The Korteweg-De Vries Equation: Some Solution Methods

I. Reduction to one Independent Variable via a Lie Symmetry:

Some interesting simple solutions may be obtained by using the ordinary Lie symmetries of the equation to reduce the number of independent variables. There are four such symmetries, which we describe below with a labelling that will be described somewhat later:

$$\vec{v}^0_0 = \partial_x\,, \quad \vec{v}^0_1 = \partial_t\,, \quad \vec{v}^1_0 = t\partial_x - \partial_u\,, \quad \vec{v}^1_1 = x\partial_x + 3t\partial_t - 2u\partial_u\,. \tag{1.1}$$

For instance, we may use the first two of these symmetries, combined with a constant w in between, namely $-4w^2\partial_x + \partial_t$, to look for solutions of the form $u(x,t) = v(x + 4w^2 t) \equiv v(\xi)$, where the choice of the constant parameter between them, namely $4w^2$ was made because it simplifies the final form of the solution. Inserting this into our equation we obtain the following, which we then integrate once quite easily, where A is a constant of integration:

$$v''' + vv' - 4w^2 v' = 0 \quad \Longrightarrow v'' + \tfrac{1}{2}v^2 - wv = A\,. \tag{1.2}$$

Multiplying this equation by $2v'$ we may integrate once again to obtain

$$(v')^2 + v^3/3 - 4w^2 v^2 = 2Av + B\,. \tag{1.3}$$

At this point the general solution, since v' is the square root of a cubic polynomial, is just the Weierstrass elliptic function, $\wp(\xi + \delta)$, for some arbitrary constant δ, while the parameters of the Weierstrass function depend on our constants A and B. On the other hand, were we to insist that the solution and its first and second derivative should all vanish as we let ξ go toward infinity, then it would force A and B to vanish, which simplifies the solution to a hyperbolic function:

$$(v')^2 + v^3/3 - 4w^2 v^2 = 0 \implies \quad v = 12w^2\,\mathrm{sech}^2[w\xi + \delta] \text{ or } \; u(x,t) = 12w^2\,\mathrm{sech}^2[w(x + 4w^2 t) + \delta]\,. \tag{1.4a}$$

This is the "famous" 1-soliton solution of the KdV equation, which is a travelling wave that does not change its shape as it travels, even though the defining equation is both nonlinear and dispersive. What happens of course is the effects of those two "problems" cancels out in this case. Notice that even the constant 12 in front of the function is important in order for this to be desired solution. If we go somewhat more generally, and simply require that v be bounded as ξ goes to infinity, then we may re-write the cubic polynomial in our equation in terms of its roots:

$$v^3 - 12w^2 v^2 - 6Av - 6B = (v - r_1)(v - r_2)(v - r_3)\,, \qquad r_3 > r_2 > r_1\,, \tag{1.4b}$$

in which case the (more general) solution is given by

$$v = (r_3 - r_2)\,\mathrm{cn}^2(\lambda\xi + \delta) + r_2\,. \tag{1.4c}$$

Going now to a different symmetry, let us consider the reduction generated by the linear combination of symmetries involving two different ones, namely the symmetry $\vec{h} \equiv t\partial_x - \partial_u + (1/b)\partial_t$, where b is a constant. This time the invariants of this symmetry, i.e., the solutions of $\vec{h}f(x,t,u) = 0$ are given by

$$y = x - \tfrac{1}{2}bt^2\,, \quad v = u + bt\,, \tag{1.5a}$$

so that we should look for solutions for $v = v(y)$. This means we should insert into our equation the form $u = -bt + v(x - \frac{1}{2}bt^2)$. This gives us the following:

$$
\begin{aligned}
&u_x = v_y \; , \; u_{xxx} \; = v_{yyy} \; , \; u_t = -b - btv_y \; , \\
\Longrightarrow \quad & 0 = u_{xxx} + uu_x - u_t = v_{yyy} + (-bt + v)v_y + b + btv_y = v_{yyy} + vv_y + b \; .
\end{aligned}
\tag{1.5b}
$$

This may be easily integrated once:

$$
v_{yy} + \tfrac{1}{2}v^2 + by + c = 0 \; . \tag{1.5c}
$$

This is a fairly famous equation, with solution often referred to as *the first Painlevé transcendent*. The solutions are meromorphic in the entire complex plane, and are not expressible in terms of other, more standard functions. The entire set of 6 Painlevé transcendents are the only solutions to second-order ode's that have no "moveable" branch points and may not be expressed in terms of already-well-known functions, i.e., those which are solutions of second-order, linear ode's, or elliptic functions.

Lastly, let us note that the KdV equation has a scale invariance, which may be characterized by the following:

$$
(x, t, u) \Longrightarrow (\lambda x, \lambda^3 t, \lambda^{-2} u) \; , \tag{1.6a}
$$

which is induced by the most complicated of the symmetry vectors above, $\vec{v}_1^1$. Its invariants may be written in the form

$$
\begin{gathered}
y = -t^{-1/3}x \; , \quad v = t^{2/3}u \; , \\
u_x = -t^{-1}v_y \; , \quad u_{xxx} = -t^{-5/3}v_{yyy} \; , \quad u_t = -\tfrac{1}{3}t^{-5/3}(2v + yv_y) \; , \\
\Longrightarrow \quad 0 = -u_{xxx} - uu_x + u_t = t^{-5/3}\left[v_{yyy} + vv_y - \tfrac{2}{3}v - \tfrac{1}{3}yv_y\right] \; .
\end{gathered}
\tag{1.6b}
$$

