the exponents of the $e^{-i k x}$ terms usually integrated against.) Now $\psi_{+}$is supposed to be a solution to $\psi_{x x}+\left(k^{2}-V\right) \psi=0$; what if we ignore $V$ for a moment? Then, $\psi_{x x}+k^{2} \psi=0$ is what we get when we Fourier transform a wave equation in $x$ and another variable-usually $t$, but here called $z$-such as $\phi_{x x}-\phi_{z z}=0$. We want $\psi_{+} \sim e^{i k x}$ as $x \rightarrow+\infty$, so if we think of $\psi_{+}$as a Fourier transform, a good first try would be the Fourier transform (with the sign in the exponent of $e^{-i k z}$ flipped to get $e^{i k z}$-see previous parenthetical comment)

$$
\psi_{+} \approx \int_{-\infty}^{\infty} \delta(x-z) e^{i k z} d z=e^{i k x}
$$

When we're not ignoring $V$, however, there is no way this will generally be exact.
We add some unknown function $K(x, z)$ to the delta function we're taking the Fourier transform of to get the true $\psi_{+}$:

$$
\psi_{+}=\int_{-\infty}^{\infty}(\delta(x-z)+K(x, z)) e^{i k z} d z=e^{i k x}+\int_{-\infty}^{\infty} K(x, z) e^{i k z} d z
$$

An additional restriction must be placed on $K(x, z)$ to avoid the possibility that the integral on the right will mess up the statement " $\psi_{+} \sim e^{i k x}$ as $x \rightarrow+\infty$." One starting point is to make $K(x, z)=0$ whenever $x \rightarrow+\infty$ for fixed $z$, and an appropriate way to implement this is to make $K(x, z)=0$ whenever $x>z$. Then, since $K(x, z)$ can be nonzero only for $z \geq x$, the lower limit of integration on the integral becomes $x$ :

$$
\psi_{+}=e^{i k x}+\int_{x}^{\infty} K(x, z) e^{i k z} d z
$$

Now we force this to really be a solution to $\psi_{x x}+\left(k^{2}-V\right) \psi=0$ by calculating $\left(\psi_{+}\right)_{x x}+$ ( $\left.k^{2}-V\right) \psi_{+}$from the above equation for $\psi_{+}$in terms of $K(x, z)$. We omit these calculations, which may be found on pages 49 and 50 of [6], but state the results. After integrating by
parts twice in the integral for $\psi_{+}$, one can write the equation as

$$
0=\left(\psi_{+}\right)_{x x}+\left(k^{2}-V\right) \psi_{+}=-e^{i k x}\left(V+2 \frac{d \hat{K}}{d x}\right)+\int_{x}^{\infty}\left(K_{x x}-K_{z z}-V(x) K\right) e^{i k z} d z
$$

Above, $\hat{K}(x) \equiv K(x, x)$. In turn, we can force this equation (and everything else) to work out by requiring $K_{x x}-K_{z z}-V(x) K=0$ for $z>x$, demanding that $V(x)+2 \frac{d \hat{K}}{d x}=0$, and also enforcing decay of $K(x, z), K_{z}(x, z)$ as $z \rightarrow+\infty$ to more effectively guarantee that $K$ does not interfere with $\psi_{+} \sim e^{i k x}$ as $x \rightarrow+\infty$. This problem turns out to always have a unique solution $K$ by general theory, even though general theory doesn't explain how to calculate $K$ yet. The main upshot is that once we know $K$, we can get $V$ from $V(x)=-2 \frac{d \hat{K}}{d x}$.

All that remains is to shift things around to unearth the integral equation that $K(x, z)$ satisfies-the Marchenko integral equation. Substituting $e^{i k x}+\int_{x}^{\infty} K(x, z) e^{i k z} d z$ for $\psi_{+}$in the equation $\hat{\psi}=\psi_{+}^{*}+b(k) \psi_{+}$gives a result that, after rearrangement, can be written as

$$
\int_{-\infty}^{\infty} K(x, z) e^{-i k z} d z=\hat{\psi}-e^{-i k x}-b(k) e^{i k x}-b(k) \int_{-\infty}^{\infty} K(x, z) e^{i k z} d z
$$

We can extend the lower limits of integration from $x$ down to $-\infty$ because $K(x, z)$ is zero there anyway-there is no additional contribution to the integrals. The left side is the Fourier transform of $K$, so we can inverse Fourier transform both sides to obtain an equation for $K$. This is the Marchenko integral equation, although it requires a little housekeeping to make presentable.

$$
K(x, z)=\frac{1}{2 \pi} \int_{-\infty}^{\infty}\left[\hat{\psi}(x)-e^{-i k x}-b(k) e^{i k x}-b(k) \int_{-\infty}^{\infty} K(x, y) e^{i k y} d y\right] e^{i k z} d k
$$

First, we define $F_{0}(x) \equiv \frac{1}{2 \pi} \int_{-\infty}^{\infty} b(k) e^{i k x} d k$ to be the inverse Fourier transform of the reflection coefficient $b(k)$. Moving $b(k)$ inside the $y$-integral (for notational convenience we're changed the dummy variable to $y$ to avoid confusion) and swapping the order of integration allows us to write the mostly cleaned-up equation

$$
K(x, z)=\frac{1}{2 \pi} \int_{-\infty}^{\infty}\left(\hat{\psi}-e^{-i k x}\right) e^{i k z} d k-F_{0}(x+z)-\int_{-\infty}^{\infty} K(x, y) F_{0}(y+z) d y
$$

Right around this point-if not earlier-would be a good time to analytically extend $a(k)$ and $b(k)$ into the complex plane. They should really have been continued to the real line before now, in order for their inverse Fourier transforms to make sense.

The last piece of the puzzle is the inverse Fourier transform left standing, the integral over the real line with respect to $k$ of the integrand $\left(\hat{\psi}-e^{-i k x}\right) e^{i k z}$. Clearly the result should be a function of $x$ and $z$, but the precise answer is unclear. However, on pages 51 through 55 of [6], this integral is calculated by the residue theorem. In fact, $\hat{\psi}$ has $N$ poles-they are the poles of $a(k)$-in the upper half plane of the form $i \kappa_{n}$; this is where we are using the rest of
our scattering data. The integral is evaluated as

$$
\int_{-\infty}^{\infty}\left(\hat{\psi}-e^{-i k x}\right) e^{i k z} d k=2 \pi i \sum_{n=1}^{N} R_{n}
$$

where $R_{n}$, the residue at $i \kappa_{n}$, is given by

$$
R_{n}=i\left(F_{n}(x+z)+\int_{x}^{\infty} K(x, y) F_{n}(y+z) d y\right)
$$

where $F_{n}(x) \equiv c_{n}^{2} e^{-\kappa_{n} x}$. This residue calculation is omitted. We conclude from this that

$$
\frac{1}{2 \pi} \int_{-\infty}^{\infty}\left(\hat{\psi}-e^{-i k x}\right) e^{i k z} d k=-\sum_{n=1}^{N}\left(F_{n}(x+z)+\int_{x}^{\infty} K(x, y) F_{n}(y+z) d y\right)
$$

