Topic 2: The Korteweg-de Vries Equation

This is a lecture about some of the properties of the Korteweg-de Vries equation, and its role in the history of the subject of soliton theory.

The KdV equation. Symmetry properties. An equation for $u(x,t)$ is said to be a Korteweg-de Vries (KdV) equation (Korteweg and de Vries, 1895) if it is of the form

$$u_t + \alpha u u_x + \beta u_{xxx} = 0,$$

for some constants α and β. The coefficients are somewhat arbitrary, in the sense that we can change them into others, or get rid of them altogether (that is, make them all equal to one), by choosing appropriate units for the equation. The first thing we can try to do is to choose a new spatial scale, so that for some constant A we have

$$x = AX$$

defining some new length scale X that goes through one unit when x goes through A units. Similarly we can pick some constant B and set

$$t = BT$$

defining some new time scale T. Then, using the chain rule

$$\frac{\partial}{\partial x} = \frac{\mathrm{d}X}{\mathrm{d}x} \cdot \frac{\partial}{\partial X} = \frac{1}{A}\frac{\partial}{\partial X}$$

and

$$\frac{\partial}{\partial t} = \frac{\mathrm{d}T}{\mathrm{d}t} \cdot \frac{\partial}{\partial T} = \frac{1}{B}\frac{\partial}{\partial T}$$

so the equation becomes

$$\frac{1}{B}u_T + \frac{\alpha}{A}u u_X + \frac{\beta}{A^3}u_{XXX} = 0,$$

or after multiplying through by B,

$$u_T + \frac{\alpha B}{A}u u_X + \frac{\beta B}{A^3}u_{XXX} = 0.$$

With two free scaling constants A and B we get the idea that we might be able to choose them so the coefficients of uu_X and u_{XXX} both become 1. However, since B/A and B/A^3 always have the same sign, this is only possible if α and β already have the same sign. In general the best we can hope for is to arrange that α is replaced with ± 1 and β is replaced with 1 by setting

$$\frac{\beta B}{A^3} = 1 \quad \text{and} \quad \left|\frac{\alpha B}{A}\right| = 1.$$

Taking also the absolute value in the first equation and solving for $|A|$ and $|B|$ gives

$$|A| = \sqrt{\left|\frac{\beta}{\alpha}\right|} \quad \text{and} \quad |B| = \frac{1}{|\alpha|}\sqrt{\left|\frac{\beta}{\alpha}\right|}.$$

To get a solution of the original equations, let's assume that $B > 0$ (which preserves the direction of time) and thus

$$A = \operatorname{sgn}(\beta)\sqrt{\left|\frac{\beta}{\alpha}\right|} \quad \text{and} \quad B = \frac{1}{|\alpha|}\sqrt{\left|\frac{\beta}{\alpha}\right|}.$$

It is easy to check that these scale factors render the KdV equation in the form

$$u_T + \operatorname{sgn}(\alpha\beta) u u_X + u_{XXX} = 0$$

as desired.

Now, the KdV equation is a nonlinear equation, and as such it is *not* left unchanged by scaling of u. Therefore, we may define a further rescaling by choosing C to be some unit for u, and then setting

$$u = CU$$

defining some new dependent variable $U(X,T)$. The differential equation then becomes

$$U_T + \frac{\alpha BC}{A} U U_X + \frac{\beta B}{A^3} U_{XXX} = 0.$$

Clearly, the scaling parameters A, B, and C all individually affect the coefficients in the equation. However in combination their effects can cancel out. Indeed, suppose we wanted to choose A, B, and C now so that the coefficients in the equation are left unchanged. Then we would insist that

$$\frac{BC}{A} = 1 \quad \text{and} \quad \frac{B}{A^3} = 1.$$

This is only two equations on three unknowns, so we could choose $A = \lambda$ to be an arbitrary nonzero number, and then solve for $B = \lambda^3$ and $C = \lambda^{-2}$. This shows that whenever $u(x,t)$ is a solution of

$$u_t + \alpha u u_x + \beta u_{xxx} = 0,$$

then so is $\lambda^2 u(\lambda x, \lambda^3 t)$ for any constant $\lambda \neq 0$. Such a family of transformations that preserves the equation but modifies the solution is called a *symmetry group* of the equation. For obvious reasons this kind of symmetry group is called a *scaling symmetry group* or sometimes a *similarity transformation group*. Note that contained in this group is the discrete transformation group associated with $\lambda = \pm 1$. This amounts to the observation that whenever $u(x,t)$ is a solution, so is $u(-x,-t)$.

Another kind of symmetry group of the KdV equation involves going into a frame of reference moving with constant velocity c with respect to the fixed frame. This means that instead of using the independent variables x and t we want to use the independent variables

$$\xi = x - ct, \qquad \tau = t.$$

Then, since by the chain rule

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

and

$$\frac{\partial}{\partial t} = \frac{\partial \xi}{\partial t} \cdot \frac{\partial}{\partial \xi} + \frac{\partial \tau}{\partial t} \cdot \frac{\partial}{\partial \tau} = -c \frac{\partial}{\partial \xi} + \frac{\partial}{\partial \tau},$$

the KdV equation becomes

$$u_\tau + (\alpha u - c)\, u_\xi + \beta u_{\xi\xi\xi} = 0.$$

This has not left the equation invariant. But now we may notice that if we add a constant to u:

$$u = U + \frac{c}{\alpha}$$

then we arrive at

$$U_\tau + \alpha U U_\xi + \beta U_{\xi\xi\xi} = 0.$$

Therefore, whenever $u(x,t)$ is a solution of

$$u_t + \alpha u u_x + \beta u_{xxx} = 0,$$

then so is $u(x - ct, t) + c/\alpha$. This kind of symmetry group is called a *translational symmetry group* or a *Galilean transformation group*.

Origin of the KdV equation. The period 1834–1895. In the early days of mathematical modeling of water waves, it was assumed that the wave height was small compared to the water depth, which leads to linear dispersive equations a representative model of which is

$$u_t + u_{xxx} = 0.$$

Such equations are somewhat satisfying in this regard because they have solutions that resemble waves traveling along with constant speed and fixed profile along the water surface, just like one sees in nature. To find them, we go into a moving frame with speed c by introducing new variables $\xi = x - ct$ and $\tau = t$, and then we seek solutions of the resulting equation that are independent of τ:

$$u = f(\xi) \qquad \text{implies} \qquad -cf'(\xi) + f'''(\xi) = 0.$$

Integrating once,

$$f''(\xi) - cf(\xi) = B$$

where B is an integration constant. If $c = -k^2 < 0$, then f will be bounded as a function of ξ, and we get

$$f(\xi) = A\cos(k(\xi - \xi_0)) + \frac{B}{k^2} = A\cos(k(\xi - \xi_0)) + D,$$

where D is another arbitrary constant. Thus, the family of bounded traveling wave solutions to this linear equation is exhausted by periodic sinusoidal wave shapes. The speed of propagation depends on the wavenumber k, since $c = -k^2$. (Note that these waves all propagate to the left only; the apparent broken symmetry is explained by the fact that the equation $u_t + u_{xxx} = 0$ is derived from water wave theory in a moving coordinate frame.) There is a lot that can be built out of these waves by superposition (it is a linear equation) but the basic point of view at this time in history was that the only waves that travel along at a constant speed without changing their form were periodic trains of waves.