The method toward the solution of this equation is not particularly clear; therefore, one may apply a so-called Miura transformation to it. We first define a new variable, w, that is a "potential" for v, and then determine the equation that w must satisfy, and then show how to resolve that equation:

$$
\begin{aligned}
v \; &= w_y - \tfrac{1}{6}w^2 \; , \\
0 &= v_{yyy} + vv_y - \tfrac{2}{3}v - \tfrac{1}{3}yv_y \\
&= w_{yyyy} - \tfrac{1}{3}[ww_{yyy} + w(w_y)^2] - \tfrac{1}{18}[3w^2w_{yy} - w^3w_y] - \tfrac{1}{9}[3yw_{yy} - yww_y] - \tfrac{2}{3}w_y + \tfrac{1}{9}w^2 \\
&\quad = (\partial_y - \tfrac{1}{3}w)(w_{yyy} - \tfrac{1}{6}(w^2 + 2y)w_y - \tfrac{1}{3}w = 0 \; .
\end{aligned}
\tag{1.6c}
$$

The last equation in this listing, the third-order one, may now be integrated once to give us

$$
w_{yy} = \tfrac{1}{18}w^3 + \tfrac{1}{3}yw + a \; , \tag{1.6d}
$$

where a is a constant of integration. This is now the equation satisfied by the *second Painlevé transcendent*, which is similar to, but somewhat more complicated than the first one.

II. Some Uses of Generalized Symmetries:

When using a symmetry of the type we have been considering, namely one which involves derivatives along t, x, and u, we also need to be able to understand how such a vector acts on some coordinate, in the jet bundle, such as u_x or u_{xx}. The original vector field lies in the tangent bundle to $J^0(M, N)$, where M denotes the space of independent variables, $\{x, t\}$ in this case, and N the space of dependent variables,

u in this case. In order, however, to describe the pde, we need to prolong the given vector field to higher order jet bundles, being careful that it maintains the relationship of these higher variables to one another and their antecedents back in $J^0(M,N)$. This is usually done via the *method of characteristics* which allows us to characterize a vector field in the tangent space to a jet bundle for some particular pde in terms of a single scalar function. It is usual to describe the symmetry vector $\tilde{v}$ in terms of an n-tuple of characteristic functions defined over some jet bundle, $J^\ell(M,N)$, which we label as F^α, where there is such a function for each dimension of the space of dependent variables:

$$
\begin{aligned}
\tilde{v} \;&\equiv\; v^b\partial_{x^b} + v^\alpha\partial_{z^\alpha} + v^\alpha_a\partial_{z^\alpha_a} + v^\alpha_{ab}\partial_{z^\alpha_{ab}} + v^\alpha_{abc}\partial_{z^\alpha_{abc}} + \;\dots \\
&= \sum_{\sigma=0}^{\infty}(D_\sigma F^\alpha)\partial_{z^\alpha_\sigma} + v^b\{\partial_b + z^\alpha_b\partial_\alpha + z^\alpha_{bc}\partial_{z^\alpha_c} + \dots\} \\
&= \sum_{\sigma=0}^{\infty}(D_\sigma F^\alpha)\partial_{z^\alpha_\sigma} + v^b D_b\,; \qquad\qquad v^b\delta^\alpha_\beta = -F^\alpha_{,z^\beta_b}\;.
\end{aligned}
\tag{2.1}
$$

Here the first line simply gives names to the various coefficients in an arbitrary vector field over $J^\infty(M,N)$. The third line gives us a definition of what is called *the total derivative*, D_a, over that jet bundle. Lastly we should note that the summation indicated by σ is over all possible powers of derivatives, with respect to all variables, including the possibility of no derivative at all, in which case D_0 is just the identity operator.

If we view our k-th order pde on some higher jet, at least $J^{k+1}(M,N)$, we see that $\forall\ i, D_i$ is a symmetry of the pde since D_iF^A is already “a part” of (the prolongation of) the pde. However, of course this is a valid statement about **any** pde, and therefore not very interesting! Such symmetries are called trivial and we will try to ignore trivial symmetries, which means that we will consider equivalent any two symmetries if their difference is a linear combination of trivial ones. Therefore, the general section-preserving symmetry, in Eq. (2.1), is equivalent to the one without the extra term, v^iD_i!

Symmetries for which $v^i \equiv 0$ are referred to as evolutionary symmetries and we may see that every (section-preserving) symmetry is equivalent to an evolutionary one, written simply as

$$
\tilde{v} = \sum_\sigma (D_\sigma F^\alpha)\partial_{z^\alpha_\sigma} \equiv \text{З}_F \quad , \tag{2.2}
$$

where the notation with the Russian letter, З, is intended to emphasize that this is an evolutionary symmetry. If one wants to return to the equivalent symmetry where $v^i \neq 0$, since we of course have the characteristic F^α, we may calculate from $v^i\delta^\alpha_\beta = -F^\alpha_{,z^\beta_i}$ and proceed!

