Strings, Membranes, and Such

Philip Zucker

December 16, 2011

1 Introduction

Strings and Membranes: Why should we care? In terms of direct applications to building better washing machines, many would say we don't. This is an attitude that is common towards most physical models and almost all mathematical techniques that the eggheads and four-eyes of the world churn out. But upon reflection, even if brighter brights and whiter whites are your goal, these things are worth looking into. A washing machine or a car or steel beam or even the electromagnetic field is essentially a big mixed-up membrane in essence and can be simulated on a computer as such. The focus on strings and membranes is just because they are a simple testing ground for our mathematical techniques of solving partial differential equations. A string really can't have much simpler dynamics. What could we cut out? We could perhaps cut the string up into a finite number of balls and springs. Typically that's how the subject is introduced. Honestly, such a situation is more complicated, not less. Figuring out how to deal with 100 balls and springs and the indexes they bring with them is a tad tougher than an infinite number of them. Go figure.We could make the string inflexible, but then it has no dynamics at all! Ignoring washing machines engineers for the moment and focusing on the needs of physicists, you better have a pretty darn good grasp of classical fields if you want to doggy paddle your way through modern physics. So let's do it.

2 Strings

We've got a rubber band stretched between two posts. I will use the term string, rubber band, and spring interchangeably, because a dog is a cat as far as its properties of elasticity and vibrational modes go. If we want to make our string move, we can appeal to the least action principle and Lagrangian mechanics.

The Least Action principle is a very strange stance to take on dynamics, in that you assume knowledge of the initial and final configuration of some system and how long it took to do it. But how exactly did it do it? The answer takes the form of a game. We give points for certain qualities of paths and take away points for others and try and figure out which path gets the most points. Pretend you are watching a fat child. He needs to walk to Arby's in twenty

minutes to make the breakfast deadline. The fatter he is, the less he'll like to go fast or divert from his path. Going slow at any point along his path means he'll have to go faster elsewhere to make the twenty minute deadline. So you expect uniform speed. However, perhaps there are parks along the way where he'd love to dawdle and smell the flowers. By assigning numerical weights to all these factors (hungriness, love of flowers, etc.) you can calculate which path he will take to.

Now, our fat boy shall become a string. We start by saying the action S is the time integral of the Lagrangian, which is given by the kinetic energy minus the potential energy.

$$S = \int L dt$$

$$L = T - U$$

Let's write down the expression for the potential energy of the system. The energy contained in your everyday spring is

$$U = \frac{1}{2}k(x - x_0)^2$$

This is great, except that our rubber band may not be stretched uniformly. So let's look at the big rubber band as a bunch of little rubber bands tied together and add up the energy of each. The length of a baby band is given by the Pythagorean theorem, a formula you may recognize from the arc length of a curve y(x)

$$ds = \sqrt{dx^2 + dy^2} = dx\sqrt{1 + (\frac{dy}{dx})^2}$$

The baby band also has a equilibrium length. Here we'll already start the approximations.We'll assume a uniform equilibrium length, (i.e. one that isn't a function of where we are along the rubber band). For stretching that is small in the horizontal direction, this is pretty much true.

$$ds_0 = \frac{L_0}{L}dx$$

We assign the constant in front of the dx so that the sum of all the equilibrium lengths equals the total equilibrium length L_0

$$\int_0^L ds_0 = \frac{LL_0}{L} = L_0$$

Another factor that we must take into account is that shorter springs are stiffer. If you cut a spring in half, it takes twice the force to stretch it the same distance, i.e. the spring constant is doubled. Cut in thirds, the force is tripled. We can represent that by writing

$$dk = \frac{Lk}{dx}$$

Where dk is the spring constant of one of our small springs. Amusingly for small spring lengths, dk is not small, but will blow up. Luckily, one we put our whole expression together, the small stretching sizes of the little springs will avert this problem.

$$\begin{aligned} U &= \sum \frac{1}{2} dk(ds - ds_0)^2 = \sum \frac{1}{2} \frac{Lk}{dx}(dx\sqrt{1 + (\frac{dy}{dx})^2} - \frac{L_0}{L}dx)^2 \\ &= \int_0^L \frac{1}{2} Lkdx(1 + (\frac{dy}{dx})^2 - \frac{2L_0}{L}\sqrt{1 + (\frac{dy}{dx})^2} + \frac{L_0^2}{L^2}) \end{aligned}$$