In 1834 an event occurred that began to change this point of view. John Scott-Russell's account of his accidental observation of what he called a "great wave of translation" and what turned out to be the solitary wave solution of the KdV equation is best described by his own words, written to the British Association in 1844 as part of his "Report on Waves":

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

Scott-Russell's "great wave of translation" was recreated in 1995 at the naming ceremony of the Scott-Russell aqueduct on the Union canal connecting Glasgow and Edinburgh where Scott-Russell had made his original observation (see photo on the class webpage).

The water wave theory needed to be corrected to be able to reproduce the solitary wave of Scott-Russell, and the essential modification required was to throw out the assumption that the wave height was small compared with the water depth (it was still assumed that the wave was long compared with the depth, however). This modification makes the equations of motion *nonlinear*. In 1895, Korteweg and de Vries rediscovered[1] a model of water waves in this situation the equation that bears their name:

$$u_t + uu_x + u_{xxx} = 0.$$

We are assuming that we have used scaling to make the coefficients all one. This equation has within it the solitary wave of Scott-Russell. Again we look for traveling waves $u = f(\xi = x - ct)$:

$$-cf'(\xi) + f(\xi)f'(\xi) + f'''(\xi) = 0.$$

Noting that $ff' = (f^2)'/2$, we can integrate one time to get

$$-cf(\xi) + \frac{1}{2}f(\xi)^2 + f''(\xi) = A$$

where A is an integration constant. If we want to capture a wave profile that is a "bump" on a constant level background (the resting height of water in the channel), then taking into account the Galilean symmetry group we may assume that the background level is zero, and then we are seeking a traveling wave profile f that decays to zero with all of its derivatives as $\xi \to \pm\infty$. This implies that we should choose $A = 0$ for the integration constant. The wave profile equation thus takes the form

$$f'' + \left[\frac{1}{2}f^2 - cf\right] = 0,$$

[1] Boussinesq wrote it down earlier, in the 1870's.

which we may view as a nonlinear oscillator equation with a nonlinear "restoring force". The oscillator is at equilibrium when the velocity is zero ($f' = 0$) and when the force is zero. There are two equilibria for f: $f = 0$ and $f = 2c$. Assuming $c > 0$, we have a saddle point at $f = 0$ and a center point at $f = 2c > 0$. If $c < 0$, then it is the other way around, with a saddle point at $f = 2c < 0$ and a center point at $f = 0$. The phase portrait of this nonlinear oscillator in the (f, f')-plane exhibits a distinguished orbit, namely an arc that connects the saddle point to itself after passing once around the center point. Such an orbit is called a *homoclinic orbit*. It represents a solution of the differential equation that decays to the constant value of f associated with the saddle point as $\xi \to \pm\infty$. It is easy to see that all other non-equilibrium solutions of the differential equation are either periodic functions of ξ (these correspond to the closed orbits bounded by the homoclinic orbit) or are unbounded functions of ξ. It is a good exercise to confirm these facts with sketches of the phase portrait, and to do this it is useful to see that there is a conserved energy, that is, upon multiplying by f' we can integrate once more:

$$\frac{1}{2}(f')^2 + \frac{1}{6}f^3 - \frac{c}{2}f^2 = E$$

where E is an integration constant. Picking different values of E gives different implicit curves in the phase plane which contain the various orbits.

It should be clear that the only candidate for a solitary wave f that decays to zero as $\xi \to \pm\infty$ is to have an orbit homoclinic to the origin in the phase portrait, and this is only possible if $c > 0$ in which case it is the origin that is the saddle point. To find the solution corresponding to this orbit under the assumption that $c > 0$, observe that we must now choose $E = 0$ because we need the level curve through the origin $(f, f') = (0, 0)$. By inspection of the phase portrait for $c > 0$, it is clear that the homoclinic orbit lies in the right half of the (f, f') phase plane (and it is not hard to prove this rigorously) so there is no harm in writing $f = g^2$. Therefore, the energy conservation law for $E = 0$ takes the form

$$2g^2(g')^2 + \frac{1}{6}g^6 - \frac{c}{2}g^4 = 0$$

or

$$(g')^2 + \frac{1}{12}g^4 - \frac{c}{4}g^2 = 0$$

which is the homoclinic orbit equation for the Duffing oscillator, and can be solved in view of the identities:

$$\operatorname{sech}^2(x) + \tanh^2(x) = 1, \qquad \frac{\mathrm{d}}{\mathrm{d}x}\operatorname{sech}(x) = -\operatorname{sech}(x)\tanh(x).$$

Thus, assuming $g(\xi) = a\operatorname{sech}(b(\xi - \xi_0))$, and substituting,

$$a^2b^2\operatorname{sech}^2(b(\xi-\xi_0))\tanh^2(b(\xi-\xi_0)) + \frac{1}{12}a^4\operatorname{sech}^4(b(\xi-\xi_0)) - \frac{c}{4}a^2\operatorname{sech}^2(b(\xi-\xi_0)) = 0$$

or

$$b^2\tanh^2(b(\xi-\xi_0)) + \frac{a^2}{12}\operatorname{sech}^2(b(\xi-\xi_0)) - \frac{c}{4} = 0.$$

Therefore if $b^2 = a^2/12 = c/4$ then we have found a solution. The solitary wave of Scott-Russell is

$$u(x,t) = a^2\operatorname{sech}^2(b(x-ct-x_0)) = 3c\operatorname{sech}^2\left(\frac{\sqrt{c}}{2}(x-ct-x_0)\right)$$

for any wave speed $c > 0$. Note that while the linear waves (bounded traveling wave solutions of $u_t + u_{xxx} = 0$) travel to the left only, the solitary waves (localized traveling wave solutions of $u_t + uu_x + u_{xxx} = 0$) travel to the right only.

The solitary wave can exist in the KdV equation because this equation contains effects of both nonlinearity and dispersion, and these effects are dynamically balanced by the solitary wave solution. We recall once again the analogy of the kids walking on the trampoline.

