

The derivation of the nonlinear Schrödinger equation

Jordan Bell

jordan.bell@gmail.com

Department of Mathematics, University of Toronto

April 3, 2014

In these notes I am following Peter D. Miller, *Applied Asymptotic Analysis*.

The sine-Gordon equation is

$$\frac{\partial^2 \varphi}{\partial t^2} - \frac{\partial^2 \varphi}{\partial x^2} + \sin \varphi = 0.$$

$\varphi(x, t) = 0$ is a solution. Often, when we have a complicated problem we perturb a solution to find something that is close to a solution. Since 0 is a solution, this means that we perturb 0 to get something that is close to a solution of the sine-Gordon equation. In other words, we are going to look for a function that approximately solves the sine-Gordon equation and that has small amplitude. So we assume that φ is of the form $\varphi = \epsilon u$. $u = u(x, t, \epsilon)$.

Then we get the equation

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} + \frac{1}{\epsilon} \sin(\epsilon u) = 0, \quad (1)$$

or

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} + u - \frac{\epsilon^2}{6} u^3 = O(\epsilon^4).$$

We let $T_0 = t, T_1 = \epsilon t, T_2 = \epsilon^2 t$, and $X_0 = x, X_1 = \epsilon x$. This is called a multiple scale decomposition of our time and space variables.

$$u(x, t, \epsilon) = U(X_0, X_1, T_0, T_1, T_2, \epsilon).$$

Using the chain we get get the following:

$$\begin{aligned}\partial_t^2 u &= \partial_{T_0}^2 U + 2\epsilon \partial_{T_0} \partial_{T_1} U + \epsilon^2 (\partial_{T_1}^2 U + 2\partial_{T_0} \partial_{T_2} U) \\ \partial_x^2 u &= \partial_{X_0}^2 U + 2\epsilon \partial_{X_0} \partial_{X_1} U + \epsilon^2 \partial_{X_1}^2 U.\end{aligned}$$

We want to get an asymptotic expansion of U in powers of ϵ .

$$\begin{aligned}U(X_0, X_1, T_0, T_1, T_2, \epsilon) &= U_0(X_0, X_1, T_0, T_1, T_2) \\ &+ \epsilon U_1(X_0, X_1, T_0, T_1, T_2) \\ &+ \epsilon^2 U_2(X_0, X_1, T_0, T_1, T_2) + O(\epsilon^3).\end{aligned}$$

Let's now replace u with U in (1) and compare powers of ϵ . For the constant terms,

$$\frac{\partial^2 U_0}{\partial T_0^2} - \frac{\partial^2 U_0}{\partial X_0^2} + U_0 = 0.$$

A solution of this $U_0 = Ae^{i\theta} + A^*e^{-i\theta}$, where $A = A(X_1, T_1, T_2)$, $\theta = kx - \omega t = kX_0 - \omega T_0$, and k and ω satisfy the dispersion relation $\omega^2 = k^2 + 1$.

Comparing coefficients of ϵ ,

$$\begin{aligned}\frac{\partial^2 U_1}{\partial T_0^2} - \frac{\partial^2 U_1}{\partial X_0^2} + U_1 &= -2\frac{\partial^2 U_0}{\partial T_0 \partial T_1} + 2\frac{\partial^2 U_0}{\partial X_0 \partial X_1} \\ &= 2i\left(\omega \frac{\partial A}{\partial T_1} + k \frac{\partial A}{\partial X_1}\right)e^{i\theta} + \text{c.c.}\end{aligned}$$

(c.c. means the complex conjugate of the preceding)

To avoid "resonance" (which I hope to explain more fully later), we want the right hand side to be 0, so we want

$$\frac{\partial A}{\partial T_1} + \frac{k}{\omega} \frac{\partial A}{\partial X_1} = 0.$$

The solution of this is $A = f(X_1 - \frac{k}{\omega}T_1)$, for any function f .

Taking ω as a function of k , $\omega' = \frac{k}{\omega}$.

With this choice of A , we can choose $U_1 = 0$.

Now we compare the coefficients of ϵ^2 . Using the fact that $U_1 = 0$ we get

$$\begin{aligned}\frac{\partial^2 U_2}{\partial T_0^2} - \frac{\partial^2 U_2}{\partial X_0^2} + U_2 &= -\frac{\partial^2 U_0}{\partial T_1^2} - 2\frac{\partial^2 U_0}{\partial T_0 \partial T_2} + \frac{\partial^2 U_0}{\partial X_1^2} + \frac{1}{6}U_0^3 \\ &= \left(-\frac{\partial^2 A}{\partial T_1^2} + 2i\omega \frac{\partial A}{\partial T_2} + \frac{\partial^2 A}{\partial X_1^2} + \frac{1}{2}A^2 A^*\right)e^{i\theta} \\ &+ \frac{1}{6}A^3 e^{3i\theta} + \text{c.c.}\end{aligned}$$

Again we want to avoid “resonance”, so we want the coefficient of $e^{i\theta}$ to be 0. So we want

$$-\frac{\partial^2 A}{\partial T_1^2} + 2i\omega \frac{\partial A}{\partial T_2} + \frac{\partial^2 A}{\partial X_1^2} + \frac{1}{2}A^2 A^* = 0.$$

We do another change of variables: $\xi = X_1 - \omega' T_1$ and $\tau = T_1$. The nonresonance condition for U_1 is equivalent to $\frac{\partial A}{\partial \tau} = 0$.