Let's do some hand waving approximations. We'll use the fact that the stretching is small in the y direction to approximate that ugly square root with the binomial theorem

$$\sqrt{1 + (\frac{dy}{dx})^2} \approx 1 + \frac{1}{2}(\frac{dy}{dx})^2 + \dots$$

Integrate out the things we can

$$U = \frac{1}{2}(L^2 + L_0^2 - 2L_0L)k + \frac{1}{2}k(L - L_0)\int_0^L (\frac{dy}{dx})^2 dx = \frac{1}{2}k(L - L_0)^2 + \frac{1}{2}\tau \int_0^L (\frac{dy}{dx})^2 dx$$

Interesting! The first part is clearly the energy it takes to stretch the string from L_0 to L. There is nothing we can do about this energy since the ends of the strings are fixed, so we can ignore it. It is unminimizable under the constraints of the problem. In the second term, I have noticed that $k(L - L_0)$ is just the tension in the spring according to Hooke's law, which I have labeled as τ.

The hard part is over. We found the potential of a rubber band. The kinetic energy is a cinch by comparison. Again, we break the band up into lots of little rubber bands. If we assume they only move vertically then

$$dT = \frac{1}{2}dmv^2 = \frac{1}{2}dm(\frac{dy}{dt})^2$$

The tiny bit of mass each bit of rubber band has can be written in terms of a mass per unit length of string σ

$$dm = \sigma dx$$

$$T = \int dT = \int_0^L \frac{1}{2}\sigma(\frac{dy}{dt})^2 dx$$

The action for a free string is then

$$S = \int [T - U]\, dt = \iint \left[\frac{1}{2}\sigma(\frac{dy}{dt})^2 - \frac{1}{2}\tau(\frac{dy}{dx})^2\right] dxdt$$

From this we get the differential equation

$$\sigma\frac{d^2y}{dt^2} - \frac{d}{dx}(\tau\frac{dy}{dx}) = 0$$

In our particular derivation, we've been assuming $\tau$is constant for simplicity. It is possible to avoid this constraint and still get the same equations up to this point. Here we utilize the fact $\tau$is constant to pull it out of the derivative and rearrange some terms.

$$\frac{\sigma}{\tau}\frac{\partial^2 y}{\partial t^2} - \frac{\partial^2 y}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 y}{\partial t^2} - \frac{\partial^2 y}{\partial x^2} = 0$$

By jove, we've done it.

3 Sturm-Liouville

What we now have on our hands is a pretty neat equation, but what the heck do we do with it? Ordinary differential equations that have only one variable aren't trivial, adding another to the mix isn't going to make life easy.

Here's the physics rule of thumb: If you are ever presented with a problem, a good first course of action is to find a way to make it as much like the simple harmonic oscillator as possible. If you approximate enough and make enough assumptions, you can usually find one somewhere. We can dignify this (although we might attempt it even if we couldn't) for our partial differential equation by using the separation of variables, which means we assume that the solution is of the form

$$y(x,t) = X(x)T(t)$$

Why can we assume this? If this is true, then we can pull some nice tricks. Let's just be optimistic. Optimism solves all problems. We proceed to do this for our string equation

$$\sigma\frac{\partial^2 y}{\partial t^2} - \frac{\partial}{\partial x}(\tau\frac{\partial y}{\partial x}) = 0$$

Let's put in our guess and slide functions out of partial derivatives where we can.

$$\frac{\partial^2 X(x)T(t)}{\partial t^2} - \frac{1}{\sigma}\frac{\partial}{\partial x}(\tau\frac{\partial X(x)T(t)}{\partial x}) = \sigma X(x)\frac{\partial^2 T(t)}{\partial t^2} - \frac{1}{\sigma}T(t)\frac{\partial}{\partial x}(\tau\frac{\partial X(x)}{\partial x}) = 0$$

Divide out everything by y.

$$\frac{1}{T(t)}\frac{\partial^2 T(t)}{\partial t^2} - \frac{1}{\sigma}\frac{1}{X(x)}\frac{\partial}{\partial x}(\tau\frac{\partial X(x)}{\partial x}) = 0$$

The first term is a function of only t and the second is a function of only x, yet the two added together are zero? That's like filling a barrel with tortoises and then emptying it by only taking out rabbits. That'll never work! So the two both have to be equal to some constant

$$\frac{1}{T(t)}\frac{\partial^2 T(t)}{\partial t^2} = -\omega^2$$

This one is pretty easy, if you've seen the trick.

$$T(t) = C\sin(\omega t + \theta_0)$$

How about the other one?

$$\frac{1}{\sigma}\frac{1}{X(x)}\frac{\partial}{\partial x}(\tau\frac{\partial X(x)}{\partial x}) = \omega^2$$

$$-\sigma\omega^2 X(x) - \frac{d}{dx}(\tau\frac{dX(x)}{dx}) = 0$$

Sturm-Liouville Theory gives you the tools you need to solve this linear second-order differential equation. It is super mathy, but therein lies its power and I would be doing you and it a disservice to try to avoid the mathiness. I don't have time to explain the theory to a level of completeness, so instead I'll hit some of the more interesting highlights.