The twentieth century. The experiments of Fermi-Pasta-Ulam-Tsingou and Zabusky-Kruskal. In the 1950's, Enrico Fermi, John Pasta, Stan Ulam, and Mary Tsingou studied the question of how *thermalization* occurs in solid-state physics. A simple model for a solid material is a chain of identical particles (of mass m) connected to nearest neighbors by identical springs (with restoring force $F(\Delta)$ where Δ is the length of the spring). Thus, Newton's second law says

$$m\frac{\mathrm{d}^2 x_n}{\mathrm{d}t^2} = F(\Delta_0 + x_{n+1} - x_n) - F(\Delta_0 + x_n - x_{n-1}), \quad x_n = x_n(t), \quad n \in \mathbb{Z}$$

where Δ_0 is the equilibrium displacement of the particles. In the linear case, Hooke's law says that for some constant $k > 0$, $F(\Delta) = k\Delta$, and the system becomes

$$m\frac{\mathrm{d}^2 x_n}{\mathrm{d}t^2} = k(x_{n+1} - 2x_n + x_{n-1}).$$

The problem appears complicated, but it is easily solved by Fourier theory. Multiply through by $\mathrm{e}^{-\mathrm{i}pn}$ and sum over n assuming that $x_n(t)$ decays to zero for large n:

$$m\frac{\mathrm{d}^2}{\mathrm{d}t^2}\sum_{n=-\infty}^{\infty} x_n(t)\mathrm{e}^{-\mathrm{i}pn} = 2k(\cos(p) - 1)\sum_{n=-\infty}^{\infty} x_n(t)\mathrm{e}^{-\mathrm{i}pn}.$$

The sums are the Fourier coefficients:

$$\hat{x}(p,t) := \sum_{n=-\infty}^{\infty} x_n(t)\mathrm{e}^{-\mathrm{i}pn} \quad \Longrightarrow \quad m\frac{\mathrm{d}^2\hat{x}}{\mathrm{d}t^2}(p,t) = 2k(\cos(p) - 1)\hat{x}(p,t).$$

It is easy to solve this ODE in closed form for the Fourier coefficients:

$$\hat{x}(p,t) = A(p)\cos(\omega(p)t) + B(p)\sin(\omega(p)t)$$

where $A(p)$ and $B(p)$ are complex constants, and the frequency is

$$\omega(p) = \sqrt{\frac{2k}{m}(1 - \cos(p))}.$$

Since $x_n(t)$ is real-valued, $A(-p) = A(p)^*$ and $B(-p) = B(p)^*$ hold. The time-independent quantity $|A(p)|^2 + |B(p)|^2$ is usually interpreted as the *energy in mode p* (as the total squared L^2 norm is the sum or integral of these quantities by the Plancherel Theorem). There are two observations here:

- On the basis of physical reasoning of statistical thermodynamics, one expects that as the system evolves in time, it should evolve toward *equipartition* in which the energy is divided more or less equally among all of the modes.
- But the equation of motion shows that all $\hat{x}(p,t)$ are *uncoupled* from each other; indeed this is the very definition of what is meant by a *normal mode*. Therefore unless the initial condition is already in equipartition, the system can never evolve into such a state!

Fermi, Pasta, Ulam, and Tsingou sought to explain the observed equipartition by breaking the feature of the simple Hooke's law model that leads to the decoupling of the modes. They made the springs nonlinear. They used an expression of the form

$$F(\Delta) = k\Delta + \alpha\Delta^3$$

and performed numerical integration of the ODEs for the $x_n(t)$ on one of the first computers, the MANIAC at Los Alamos. They used an initial condition in which all of the energy was in the first few Fourier modes. At a sequence of times $t = t_k$ they computed the Fourier coefficients $\hat{x}(p,t_k)$ from the output of their program. As expected, the energy leaked from mode to mode. But what surprised them was that after some time (not too long) most of the energy spilled back into the first few modes! It was as if the nonlinearity they put into the problem only "temporarily" thermalized the material.

Later in 1965, Zabusky and Kruskal revisited this problem from the point of view of partial differential equations. They assumed that the particle positions $x_n(t)$ could be considered to be samples of a smooth function $v(x,t)$: $x_n(t) = v(nh,t)$ for some small lattice spacing h. Taking the limit $h \to 0$ in an appropriate frame of reference moving with constant velocity, they derived as a *continuum limit* equation for a function $u(x,t)$ related to v the KdV equation

$$u_t + uu_x + \epsilon^2 u_{xxx} = 0.$$

where ϵ is proportional to h. They carried out numerical simulations of this partial differential equation with initial data similar to FPUT's.

You can carry out Zabusky and Kruskal's numerical experiment and also view the dynamics corresponding power spectra $A(p)^2 + B(p)^2$ using the Mathematica notebook `ZabuskyKruskal-651-W18.nb`.

The pulses shed by the evolution were identified in shape by Zabusky and Kruskal as the solitary waves we obtained earlier. But they also noted the fact that these shapes survived multiple collisions with each other, and it was they who coined the term *solitons* to describe these robust objects.

Mathematical structure within the KdV equation.

Local conservation laws. The KdV equation

$$u_t + uu_x + u_{xxx} = 0,$$

conserves the *mass functional*

$$m[u] := \int_{-\infty}^{\infty} u(x)\,\mathrm{d}x.$$

This just represents the area of water under the free surface of the fluid. The best way to see that the mass is conserved is to notice that the KdV equation can written in the form

$$u_t + \left[\frac{1}{2}u^2 + u_{xx}\right]_x = 0.$$

Such an equation, of the form $D_t + F_x = 0$, is called a *local conservation law.* The *conserved local density* is D and its *flux* is F. Every local conservation law implies, under appropriate boundary conditions, the conservation of an integral of D:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_a^b D(x)\,\mathrm{d}x + F(b) - F(a) = 0$$

so if F vanishes or takes equal values at the boundary, the integral of D is conserved. Similarly, if u satisfies periodic boundary conditions on some finite interval. It is also easy to show that u^2 is a conserved local density for KdV: begin by multiplying the KdV equation through by $2u$ after which it becomes

$$(u^2)_t + 2u^2u_x + 2uu_{xxx} = 0.$$

Now we want to show that the last two terms comprise the x-derivative of some expression F in u and its derivatives. The first term is clearly a perfect x derivative:

$$2u^2u_x = \left(\frac{2}{3}u^3\right)_x,$$

but (in my opinion) the term $2uu_{xxx}$ is less obvious. But it too is a perfect derivative:

$$\begin{aligned} 2uu_{xxx} &= 2uu_{xxx} + 2u_xu_{xx} - 2u_xu_{xx} \\ &= (2uu_{xx})_x - \left(u_x^2\right)_x \\ &= \left(2uu_{xx} - u_x^2\right)_x. \end{aligned}$$

So we have found the local conservation law

$$\left(u^2\right)_t + \left(\frac{2}{3}u^3 + 2uu_{xx} - u_x^2\right)_x = 0$$

which implies that under suitable boundary conditions the integral of u^2 is a conserved quantity. Further conserved local densities also exist, but they must involve x-derivatives of u.