The nonresonance condition for U_2 is hence

$$2i\omega \frac{\partial A}{\partial T_2} + \left(1 - \frac{k^2}{\omega^2}\right) \frac{\partial^2 A}{\partial \xi^2} + \frac{1}{2}A^2 A^* = 0.$$

This can be written as

$$i \frac{\partial A}{\partial T_2} + \frac{\omega''}{2} \frac{\partial^2 A}{\partial \xi^2} + \beta |A|^2 A = 0,$$

where $\beta = \frac{1}{4\omega}$.

The following notes follow the chapter on multiple scale analysis in Carl M. Bender, Steven A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers: Asymptotic Methods and Perturbation Theory*.

For the differential equation

$$\frac{d^2 y}{dt^2}(t) + y(t) = \cos(\omega t), \quad \omega \in \mathbb{R},$$

if $|\omega| \neq 1$ then the general solution is

$$y(t) = A \cos t + B \sin t + \frac{\cos(\omega t)}{1 - \omega^2},$$

while if $|\omega| = 1$ then the general solution is

$$y(t) = A \cos t + B \sin t + \frac{1}{2}t \sin t.$$

We interpret the inhomogeneous term as an external driving force on the system. In the first case the driving force is out of phase with the system (which has the resonant frequencies $\omega = 1$ and $\omega = -1$). In the second case the driving force is in phase with the system. In the first case the solution is a bounded function of t , while in the second case the solution is not a bounded function of t .

The term $\frac{1}{2}t \sin t$ in the second case is called a *secular term*. Secular variation is a general notion that refers to any behaviour that is not periodic

in a system, at least with respect to some chosen time scale. We would probably not think of the Earth orbiting the Sun as secular variation in the Earth's position. In summary: secular=something that changes in the long term.¹

In the equation $Ly = f$, where L is a linear differential operator with constant coefficients, there will be secular terms in the general solution if there is some term in f (written as a sum of trigonometric functions) that L sends to 0.

1 Duffing's equation

Consider Duffing's equation:

$$\frac{d^2y}{dt^2} + y + \epsilon y^3 = 0, \quad y(0) = 1, y'(0) = 0.$$

Duffing's equation arises if we do the substitution $u = \epsilon^{1/2}y$ in $\frac{d^2u}{dt^2} + u + u^3 = 0$.

We attempt to get a perturbation solution to this equation.

$$y(t) = \sum_{n \geq 0} \epsilon^n y_n(t),$$

where $y_0(0) = 1, y_0'(0) = 0$ and for $n \geq 1$ we have $y_n(0) = 0, y_n'(0) = 0$. We substitute this into Duffing's equation and compare like powers of ϵ to get:

$$\begin{aligned} y_0'' + y_0 &= 0, \\ y_1'' + y_1 + 1 &= -y_0^3, \\ y_2'' + y_2 &= -3y_0^2 y_1, \\ &\dots \end{aligned}$$

$y_0(t) = \cos t$. Now we find $y_1(t)$. It ends up being

$$y_1(t) = \frac{1}{32} \cos 3t - \frac{1}{32} \cos t - \frac{3}{8} t \sin t.$$

$y_1(t)$ contains a secular term; we expected this because the differential equation for y_1 had the forcing $\frac{3}{4} \cos t$ and 1 is a resonant frequency for the system.

¹The meaning of "secular" here is distinct from its typical meaning of "worldly" (in contrast to spiritual). "secular" comes from "saeculum", which means "an age" or "a lifetime"; cf. the Roman Secular Games.

Therefore a perturbation solution of Duffing's equation is

$$y(t) = \cos t + \epsilon \left(\frac{1}{32} \cos 3t - \frac{1}{32} \cos t - \frac{3}{8} t \sin t \right) + O(\epsilon^2).$$

That is, for each fixed t the difference between $y(t)$ and $y_0(t) + \epsilon y_1(t)$ is $O(\epsilon^2)$ as $\epsilon \rightarrow 0$. The larger we pick t the smaller we will have to pick ϵ , so there is no choice of ϵ that will work for all t .

Even though our approximation $y_0 + \epsilon y_1$ of y is unbounded, we can show that in fact y is bounded.

If we multiply the Duffing equation by the integrating factor $\frac{dy}{dt}$ and integrate in t we get

$$\frac{1}{2} \left(\frac{dy}{dt} \right)^2 + \frac{1}{2} y^2 + \frac{1}{4} \epsilon y^4 = C.$$

$y(0) = 1$ and $y'(0) = 0$, so $C = \frac{1}{2} + \epsilon \frac{1}{4}$. It follows that $\frac{1}{2} y(t)^2 \leq C$ for all t , and therefore $|y(t)| \leq \sqrt{2C} = \sqrt{1 + \frac{\epsilon}{2}}$.

