The Korteweg-De Vries Equation: A Derivation

We begin with the standard "conservation" equations for fluid motion:

$$\begin{aligned} \partial_t\rho + \nabla\cdot(\rho\vec{v}) &= 0\ , \\ \rho(\partial_t + \vec{v}\cdot\nabla)\vec{v} &= -\nabla P + \vec{f}\ , \end{aligned} \tag{1}$$

where ρ is the density and $\vec{v}$ the velocity of the fluid, while P is the internal pressure and $\vec{f}$ is the external force density. The first one is a statement about conservation of mass, the second one about conservation of momentum. We next assume that our fluid is **incompressible** and **irrotational**, which gives us the additional constraints:

$$\nabla\rho = 0\ , \quad \partial_t\rho = 0\ , \qquad \nabla\times\vec{v} = 0\ . \tag{2}$$

The last equation allows us to consider the velocity in terms of some potential, and insertion of that form into the first of our hydrodynamic equations requires that this potential satisfy Laplace's equation:

$$\vec{v} = \nabla\phi\ , \qquad \nabla^2\phi = 0\ , \tag{3}$$

with the momentum density equation still to be satisfied as well. Considering that one further, we first note that we intend to consider the case when the external force is that caused by gravity, so that we will set $\vec{f} = -\rho g\,\hat{y}$. Next we will note the following 3-vector identity, which is quite useful for irrotational vectors:

$$\begin{aligned} \vec{v}\times(\nabla\times\vec{v}) &= -(\vec{v}\cdot\nabla)\vec{v} + \tfrac{1}{2}\nabla(\vec{v}^{\,2})\ , \\ \text{Proof:}\ \ [\vec{v}\times(\nabla\times\vec{v})]^\ell &= \epsilon^{\ell mi}v_m\epsilon_{ijk}\partial^j v^k = \delta^{\ell m}_{jk}v_m\partial^j v^k = v_m\partial^\ell v^m - v_m\partial^m v^\ell\ . \end{aligned} \tag{4}$$

As our velocity field has zero curl, i.e., is irrotational, this allows us the easy substitution generated by Eq. (4), which then allows the entirety of the momentum equation to be written as a gradient:

$$\nabla\left(\partial_t\phi + \tfrac{1}{2}\vec{v}^{\,2} + P/\rho + gy\right) = 0\ . \tag{5}$$

We note then that the quantity under the ∇ must then depend only on the time, which is a term that may be absorbed into our potential function, since adding a function only of the time will not affect its role, namely to determine the velocities. This gives us then two equations to determine our potential, namely Laplace's equation and the following one:

$$\partial_t\phi + \tfrac{1}{2}(\nabla\phi)^2 + P/\rho + gy = \partial_t\phi + \tfrac{1}{2}(u^2+v^2) + P/\rho + gy = 0\ , \qquad \vec{v} = \nabla\phi \equiv u\hat{x} + v\hat{y}\ . \tag{6}$$

We must now describe the problem, set up the geometry, and, in particular, determine the boundary conditions. We are interested in the (irrotational) flow of this (incompressible) water down a long channel which is so narrow as to allow us to consider it one-dimensional. We therefore set up an $\hat{x}$-axis along the length of the channel, and a $\hat{y}$-axis vertically, ignoring totally the $\hat{z}$-direction. The meaning of shallow is that the waves that we want to study should be much longer than the depth of the water; a very different way of saying this is that we are making *a long wavelength approximation.* We describe this in some detail by saying that the depth of the water at rest is given by h, a characteristic length of the waves to be searched for is given by ℓ and we assume that $\ell >> h$. On the other hand, we also want, surely, to have this irrotational flow, i.e., we want to avoid any turbulence in the motion; therefore, we propose as well to restrain these searched-for waves so that their amplitude is characterized by some length a, and we require that $a << h$. It will be useful to define two small quantities, created by these requirements:

$$\epsilon \equiv \frac{a}{h}\ , \qquad \delta \equiv \left(\frac{h}{\ell}\right)^2\ , \tag{7}$$

and we propose to treat both of these quantities as perturbations to the simple, laminar flow, and to suppose that they are both of the same (small) size. Although they appear "different," since they are different powers of their ratios, we will see that, modulo overall factors, only this quadratic power of h/ℓ actually appears in the final equations.