Constants of motion in finite-dimensional systems. Consider a finite-dimensional system of the form

$$\frac{\mathrm{d}u_j}{\mathrm{d}t} = F_j(u_1, \ldots, u_N), \quad j = 1, \ldots, N.$$

A *conserved quantity* or *constant of the motion* for this system is a function $E(u_1, \ldots, u_N)$ defined on the phase space such that whenever $u_j(t)$ satisfy the ODEs, then

$$\frac{\mathrm{d}}{\mathrm{d}t} E(u_1(t), \ldots, u_N(t)) = 0.$$

Thus, the value of the function E as t varies is equal to the value calculated on the initial conditions:

$$E(u_1(t), \ldots, u_N(t)) = E(u_1(0), \ldots, u_N(0)).$$

This means that it is consistent to restrict the system of ODEs to a level hypersurface given by

$$E(u_1, \ldots, u_N) = E_0$$

where E_0 is specified by initial conditions. The existence of one constant of motion reduces the dimension of the system of ODEs by one; in the best case we can solve the above equation for u_N in terms of $u_1, \ldots, u_{N-1}$ and E_0:

$$u_N = G_N(u_1, \ldots, u_{N-1}; E_0)$$

and then we have a smaller system governing $u_1, \ldots, u_{N-1}$ only:

$$\frac{\mathrm{d}u_j}{\mathrm{d}t} = F_j(u_1, \ldots, u_{N-1}, G_N(u_1, \ldots, u_{N-1}; E_0)) = \tilde{F}_j(u_1, \ldots, u_{N-1}; E_0), \quad j = 1, \ldots, N-1,$$

which determines $u_1(t), \ldots, u_{N-1}(t)$ from initial conditions and then we get $u_N(t)$ as a function of these and E_0. The process can be repeated. If we are lucky enough to find $N-1$ constants of motion $E_1, \ldots, E_{N-1}$, then the system will be reduced to a single autonomous nonlinear first-order differential equation:

$$\frac{\mathrm{d}u_1}{\mathrm{d}t} = \tilde{F}_1(u_1; E_{1,0}, E_{2,0}, \ldots, E_{N-1,0})$$

which can be solved to determine t as a function of u_1 by separation of variables. Modulo inverting the function $t(u_1)$ to find $u_1(t)$, this procedure solves the system of differential equations in closed form.

Finite-dimensional Hamiltonian systems. Completely integrable systems. If there is more structure, then fewer conserved quantities are needed to solve the problem. A *Hamiltonian* system is one of dimension $2N$ that can be written in the form

$$\frac{\mathrm{d}\vec{u}}{\mathrm{d}t} = \mathbf{J}\nabla H$$

for some function $H(\vec{u})$ called the *Hamiltonian function*; here $\mathbf{J}$ is (for the moment) the matrix (broken into four $N \times N$ blocks)

$$\mathbf{J} = \begin{pmatrix} 0 & \mathbb{I} \\ -\mathbb{I} & 0 \end{pmatrix}.$$

We can write easily an equation for any function $F(u_1, \ldots u_N)$ specified on the phase space, just by using the chain rule:

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \left\langle \nabla F, \frac{\mathrm{d}\vec{u}}{\mathrm{d}t} \right\rangle = \langle \nabla F, \mathbf{J}\nabla H \rangle =: \{F, H\}.$$

The angle brackets are the usual inner (dot) product, and the expression on the right-hand side defines the *Poisson bracket* between two functions F and H specified on the phase space.

It is easy to see that the Poisson bracket of two suitable functions on the phase space is again a function defined on the phase space. For example, suppose $N = 1$, so that $\vec{u} = (x, y)^{\mathsf{T}} \in \mathbb{R}^2$. Taking $F(\vec{u}) := x^2 + 3xy$ and $G(\vec{u}) = x + y^3$, we get

$$\nabla F = \begin{pmatrix} 2x + 3y \\ 3x \end{pmatrix}, \quad \nabla G = \begin{pmatrix} 1 \\ 3y^2 \end{pmatrix}, \quad \text{and so} \quad \mathbf{J}\nabla G = \begin{pmatrix} 3y^2 \\ -1 \end{pmatrix}.$$

Taking the dot product of ∇F and $\mathbf{J}\nabla G$ then gives $\{F, G\} = 6xy^2 + 9y^3 - 3x$. Clearly in the case of $N = 1$, the effect of multiplying by $\mathbf{J}$ is just to rotate ∇G by 90°. The Poisson bracket of two functions on the plane is therefore zero at some point if and only if their gradient vectors are parallel. There are, however,

more ways for a Poisson bracket to vanish at a point when we consider higher-dimensional phase spaces with $N > 1$. This is key for constructing integrable systems, as we will see.

The Poisson bracket operation has the following elementary properties:

- Skew symmetry: $\{F, G\} = -\{G, F\}$.
- Bilinearity: $\{c_1F_1 + c_2F_2, G\} = c_1\{F_1, G\} + c_2\{F_2, G\}$ where c_1 and c_2 are arbitrary constants. (Similarly replacing G by $c_1G_1 + c_2G_2$ by skew symmetry.)
- Leibniz rule: $\{F_1F_2, G\} = F_1\{F_2, G\} + F_2\{F_1, G\}$.
- Jacobi identity: $\{F, \{G, H\}\} + \{H, \{F, G\}\} + \{G, \{H, F\}\} = 0$.

These follow immediately from our formula for $\mathbf{J}$, but in fact we can replace this $\mathbf{J}$ with any matrix function $\mathbf{J} = \mathbf{J}(\vec{u})$ as long as the skew symmetry and Jacobi identity properties hold (the bilinear property and Leibniz rule hold automatically).

In a Hamiltonian system a function $F(u_1, \ldots, u_N)$ is a conserved quantity if

$$\{F, H\} = 0$$

as this is equivalent to $\mathrm{d}F/\mathrm{d}t = 0$. We say that F "Poisson-commutes" with H. Note that by skew-symmetry, we have $\{H, H\} = 0$, so the Hamiltonian function itself is always a conserved quantity. Every conserved quantity defines a symmetry group for the solutions of a Hamiltonian system. Suppose F is conserved. Then we can find (at least locally) a vector function $\vec{u}$ of two variables t and s such that both

$$\frac{\partial u_j}{\partial t} = \{u_j, H\} \qquad \text{and} \qquad \frac{\partial u_j}{\partial s} = \{u_j, F\}$$

because the mixed partials commute: taking $\partial/\partial s$ of the first equation and subtracting $\partial/\partial t$ of the second equation should give zero, and indeed:

$$\begin{aligned}
\frac{\partial}{\partial s}\{u_j, H\} - \frac{\partial}{\partial t}\{u_j, F\} &= \{\{u_j, H\}, F\} - \{\{u_j, F\}, H\} \\
&= -\{\{H, u_j\}, F\} - \{\{u_j, F\}, H\} \\
&= \{\{F, H\}, u_j\} \\
&= 0.
\end{aligned}$$