On the other hand, it should definitely be noted that З_F usually has infinitely many components and is therefore not as nice to work with, even though the formula looks simpler. For instance, for a simple translation symmetry in the independent variable x, we would have

$$
\tilde{v} = -\partial_x \;\Rightarrow F = z_x\,, \;\Longrightarrow \text{З}_{z_x} = z_x\partial_z + z_{xx}\partial_{z_x} + z_{xt}\partial_{z_t} + \dots \;=\; D_x - \partial_x \quad .
$$

We resolve most of this difficulty by agreeing to talk about evolutionary symmetries principally in terms of their characteristics, which has the advantages to be described now. In order to do this we need two things. We need a method of determining the commutators via the characteristics, and, even more importantly, we need a method of determining **directly** the appropriate characteristics for a given pde—the characteristics of its symmetries. We first consider the determination of commutators directly from their characteristics: if $\mathfrak{Z}_\eta$ and $\mathfrak{Z}_\psi$ are two evolutionary symmetries, then their commutator should of course again be an evolutionary vector field for some characteristic χ:

$$[\mathfrak{Z}_\eta, \mathfrak{Z}_\psi] = \left[\sum_\sigma (D_\sigma \eta)\partial_{z_\sigma}, \sum_\tau (D_\tau \psi)\partial_{z_\tau}\right] \equiv \mathfrak{Z}_\chi \quad , \tag{2.3}$$

where χ is defined by

$$\chi^\alpha \equiv \{\eta, \psi\}^\alpha \equiv \mathfrak{Z}_\eta(\psi^\alpha) - \mathfrak{Z}_\psi(\eta^\alpha) = \sum_\sigma \left\{ D_\sigma(\eta^\beta)\psi^\alpha_{,z^\beta_\sigma} - D_\sigma(\psi^\beta)\eta^\alpha_{,z^\beta_\sigma} \right\} \quad . \tag{2.4}$$

This gives a reasonably simple map between all evolutionary vector fields and the space of n-tuples of functions that maps the Lie bracket of the fields into this Poisson-like bracket of the n-tuples of functions.

We may then use this mechanism to determine generalized symmetries, i.e., symmetries other than simple prolongations of Lie symmetries to higher order jets. To describe how to determine, and define, those symmetries we must first note that a single k-th order pde may be written in terms of a choice of coordinates on $J^k(M, N)$. Doing this defines a 1-codimensional variety on that bundle which may be thought of as the "hyperplane" which is the solution manifold for our pde; describe that single constraint on the jet variables by the function $Q = 0$. We need appropriate coordinates on that variety, and acquire them most easily by simply resolving that function Q for one of the original coordinates, and using it to re-define that coordinate as a function of the others. At that time we may refer to it as a co-coordinate. For instance, for the KdV equation, it is most plausible to solve it for u_t, leaving the variables $\{t, x, u, u_x, u_{xx}, u_{xxx}\}$ as coordinates on this solution manifold. We will denote this definition of the co-coordinate by putting an "overtilde" on the coordinate: $u_t = \widetilde{u}_t(t, x, u, u_x, u_{xx}, u_{xxx})$. When the pde is prolonged to higher jet variables, we must remember to keep this relationship valid. This means we will be required to look at all the original jet-bundle coordinates involving at least one t-derivative as now co-coordinates with their relationship to the coordinates determined by taking the appropriate restricted total derivative of the functional relationship for $\widetilde{u}_t$; for example, $\widetilde{u}_{xxt} = \bar{D}_x^2 \widetilde{u}_t$. Here the restricted total derivative, also called the pullback of the total derivative, is the total derivative restricted to our variety which is the "solution manifold." Again, for the KdV equation this would mean that u_t, u_{xt}, etc. are no longer coordinates so that, for instance we would have

$$\begin{aligned}
\bar{D}_x &= \partial_x + u_x\partial_u + u_{xx}\partial_{u_x} + u_{xxx}\partial_{u_{xx}} + \ \ldots \quad , \\
\bar{D}_t &= \partial_t + \widetilde{u}_t\partial_u + \widetilde{u}_{xt}\partial_{u_x} + \widetilde{u}_{xxt}\partial_{u_{xx}} + \ \ldots \quad , \\
\widetilde{u}_{xt} &= \bar{D}_x\widetilde{u}_t \ , \quad \widetilde{u}_{xxt} = \bar{D}_x^2\widetilde{u}_t \ , \quad \ldots \ .
\end{aligned} \tag{2.5}$$

The most useful mechanism to do that involves the (Vinogradov) universal linearization operator, also referred to as the Frechet linearisation operator. A generalized symmetry operator for a pde Q is determined by a characteristic function, F, that satisfies the equation

$$\bar{\mathfrak{Z}}_Q(F) \equiv \bar{L}_Q(F) = 0 \ . \tag{2.6}$$

Returning to the KdV equation, one finds that there are countably many solutions of this equation. A basis for these characteristics is then described by

$$\begin{aligned}
\psi^i_\ell &= t^i\, u_{(2\ell+1)} + O(2\ell - 1) \ , \quad i = 0, 1 \quad \ell = 0, 1, 2, 3, \ldots \ , \\
\{\psi^i_\ell, \psi^j_m\} \ &= \frac{1}{3}[(2\ell + 1)j - (2m + 1)i]\, \psi^{i+j-1}_{\ell+m-1} \ .
\end{aligned} \tag{2.7}$$

where $u_{(m)}$ denotes the coordinate which is the m-th derivative with respect to x of u when restricted down to a solution. The first several solutions are as follows, where the characteristics for the Lie symmetries are also noted:

$$
\begin{aligned}
\psi_0^0 &= u_x \ , \quad \psi_1^0 = u_{xxx} + uu_x \ , \\
\psi_0^1 &= tu_x + 1 \ , \quad \psi_1^1 = t(u_{xxx} + uu_x) + \tfrac{1}{3}xu_x + \tfrac{2}{3}u \ , \\
\psi_2^0 &= u_{(5)} + \frac{5}{3}uu_{(3)} + \frac{10}{3}u_x u_{xx} + \frac{5}{6}u^2 z_x \quad , \\
\psi_3^0 &= u_{(7)} + \frac{7}{3}uu_{(5)} + 7u_{(1)}u_{(4)} + \frac{35}{3}u_{(2)}u_{(3)} + \frac{35}{18}u^2 u_{(3)} \\
&\quad + \frac{70}{9}uu_x u_{xx} + \frac{35}{18}u_x^3 + \frac{35}{54}u^3 u_x \quad ,
\end{aligned}
\tag{2.8}
$$