As well we suppose that the surface of the water at rest is the zero for the vertical direction, so that the bottom of the channel corresponds to $y = -h$, and, **also** we take the pressure to vanish at, and near, the surface. [This gives us what is often referred to as Bernoulli's equation.] As the bottom is rigidly fixed, the water cannot move it, so that a boundary condition is surely that

$$(\vec{v})^y\Big|_{\text{bottom}} = v|_{\text{bottom}} = \frac{\partial \phi}{\partial y}(x,0) = 0\ . \tag{8}$$

We will take some amplitude for the travelling waves, $\eta = \eta(x,t)$, so that the surface of the liquid that has waves travelling on it will be given by

$$\begin{aligned} y|_{\text{surface}} &= h + \eta(x,t)\ , \\ \Longrightarrow \quad v|_{\text{surface}} &= \frac{dy}{dt}|_{\text{surface}} = \partial_t\eta + \partial_x\eta\frac{dx}{dt}|_{\text{surface}}\ , \\ \Longrightarrow \quad \phi_y|_{\text{surface}} &= \partial_t\eta + \phi_x|_{\text{surface}}\partial_x\eta\ . \end{aligned} \tag{9}$$

We may also re-write Eq. (6) at the surface, remembering that the pressure vanishes there:

$$\phi_t|_{\text{surface}} + \tfrac{1}{2}(u^2+v^2)|_{\text{surface}} + g\eta = 0\ , \tag{10}$$

where I have ignored the constant term gh.

It is of some mild interest to first resolve these equations for gravity-forced, linear waves before running ahead to obtain the (better, nonlinear) approximation that gives the KdV equation. To do that we consider first the linear approximation to all the above equations, which amounts to considering them for very small amplitudes:

$$\begin{aligned} \nabla^2\phi = 0\ , \quad \phi_y(x, y=-h) = 0\ , \\ \text{and at the surface:}\ \ \partial_t\eta - \partial_y\phi = 0 = \partial_t\phi + g\eta\ . \end{aligned} \tag{11a}$$

We first eliminate the wave amplitude function, η from the three equations other than the last one, by differentiating it and substituting, which gives

$$\partial_{tt}\phi + g\partial_y\phi = 0\ , \tag{11b}$$

and allows us to use that last equation to determine η when we have found the potential ϕ. To do that, since we know we are interested in wavelike solutions, we first propose an ansatz

$$\phi = Y(y)\sin(kx-\omega t)\ . \tag{11c}$$

The Laplace equation requires that $Y = Ae^{+ky} + Be^{-ky}$. When this is inserted into the boundary condition at the bottom, we find that $B/A = e^{-2kh}$, which then gives us a form for ϕ:

$$\phi = 2Ae^{-kh}\cosh k(y+h)\sin(kx-\omega t)\ . \tag{11d}$$

Inserting this into the requirement we had that eliminated η_t, earlier, i.e., Eq. (11b), gives us an equation for the frequency:

$$\omega^2 = gk \tanh k(y+h)|_{\text{surface}} = gk \tanh kh\ . \tag{11e}$$

Lastly, we may use that "extra" equation to determine the wave amplitude:

$$\eta = -\tfrac{1}{g}\,\phi_t = A\sqrt{(2k/g)\sinh(2kh)}\,\sin(kx-\omega t)\ . \tag{11f}$$

We see that the wave has an interesting wavelength-dependent amplitude, and a dispersion relation for its frequency that says that ω is just proportional to k, with factor between them, i.e., the speed, equal to $c_0 \equiv \sqrt{gh}$, for very long wavelengths, i.e., k near 0, but which goes more like $\sqrt{g/k} = c_0/\sqrt{kh} = 2\pi c_0\sqrt{\lambda/h}$ as the wavelengths become shorter. This tells us that the important factor, defining what it means for the wavelength to be longer, or shorter, is the ratio of the wavelength to the depth, h, of the channel; as well, of course, since this last relationship is the limit for very short wavelengths, that the long wavelength waves travel faster and with less dispersion.

Now let us go forward, with the insight created by this simpler example, and look in some detail at the next higher levels of perturbation to this problem. Returning to Eq. (6), and differentiating it along the channel, i.e., with respect to x, and evaluating at the surface of the water, we obtain

$$\phi_{xt} + \phi_x\phi_{xx} + \phi_y\phi_{xy} + g\eta_x = u_t + uu_x + vv_x + g\eta_x = 0\ . \tag{12}$$

We then take a standard sort of an approach for small amplitude deviations and expand ϕ in a power series in y:

$$\phi = \sum_0^\infty y^n\phi_n(x,t) \qquad \Longrightarrow \qquad \phi_y = \sum_1^\infty ny^{n-1}\phi_n\ . \tag{13}$$