On the first line, we used the chain-rule identity to calculate the derivatives (note that F is the "Hamiltonian" for the "time" s). Next, we used skew symmetry and bilinearity, and then the Jacobi identity to arrive at the penultimate line. The final line follows from the conservation property $\mathrm{d}F/\mathrm{d}t = 0$ and bilinearity once again. The differential equation governing the s-dependence of $\vec{u}$ gives a symmetry group parametrized by s, because if for $s = 0$ the function $\vec{u}(0, t)$ satisfies $\mathrm{d}\vec{u}/\mathrm{d}t = \mathbf{J}\nabla H$, then for each $s \in \mathbb{R}$ so does $\vec{u}(s, t)$. You can also check that the transformations $\vec{u}(0, t) \mapsto \vec{u}(s, t)$ obtained by solving the differential equation for various final times s satisfy the axioms of a group (look it up if you don't know them). The association of a one-parameter continuous (Lie) group of a Hamiltonian system with each conserved quantity is a result known as *Noether's theorem*.

The best situation to be in for a Hamiltonian system of dimension $2N$ is characterized by the Liouville-Arnold theorem:

Theorem 1 (Liouville-Arnold)**.** *Suppose that the functions $F_1, \ldots, F_N$ on $\mathbb{R}^{2N}$ satisfy*

$$\{F_j, F_k\} = 0, \qquad j, k = 1, \ldots, N$$

(we say that they are in involution). Pick $\vec{f}$ and let $M(\vec{f})$ be a common level set of the F_k:

$$M(\vec{f}) = \{\vec{u} \in \mathbb{R}^{2N} : F_k(\vec{u}) = f_k\,, \quad k = 1, \ldots, N\},$$

and suppose that the gradient vectors $\{\nabla F_1, \ldots, \nabla F_N\}$ are linearly independent at every point of $M(\vec{f})$. Then

(1) *$M(\vec{f})$ is a smooth manifold, invariant under the dynamics of the Hamiltonian system with Hamiltonian function $H = F_1$.*

(2) *If the manifold $M(\vec{f})$ is compact and connected, then it is diffeomorphic to the N-dimensional torus*

$$T^N := \{(\varphi_1, \ldots, \varphi_N) \mod 2\pi\mathbb{Z}^N\}.$$

(3) *The Hamiltonian system with Hamiltonian function H determines a globally linear motion on $M(\vec{f})$, that is, there exist coordinates $(\varphi_1, \dots, \varphi_N)$ on $M(\vec{f})$ such that $\mathrm{d}\vec{u}/\mathrm{d}t = \mathbf{J}\nabla_u H$ restricted to $M(\vec{f})$ is just*

$$\frac{\mathrm{d}\varphi_j}{\mathrm{d}t} = \omega_j(f_1, \dots, f_N), \qquad j = 1, \dots, N.$$

Such a system is said to be completely integrable.

In the case that $M(\vec{f})$ is compact and connected, the coordinates $(\varphi_1, \dots, \varphi_N)$ have the interpretation of angles, and the equations of motion can be integrated explicitly:

$$\varphi_j(t) = \varphi_j(0) + \omega_j(f_1, \dots, f_N)t \qquad \mod 2\pi, \quad j = 1, \dots, N,$$

which implies a quasiperiodic motion on the torus $M(\vec{f})$. Note the undistinguished role played by $H = F_1$ in the statement of the theorem. This implies that on the common level set $M(\vec{f})$ we may simultaneously solve all N systems $\mathrm{d}\vec{u}/\mathrm{d}t_k = \mathbf{J}\nabla F_k$, which all admit restriction to $M(\vec{f})$ and take the same form:

$$\frac{\mathrm{d}\varphi_j}{\mathrm{d}t_k} = \omega_{jk}(f_1, \dots, f_N), \qquad j, k = 1, \dots, N.$$

The simultaneous solution has the form

$$\varphi_j(t_1, \dots, t_N) = \varphi_j(0, \dots, 0) + \sum_{k=1}^{N} \omega_{jk}(f_1, \dots, f_N)t_k, \quad j = 1, \dots, N.$$

Infinite-dimensional Hamiltonian systems. In continuum-mechanical problems like those of wave propagation, we are concerned with the way that a function u of a continuous variable x evolves in time t. The Hamiltonian formalism generalizes easily to this context. The phase space now consists of functions $u(x)$ instead of finite-dimensional vectors. Functions on the phase space become *functionals* $F[u]$ returning a numerical value for each input function $u(\cdot)$. We restrict attention to functionals having the form

$$F[u] = \int_{-\infty}^{\infty} f(u, u_x, u_{xx}, \dots)\,\mathrm{d}x.$$

The component index j of the vector $\vec{u}$ with components u_j gets generalized to the continuous variable x, and in place of the gradient ∇F of a function we have the *variational derivative* of a functional written $\delta F/\delta u$ and calculated as

$$F[u + \epsilon v] = F[u] + \epsilon \int_{-\infty}^{\infty} \frac{\delta F}{\delta u}(x)v(x)\,\mathrm{d}x + O(\epsilon^2).$$

For example, if

$$F[u] := \int_{-\infty}^{\infty} \left[au(x)^3 + bu(x)u_{xx}(x)\right]\,\mathrm{d}x,$$

then integrating by parts twice for one of the terms to replace v_{xx} by v

$$\frac{\delta F}{\delta u}(x) = 3au(x)^2 + 2bu_{xx}(x).$$

For each value of x (continuum analogue of the index $j = 1, \dots, 2N$) we have a scalar-valued functional (continuum analogue of a function) of $u(\cdot)$ (continuum analogue of a point on the phase space $\mathbb{R}^{2N}$), so indeed $\delta F/\delta u(\cdot)$ is a continuum analogue of the gradient of a function F on the phase space. Finally, the matrix $\mathbf{J}(u)$ gets generalized to a linear operator $\mathcal{J}$ on the space of functions $u(x)$ such that the skew-symmetry property and Jacobi identity hold for the corresponding Poisson bracket, which is developed with respect to the inner product

$$\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)g(x)\,\mathrm{d}x.$$

So, for example, the KdV equation is a Hamiltonian system in which

$$\mathcal{J} := \frac{\partial}{\partial x},$$

That is, the KdV equation can be written in the form

$$u_t = \mathcal{J}\frac{\delta H}{\delta u} = \left(\frac{\delta H}{\delta u}\right)_x.$$

What is the Hamiltonian functional that works here? We need to find a functional H so that

$$\frac{\delta H}{\delta u} = -\frac{1}{2}u^2 - u_{xx}.$$

From our previous example, it is easy to see that we may take $a = -1/6$ and $b = -1/2$ to obtain

$$H[u] := -\int_{-\infty}^{\infty}\left[\frac{1}{6}u(x)^3 + \frac{1}{2}u(x)u_{xx}(x)\right]\,\mathrm{d}x.$$