A large number of these are written down in some other handout. They may be generated, for instance, by the recursion operator that is also given there. However, it is worth making a few comments about them. It may be seen from the commutator equation that the entire set $\{\psi_\ell^0 \mid \ell = 0, 1, 2, \ldots\}$ commutes with one another. Therefore one may construct additional pde's with flows that commute with each other. We create such things by simply considering the pde's

$$
u_t = \psi_\ell^0 = \psi_\ell^0(u_{(2\ell+1)}, u_{(2\ell-1)}, u_{(2\ell-2)}, \ldots,\ u_x, u) \ . \tag{2.9a}
$$

This countable set of pde's is usually referred to as the KdV hierarchy. As already noted they commute; therefore, one may attempt to find simultaneous solutions. [I note that since they commute, one may also think of the flow parameter along each one as a different quantity, so that sometimes one invents a countable family of parameters, $\{t_\ell\}_0^\infty$ and re-writes Eqs. (2.9a) in the form

$$
u_{t_\ell} = \psi_\ell^0 = \psi_\ell^0(u_{(2\ell+1)}, u_{(2\ell-1)}, u_{(2\ell-2)}, \ldots,\ u_x, u) \ , \tag{2.9b}
$$

so that a simultaneous solution of all the equations up to and including some particular value of ℓ may be described as depending on that many constants.

This is exactly how one finds the n-soliton solutions for $n > 1$. In particular the 1-soliton solution is obtained by insisting on a solution of the equation that also satisfies a relationship between the characteristics for the symmetries. The lowest-order version of this relationship is just that $\psi_1^0 = v\,\psi_0^0$, where v is the speed of that soliton. This is hardly surprising since $\psi_0^0 = u_x$ and $\psi_1^0 = -u_t$ for a solution of the equation itself, so that this constraint is equivalent to $u_t + vu_x = 0$, which says that the solution should be a function of the one combination of dependent variables, $x - vt$. Nonetheless, it is indeed quite surprising when one learns that this 1-soliton solution has many, related properties, and that this velocity-relationship continues for higher numbers of solitons. The 1-soliton solution is discussed in more detail in Appendix II, where it is noted that

$$
\psi_j^0\Big|_{1-\text{soliton sol'n}} = v^j\, \psi_0^0\Big|_{1-\text{soliton sol'n}} \quad . \tag{2.9c}
$$

However, there are still other interesting properties they have. Each of them is a perfect derivative, which allows us to take each of the equations above and write it as **a conservation law,** i.e., in the form of an equation with a derivative of u on one side and a different derivative, of the same u, on the other side. The first example, of course, is the original KdV equation itself, which may be written in the form

$$
\bar{D}_t u = u_t = u_{xxx} + uu_x = (u_{xx} + u^2/2)_{,x} = \bar{D}_x(u_{xx} + u^2/2) \ . \tag{2.10}
$$

Obviously we could insert a simple potential v, such that $u = v_x$ into this equation, remove some derivatives and obtain an equation that v must satisfy, which is, however, no easier to solve than the original one:

$$v_{xxx} + \tfrac{1}{2}(v_x)^2 - v_t = 0 \ , \tag{2.11}$$

which is usually referred to as **the KdV-potential, or pKdV, equation**. However this process of potentialization can also be continued infinitely often, since each of the characteristics for our higher symmetries may be written as a perfect derivative, so that we have

$$u_{t_{2j+1}} = \psi_j^0(u) \equiv \bar{D}_x\{\chi_j^0(u)\} \ , \tag{2.12}$$

The first few of them have the following explicit form as derivatives:

$$\begin{aligned}
\psi_0^0 &= \bar{D}_x(u) \quad , \\
\psi_1^0 &= \bar{D}_x(u_{xx} + \hat{u}^2) \quad , \\
\psi_2^0 &= \bar{D}_x(u_{(4)} + \frac{5}{3}uu_{(2)} + \frac{5}{6}u_{(2)}^2 + \frac{5}{18}u^3) \quad , \\
\psi_3^0 &= \bar{D}_x(u_{(6)} + \frac{7}{3}uu_{(4)} + \frac{14}{3}u_{(1)}u_{(3)} + \frac{7}{2}u_{(2)}^2 + \frac{35}{18}u^2u_{(2)} + \frac{35}{18}uu_{(1)}^2 + \frac{35}{216}u^4) \quad .
\end{aligned} \tag{2.13}$$

This allows each of the members of the hierarchy of equations, as written in Eqs. (2.9), to be considered as a conservation law; i.e., they have the form of the time derivative of something equals the space derivative of something else, just the usual form in electromagnetic theory which says that $\rho_{,t} + \nabla \cdot \vec{J} = 0$. We may therefore, for each of them associate a potential for u, defined via $u \equiv v^j_{,x}$, so that the equations take the form

$$v^j_{,t_{2j+1}} = \chi_j^0(u)\Big|_{u=v^j_{,x}} \ , \tag{2.14}$$

thereby providing a related hierarchy associated with the pKdV equation.

Since one can do this once, one might then hope, since all these equations are indeed compatible, to obtain **a single potential** to define them all. In fact one such quantity does exist, usually referred to as a τ-function, and ascribed to Sato, although Hirota had a great deal to do with it as well:

$$u = 12\partial_x\left(\frac{\tau_{,x}}{\tau}\right) = 12(\partial_x)^2 \ln \tau \ . \tag{2.15}$$

It turns out that this provides a much cleaner way to construct the n soliton solutions. This approach is discussed in another handout about the KdV equation, and gives explicit forms for solutions with arbitrary numbers of solitons.