so we may write the Marchenko integral equation as

$$
K(x, z)=-F(x+z)-\int_{x}^{\infty} K(x, y) F(y+z) d y
$$

where

$$
F(x) \equiv \sum_{n=0}^{N} F_{n}(x)
$$

we recall that $F_{0}(x)=\frac{1}{2 \pi} \int_{-\infty}^{\infty} b(k) e^{i k x} d k$. The derivation is complete.
If the Marchenko equation is any good, then when we feed it the scattering data we obtained earlier from the potential $V(x)=-2 \delta(x)$, in particular

$$
b(k)=\frac{-1}{1+i k}
$$

and $\kappa_{1}=1=c_{1}$, it will return the potential $-2 \delta(x)$ to us. We give an outline of how this comes about. To use the Marchenko equation we need $F(x)$, which is $F_{0}(x)+F_{1}(x)$ here. Knowing $\kappa_{1}=1=c_{1}$ gives $F_{1}(x) \equiv c_{1}^{2} e^{-\kappa_{1} x}=e^{-x}$. On the other hand,

$$
F_{0}(x) \equiv \frac{1}{2 \pi} \int_{-\infty}^{\infty} b(k) e^{i k x} d k=\frac{1}{2 \pi} \int_{-\infty}^{\infty} \frac{-1}{1+i k} e^{i k x} d k
$$

If $x>0$, this integral is evaluated with the residue theorem. We choose a semicircular contour which goes from $-R$ to $R$ on the real axis and then travels counterclockwise in a circular arc from $R$ back to $-R$. Evaluating the contour integral for $R>1$ and letting $R \rightarrow \infty$ yields (from writing the integrand in the form $\frac{1}{k-i}-\frac{e^{i k x}}{i}$ and setting $k=i$ in the second factor) the residue $\frac{-e^{-x}}{i}$, whence for $x>0$ we have

$$
F_{0}(x)=\frac{1}{2 \pi}(2 \pi i) \frac{-e^{-x}}{i}=-e^{-x}
$$

On the other hand, for $x<0$, a similar contour in the lower half-plane gives 0 for the integral by Cauchy's integral theorem. This means we can write $F_{0}(x)$ as

$$
F_{0}(x)=-H(x) e^{-x}
$$

where the Heaviside step function is $H(x) \equiv 1$ for $x>0$ and $H(x) \equiv 0$ for $x<0$. Thus,

$$
F(x)=F_{0}(x)+F_{1}(x)=-H(x) e^{-x}+e^{-x}=H(-x) e^{-x} .
$$

Let's look at what this means for the Marchenko equation

$$
K(x, z)+F(x+z)+\int_{x}^{\infty} K(x, y) F(y+z) d y=0
$$

We see that for $x+z>0$ the term $F(x+z)$ is 0 because of the $H(-x)$ factor in $F(x)$. Moreover, since $y \geq x$ in the integrand, $y+z>0$ for $y$ between the limits of integration, and the Marchenko equation becomes

$$
K(x, z)=0
$$

for $x+z>0$. For $x+z<0$, on the other hand, $F(x+z)=e^{-x-z}$, but the only $y \geq x$ for which $y+z<0$ (where $F(y+z)$ might be nonzero) are those satisfying $y<-z$. Thus, we can take $-z$ to be the upper limit of integration for our integral. Solving for the integral gives

$$
\int_{x}^{-z} K(x, y) e^{-y-z} d y=-K(x, z)-e^{-x-z}
$$

Multiplying both sides by $e^{z}$ gives

$$
\int_{x}^{-z} K(x, y) e^{-y} d y=-K(x, z) e^{z}-e^{-x}
$$

Here I see that the equation looks like the calculus equation $\int_{x}^{-z} a e^{-y} d y=-a e^{z}+a e^{-x}$ for a definite integral, so I guess that $K(x, y)$ is the constant $a$ that would make these two equations the same, namely $a=-1$. This shows that $K(x, z) \equiv-1$ for $x+z<0$ and $K(x, z) \equiv 0$ for $x+z>0$ is consistent with the Marchenko equation, that is, it is a solution. However, it turns out that solutions to the Marchenko equation are unique, so this answer must be the correct one. We can write our solution as

$$
K(x, z)=-H(-x-z) .
$$

Then $V(x)$ is given by $V(x)=-2 \frac{d K(x, x)}{d x}$. Clearly $\hat{K}(x) \equiv K(x, x)$ is just $-H(-2 x)=$ $-H(-x)$ since the Heaviside function is unchanged by a horizontal rescaling. We conclude, by (i) the chain rule, (ii) the fact that $H^{\prime}(x)=\delta(x)$, and (iii) $\delta(-x)=\delta(x)$, that

$$
V(x)=-2(-H(-x))^{\prime}=2 H^{\prime}(-x)(-1)=-2 \delta(-x)=-2 \delta(x),
$$

as expected.

The reader interested in working out a nontrivial example of solving a KdV initial-value problem with the Marchenko equaiton is encouraged to solve

$$
\frac{\partial u}{\partial t}-6 u \frac{\partial u}{\partial x}+\frac{\partial^{3} u}{\partial x^{3}}=0
$$

with the initial condition

$$
u(x, 0) \equiv-6 \operatorname{sech}^{2}(x)
$$

This corresponds physically to an initial profile of a solitary wave with three times the height it is "supposed" to have. The reader can carry the whole solution process through by taking the normalized Schrödinger equation and substituting $T=\tanh (x)$ to reduce it to Legendre's equation, obtaining two bound states and $b(k) \equiv 0$ (the reader may take my word for the latter). After evolving the scattering data through time using the rules stated earlier, the Marchenko equation yields $u(x, t) \equiv V_{t}(x)$. Since $b(k) \equiv 0$, the input $F(x)$ for the Marchenko equation is as sum of multiples of two exponentials, so $F(x+z)=X_{1}(x) Z_{1}(z)+X_{2}(x) Z_{2}(z)$ for the reader's choice of $X_{1}(x), Z_{1}(z), X_{2}(x)$, and $Z_{2}(z)$. From this presentation of $F(x+z)$ it is easy to see that $K(x, z)$ may be written as $K(x, z)=L_{1}(x) Z_{1}(z)+L_{2}(x) Z_{2}(z)$. Plugging this and the emphasized presentation of $F(x+z)$ into the Marchenko equation gives two linear algebraic equations in the two unknowns $L_{1}(x)$ and $L_{2}(x)$ due to the linear independence of the two exponentials in the definition of $F$, as functions of $z$. Solving these allows $K(x, z)$, and hence $u(x, t)$, to be written explicitly. In fact, I obtain

$$
u(x, t)=-12 \frac{3+4 \cosh (2 x-8 t)+\cosh (4 x-64 t)}{(3 \cosh (x-28 t)+\cosh (3 x-36 t))^{2}}
$$

The interested reader with access to a computer can then plot this exact analytic expression and animate it to watch a movie. What will be observed is a too-high ${ }^{14}$ solitary wave breaking up into a 2 -soliton. After emerging, the two bumps proceed to travel at essentially constant "eigenspeeds" indefinitely. One can also start the animation from a negative value of time to see a "collision." See Figures 8 through 15. The reason all of this can be done so nicely is that we can evolve the scattering data using simple rules and because, as I have claimed, $b(k)=0$ identically. In other words, the potential is reflectionless, an assertion examined in the next subsection.