Evaluating our constraint at the bottom, Eq. (8), gives us the important requirement that

$$\phi_1 = 0\ . \tag{14}$$

On the other hand, insertion of this sum into Laplace's equation gives us the following, generating a recursion relationship among the coefficients:

$$\sum_0^\infty y^n\left\{\phi_{n,xx} + (n+2)(n+1)\phi_{n+2}\right\} = 0 \quad \Longrightarrow \quad \phi_{n,xx} + (n+2)(n+1)\phi_{n+2} = 0\ . \tag{15}$$

As we already know that $\phi_1 = 0$, this tells us that all the odd terms vanish, and we have the straightforward form for ϕ given by

$$\begin{aligned} \phi &= \sum_{m=0}^\infty \frac{(-1)^m y^{2m}}{(2m)!} f^{(2m)}\ , \quad f \equiv \phi_0(x,t)\ , \\ &\Longrightarrow \quad \begin{cases} u = \phi_x = f_x - \tfrac{1}{2}y^2 f_{xxx} + \ \ldots\ , \\ v = \phi_y = -yf_{xx} + \tfrac{1}{6}y^3 f_{xxxx} + \ \ldots\ . \end{cases} \end{aligned} \tag{16}$$

where the notation $f^{(2m)}$ refers to the $2m$-derivative of f with respect to x.

At this point, when dealing perturbatively with a nonlinear problem, it is quite useful to first create some new, dimensionless variables, which will allow us to introduce the quantities of small order already named earlier. We scale the horizontal distance in terms of its characteristic length, ℓ. We scale the time in terms of the simplest, linear speed, $c_0 \equiv \sqrt{gh}$, from Eq. (11e) above, and the amplitude in terms of its maximum a, so that, eventually, we define the following dimensionless variables via the following scalings, where we recall that our potential has such dimensions that its derivative, with respect to either x or y, should be a velocity:

$$\begin{gathered}\overline{x} \equiv \frac{x}{\ell}\ , \quad \overline{y} \equiv \frac{y}{h}\ , \quad \overline{t} \equiv \frac{t}{\ell/c_0} \\ \overline{\eta} \equiv \frac{\eta}{a}\ , \quad \overline{\phi} \equiv \frac{h\,\phi}{a\ell\, c_0} = \frac{\phi}{\epsilon\ell c_0}\ , \qquad \Longrightarrow \quad \text{the surface is now at } \ \overline{y} = 1 + \epsilon\overline{\eta}\ , \\ \Longrightarrow \quad \begin{cases} \overline{u} = u/\epsilon c_0\ , \\ \overline{v} = (\delta/\epsilon c_0)v\ , \\ \overline{f} = f/\epsilon\ell c_0\ . \end{cases}\end{gathered} \tag{17}$$

The dimensionless versions of the equations of interest may now be worked out. It is useful to first write out the equations for the (dimensionless) velocity components in this approach, to lowest order in small quantities, where I recall that ϵ and δ are of the same order of smallness, so that the symbol O^2 means any terms of second-order in these quantities, such as $\epsilon\delta$, ϵ^2, or δ^2:

$$\begin{gathered}\overline{\phi} = \overline{f} - \tfrac{1}{2}(1+\epsilon\eta)^2\delta\overline{f}_{\overline{x}\overline{x}} + O^2\ , \\ \overline{u} = \overline{f}_{\overline{x}} - \tfrac{1}{2}\delta\overline{f}_{\overline{x}\overline{x}\overline{x}} + O^2\ , \qquad \overline{v} = -\delta\left[(1+\epsilon\overline{\eta})\overline{f}_{\overline{x}\overline{x}} - \tfrac{1}{6}\delta\overline{f}_{\overline{x}\overline{x}\overline{x}\overline{x}} + O^2\right]\ .\end{gathered} \tag{18}$$

With that available, we now re-write our constraint on the vertical liquid velocity, as given in Eq. (9), in terms of these quantities, and then divide out by the common factor $\epsilon c_0/\delta^{1/2}$, which gives

$$\overline{\phi}_{\overline{y}} = \delta(\overline{\eta}_{\overline{t}} + \epsilon\overline{\eta}_{\overline{x}}\overline{\phi}_{\overline{x}})\ . \tag{19a}$$

When our expansions for these quantities, in Eqs. (16), are inserted, a common factor of δ emerges, which we divide out, and the result is then

$$\overline{\eta}_{\overline{t}} + \epsilon\overline{\eta}_{\overline{x}}\overline{f}_{\overline{x}} + (1+\epsilon\overline{\eta})\overline{f}_{\overline{x}\overline{x}} - \tfrac{1}{6}\delta\overline{f}_{\overline{x}\overline{x}\overline{x}\overline{x}} + O^2 = 0\ . \tag{19b}$$