III. Bäcklund transformations for the KdV equation:

Although I want to talk in detail about the Bäcklund transformation for the KdV equation, there is some value in first describing such a transformation for the sine-Gordon equation, where it takes on a much simpler form. That equation is given by either of the following two equivalent forms:

$$u_{xy} = \sin u = u_{zz} - u_{tt} \ , \qquad z \equiv x + y \ , \ t \equiv x - y \ . \tag{3.1}$$

A Bäcklund transformation is basically a method to "trade-in" some particular solution for a given pde for a solution of a different pde, which could indeed even be the same one. More precisely stated it constitutes a map from the $(k-1)$-st jet bundle for some given pde, for $u \in N$, to the first jet bundle for some other pde in the same independent variables, but different dependent ones, i.e., for $v \in N'$:

$$
\begin{aligned}
&\Upsilon : J^{k-1}(M,N) \times N' \;\rightarrow\; J^1(M,N') \;, \\
&\Upsilon : \begin{cases} x^a = x^a \;,\; z^\mu \;= z^\mu \;,\; z'^A = z'^A \;, \\ \qquad\qquad z'^A_a = Z^A_a(x,z,z_\sigma,z') \;. \end{cases}
\end{aligned}
\tag{3.2}
$$

One usually can map quite trivial solutions of the first pde into interesting solutions of the second; often, then, one can insert that new one and obtain yet newer ones, in a continuing sequence. When the two pde's are actually the same, it is often referred to as an auto-Bäcklund transformation. That this map goes into J^1 means that the explicit form of the transformation equations are only first-order pde's in the new dependent variable so that it is usually fairly straightforward to solve them if the solution for the other one is given.

To proceed with the example for the sine-Gordon equation, we first begin with two unknown functions, u and v, of two independent variables, x and y, which are required to satisfy the following first-order pde's:

$$
u_x + v_x = 2\lambda \sin[(u-v)/2] \;, \qquad u_y - v_y = \tfrac{2}{\lambda}\sin[(u+v)/2] \;. \tag{3.3}
$$

We act on the first one by ∂_y and the second by ∂_x, which gives us

$$
\begin{aligned}
u_{xy} + v_{xy} &= \lambda \cos[(u-v)/2](u_y - v_y) = 2\cos[(u-v)/2]\sin[(u+v)/2] = \sin u + \sin v \;, \\
u_{yx} - v_{yx} &= \tfrac{1}{\lambda}\cos[(u+v)/2](u_x + v_x) = 2\cos[(u+v)/2]\sin[(u-v)/2] = \sin u - \sin v \;.
\end{aligned}
\tag{3.4}
$$

Adding and subtracting we see that both u and v are required to satisfy the sine-Gordon equation, $u_{xy} = \sin u$, in order for these equations to both be able to be satisfied at once. Therefore, given a solution u of the sine-Gordon, a solution of this pair of first-order equations will provide another solution v of the sine-Gordon equation, which, moreover, now contains an arbitrary constant, λ.

A simple approach, therefore, is to note that $u = 0$ is clearly a solution of the sine-Gordon equation. We insert this value into the first-order equations for v, and integrate them easily, to obtain the following non-trivial solution:

$$
\begin{aligned}
&\log\tan(v/4) = \;-\,(\lambda x + y/\lambda) \equiv -\eta \;\Longrightarrow\; \quad v = 4\,\mathrm{Tan}^{-1}(e^{-\eta}) \;, \\
&\qquad\qquad \Longrightarrow \quad \text{moves with velocity } \;\frac{1/\lambda - \lambda}{1/\lambda + \lambda} \;.
\end{aligned}
\tag{3.5}
$$

To understand the statement about the velocity, recall that x and y are null coordinates, so the solution in the more traditional Cartesian coordinates, $z = x + y$ and $t = x - y$, has the argument in the form $\eta = \lambda(z+t)/2 + \lambda^{-1}(z-t)/2 = \frac{1}{2}[(\lambda + 1/\lambda)z - (1/\lambda - \lambda)t]$. This is the famous 1-soliton solution

We now proceed to describe the Bäcklund transformation for the KdV equation, which was originally determined by Estabrook and Wahlquist. This is a transformation between the KdV equation and the pKdV equation. The Estabrook Wahlquist method involves the determination of a prolongation into the second jet space. More precisely we recall that the commutator of any two total derivative operators is zero, such as $[\bar{D}_x, \bar{D}_y] = 0$, where the total derivative operator is the prolongation of the ordinary coordinate derivatives up to $J^\infty(M,N)$. On the other hand we now want to prolong further, into $J^1(M, N')$; therefore, we append to these derivative operators an additional part that constitutes a vector field over N':

$$\left.\begin{aligned} D_x &\to D_x + \vec{F} \\ D_t &\to D_t + \vec{G} \end{aligned}\right\} \qquad [D_x + \vec{F}, D_t + \vec{G}] \neq 0 \ , \tag{3.6}$$

where the commutator vanishes only when both pde's are satisfied, i.e., when it is restricted to both varieties at once. Since, for instance, $\vec{F}$ is a vector field over N', it will have the form $\vec{F} = F^\mu \partial_{w^\mu}$, where the w^μ are the coordinates in N', i.e., the new dependent variables for the new pde's. The E-W procedure determines these prolongations in terms a set of elements of a vector space with a Lie product, realized by the commutator of vector fields. As well it determines the value for some of the Lie products in question, but not all of them. Such a vector space is called a semi-closed but otherwise free Lie algebra. If one then determines some particular vector-field realization of a closure of that Lie algebra, in terms of some number of new dependent variables, w^μ, then this is what determines the desired Bäcklund transformation.