[^9]

Figure 8. A plot of $-u$ against $x$ for $t=-0.25$.


Figure 9. A plot of $-u$ against $x$ for $t=-0.1$.


Figure 10. A plot of $-u$ against $x$ for $t=-0.05$.


Figure 11. A plot of $-u$ against $x$ for $t=0$.


Figure 12. A plot of $-u$ against $x$ for $t=0.05$.


Figure 13. A plot of $-u$ against $x$ for $t=0.1$.


Figure 14. A plot of $-u$ against $x$ for $t=0.25$.


Figure 15. A plot of $-u$ against $x$ for $t=0.5$.
5.3. When is a vertically scaled solitary wave a reflectionless potential? The skeptical reader may be forgiven for not wanting to accept my claims about certain potentials of the form $V(x) \equiv-a_{0} \operatorname{sech}^{2}(x)$ being reflectionless without explanation. To find out which choices of vertical scaling ${ }^{15} a_{0}$ imply $b(k) \equiv 0$ we must return to the associated Legendre equation and derive the asymptotics needed to calculate $b(k)$ for the hypergeometric solutions corresponding to the continuous spectrum. It will save time to instead cite a source, such as page 47 of [6], which provides such asymptotics. There, Drazin and Johnson say that it is a "fairly simple exercise to confirm" (yet I think this will depend on the reader) that when

$$
\hat{\psi}(x) \sim a(k) e^{-i k x}
$$

as $x \rightarrow-\infty$, "we obtain

$$
\hat{\psi}(x) \sim \frac{a(k) \Gamma(\tilde{c}) \Gamma(\tilde{a}+\tilde{b}-\tilde{c})}{\Gamma(\tilde{a}) \Gamma(\tilde{b})} e^{-i k x}+\frac{a(k) \Gamma(\tilde{c}) \Gamma(\tilde{c}-\tilde{a}-\tilde{b})}{\Gamma(\tilde{c}-\tilde{a}) \Gamma(\tilde{c}-\tilde{b})} e^{i k x}
$$

as $x \rightarrow+\infty$." Here, " $\tilde{a}=\frac{1}{2}-i k+\left(a_{0}+\frac{1}{4}\right)^{1 / 2}, \tilde{b}=\frac{1}{2}-i k-\left(a_{0}+\frac{1}{4}\right)^{1 / 2}$, and $\tilde{c}=1-i k$." Whether a simple exercise or not, the above result will certainly look intimidating to any reader unfamiliar with Euler's gamma function $\Gamma(x)$, so we discuss some properties of that function in a moment. Comparing this with the condition " $\hat{\psi} \sim e^{-i k x}+b(k) e^{i k x}$ as $x \rightarrow+\infty$ " shows that the coefficient of $e^{i k x}$ is $b(k)$, and $a(k)$ may be found by setting the coefficient of $e^{-i k x}$ to 1 . It turns out that to see when $b(k)$ is zero, however, it suffices to examine the effect of the bottom portion, so we write

$$
b(k) \propto \frac{1}{\Gamma(\tilde{c}-\tilde{a}) \Gamma(\tilde{c}-\tilde{b})}=\frac{1}{\Gamma\left(\frac{1}{2}-\left(a_{0}+\frac{1}{4}\right)^{1 / 2}\right) \Gamma\left(\frac{1}{2}+\left(a_{0}+\frac{1}{4}\right)^{1 / 2}\right)} .
$$

We will see how this answers our question of when $b(k) \equiv 0$ once we work out a couple of needed example problems involving the function $\Gamma(x)$ of a real variable $x$. I have tried to write these mathematical examples so that at least the ideas I think are important are clearly visible. The strategy for the second one may originate with Emil Artin. For full detail, the reader might need to verify or look up a few facts.

Example 1: Observe that, if $s(x) \equiv \sin (\pi x)$, then the double angle formula may be written in the form $s(x)=2 s\left(\frac{x}{2}\right) \cos \left(\pi \frac{x}{2}\right)=2 s\left(\frac{x}{2}\right) s\left(\frac{x+1}{2}\right)$. Is a similar relation true for the function $\Gamma(x)$ defined, for $x>0$, by $\Gamma(x) \equiv \int_{0}^{\infty} e^{-t} t^{x-1} d t$ ? To check, we write it more simply, that is, as a limit. Replacing $\infty$ with $n$ and $e^{-t}$ with $\left(1-\frac{t}{n}\right)^{n}$ simultaneously and letting $n$ become large, we find that $\Gamma(x)=\lim _{n \rightarrow \infty} \int_{0}^{n}\left(1-\frac{t}{n}\right)^{n} t^{x-1} d t=$

[^10]$\lim _{n \rightarrow \infty} \frac{n^{x} n!}{x(x+1) \ldots(x+n)}$, where in passing to the final equality we have integrated by parts $n$ times. This explicit limit is Gauss's definition of the gamma function. It has the advantage of converging whenever $x$ isn't a nonpositive integer, so it defines $\Gamma(x)$ for such $x$ even if $x<0$. To compare $\Gamma(x)$ with $s(x)$ we consider, instead of $2 \Gamma\left(\frac{x}{2}\right) \Gamma\left(\frac{x+1}{2}\right)$, the quantity $2^{x} \Gamma\left(\frac{x}{2}\right) \Gamma\left(\frac{x+1}{2}\right)$ for fixed $x>0$. It equals
$$
\lim _{n \rightarrow \infty} 2^{x} \frac{n^{\frac{x}{2}} n^{\frac{x+1}{2}}(n!)^{2}}{\left[\left(\frac{x}{2}\right)\left(\frac{x}{2}+1\right) \ldots\left(\frac{x}{2}+n\right)\right]\left[\left(\frac{x+1}{2}\right)\left(\frac{x+1}{2}+1\right) \ldots\left(\frac{x+1}{2}+n\right)\right]}
$$