Note that this gives us a third conserved local density (after u and u^2) since H is automatically a conserved quantity and it is of the form

$$H[u] = \int_{-\infty}^{\infty} D(u, u_x, \dots)\,\mathrm{d}x \qquad \text{with} \qquad D(u, u_x, \dots) := -\frac{1}{6}u^3 - \frac{1}{2}uu_{xx}.$$

Note that for the KdV equation, the Poisson bracket takes the form

$$\{F, G\} := \int_{-\infty}^{\infty}\frac{\delta F}{\delta u}(x)\frac{\mathrm{d}}{\mathrm{d}x}\frac{\delta G}{\delta u}(x)\,\mathrm{d}x.$$

So, for example if

$$F[u] := \int_{-\infty}^{\infty} u(x)^4\,\mathrm{d}x$$

and $G[u] = H[u]$, then since $\delta F/\delta u(x) = 4u(x)^3$, we have

$$\{F, G\} = \int_{-\infty}^{\infty}\left[4u(x)^3\frac{\mathrm{d}}{\mathrm{d}x}\left(-\frac{1}{2}u(x)^2 - u_{xx}(x)\right)\right]\,\mathrm{d}x = \int_{-\infty}^{\infty}\left[-4u(x)^4u_x(x) - 4u(x)^3u_{xxx}(x)\right]\,\mathrm{d}x$$

which again has the form of a functional defined on the phase space of suitable functions u. We might observe that if u decays rapidly as $x \to \pm\infty$ then the first term integrates to zero. The second term generally does not, so the Poisson bracket defines a nonzero functional on the phase space in this case. Clearly F is not conserved by the KdV equation because $\{F, H\} \neq 0$.

A good reference for Hamiltonian properties of soliton equations is the book *Hamiltonian Methods in the Theory of Solitons* by Faddeev and Takhtajan (Springer, 1987).

Could the KdV equation be a completely integrable system? The fact that the heights of the solitons in Zabusky and Kruskal's simulations were preserved throughout the collisions suggested to them that there were many conserved quantities in the KdV equation. Indeed, from looking at the numerical simulations of KdV one gets the idea that (symmetric functions of) the heights of the various solitons could be harvested as conserved quantities because:

- They can be measured from an initial condition (at least one in which the solitons are very far from one another).
- They are apparently constants of the motion (at least if one doesn't try to measure them during the complicated nonlinear interactions).

Since in principle an initial condition could be rigged that contains an arbitrary number of well-separated solitons, it starts to appear that maybe the KdV equation supports an infinite number of conserved quantities.

This led to a sudden, intense interest in the local conservation laws consistent with the KdV equation. By brute force, several more than the first three we have already seen were obtained (in other words, just by clever guessing). In fact, Zabusky and Kruskal themselves had found a few more, and later Robert Miura was able to find a total of ten local conservation laws consistent with KdV. Of course this is not enough to prove anything, but soon a general method was discovered, which now goes by the name of the *Gardner transform*:

$$u = w + \epsilon w_x + A\epsilon^2 w^2,$$

where ϵ and A are arbitrary parameters. Note that $\epsilon = 0$ gives the identity transformation. We want to substitute this into KdV

$$u_t + uu_x + u_{xxx} = 0$$

and find out what equation w satisfies. So, by direct calculation,

$$u_t = w_t + \epsilon w_{xt} + 2A\epsilon^2 w w_t = \left[1 + \epsilon\frac{\partial}{\partial x} + 2A\epsilon^2 w\right] w_t,$$

and

$$\begin{aligned} uu_x + u_{xxx} &= ww_x + w_{xxx} + \epsilon\left(ww_{xx} + w_x^2 + w_{xxxx}\right) + \epsilon^2\left(3Aw^2 w_x + 2Aww_{xxx} + (1+6A)w_x w_{xx}\right) \\ &\quad + \epsilon^3\left(2Aww_x^2 + Aw^2 w_{xx}\right) + 2A^2\epsilon^4 w^3 w_x \\ &= \left[1 + \epsilon\frac{\partial}{\partial x} + 2A\epsilon^2 w\right]\left[ww_x + w_{xxx} + A\epsilon^2 w^2 w_x\right] + \epsilon^2(1+6A)w_x w_{xx}. \end{aligned}$$

Now pick $A = -1/6$. Then

$$u_t + uu_x + u_{xxx} = \left[1 + \epsilon\frac{\partial}{\partial x} - \frac{\epsilon^2}{3}w\right]\left[w_t + ww_x + w_{xxx} - \frac{\epsilon^2}{6}w^2 w_x\right].$$

Hence, if w satisfies the *Gardner equation*

$$w_t + ww_x - \frac{\epsilon^2}{6}w^2 w_x + w_{xxx} = 0,$$

then $u = w + \epsilon w_x - \epsilon^2 w^2/6$ satisfies KdV (but not necessarily vice-versa). Note that the Gardner equation can also be written as a local conservation law:

$$w_t + \left[\frac{1}{2}w^2 - \frac{\epsilon^2}{18}w^3 + w_{xx}\right]_x = 0.$$

But furthermore, the Gardner transform may be inverted to express w in terms of u, at least at the level of a formal series:

$$w \sim \sum_{n=0}^{\infty} \epsilon^n w_n.$$

By comparing powers of ϵ in the Gardner transform, the coefficients w_n evidently have to satisfy

$$w_0 = u\,, \qquad w_1 = -w_{0,x} = -u_x\,, \qquad w_2 = -w_{1,x} + \frac{1}{6}w_0^2 = u_{xx} + \frac{1}{6}u^2,$$

and for $n \geq 3$,

$$w_n = -w_{n-1,x} + \frac{1}{3}uw_{n-2} + \frac{1}{6}\sum_{k=1}^{n-3} w_k w_{n-2-k}.$$

Substituting the formal expansion of w in powers of ϵ into the Gardner equation, we see by collecting powers of ϵ that what we really have is a *generating function for local conservation laws*. Indeed the generating function for conserved local densities is w itself:

$$D(\epsilon) = w \sim \sum_{n=0}^{\infty} \epsilon^n D_n$$

with

$$D_0 = u, \qquad D_1 = -u_x, \qquad D_2 = u_{xx} + \frac{1}{6}u^2,$$

and so on. The generating function for the corresponding fluxes is

$$F(\epsilon) = \frac{1}{2}w^2 - \frac{\epsilon^2}{18}w^3 + w_{xx} \sim \sum_{n=0}^{\infty} \epsilon^n F_n$$

with

$$F_0 = \frac{1}{2}w_0^2 + w_{0,xx} = \frac{1}{2}u^2 + u_{xx}, \qquad F_1 = w_0 w_1 + w_{1,xx} = -uu_x - u_{xxx},$$

and so on. Clearly, an arbitrary number of local conservation laws can be generated in this way. They are independent, in the sense that at each level, higher and higher powers of u are introduced. However, only half of them are nontrivial; note that whenever one has a local conservation law of the form

$$D_t + F_x = 0,$$

where additionally it turns out that D has the form

$$D = \frac{\partial E}{\partial x}$$

where E is a polynomial in u and its x derivatives, then applying the fundamental theorem of calculus to $D_t + F_x = 0$ gives as before that

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{-\infty}^{\infty} D \, \mathrm{d}x = 0,$$

but furthermore the integral itself evaluates to zero for fields u satisfying suitable boundary conditions. A nontrivial local conservation law is therefore one for which the conserved local density D is not a perfect derivative.