Without giving the derivation, which is in fact given in the long paper online, we state that the result for the KdV is that there should be four vector fields, $\{\vec{\alpha}, \vec{\beta}, \vec{\gamma}, \vec{\delta}\}$, which must satisfy a set of constraints on their commutators, i.e., their Lie products:

$$\begin{aligned} [\vec{\alpha}, \vec{\beta}] &= 0 = [\vec{\alpha}, \vec{\gamma}] = [\vec{\beta}, [\vec{\beta}, [\vec{\beta}, \vec{\gamma}]]] = [\vec{\gamma}, \vec{\delta}] \ , \\ [\vec{\alpha}, \vec{\delta}] &- [\vec{\beta}, \vec{\gamma}] - 3[\vec{\beta}, [\vec{\gamma}, [\vec{\beta}, \vec{\gamma}]]] = 0 = [\vec{\beta}, \vec{\delta}] + [\vec{\gamma}, [\vec{\gamma}, [\vec{\beta}, \vec{\gamma}]]] \ ; \\ \vec{F} &= -\tfrac{1}{2}\vec{\alpha} u^2 - \vec{\beta} u - \vec{\gamma} \quad , \\ \vec{G} &= -(\vec{\alpha} z + \vec{\beta}) u_{xx} + \tfrac{1}{2}\vec{\alpha}(u_x)^2 + [\vec{\beta}, \vec{\gamma}] u_x - \tfrac{1}{3}\vec{\alpha} u^3 \\ &\qquad - \tfrac{1}{2}\{\vec{\beta} - [\vec{\beta}, [\vec{\beta}, \vec{\gamma}]]\} u^2 - [\vec{\gamma}, [\vec{\gamma}, \vec{\beta}]] u - \vec{\delta} \quad . \end{aligned} \tag{3.7}$$

When one determines a set of vector field representations for these 4 fields, satisfying the relationships in Eqs. (3.7), then $\vec{F}$ is a prolongation of ∂_x into the tangent bundle over N' and therefore has the form

$$\vec{F} = F^\mu_x \partial_{w^\mu} = w^\mu_x \partial_{w^\mu} \ , \qquad \vec{G} = G^\mu_t \partial_{w^\mu} = w^\mu_t \partial_{w^\mu} \ . \tag{3.8}$$

The form of these coefficients, as functions of the appropriate variables, then constitutes the desired Bäcklund transformations. Notice that the number of new dependent variables, w^μ, is not determined a priori; that number is determined in terms of the representation one finds for the partial Lie algebra that is required. We also note that the integrability conditions on this set are already satisfied—by virtue of the KdV equation—so such w^μ's will exist.

There are various ways one can attempt to determine such a representation; however, they can basically be characterized in only two really different ways: the first is to "guess" functions of some number of w^μ's satisfying these relations. The second is to create an algebra satisfying only these relations and use that knowledge to determine what functions of w^μ the F^μ and G^μ are. In either case, having determined how many w^μ's one will consider, and what functions they are, we then have found our "potentials." (Since they were found by this method whereby they depend on themselves, they are actually called *pseudopotentials*.)

For a first example, I make the simplest choice for these functions. Although it is the simplest, it nonetheless will demonstrate several interesting things one may do with them! We first decide to look only for one such pseudopotential, so that the number of w^μ's is simply 1. The simplest realization in one dimension is then given by the following homomorphism:

$$
\begin{aligned}
&\vec{\alpha} &&\to 0 \quad , \\
&\vec{\beta} &&\to \hat{\partial}_w \Longleftarrow \quad \tfrac{1}{2}\vec{\mathbf{f}} \, , \\
&\vec{\gamma} &&\to \tfrac{1}{3}w^2\partial_w \Longleftarrow \quad -\tfrac{1}{3}\vec{\mathbf{e}} \, , \\
&\tilde{\delta} &&\to 0 \quad , \\
\Longrightarrow \quad &[\vec{\gamma},\vec{\beta}] &&\to -\tfrac{1}{3}w\partial_w \Longleftarrow \quad -\tfrac{1}{6}\vec{\mathbf{h}} \, ,
\end{aligned}
\tag{3.9}
$$

where $\{\vec{\mathbf{h}},\vec{\mathbf{e}},\vec{\mathbf{f}}\}$ are (a 1-dimensional) vector-field realization of the standard Chevalley generators for $\boldsymbol{sl}(2)$. To make this perhaps a little more familiar, note that if we make the following new names,

$$
\begin{aligned}
\mathbf{H} &\equiv 6[\vec{\gamma},\vec{\beta}] &&\to -2w\partial_w \quad , \\
\mathbf{E}_+ &\equiv -2\vec{\beta} &&\to -\partial_w \quad , \\
\mathbf{E}_- &\equiv +3\vec{\gamma} &&\to w^2\partial_w ,
\end{aligned}
\tag{3.10a}
$$

then the resulting objects have rather simple commutation relations:

$$
\begin{aligned}
[\mathbf{H}\,,\mathbf{E}_\pm] &= \pm 2\,\mathbf{E}_\pm \quad , \\
[\mathbf{E}_+\,,\mathbf{E}_-] &= \mathbf{H} \quad ,
\end{aligned}
\tag{3.10b}
$$

analogous to the commutation relations for i times the $\hat{z}$- and $\pm$-components of ordinary angular momentum. This allows us to see yet another (interesting) role that the rotation group plays.