This does not look like $\Gamma(x)$. It does simplify a little, to $\lim _{n \rightarrow \infty} \frac{(2 n)^{x}(n!)^{2} \sqrt{n}}{x(x+1) \ldots(x+2 n)} \cdot \frac{2^{2 n+2}}{x+2 n+1}$. This still doesn't look like $\Gamma(x)$, but if we take the sequence whose limit Gauss defines to be $\Gamma(x)$ and consider it only for even integers $m$ (a subsequence, $m=$ $2,4,6, \ldots, 2 n, \ldots)$ then we see that
$2^{x} \Gamma\left(\frac{x}{2}\right) \Gamma\left(\frac{x+1}{2}\right)=\lim _{n \rightarrow \infty} \frac{(2 n)^{x}(2 n)!}{x(x+1) \ldots(x+2 n)} \cdot \frac{(n!)^{2} 2^{2 n}}{(2 n)!\sqrt{n}} \cdot \frac{4 n}{x+2 n+1}=\Gamma(x) \cdot \sqrt{\pi} \cdot 2$,
where

$$
\lim _{n \rightarrow \infty} \frac{(n!)^{2} 2^{2 n}}{(2 n)!\sqrt{n}}=\sqrt{\pi}
$$

is a standard limit computed by showing that the ratio of the integrals of $\sin ^{2 n}(x)$ and $\sin ^{2 n+1}(x)$ between 0 and $\pi / 2$ goes to 1 as $n \rightarrow \infty$. Thereby, a limit for $\pi$ is obtained whose square root yields the above limit (this method of obtaining the limit for $\sqrt{\pi}$ due to Wallis is a substantial exercise in calculus textbooks like [17]). Anyway, we have obtained Legendre's famous "duplication" or "half-angle" formula for the gamma function, often written

$$
\sqrt{\pi} \Gamma(x)=2^{x-1} \Gamma\left(\frac{x}{2}\right) \Gamma\left(\frac{x+1}{2}\right)
$$

It is amusing that when $x=1$ in the above formula, and after verifying that $\Gamma(1)=1$, we can substitute $t=u^{2}$ in the integral with which we originally defined the gamma function. This yields the cute identity $\sqrt{\pi}=\Gamma\left(\frac{1}{2}\right)=\int_{-\infty}^{\infty} e^{-u^{2}} d u$.
Example 2: We we wish to find out if it would be possible for us to graph the function $f(x) \equiv \Gamma(x) \Gamma(1-x) \sin (\pi x)$ by hand. Unfortunately, it appears that $f(x)$ is not defined when $x$ is an integer, as $\Gamma(x)$ is not defined for nonpositive integers. However, the reader may verify from Gauss's limit for $\Gamma(x)$ that $\Gamma(1 \pm x)=( \pm x) \Gamma(x)$. Then, we may write $\Gamma(x)=\frac{\Gamma(1+x)}{x}$ and use the Taylor series for the sine function to write
$f(x)=\frac{\Gamma(1+x)}{x} \Gamma(1-x) \sin (\pi x)=\Gamma(1+x) \Gamma(1-x)\left(\pi-\frac{\pi^{3} x^{2}}{3!}+\frac{\pi^{5} x^{4}}{5!}-\ldots\right)$.
The right side is defined at $x=0$, and letting $f(0)=\pi$ ensures that $f$ is continuous there, as $\Gamma(x) \rightarrow 1$ when $x \rightarrow 1$. What about integers $n \neq 0$ ? We don't have to worry
about them because $f$ is a periodic function (this is what drew it to my attention in the first place). Indeed, both $\Gamma(x) \Gamma(1-x)$ and $\sin (\pi x)$ are functions $h(x)$ which satisfy $h(x+1)=-h(x)$ (antiperiodic of period 1). E.g., replacing $x$ with $x+1$ in $h(x) \equiv \Gamma(x) \Gamma(1-x)$ gives $h(x+1)=\Gamma(x+1) \Gamma(1-(x+1))$, which can be written as

$$
x \Gamma(x) \Gamma(-x)=-\Gamma(x)(-x) \Gamma(-x)=-\Gamma(x) \Gamma(1-x)=-h(x) .
$$

Thus, the product function $f(x)$ satisfies $f(x+1)=(-1)^{2} f(x)=f(x)$ (periodic of period 1). Due to periodicity, letting $f(n)=\pi$ for all integers makes $f$ a continuous (in fact smooth) periodic function of period 1 on the whole real line. Does this function have any zeros? By periodicity, it suffices to check the interval ( 0,1 ), and since $\sin (\pi x)$ is positive there and $\Gamma(x) \Gamma(1-x)$ is positive there (the integrals are positive), it follows that $f(x)>0$ for all $x$. To see how the values of $f$ relate to each other in $(0,1)$, we can multiply the double-angle formula $s(x)=2 s\left(\frac{x}{2}\right) s\left(\frac{1+x}{2}\right)$ for $s(x) \equiv \sin (\pi x)$ with the Legendre relations for $\Gamma$ at both $x$ and $1-x$. Equivalently, multiplying $f(x)$ by $\pi$ to write $\pi f(x)=\sqrt{\pi} \Gamma(x) \sqrt{\pi} \Gamma(1-x) s(x)$ for computational convenience and expanding multiplicatively gives

$$
\pi f(x)=2^{x-1} 2^{(1-x)-1} \Gamma\left(\frac{x}{2}\right) \Gamma\left(\frac{1+x}{2}\right) \Gamma\left(\frac{1-x}{2}\right) \Gamma\left(\frac{1+(1-x)}{2}\right) 2 s\left(\frac{x}{2}\right) s\left(\frac{1+x}{2}\right)=f\left(\frac{x}{2}\right) f\left(\frac{1+x}{2}\right) .
$$

What is this telling us? To change it into something more comprehensible, we apply logarithms to both sides to transmute the multiplication to addition. Letting $g(x) \equiv$ $\log _{e}(f(x))$, we have

$$
\log _{e}(\pi)+g(x)=g\left(\frac{x}{2}\right)+g\left(\frac{x+1}{2}\right) .
$$

We can use this growth relation to get an estimate for how the graph bends, which is determined by $g^{\prime \prime}(x)$. Taking derivatives twice and applying the chain rule, we get

$$
g^{\prime \prime}(x)=\frac{1}{4}\left(g^{\prime \prime}\left(\frac{x}{2}\right)+g^{\prime \prime}\left(\frac{x+1}{2}\right)\right) .
$$

To obtain an estimate for the bending as a function of $x,\left|g^{\prime \prime}(x)\right|$, we first let $M=$ $\max \left|g^{\prime \prime}(x)\right| \geq 0$ be the maximum magnitude of $g^{\prime \prime}(x)$ on the real line. (As $\left|g^{\prime \prime}(x)\right|$ may be shown to be continuous, it attains a maximum value on $[0,1]$; a one-sentence proof that continuous functions on closed intervals attain maxima may found in [8]. By periodicity, this maximum is also the maximum over the real line.) For some $x_{0}$, we have $\left|g^{\prime \prime}\left(x_{0}\right)\right|=M \geq\left|g^{\prime \prime}(x)\right|$ for all $x$. Estimating, we find

$$
\left|g^{\prime \prime}(x)\right| \leq \frac{1}{4}(M+M)=\frac{M}{2} .
$$

Thus, the bending of the graph never exceeds half of the maximum bending of the graph. At $x=x_{0}$ this says

$$
M=\left|g^{\prime \prime}\left(x_{0}\right)\right| \leq \frac{M}{2}
$$

If $M>0$, then $M>M / 2$, which is inconsistent with this estimate. As $M \geq 0$, we are forced to conclude that $M=0$, so $g^{\prime \prime}(x)=0$ for all $x$. Thus, the graph of $g$ experiences no bending, and must be a straight line. But $g$ is periodic! The only way that a straight line can be a graph of a periodic function is if it is horizontal. In other words, the function $g(x)=\log _{e}(f(x))$, and hence $f(x)$, is constant (something we can graph!). As $f(0)=\pi$, it follows that for all $x$,

$$
\Gamma(x) \Gamma(1-x) \sin (\pi x)=f(x)=\pi .
$$

Defining $u$ by $x=\frac{1}{2}-u$, this can be written alternatively as Euler's reflection formula,

$$
\Gamma\left(\frac{1}{2}-u\right) \Gamma\left(\frac{1}{2}+u\right)=\frac{\pi}{\sin \left(\frac{\pi}{2}-\pi u\right)}=\frac{\pi}{\cos (\pi u)} .
$$