In any case, there is now clearly an infinite number of independent local conservation laws. The question now arises: is this enough to solve the KdV equation?

The breakthrough. The inverse-scattering transform. The mathematical breakthrough came in 1967, in work of Gardner, Greene, Kruskal, and Miura. The existence of an infinite number of local conservation laws for KdV suggested that it might be an infinite-dimensional analogue of a completely integrable Hamiltonian system. This would mean that the KdV equation could somehow be solved, in an analogue of obtaining and solving the elementary equations for the angular variables $(\varphi_1, \ldots, \varphi_N)$ on the torus $M(\vec{f})$.

There is another nonlinear equation, very similar in appearance to KdV, that can be solved exactly (this had been known for at least 50 years or so):

$$u_t + uu_x - 3u_{xx} = 0$$

which is called *Burgers' equation*. (The coefficient of 3 is not important; it can be scaled out; but we keep it to make a formal connection with KdV below.) It is not Hamiltonian. Indeed, the u_{xx} term represents diffusion rather than dispersion. This nonlinear equation can be solved because it can be linearized through the *Cole-Hopf transform*: one introduces a new dependent variable ψ related to u by the formula

$$u = -6\frac{\psi_x}{\psi}.$$

Substituting this into Burgers' equation one finds

$$\left(\frac{\psi_x}{\psi^2} - \frac{1}{\psi}\frac{\partial}{\partial x}\right)(\psi_t - 3\psi_{xx}) = 0$$

and therefore a solution of Burgers' equation is obtained whenever ψ satisfies

$$\psi_t - 3\psi_{xx} = 0$$

This is just the heat (or diffusion) equation. It is linear, and can be solved explicitly by Fourier transforms. The Cole-Hopf transform therefore *linearizes* Burgers' equation.

The story is told that Gardner, Greene, Kruskal, and Miura were looking for a generalization of the Cole-Hopf transform that might be applicable to the KdV equation. They were working on the premise that since KdV involves more derivatives, then so should the transform. So they tried

$$u = -6\frac{\psi_{xx}}{\psi}.$$

This did not work. So they thought to take advantage of the Galilean symmetry group of the KdV equation which allows one to add a constant to u at the cost of going into a moving frame of reference. Therefore they considered instead

$$u + \lambda = -6\frac{\psi_{xx}}{\psi}.$$

Then they started to think that maybe they should view this as an equation to be solved for ψ given u instead of the other way around. Thus they wrote on the blackboard

$$-6\psi_{xx} - u\psi = \lambda\psi. \tag{1}$$

They were staring at the equation (1) on the blackboard, when somebody walked by in the corridor and said: "Hey, are you guys doing quantum mechanics in there?" Of course the equation on the blackboard was just Schrödinger's equation from quantum mechanics describing the wavefunction ψ corresponding to a particle of energy λ in moving in the potential $V = -u$ (with $\hbar^2/(2m) = 6$).

Suppose that V decays rapidly to zero as $|x| \to \infty$. Then the *spectrum* of Schrödinger's equation is (by one definition) the collection of values of λ in the complex plane for which there is a solution $\psi(x)$ that is bounded uniformly for all x. Moreover, we divide the spectrum qualitatively into two parts:

- The *discrete spectrum* or *point spectrum* consists of values λ for which there exists a solution $\psi(x)$ that beyond being bounded in x actually decays to zero as $|x| \to \infty$ in such a sense that $\psi \in L^2(\mathbb{R})$, that is, $|\psi|^2$ is integrable over $\mathbb{R}$. For suitable $V(x)$ the discrete spectrum for the Schrödinger operator consists of a finite number N of negative real values of λ: $\lambda = -\kappa_n^2$, $n = 1, 2, \ldots, N$. The values of λ in the point spectrum are the *eigenvalues* of the Schrödinger operator with potential V. If $\lambda = -\kappa_n^2$ is an eigenvalue, and the corresponding $L^2(\mathbb{R})$ solution is called $\psi_n(x)$ (an *eigenfunction*) then we may assume $\psi_n(x)$ is a real-valued function, and then of course $K\psi_n(x)$ is another eigenfunction for any real K (Schrödinger's equation is linear in ψ), so by choosing K properly we may assume that ψ_n is *normalized*, *i.e.,*
$$\int_{-\infty}^{\infty} \psi_n(x)^2 \, \mathrm{d}x = 1.$$
The normalized eigenfunction has asymptotic behavior for large positive x given by
$$\psi_n(x) = c_n \mathrm{e}^{-\kappa_n x/\sqrt{6}}(1 + o(1)),$$
as $x \to +\infty$. The constant c_n is determined up to a sign (by the obvious symmetry $\psi \mapsto -\psi$ we assume without loss of generality that $c_n > 0$, making it unique), and is called the *norming constant* associated with the eigenfunction ψ_n or the eigenvalue $\lambda = -\kappa_n^2$.
- The *continuous spectrum* is (for our purposes) the rest of the spectrum. For suitable $V(x)$ decaying to zero with large x, the continuous spectrum for the Schrödinger operator consists of all nonnegative real $\lambda = k^2$, $k \in \mathbb{R}$. For each nonzero $k \in \mathbb{R}$, there is a solution of the Schrödinger equation that has the asymptotic form
$$\psi(x; k) = \mathrm{e}^{\mathrm{i}kx/\sqrt{6}} + r(k)\mathrm{e}^{-\mathrm{i}kx/\sqrt{6}} + o(1)$$
as $x \to -\infty$ and
$$\psi(x; k) = t(k)\mathrm{e}^{\mathrm{i}kx/\sqrt{6}} + o(1)$$
as $x \to +\infty$. Interpreting $\mathrm{e}^{\mathrm{i}kx/\sqrt{6}}$ as a wave propagating to the right and $\mathrm{e}^{-\mathrm{i}kx/\sqrt{6}}$ as a wave propagating to the left, it is natural to call $r(k)$ the *reflection coefficient* and $t(k)$ the *transmission coefficient* for the Schrödinger operator with potential V.