In the same way our other boundary condition at the surface, Eq. (10), when these re-definitions are inserted, and when it is divided out by the common factor ag, gives us

$$\overline{\phi}_{\overline{t}} + \tfrac{1}{2}\epsilon[(\overline{\phi}_{\overline{x}})^2 + (\overline{\phi}_{\overline{y}})^2/\delta] + \overline{\eta} = 0\ . \tag{20a}$$

Again when the expansions of these quantities are inserted, we obtain

$$\overline{f}_{\overline{t}} - \tfrac{1}{2}\delta\overline{f}_{\overline{x}\overline{x}\overline{t}} + \tfrac{1}{2}\epsilon(\overline{f}_{\overline{x}})^2 + \overline{\eta} + O^2 = 0\ . \tag{20b}$$

As previously, when we looked at the linear problem, it is useful to go ahead and find how this equation changes along the channel. Therefore, we differentiate it with respect to x, and then re-write both that equation and the previous Eq. (19b) using $w \equiv \overline{f}_{\overline{x}}$. At the same time, since all our variables are now in this dimensionless form, we stop writing explicitly the overbars, remembering from now on that all our variables are in fact dimensionless and, were we to want to go backwards we would have to use Eqs. (17):

$$\begin{gathered}w_t - \tfrac{1}{2}\delta w_{xxt} + \epsilon w w_x + \eta_x = 0\ , \\ \eta_t + \epsilon w\eta_x + (1+\epsilon\eta)w_x - \tfrac{1}{6}\delta w_{xxx} = 0\ ,\end{gathered} \tag{21}$$

where, lastly, we have also stopped writing down the reminder that all these equations are correct only to second-order, in ϵ and δ, separately.

We now need to resolve this complicated pair of equations into something more manageable. The clear beginning is to make sure that the equations are satisfied to zero-th order. Therefore, ignoring both the ϵ- and the δ-terms, the current pair of equations reduces simply to

$$w_t + \eta_x = 0 = w_x + \eta_t \ . \tag{22a}$$

We may separate the unknown functions in this pair of equations, which then implies that

$$w_{tt} = w_{xx} \qquad \text{and} \qquad \eta_{xx} = \eta_{tt} \ . \tag{22b}$$

In other words, at the lowest order, it is necessary that both w and η satisfy waves equations that have (dimensionless) speed 1, as already expected by the method which we used to make the variables dimensionless. The next requirement, then, is an ansatz based on the thought that w and η seem fairly similar, and that their differences are probably somewhere in, at least, first-order terms. Therefore, we now suppose that there may exist a solution for w in terms of η plus small terms:

$$w \equiv \eta + \epsilon F + \delta G + O^2 \ , \tag{23}$$

where F and G depend on x and t, and it is important to recall that the physical meanings of ϵ and δ are quite different, so that, in principle even though they are the same approximate size they are quite different and therefore must be studied separately. It is for this reason that we write an additional, first-order term for the difference between u and η for each of them. Since we must still satisfy Eqs. (22), we insert this ansatz into them, which implies that $\eta_t + \eta_x$ is of first order in ϵ and δ. This will be important because we will soon find a term in our equations which is of the form $\delta(\eta_x - \eta_t)$, which we see must be the same as $\delta(2\eta_x) + O^2$. On the other hand, we have already seen that to lowest order we must have $w_t + w_x$ to vanish; therefore, we must impose on our "extra terms," i.e., F and G, which are already multiplied by a small parameter, that they also have this property: $F_x + F_t = O^1$ and $G_x + G_t = O^1$. With these provisos in hand, we now insert this ansatz into our two equations, Eqs. (21), and note that the purpose is to determine consistently the functions F and G in such a way that both our equations are satisfied, at least to the order at which they have been written. These equations then take the following form:

$$\begin{aligned} \eta_t + \eta_x + \epsilon(F_t + \eta\eta_x) + \delta(G_t - \tfrac{1}{2}\eta_{xxt}) &= 0 \ , \\ \eta_x + \eta_t + \epsilon(F_x + 2\eta\eta_x) + \delta(G_x - \tfrac{1}{6}\eta_{xxx}) &= 0; \ , \end{aligned} \tag{24}$$

where of course we have now dropped any explicit statement that there are higher-order terms which are being ignored.