Rearranging to apply to the problem of reflectionless $\operatorname{sech}^{2}(x)$ potentials, we get

$$
\frac{1}{\Gamma\left(\frac{1}{2}-u\right) \Gamma\left(\frac{1}{2}+u\right)}=\frac{\cos (\pi u)}{\pi}
$$

Returning to our reflection coefficients $b(k)$, we apply Euler's reflection formula to obtain

$$
b(k) \propto \frac{1}{\Gamma\left(\frac{1}{2}-\left(a_{0}+\frac{1}{4}\right)^{1 / 2}\right) \Gamma\left(\frac{1}{2}+\left(a_{0}+\frac{1}{4}\right)^{1 / 2}\right)}=\frac{\cos \left(\pi\left(a_{0}+\frac{1}{4}\right)^{1 / 2}\right)}{\pi}
$$

I have claimed that $V_{a_{0}=2}(x) \equiv-2 \operatorname{sech}^{2}(x)$ and $V_{a_{0}=6}(x) \equiv-6 \operatorname{sech}^{2}(x)$ are reflectionless potentials. This follows immediately from the proportionalilty relation above. Plugging in $a_{0}=2$ and $a_{0}=6$ make $\left(a_{0}+\frac{1}{4}\right)^{1 / 2}$ equal to $\frac{3}{2}$ and $\frac{5}{2}$ respectively, and since $\cos \left(\frac{3 \pi}{2}\right)=$ $0=\cos \left(\frac{5 \pi}{2}\right)$ the reflection coefficient is proportional to 0 in these cases, that is, $b(k)=0$ identically. See Figure 16. In general, we see that $V_{a_{0}}(x) \equiv-a_{0} \operatorname{sech}^{2}(x)$ gives rise to a reflectionless potential when, for some nonnegative integer $N$,

$$
\pi\left(a_{0}+\frac{1}{4}\right)^{1 / 2}=\pi\left(N+\frac{1}{2}\right),
$$

or when $a_{0}+\frac{1}{4}=\left(N+\frac{1}{2}\right)^{2}$, that is, $a_{0}=N(N+1)$. In retrospect this is not too surprising, due to the connection we've seen between the $\operatorname{sech}^{2}(x)$ potential and the associated Legendre equation. As is well known, the ordinary Legendre equation has eigenvalues $N(N+1)$, and one could perhaps quote results from its theory that reduce the possibilites for reflectionless potential scalings to these. Nevertheless, I personally have found it satisfying to feel out these historically well-traveled connections with bare hands.


Figure 16. The $k$-independent proportionality factor $\frac{\cos \left(\pi\left(a_{0}+\frac{1}{4}\right)^{1 / 2}\right)}{\pi}$ for the scattering coefficient, plotted against $a_{0}$. This factor is seen to be zero for $a_{0}=2$ and $a_{0}=6$, with the next reflectionless potential-for the 3 -soliton solution-corresponding to $a_{0}=12$. These special $a_{0}$ are of the form $N(N+1)$.
5.4. Why is the discrete spectrum time-independent for KdV potentials? Finally, we come to a real mystery. The scattering and inverse scattering, though necessarily elaborate, always work in reasonable situations; this has little to do with the KdV equation. The fact that explicit formulas can be obtained for $N$-soliton solutions to the KdV equation is interesting and surprising to me, but is ultimately explained by a study in special functions. The most mysterious thing here, I think, is that the inverse scattering methodspecifically that the scattering data evolves through time according to simiple rules-exists at all. There's no mistaking the fact that this has to do with the potentials $V_{t}(x)$ forming a solution $u(x, t) \equiv V_{t}(x)$ for the KdV equation specifically. Thus, any ability we have to apply what we've learned here to other equations and contexts will depend on how well we are able to explain the scattering data evolution.

At the root of understanding this evolution is the fact that the discrete spectrum $\Lambda_{t}$ of the normalized Schrödinger operator with potential $V_{t}(x)$ is time-independent, that is, $\Lambda_{t}=\Lambda_{0}$. Recall that it is only after we know this that we can define $c_{n}(t)$ for all $t$ by $\psi_{n}(x) \sim c_{n}(t) e^{-\kappa_{n} x}$ as $x \rightarrow \infty$ for appropriate time-dependent $\psi_{n}$ solving the normalized Schrödinger equation with potential $V_{t}(x)$. We could allow the $\kappa_{n}$ to depend on $t$ and attempt in this way to find
$u(x, t)$ that solve other equations, but it might well be that we obtain a different sort of formula for $u(x, t)$ at every time $t$ (but I couldn't solve Marchenko when $\kappa_{n}$ varies anyway). It is awfully convenient that $\Lambda_{t}=\Lambda_{0}$.

To properly understand the spectrum, we should phrase our question in terms of linear operators. We define the normalized linear Schrödinger operator $L$ by $L v \equiv-\frac{\partial^{2} v}{\partial x^{2}}+u v$, where $u$ is our potential and subscripts denote differentation. When we allow $u=u(x, t)$ to vary with time, $L$ becomes indexed by $t$. Then our question is, "Why is the discrete spectrum of this symmetric operator time-independent when $u$ solves the KdV equation?"

The evolution of $\Lambda_{t}$ should be determined somehow by the evolution equation giving $\frac{d L}{d t}$. But calculating this quantity, $\dot{L}$, (noting that the time derivative of the operator $\frac{\partial}{\partial x}$ is 0 as the operator itself is time-independent) gives

$$
\dot{L} v=\frac{\partial u}{\partial t} v
$$

which by itself explains nothing: I can use the KdV equation to write $\frac{\partial u}{\partial t}=6 u \frac{\partial u}{\partial x}-\frac{\partial^{3} u}{\partial x^{3}}$ and then I am stuck. This is a similar problem conceptually to how a solitary wave $\eta(x, t)=$ $\eta(x-c t)$ solving the KdV equation satisfies $\frac{\partial \eta}{\partial t}+\frac{\partial(c \eta)}{\partial x}=0$, and its governing equation even reduces to this when you plug in $\eta$, yet the linear operator $K w \equiv \frac{\partial w}{\partial t}+\frac{\partial(c w)}{\partial x}$ is no KdV equation. The information we want is not at the surface.