This gave the four an idea they didn't have before. Why not calculate the eigenvalues λ assuming that the potential $V = -u$ depends not only on x but on some external parameter t, such that $u(x, t)$ solves the KdV equation? The idea is that for each fixed t one has a potential function V and tries to find the eigenvalues λ, which in general depend on t through the parametric dependence in V. When $u(x, t)$ changes according to KdV, the eigenvalues should change in some fashion, and the aim was to find out how.

Here is the miracle they found: the eigenvalues did not change at all. As long as the t dependence of $u(x, t)$ is obtained from the KdV equation, the eigenvalues λ do not depend on t. Considered as a functional of u, each eigenvalue $\lambda = -\kappa_n^2$ is a constant of motion of the KdV equation! Moreover, they used the KdV equation to calculate how the associated norming constants c_n would vary as the parameter t is changed. They found
$$c_n(t) = c_n(0)e^{\alpha\kappa_n^3 t}, \quad \alpha := \frac{2}{3\sqrt{6}}.$$
Thus the norming constants *did* depend on t, but in a very simple and explicit way.

Excited by this result, they also calculated the time dependence of the reflection coefficient $r(k)$. They found

$$r(k;t) = r(k;0)\mathrm{e}^{-2\mathrm{i}\alpha k^3 t}$$

In particular, this means that $|r(k;t)|$ is independent of t for all real k. For each real k the modulus of the reflection coefficient, viewed as a functional of u, is a constant of motion of the KdV equation! The time dependence of $r(k;t)$, is, however, completely simple and explicit.

I'm not sure if I really believe that it happened this way, although it makes for a great story. I think it is just as likely that Gardner, Greene, Kruskal, and Miura arrived at the Schrödinger equation (1) by considering the Gardner transform $u = w + \epsilon w_x - \frac{1}{6}\epsilon^2 w^2$ as an equation to be solved for w given u. Thus, w satisfies a first-order quadratically-nonlinear differential equation; such differential equations are known as *Riccati equations*. It is classical, and certainly would have been known to the four researchers that a Riccati equation can be mapped to a second-order *linear* differential equation by a substitution of the form $w = k\phi_x/\phi$ for a suitable constant k. Thus,

$$u = w + \epsilon w_x - \tfrac{1}{6}\epsilon^2 w^2 \quad \text{becomes} \quad u = k\frac{\phi_x}{\phi} + \epsilon k\frac{\phi_{xx}\phi - \phi_x^2}{\phi^2} - \frac{\epsilon^2 k^2}{6}\frac{\phi_x^2}{\phi^2},$$

so choosing $k = -6/\epsilon$ the terms involving ϕ_x^2 cancel and after multiplying through by ϕ we are left with a linear equation:

$$-6\phi_{xx} - \frac{6}{\epsilon}\phi_x - u\phi = 0.$$

Finally, there is a standard trick for removing the first-order term from such an equation, namely via a substitution of the form $\phi = \psi \mathrm{e}^{\beta x}$ for some constant β to be determined. Since then $\phi_x = (\psi_x + \beta\psi)\mathrm{e}^{\beta x}$ and $\phi_{xx} = (\psi_{xx} + 2\beta\psi_x + \beta^2\psi)\mathrm{e}^{\beta x}$, substituting and cancelling $\mathrm{e}^{\beta x}$ gives

$$-6\psi_{xx} - \left(12\beta + \frac{6}{\epsilon}\right)\psi_x - \left(u + 6\beta^2 + \frac{6\beta}{\epsilon}\right)\psi = 0$$

so choosing $\beta = -1/(2\epsilon)$ yields (1) where $\lambda = -6\beta^2 = -3/(2\epsilon^2)$. Regarding the inspiration to look at how the spectrum evolves under KdV, it is likely that someone on the team was aware that the soliton solutions u of KdV were revealed by Kay and Moses in 1956 to be special cases of so-called *reflectionless potentials* for the Schrödinger equation and this was motivation enough to ask how the spectrum varies for more general solutions $u(x,t)$. But no matter how they arrived at the idea, it has to be appreciated as a remarkable thing.

The linear-in-t exponents $-2\alpha k^3 t$ and $\alpha\kappa_n^3 t$ should be viewed as being in complete analogy with the linear flow of the variables φ_n on the invariant simultaneous level set of all of the constants of motion of a completely integrable system. The machinery is all here now: an arbitrary number of conserved functionals (eigenvalues $\lambda = -\kappa_j^2$ and $|r(k)|$, viewed as functionals of u) and corresponding linear motion of phases (logarithms of norming constants and $\arg(r(k))$). What we have learned is that all of these objects have an additional *spectral* interpretation in terms of the Schrödinger operator with potential $V = -u$.

The only thing missing is an answer to the following question: if the eigenvalues $\lambda = -\kappa_n^2$ and corresponding norming constants $c_n(t)$ are known along with the reflection coefficient $r(k;t)$ at some time $t > 0$ according to the above formulæ (which only require that we can calculate them for $t = 0$, that is, from the given initial condition $u(x,0)$ for the KdV equation), can we somehow get back from them the corresponding potential $V = -u(x,t)$? If so, then we have found the solution of the initial-value problem for the KdV equation. The answer is affirmative, as the *spectral data* we know turns out to be exactly what is required to solve the problem of *inverse scattering*. That is, there is a procedure for determining the potential V of a Schrödinger operator given the discrete spectrum, norming constants, and reflection coefficient. This had been known in the physics literature since the 1950's, due to contributions by Gelfan'd, Levitan, and Marčenko. This procedure involves solving a linear integral equation.

These are then the steps for solving the KdV equation:

(1) Given the initial condition $u(x,0)$, consider $V = -u$ as a potential function in Schrödinger's equation, and calculate the discrete spectrum $\lambda = -\kappa_n^2$, associated norming constants $c_n = c_n(0)$, and for $k \in \mathbb{R} \setminus \{0\}$ the reflection coefficient $r(k) = r(k;0)$.

(2) Introduce explicitly dependence of these objects on t: the eigenvalues $\lambda = -\kappa_n^2$ remain fixed, while $c_n(t) = c_n(0)\mathrm{e}^{\alpha\kappa_n^3 t}$ and $r(k;t) = r(k;0)\mathrm{e}^{-2\mathrm{i}\alpha k^3 t}$.

(3) Carry out the procedure of inverse-scattering to determine $V = -u(x,t)$ from this "time-evolved" spectral data.

Step 1 is sometimes called the *direct scattering problem* and step 3 is the *inverse-scattering problem*. The whole method is called the *inverse-scattering transform* by direct analogy with the Fourier transform for solving linear problems.

Note that it turns out that there is no "change of variables" formula like the Cole-Hopf transform for linearizing the KdV equation. The KdV equation can indeed be linearized, but it is with the help of the spectral theory of the Schrödinger operator, and not by means of an explicit change of variables formula.