2 Multiple scale analysis

We are going to introduce multiple scale analysis using the example of the Duffing equation.

Assume that

$$y(t) = \sum_{n \geq 0} \epsilon^n Y_n(t, \tau).$$

$\tau = \epsilon t$, but we treat t and τ as independent.²

Using the chain rule,

$$\frac{dy}{dt} = \left(\frac{\partial Y_0}{\partial t} + \frac{\partial Y_0}{\partial \tau} \frac{d\tau}{dt} \right) + \epsilon \left(\frac{\partial Y_1}{\partial \tau} + \frac{\partial Y_1}{\partial t} \frac{\partial \tau}{\partial t} \right) + O(\epsilon^2).$$

Using $\frac{d\tau}{dt} = \epsilon$ this becomes

$$\frac{dy}{dt} = \frac{\partial Y_0}{\partial t} + \epsilon \left(\frac{\partial Y_0}{\partial \tau} + \frac{\partial Y_1}{\partial t} \right) + O(\epsilon^2). \quad (2)$$

Differentiating with respect to t gives

$$\frac{d^2 y}{dt^2} = \frac{\partial^2 Y_0}{\partial t^2} + \epsilon \left(2 \frac{\partial^2 Y_0}{\partial t \partial \tau} + \frac{\partial^2 Y_1}{\partial t^2} \right) + O(\epsilon^2)$$

²I need to understand and explain better this artifice of introducing τ .

Now we substitute the expressions we have found for y and $\frac{d^2y}{dt^2}$ into the Duffing equation and compare powers of ϵ . We get

$$\begin{aligned}\frac{\partial^2 Y_0}{\partial t^2} + Y_0 &= 0, \\ \frac{\partial^2 Y_1}{\partial t^2} + Y_1 &= -Y_0^3 - 2\frac{\partial^2 Y_0}{\partial \tau \partial t}, \\ &\dots\end{aligned}$$

A real valued solution of the equation for Y_0 is of the form $Y_0(t, \tau) = A(\tau)e^{it} + A(\tau)^*e^{-it}$, where $A(\tau)$ is any complex function of τ .

We determine $A(\tau)$ by stipulating that no secular terms appear in the forcing term in the equation for Y_1 . Using our expression for $Y_0(t, \tau)$ we can write the forcing term as

$$\begin{aligned}& -A^3 e^{3it} - 3A^2 A^* e^{it} - 3A(A^*)^2 e^{-it} - (A^*)^3 e^{-3it} - 2i\frac{dA}{d\tau}e^{it} + 2i\frac{dA^*}{d\tau}e^{-it} \\ &= -A^3 e^{3it} - (A^*)^3 e^{-3it} + e^{it}\left(-3A^2 A^* - 2i\frac{dA}{d\tau}\right) + e^{-it}\left(-3A(A^*)^2 + 2i\frac{dA^*}{d\tau}\right)\end{aligned}$$

Since e^{it} and e^{-it} satisfy the homogeneous equation for Y_1 , to avoid secular terms we need the coefficients of e^{it} and e^{-it} in the forcing term both to be 0. That is, we need that

$$-3A^2 A^* - 2i\frac{dA}{d\tau} = 0$$

and

$$-3A(A^*)^2 + 2i\frac{dA^*}{d\tau} = 0.$$

These two equations are complex conjugates. Write $A(\tau)$ in polar coordinates as $A(\tau) = R(\tau)e^{i\theta(\tau)}$. The first of the above two equations becomes

$$-3R^3 - 2i\frac{dR}{d\tau} + 2R\frac{d\theta}{d\tau} = 0.$$

Equating the real parts of the two sides we get

$$-3R^3 + 2R\frac{d\theta}{d\tau} = 0$$

and equating the imaginary parts we get

$$\frac{dR}{d\tau} = 0.$$

So $R(\tau) = R(0)$, and $\theta = \frac{3}{2}R(0)^2\tau + \theta(0)$. Thus $A(\tau) = R(0)e^{i\frac{3R(0)^2\tau}{2} + i\theta(0)}$.

Now we put this into $Y_0(t, \tau)$ to get

$$Y_0(t, \tau) = 2R(0) \cos\left(\frac{3R(0)^2\tau}{2} + \theta(0) + t\right).$$

Using (2) and the initial conditions $y(0) = 1$ and $y'(0) = 0$, we get $Y_0(0, 0) = 1$ and $\frac{\partial Y_0}{\partial t}(0, 0) = 0$. It follows that $\theta(0) = 0$ (or any integer multiple of 2π), and that $R(0) = \frac{1}{2}$. Therefore

$$Y_0(t, \tau) = \cos\left(\frac{3}{8}\tau + t\right).$$

And since $\tau = \epsilon t$,

$$y(t) = \cos\left(t\left(1 + \frac{3}{8}\epsilon\right)\right) + O(\epsilon^2).$$