There are many strategies before us, and none look promising. Earlier, we wrote the replacement for $c, ~ \varpi$, in terms of $\eta$ to obtain the correct form of the governing equation. Our problem here can perhaps be solved similarly by writing $\dot{L}$ in some appropriate form, possibly in terms of $L$. Yet with no guide as to what to expect, this turns out to be an onerous task. A much more direct approach to showing the discrete spectrum is time-independent is to somehow show, starting from the eigenvalue equation $L v=\lambda v$, that $\frac{d \lambda}{d t}=0$ identically. For this it suffices, in principle, to differentiate the eigenvalue equation and solve for $\frac{d \lambda}{d t}$. We will be able to calculate this quantity, $\dot{\lambda}$, in terms of things known to us, but only if we know $\dot{v}$, the time evolution for the corresponding eigenfunction of $L$. We can find out what this is, but that involves writing out the equation $L v=\lambda v$ explicitly, solving for $u$, and plugging that into the KdV equation-quite a mess. Worse, $\dot{\lambda}$ will get in the way-and that's what we are trying to find. In fact, it looks like we would be better off going through that process and solving for $\dot{\lambda}$ instead of $\dot{v}$, finding an ad hoc argument (integrating over space and invoking the supposition that $u(x, 0) \rightarrow 0$ as $|x| \rightarrow \infty$, maybe more than once) to show $\dot{\lambda}=0$ afterwards. It appears that this is what Gardner, Greene, Kruskal, and Miura actually did in [10]. The point is that, yes, we can write the time evolution operator $M v \equiv \frac{d v}{d t}$ for $v$ in such a way that only space-variable operations appear, but we would probably need $\frac{d \lambda}{d t}$ first, so we do not come out ahead. At this point it seems clear that I have to roll up my
sleeves and apply the quotient rule to $u=\frac{\psi_{x x}+\lambda \psi}{\psi}$, or at least the sum, product, and power rules to $u=\psi^{-1} \psi_{x x}+\lambda$.

Instead, I look in the standard graduate text by Evans (which does not treat inverse scattering-so no spoilers) on partial differential equations in the hopes of finding something useful. I quote an entire exercise statement from page 247 of [7].
(Lax pairs) Assume that $\{L(t)\}_{t \geq 0}$ is a family of symmetric linear operators on some real Hilbert space $H$, satisfying the evolution equation

$$
\dot{L}=[B, L]=B L-L B
$$

for some collection of operators $\{B(t)\}_{t \geq 0}$. Suppose also that we have a corresponding family of eigenvalues $\{\lambda(t)\}_{t \geq 0}$ and eigenvectors $\{w(t)\}_{t \geq 0}$ :

$$
L(t) w(t)=\lambda(t) w(t)
$$

Assume that $L, B, \lambda$, and $w$ all depend smoothly upon the time parameter $t$. Show that

$$
\dot{\lambda} \equiv 0
$$

Below I include my solution to this exercise, which will indeed apply to the $L$ we have chosen. In fact, I leave it to the interested reader to find a $B=B(t)$ that works. Start from the ansatz $B v=a v_{x x x}+b u v_{x}+c u_{x} v$, with $a, b$, and $c$ constant. The patient reader can show that for a suitable choice of these constants, we have

$$
[B, L] v=\left(6 u \frac{\partial u}{\partial x}-\frac{\partial^{3} u}{\partial x^{3}}\right) v
$$

so that we obtain $\dot{L} v=\frac{\partial u}{\partial t} v=[B, L] v$ identically in $v$-and hence a time-independent spectrum-when $u$ solves the KdV equation. This, at least, provides one answer to the question of this subsection. The pair ( $B, L$ ) is called a "Lax pair" for the KdV equation. A systematic search for soliton equations-each with their own "inverse scattering transform"may be initiated by writing down all of the symmetric differential operators $L$ you can think of and seeing, for each, if there is a differential operator $B$ whose commutator with $L$ gives an operator which is multiplication by the right side of an evolution equation of the form $\frac{\partial u}{\partial t}=F\left(u, \frac{\partial u}{\partial x}, \frac{\partial^{2} u}{\partial x^{2}}, \ldots\right)$. Yet little will be found. If one is willing to use matrices of differential operators in defining $L$, however, then several famous soliton equations can be found this way: the Boussinesq equation, the sine-Gordon equation, and the nonlinear Schrödinger equation in one spatial dimension (though this last requires significantly more work to find this way), these last two being introduced by [5] in appropriate physical contexts.

Note that my proof of the Lax pairs theorem makes no use of completeness, so the Lax pairs construction originating in [18]-Lax's paper on solitary waves, begun around the time of publication of the paper of Gardner, Greene, Kruskal, and Miura-will apply in any inner
product space. The proof for Hermitian operators in complex inner product spaces is entirely analogous because all of their eigenvalues are real.

Proof: The equation $\lambda(t) w(t)=L(t) w(t)$ is differentiated with respect to $t$. This yields

$$
\begin{aligned}
\dot{\lambda} w+\lambda \dot{w}=\frac{d}{d t}(\lambda(t) w(t)) & =\frac{d}{d t}(L(t) w(t)) \\
& =\dot{L} w+L \dot{w} \\
& =(B L-L B) w+L \dot{w} \\
& =B L w-L B w+L \dot{w}
\end{aligned}
$$

This implies that

$$
\dot{\lambda} w=B L w-L B w+L \dot{w}-\lambda \dot{w}
$$

Since $\langle L(t) u, v\rangle=\langle u, L(t) v\rangle$ for all $t, u, v$, calculating the quantity $\dot{\lambda}\langle w, w\rangle$ obtains

$$
\begin{aligned}
\dot{\lambda}\langle w, w\rangle=\langle\dot{\lambda} w, w\rangle & =\langle B L w-L B w+L \dot{w}-\lambda \dot{w}, w\rangle \\
& =\langle B L w, w\rangle-\langle L B w, w\rangle+\langle L \dot{w}, w\rangle-\langle\lambda \dot{w}, w\rangle \\
& =\langle B \lambda w, w\rangle-\langle B w, L w\rangle+\langle\dot{w}, L w\rangle-\lambda\langle\dot{w}, w\rangle \\
& =\langle\lambda B w, w\rangle-\langle B w, \lambda w\rangle+\langle\dot{w}, \lambda w\rangle-\lambda\langle\dot{w}, w\rangle \\
& =\lambda\langle B w, w\rangle-\lambda\langle B w, w\rangle+\lambda\langle\dot{w}, w\rangle-\lambda\langle\dot{w}, w\rangle \\
& =0 .
\end{aligned}
$$

Since $w(t)$ is an eigenvector for each $t, w(t) \neq 0$ for each $t$ by definition, so $\langle w(t), w(t)\rangle$ never vanishes. Dividing by $\langle w(t), w(t)\rangle$ in the equation $\dot{\lambda}(t)\langle w(t), w(t)\rangle=0$ gives $\dot{\lambda} \equiv 0$.
5.5. What is the inspiration for Lax pairs? I do not know the answer to this question. Lax in [18] certainly speaks of various unitarily equivalent operators, so he may have had similar matrices in the back of his mind. One simple way to generate a family of matrices $L(t)$ which all have the same set of eigenvalues is to take any constant matrix $C$ and any family of invertible matrices $A(t)$. Then

$$
L(t) \equiv A(t) C A(t)^{-1}
$$

has the same eigenvalues as $C$ for all $t$. As shown in [14] by explicit calculation, this implies that

$$
\dot{L}(t)=[B(t), L(t)]
$$

where

$$
B=\dot{A 5} A^{-1} .
$$

Someone aware of this result might be inspired to ask if the converse is true: if this evolution equation holds for some $B(t)$, then must the set of eigenvalues of $L(t)$ be time-independent? This would then lead to the Lax pairs theorem.