Inserting these values into Eq. (4.41), we obtain the following explicit Bäcklund transformation, between the KdV and another pde:

$$
\begin{aligned}
\Rightarrow \quad w_{,x} = F^1 &= -\tfrac{1}{3}w^2 - \tfrac{1}{2}u \Longleftarrow \quad -\tfrac{1}{2}\widetilde{\mathbf{f}}u + \tfrac{1}{3}\vec{\mathbf{e}} \, , \\
w_{,t} = G^1 &= -\tfrac{1}{2}u_{xx} + \tfrac{1}{3}u_x w - \tfrac{1}{6}u^2 - \tfrac{1}{9}w^2 u \Longleftarrow \quad -\tfrac{1}{2}\vec{\mathbf{f}}u_{xx} + \tfrac{1}{6}\vec{\mathbf{h}}u_x - \tfrac{1}{6}\vec{\mathbf{f}}u^2 + \tfrac{1}{9}\vec{\mathbf{e}}u \, .
\end{aligned}
\tag{3.11}
$$

The integrability condition for the existence of a simultaneous solution to this pair of equations is of course just the anticipated KdV equation:

$$
w_{,xt} - w_{,tx} = \tfrac{1}{2}\left\{u_{xxx} + uu_x - u_t\right\} \quad , \tag{3.12}
$$

as was expected. However, one may also eliminate the z's from Eqs. (3.10). This is done simply by resolving the $w_{,x}$ equation which gives us the (new) pde for the new dependent variable, w:

$$
w_{xxx} - \tfrac{2}{3}w^2 w_x - w_t = 0 \quad , \tag{3.13}
$$

which is referred to as the "modified KdV" equation, or mKdV equation.

A more complicated, but more interesting and useful, realization is found by modifying this representation only slightly, with the insertion of a constant parameter, λ:

$$
\begin{aligned}
&\vec{\alpha} &&\to 0 \quad , \\
&\vec{\beta} &&\to \tfrac{1}{2}\partial_w \Longleftarrow \quad \tfrac{1}{2}\vec{\mathbf{f}}; , \\
&\vec{\gamma} &&\to \tfrac{1}{3}(w^2-\lambda)\partial_w \Longleftarrow \quad -\tfrac{1}{3}(\vec{\mathbf{e}}+\lambda\vec{\mathbf{f}}) \, , \\
&\vec{\delta} &&\to \tfrac{4}{9}\lambda\vec{\gamma} \Longleftarrow \quad -\tfrac{4}{27}(\lambda\vec{\mathbf{e}}+\lambda^2\vec{\mathbf{f}}) \, , \\
\Longrightarrow \quad &[\vec{\gamma},\vec{\beta}] &&\to -\tfrac{1}{3}w\partial_w \Longleftarrow \quad -\tfrac{1}{6}\vec{\mathbf{h}} \, ,
\end{aligned}
\tag{3.14}
$$

which then transforms a solution into a family of solutions, parameterized by λ. In this case the new formulation for the generators $\widetilde{\mathbf{F}}$ and $\widetilde{\mathbf{G}}$ is

$$\begin{aligned}
\vec{\mathbf{F}} &= -\tfrac{1}{2}\vec{\mathbf{f}}u + \tfrac{1}{3}(\vec{\mathbf{e}} + \lambda\vec{\mathbf{f}}) , \\
\vec{\mathbf{G}} &= -\tfrac{1}{2}\vec{\mathbf{f}}u_{xx} + \tfrac{1}{6}\vec{\mathbf{h}}u_x - \tfrac{1}{6}\vec{\mathbf{f}}u^2 - \tfrac{1}{9}(-\vec{\mathbf{e}} + \lambda\vec{\mathbf{f}})u + \tfrac{4}{27}\lambda(\vec{\mathbf{e}} + \lambda\vec{\mathbf{f}}) . \\
\Longrightarrow \quad w_{,x} &= F^1 = -\tfrac{1}{3}(w^2 - \lambda) - \tfrac{1}{2}u , \\
w_{,t} &= G^1 = -\tfrac{1}{2}u_{xx} + \tfrac{1}{3}u_x w - \tfrac{1}{6}u^2 - \tfrac{1}{9}(w^2 + \lambda)u - \tfrac{4}{27}\lambda(w^2 - \lambda) .
\end{aligned} \tag{3.15a}$$

As before, if we insist that $w_{,xt} = w_{,tx}$ this gives us the straightforward requirement that the associated u must satisfy the KdV equation. Contrariwise, if we resolve the $w_{,x}$ equation for u and insert that into the $w_{,t}$ equation, then we find that w now must satisfy an extended version of the mKdV equation, although we can see that it is equivalent to the original one if viewed from a moving frame:

$$\begin{aligned}
&-w_t + w_{xxx} - \tfrac{2}{3}(w^2 - \lambda)w_x = 0 \\
&\quad \Longrightarrow -h_t + h_{xxx} - \tfrac{2}{3}h^2 h_x = 0 \text{ with } h(x,t) \equiv w(x, t - \tfrac{2}{3}\lambda x) .
\end{aligned} \tag{3.15b}$$

Another relevant consequence of this form is to re-consider the equation for w_x in Eqs. (3.15a), which is of the form of a Ricatti equation. The standard approach for such equations is to linearize them via the transformation

$$\psi_{xx} + \tfrac{1}{6}(u - \tfrac{2}{3}\lambda)\psi = 0 , \qquad w \equiv 3(w_x/w) . \tag{3.15c}$$

This gives us the time-independent Schrödinger equation, with energy proportional to λ and potential energy proportional to u. Since $u = u(x,t)$, one could consider the question as to whether the associated eigenvalues, i.e., the values of λ might also depend on the t, which would just play the role of a parameter in this equation. As it turns out, the **requirement** that λ are independent of the parameter t in $u(x,t)$ is exactly the requirement that u satisfy the KdV equation.