The next thing to do is to subtract these two equations, which cancels out the zero-th order terms, although we do, eventually, need to ensure that we have satisfied both of them:

$$\epsilon(F_x - F_t + \eta\eta_x) + \delta(G_x - G_t - \tfrac{1}{6}\eta_{xxx} + \tfrac{1}{2}\eta_{xxt}) = 0 \ . \tag{25a}$$

However, our requirements in the previous paragraph relate, at appropriate orders, the x and t derivatives that appear above. As well, we recall that ϵ and δ should be treated independently—which is why we inserted two multipliers for them in our first-order ansatz above. Therefore we now end up with two rather simple equations that determine F and G, modulo ...

$$\begin{aligned} 2F_x = -\eta\eta_x = -\tfrac{1}{2}(\eta^2)_x \ , &\qquad 2G_x = \tfrac{2}{3}\eta_{xxx} \ , \\ \Longrightarrow \quad F = -\tfrac{1}{4}\eta^2 \ , &\qquad G = \tfrac{1}{3}\eta_{xx} \ , \\ \Longrightarrow \quad w = \eta - \tfrac{1}{4}\epsilon\eta^2 &+ \tfrac{1}{3}\delta\eta_{xx} + O^2 \ . \end{aligned} \tag{25b}$$

We lastly insert this form for w into either one of the two original equations in the set Eqs. (21); we choose the second one, which gives us the form:

$$\eta_t + \eta_x + \tfrac{3}{2}\epsilon\eta\eta_x + \tfrac{1}{6}\delta\eta_{xxx} \ . \tag{26}$$

This is essentially the KdV equation; however, we still need to do some "prettying." We begin by re-scaling, one last time, the variables, to eliminate the first-order small coefficients Since the equation is nonlinear and involves three quantities for which we may choose new scales, there are quite a few ways to do this, depending on the order. Perhaps the most straightforward is to first eliminate the (linear) η_x term. We do this as follows, by translating η by a constant:

$$\eta_x + \tfrac{3}{2}\epsilon\eta\eta_x = \eta_x(1 + \tfrac{3}{2}\epsilon\eta) = \tfrac{3}{2}\epsilon\eta_x(\eta + \tfrac{2}{3}/\epsilon) = \tfrac{3}{2}\epsilon(\eta + \tfrac{2}{3}/\epsilon)_x(\eta + \tfrac{2}{3}/\epsilon) \equiv \tfrac{3}{2}\epsilon\sigma_x\sigma \ , \tag{27}$$

where $\sigma \equiv \eta + \frac{2}{3}/\epsilon$ is simply the original wave amplitude translated by this "large" constant value. The equation then becomes

$$\sigma_t + \tfrac{3}{2}\epsilon\sigma\sigma_x = \tfrac{1}{6}\delta\sigma_{xxx} \ . \tag{28}$$

Next, we choose to re-scale σ and t as follows, and then dividing the overall equation by δ:

$$\left.\begin{aligned} \sigma &= (\delta/\epsilon)\rho \ , \\ t &= \tau/\delta \ , \end{aligned}\right\} \Longrightarrow \rho_\tau + \tfrac{3}{2}\rho\rho_x + \tfrac{1}{6}\rho_{xxx} = 0 \ . \tag{29}$$

This is a perfectly reasonable form for our equation. On the other hand, there are still many things that are done to make it even "cleaner"; unfortunately, at this point there is no consensus at all concerning what should be done. I will take the one which I prefer, noting the others. One of the options is to change, or not to change, the sign of x (or τ) so that the wave motion is toward increasing values of x instead of the other way around. An entirely different sort of option is to re-scale one last time the wave amplitude ρ so as to arrange the numerical constants in the equation according to one's liking. It is true that if we arrange the nonlinear term to have the coefficient 6 then other things become simpler; on the other hand, the coefficient 1 is also quite nice. Lastly, some people really want to have a coefficient 4 multiplying the term which is a time derivative. I will then choose to re-scale one last time with

$$\left.\begin{aligned} \tau &\equiv \ -6t \ , \\ \rho &\equiv u/9 \ , \end{aligned}\right\} \Longrightarrow \quad u_t = uu_x + u_{xxx} \ . \tag{30}$$

It is also perhaps worthwhile, and certainly interesting, to recall the relationship of these latest variables to the original physical ones:

$$\eta = \frac{1}{9\epsilon}(\delta u - 6) \ , \quad t = -\frac{6\ell^2}{\delta c_0}t \ , \tag{31}$$

where in the second equation the symbol t on the left is the physical time, while the one on the right is simply the dimensionless variable we are using in our form of the equation.