I would like to offer one other possible source of inspiration, which I found on my ownthough it is hopefully well known. My idea begins more generally: suppose the eigenvalues of any linear operator $L(t)$ are time-independent, so that for eigenvalue equations of the form $L v=\lambda v$, the eigenvalues $\lambda$ are truly constant. Now we differentiate both sides, obtaining

$$
\dot{L} v+L \dot{v}=\lambda \dot{v}
$$

Thinking physically, we define the time-evolution operator $M$ for the eigenfunctions by $M v \equiv \dot{v}$. Substituting $M v$ for $\dot{v}$ in the above equation gives

$$
\dot{L} v+L M v=\lambda M v
$$

Now slip $\lambda$ inside $M$ :

$$
\dot{L} v+L M v=M \lambda v=M L v
$$

(this is an equation basically of the form $\left.f^{\prime} g+f g^{\prime}=(f g)^{\prime}\right)$. Finally, we transpose $L M v$ :

$$
\dot{L} v=M L v-L M v=[M, L] v
$$

If the eigenfunctions form a basis or complete set, then this equation will be true for all elements $v$ of the vector space by writing arbitary elements as linear combinations of basis elements. Thus,

$$
\dot{L}=[M, L],
$$

where $M$ is the time evolution operator.

## Concluding Remarks: Fiber Optics

One of the applications of soliton theory provides an amusing epilogue to the story of Russell and his interest in solitary waves. As was done in the nineteenth century with the assistance of ships such as Russell's The Great Eastern, cables are again being laid under the ocean for communication between Europe and North America. Though Russell's boat is not being used, the theory arising from his work still plays a crucial role. The cables this time are not electronic but optical.

As Hasegawa and Tappert predicted in 1973, see [12], utilization of the nonlinear dependence of the index of refraction on intensity "makes possible the transmission of picosecond optical pulses without distortion in dielectric fiber waveguides with group velocity dispersion." Moreover, "numerical simulations show that above a certain threshold power level such pulses are stable under the influence of small perturbations, large perturbations, white
noise, or absorption." The information in the optical fiber is carried in the form of solitonssolitary waves of light. Just as Russell's wave on the canal sped on "without change of form," optical solitons are remarkably effective in preserving communications data sent over long distances.

Yet according to [20], "Their result was, for the most part, regarded as the satisfaction of idle mathematical curiousity with few application possibilities." Instead, the 1970's saw heavy investment in linear fiber optics, and data losses proved a major obstruction. It was not until the 1980's that Hasegawa and Tappert's predictions for nonlinear optics were tested. As [20] states, "The results catalyzed explosive growth in nonlinear optics." Information on this subject can be readliy acquired from Boyd's text [3]; see especially Chapter 8 , and also check out Raman scattering in Chapter 10-it can be used to boost energy to compensate for long-distance loss.

Today, the idea that weak intrinsic optical nonlinearity of the glass core in a fiber can be balanced with weak waveguide dispersion-in a manner leading to solitons-is well established, and exploited to the hilt. As the Fiber Optic Reference Guide [11] says, "The ability of soliton pulses to travel on the fiber and maintain their launch wave shape makes solitons an attractive choice for very long distance, high data rate fiber optic transmission systems."

## Further Directions

Here are two further directions for study not explored here. One is the application of soliton equations to biological systems. For example, the KdV equation can be used as a rough model for blood pressure pulses. How well does this work? What other biological systems can be modeled by soliton equations? Known examples include energy transfer in proteins and DNA fluctuations, but there are probably many more. Another direction for further study is the examination of generalizations of various soliton equations ( KdV , nonlinear Schrödinger, etc.) to more than one spatial dimension. For example, the soliton resolution conjecture-of current research interest-says that "most" finite-energy initial conditions for the three-dimensional nonlinear Schrödinger equation exhibit behavior very much like that for the one-dimensional equation-initial profiles break up into a family of soliton-like waves plus "radiation." Yet the nonlinear Schrödinger equation in dimensions greater than one is not a soliton equation. How can this conjecture be reconciled with the rather non-KdV phenomenon of "soliton collapse" that occurs in two dimensions for nonlinear optics and three dimensions for plasma theory?

## Photograph Acknowledgements

The photograph in Figure 1 is from a meeting at Heriot-Watt University of Edinburgh in 1995. Scientists tried to recreate a solitary wave in the Union Canal, where John Scott

Russell had first observed the phenomenon. The picture is due to Chris Eilbeck and HeriotWatt University. The photograph in Figure 6 of the collision of two small-amplitude solitons (a 2-soliton) is from the west coast of the US, on a beach in Oregon. The picture was taken by Terry Toedtemeier in 1978.

## Further Reading

I have found three texts to be particularly helpful. From the viewpoint of pure mathematics, the text [14] by Kasman has been invaluable; it is perhaps the most elementary introduction, and its third chapter greatly influenced the opening section of this essay. However, being only a "glimpse" of solition theory, some of the essentials-the physics and the inverse scattering transform-are left out. For physical context and an introduction to hydrodynamic and topological solitons, see [5] by Dauxois and Peyrard. This book devotes a chapter to the modeling process for ion acoustic waves in a plasma, and also includes many sample applications from solid state physics and molecular biology. Finally, for actually using the inverse scattering transform, the applied mathematics perspective is most natural. The book [6] by Drazin and Johnson is a concise and effective introduction to this area.