At this point, we push forward with these transformations, to acquire the celebrated *nonlinear superposition principle* for the KdV equation, noting that some other nonlinear, integrable equations do also have such relations. We first note that if w is a solution of the mKdV equation, then $-w$ is also a solution; using the Bäcklund transformations above, and inserting these two distinct solutions of the mKdV equation, we may generate two distinct solutions of the regular KdV equation:

$$u_1 \equiv \lambda + w_x - \tfrac{1}{6}w^2 , \qquad u_2 \equiv \lambda - w_x - \tfrac{1}{6}w^2 . \tag{3.16}$$

We next insert the potential, v for u, discussed for the pKdV equation, near Eq. (2.11), and then subtract and add these two equations, integrating the one that has the difference of $u_1 - u_2 = (v_1 - v_2)_{,x}$, to obtain the following pair:

$$v_1 - v_2 = -4w , \qquad (v_1 + v_2)_{,x} = -\tfrac{4}{3}(w^2 - \lambda) . \tag{3.17}$$

We next insert this form for w into the mKdV equation that it must satisfy, but carefully replacing the quadratic term, $w^2 - \lambda$ with its form in terms of v_1 and v_2 from the second of Eqs. (3.17). This results in the following:

$$(v_1 - v_2)_{,t} = (v_1 - v_2)_{,xxx} - \tfrac{1}{2}[(v_{1,x})^2 - (v_{2,x})^2] . \tag{3.18a}$$

Inserting the form of w from the first of Eqs. (3.17) into the second one there, we obtain another relevant equation:

$$(v_1 + v_2)_{,x} = \tfrac{4}{3}\lambda - \tfrac{1}{12}(v_1 - v_2)^2 . \tag{3.18b}$$

The pair of equations (3.18) then gives us a Bäcklund transformation between two solutions of the pKdV equation itself, while our earlier version gave us a Bäcklund transformation between a solution of the

KdV equation and the mKdV equation. I note that, just as before, if I insert, for instance, the solution $v_2 = 0$ into this pair of equations they may be solved for v_1 and that solution will be the 1-soliton solution again, which in terms of v is proportional to the hyperbolic tangent, the first integral of the square of the hyperbolic secant.

We now have enough material, at least almost, to determine the famous *nonlinear superposition principle* for the solutions of this equation. We first suppose that v_0 is some given, specific solution of the pKdV equation, and then resolve the Bäcklund transformation above to determine v_1 from this v_0 and some specific (spectral) constant choice for λ_1. Next we perform that operation again, obtaining v_2, using the same v_0 but now a different choice of constant, λ_2, which allows us to write down the following pair of equations, from Eqs. (3.18b):

$$\begin{aligned} (v_1 + v_0)_{,x} &= \tfrac{4}{3}\lambda_1 - \tfrac{1}{12}(v_1 - v_0)^2 \ , \\ v_2 + v_0)_{,x} &= \tfrac{4}{3}\lambda_2 - \tfrac{1}{12}(v_2 - v_0)^2 \ . \end{aligned} \tag{3.19a}$$

Next we construct v_{12} via the Bäcklund transformation, but using the beginning solution v_1 and constant λ_2. Likewise we construct v_{21} via the Bäcklund transformation, but using the beginning solution v_2 and constant λ_1. This process gives us the (similar) equations

$$\begin{aligned} (v_{12} + v_1)_{,x} &= \tfrac{4}{3}\lambda_2 - \tfrac{1}{12}(v_{12} - v_1)^2 \ , \\ v_{21} + v_2)_{,x} &= \tfrac{4}{3}\lambda_1 - \tfrac{1}{12}(v_{21} - v_2)^2 \ . \end{aligned} \tag{3.19b}$$

Using the theorem of "permutability" originally proven by Bäcklund, which says that v_{21} and v_{12}, created in this way, must be the same, we may subtract the first pair of equations, subtract the second pair of equations, and resolve them for v_{12} in the following form:

$$v_{12} - v_0 = 16\frac{\lambda_1 - \lambda_2}{v_1 - v_2} \ . \tag{3.20}$$

You might try to see what you can do with creating a similar way of looking at the commutation relations with λ, in Eq. (3.14), taking perhaps the hint that $\frac{9}{4}\vec{\delta}$ and $\frac{9}{4}[\vec{\gamma}, [\vec{\gamma}, [\vec{\gamma}, \vec{\delta}]]]$, being linear in λ, could be used to create infinite series of powers of λ with coefficients that, otherwise, look much like the angular-momentum-like ones in Eq. (3.10a). I note that the following one is not unique; therefore, you may still consider trying your own hand. Nonetheless, if you take the mapping that sends

$$(-\lambda)^m \, \vec{\mathbf{f}} \equiv F^{(m)} \ , \quad (-\lambda)^m \, \vec{\mathbf{e}} \equiv E^{(m)} \ , \quad (-\lambda)^m \, \vec{\mathbf{h}} \equiv H^{(m)} \ , \tag{3.21}$$

and take the set of all these for all non-negative, integer values of m, then this will generate the infinite-dimensional Lie algebra which is usually called the centerless, positive loop algebra, of entire maps from the complex plane into $\boldsymbol{SL}(2,\mathbb{C})$, which has the product rules

$$[H^{(m)}, E^{(n)}] = 2E^{(m+n)} \ , \qquad [H^{(m)}, F^{(n)}] = -2F^{(m+n)} \ , \qquad [E^{(m)}, F^{(n)}] = H^{(m+n)} \ . \tag{3.22}$$