## References

[1] Airy, G.B. "Tides and Waves," Encyc. Metrop., 192 (1845), pp. 241-396.
[2] Boussinesq, J. "Théorie des ondes et des remous qui se propagent de long d'un canal rectangulaire horizontal, en communiquant au liquide continu dans ce canal des vitesses sensiblement pareilles de la surface au fond," J. Math. Pures Appl., 17 (1872), pp. 55-108.
[3] Boyd, R.W. Nonlinear Optics, Third Edition, Elsevier, Burlington, MA, 2008.
[4] Currie, I.G. Fundamental Mechanics of Fluids, Fourth Edition, CRC Press, Boca Raton, FL, 2013.
[5] Dauxois, T. and Peyrard, M. Physics of Solitons, Cambridge University Press, Cambridge, UK, 2006.
[6] Drazin, P.G. and Johnson, R.S. Solitons: an introduction, Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge, UK, 1989.
[7] Evans, L.C. Partial Differential Equations, Second Edition, Graduate Studies in Mathematics, Volume 19, American Mathematical Society, Providence, RI, 2010.
[8] Ferguson, S.J. "A One-Sentence Line-of-Sight Proof of the Extreme Value Theorem," American Mathematical Monthly, 121 (2014), p. 331.
[9] Fermi, E., Pasta, J., and Ulam, S. "Studies of Nonlinear Problems. I," in Nonlinear Wave Motion, Lectures in Applied Mathematics, vol. 15, Amer. Math. Soc., 1974, pp. 143-145.
[10] Gardner, C.S., Greene, J.M., Kruskal, M.D., and Miura, R.M. "Method for solving the Korteweg-de Vries equation," Physics Rev. Lett., 19 (1967), pp. 1095-1097.
[11] Goff, D.R. and Hansen, K.S. Fiber Optic Reference Guide: a practical guide to communications technology, Focal Press, 2002.
[12] Hasegawa, A. and Tappert, F. "Transmission of stationary nonlinear optical pulses in dispersive dielectric fibers. I. Anomalous dispersion," Applied Physics Letters, 23 (1973), pp. 142-144.
[13] de Jager, E.M. "On the Origin of the Korteweg-de Vries Equation," arXiv:math/06602661v2 [math.HO] 8 Dec 2011.
[14] Kasman, A. Glimpses of Soliton Theory, Student Mathematical Library no. 54, American Mathematical Society, Providence, RI, 2010.
[15] Korteweg, D.J. and de Vries, G. "On the change of form of long waves advancing in a rectangular channel and on a new type of long stationary wave," Philos. Mag., 39 (1895), pp. 422-443.
[16] Kruskal, M.D., and Zabusky, N.J. "Interactions of 'solitons' in a collisionless plasma and the recurrence of initial states," Physics Rev. Lett., 15 (1965), pp. 240-243.
[17] Lang, S. Short Calculus, Springer, New York, NY, 2002.
[18] Lax, P.D. "Integrals of nonlinear equations of evolution and solitary waves," Communications on Pure and Applied Mathematics, 21 (1968), pp. 467-490.
[19] Metropolis, N. "The Age of Computing: a personal memoir," Daedalus, 121 (1992), pp. 119-130.
[20] National Research Council. Nonlinear Science, The National Academies Press, Washington, D.C., 1997.
[21] Rayleigh, (Strutt, J.W.) "On Waves," Phil. Mag., 1 (1876), pp. 257-271.
[22] Russell, J.S. "Report on Waves," Report of the fourteenth meeting of the British Association for the Advancement of Science, York, September 1844 (London 1845), pp. 311-390.
[23] Schwabl, F. Statistical Mechanics, Second Edition, Springer, Berlin, Germany, 2006.
[24] Stokes, G.G. "On the theory of oscillatory waves," Transactions of the Cambridge Philosophical Society, 8 (1847), pp. 441-455.
[25] Ulam, S. Collected Papers of Enrico Fermi, University of Chicago Press, Chicago, IL, 1965.


[^0]:    ${ }^{1}$ Boussinesq does not introduce this $\zeta$ abbreviation. I've introduced it to make his work easier to follow.

[^1]:    ${ }^{2}$ Boussinesq denotes the wave velocity by $\omega$. This is an unfortunate choice for velocity as $\omega$ is frequently used to denote angular frequency. To avoid confusion, we have denoted it by $\varpi$ instead.
    ${ }^{3}$ This result, like the existence of solitary waves themselves, initially surprised me. To clarify, let us define the nonlinear opeator $Z$ by $Z \eta \equiv \zeta=\frac{3 \eta^{2}}{2 h}+\frac{h^{2}}{3} \frac{\partial^{2} \eta}{\partial x^{2}}$. In terms of $Z$, our result is that $Z \eta=\frac{\eta 0}{h} \eta$, so we have the solution to a nonlinear eigenvalue problem.

[^2]:    ${ }^{4}$ In terms of the $\delta$ and $\varepsilon$ introduced in the discussion of Rayleigh's function $\eta(x)$, the chief assumption of this thorough derivation is that $\delta^{2} \approx \varepsilon$. This might appear to be overly restrictive, but in practice it is not because $\eta_{0}$ is left free to settle to the appropriate height.
    ${ }^{5}$ Here we neglect surface tension. Actually, in their original derivation, Korteweg and de Vries account for surface tension, but their coefficients are more complicated. They also do not assume that all relevant functions vanish at $\pm \infty$, so their derivation is more general than Boussinesq's investigation in that sense.

[^3]:    ${ }^{6}$ Boussinesq did not do this in [2]. However, in one of his later papers the Korteweg-de Vries equation does appear in a footnote, apparently the result of differentiating the equation $\psi=0$ with respect to $x$.

[^4]:    ${ }^{7}$ The Fermi-Pasta-Ulam paper only gives the equations of motion, but I'm giving the Hamiltonian because I find it a bit more transparent. Thus, the paper speaks of a quadratic force where I would speak of a cubic term in the Hamiltonian. The coefficients $K$ and $\frac{K \alpha}{3}$ are written to match those in [16].

[^5]:    ${ }^{8}$ In [9] they are given directly by $a_{k}=\sum_{i} \sin \frac{i k \pi}{64}$ without normalization. For the normal modes we follow the notation and discussion of [5].
    ${ }^{9}$ By the way, a common nickname for the mechanically-analyzing, numerically-integrating computer at Los Alamos (MANIAC) was "Metropolis And von Neumann Invent Awful Computer."
    ${ }^{10}$ The computer programming was done by Mary Tsingou; although acknowledged in the paper, she was not considered an author because she was not involved in its writing (though neither was Fermi-being by that time deceased).

[^6]:    ${ }^{11}$ In detail, they considered the initial value problem for $u_{t}+u u_{x}+\delta^{2} u_{x x x}=0$, with periodic boundary conditions-a numerically stable problem which we have not considered here. They solved the equation with $u(x, 0)=\cos (\pi x), 0 \leq x \leq 2$, for $u, u_{x}$, and $u_{x x}$ periodic on $[0,2]$ for all $t$; and they chose $\delta=0.022$.

[^7]:    ${ }^{12}$ Periodicity was also enforced in their numerical simulations. This turned out to be more natural than holding the ends fixed as in the original experiments.

[^8]:    ${ }^{13}$ One is reminded vaguely of how some seventeenth century astronomers were dissatisfied with the "imperfection" of the elliptical orbits of the planets-in comparison with circular orbits-until they perceived the seeming perfection of the universal law of gravitation at the source.

[^9]:    ${ }^{14}$ To quote Dr. Don Gurnett, it is "two-high."

[^10]:    ${ }^{15}$ I hope the reader will handle with equanimity the fact that the letter $a$ is used three times below, as the transmission coefficient $a(k)$, an abbreviation $\tilde{a}$ of Drazin and Johnson, and as the vertical scaling $a_{0}$. Thank goodness for subscripts.