4 Eigenvalues and Eigenfunctions

The most general Sturm-Liouville is given by

$$-\frac{d}{dx}(p(x)\frac{dy}{dx}) - q(x)y = \lambda w(x)y$$

In other words, we're looking for the very special functions that fit the boundary conditions of our problem that you can apply the differential equation to, and after turning all the cranks on your product rules and chain rules, you get back the same function just multiplied by some constant.

$$Lf = \lambda w(x)f$$

Sounds doable. Here's the weird part: There is a countable set of such functions, one per each λ. Verging on unbelievably weird, there are enough of these functions that you can write down any function your sweet little heart desires as the sum of them (assuming your name isn't Weierstrass or Hilbert or something). And this is not one of those "It exists but couldn't be found with

a computer the size of the universe" mathematical facts. We can construct this representation pretty easily, thanks to the orthogonality of the eigenfunctions.

Why are the eigenvalues discrete? The "avoiding really thinking about it" answer is boundary conditions. True in a sense. Without the boundary conditions, there would not be discrete eigenvalues. For me, this is an insufficient answer. Let's try to do a bit better. Most of the following is pulled from Reference [7].

For the simple case of $p(x) = 1$, $q(x) = 0$, $w(x) = 1$ the solutions take the form of sine and cosine functions. This is not true for any choice of $p(x)$, $q(x)$, and $w(x)$ but there are certain aspects of these solutions that extend to every Sturm-Liouville problem.

1. The nth eigenvalue $\lambda_n$has n zeros in between the endpoints.
2. The functions reaches maximum and minimum only once between each zero.

Since we know the solution is kind of sine-like, let's guess a solution that kind of looks like one, with some parameter functions. We've seen some solutions in a book somewhere that have unequally spaced zeros or we can guess that this might be the case, so we know that the function inside the sine could be more complicated than just a linear function of x. In a similar manner, we know the total amplitude can pull some tricks, so the amplitude out front might be a function of x as well.

$$u(x) = r(x)\sin(\theta(x))$$

$$u'(x) = r'(x)\sin(\theta(x)) + r(x)\cos(\theta(x))\theta'(x)$$

But we have some wiggle room now. Two totally free functions is more than is necessary to represent one function. We have a freedom to choose another constraint if we can think of one that is useful. This is all similar reasoning to the method of variation of parameters for solving linear inhomogeneous differential equations or picking a particular gauge for solving electromagnetism problems. If that sentence is meaningless to you, don't worry about it. The constraint that a man named Prufer picked is

$$P(x)u'(x) = r(x)cos(\theta(x))$$

Now let's differentiate the identity

$$\cot(\theta) = \frac{Pu'}{u}$$

$$-\csc^2(\theta)\theta' = \frac{(Pu')'}{u} - \frac{Pu'^2}{u^2} = \frac{-Qu}{u} - \frac{1}{P}\cot^2(\theta)$$

Manipulate this a bit and you can get

$$\frac{d\theta}{dx} = Q(x)\sin^2(\theta) + \frac{1}{P(x)}\cos^2(\theta)$$

Another equation

$$\frac{dr}{dx} = \frac{1}{2}(\frac{1}{P} - Q)r\sin(2\theta)$$

Can also be derived to determine the $r(x)$ variable, but it isn't as interesting. We took a second order linear differential equation and got two coupled first order awful nonlinear equations. Why is this good? Look at that first one again. There is no $r(x)$ dependance. That first equation is an uncoupled nonlinear first order differential equation. That isn't so bad at all!

From this first equation, we can determine $\theta(x)$ pretty easily. Now that the boundary requirements create a discrete set of solutions is fairly straightforward, just like it was for for the simple sine solution case. If we want the boundary conditions to be 0, we need $\theta(x)$ to be an integer multiple of π at $\theta(L)$.

5 Variational Principle

A tool of great calculational usefulness is the variational principle. For any Sturm-Liouville problem there is a lowest eigenvalue λ_0 and a corresponding eigenfunction $y_0(x)$. We can find a very good estimate of this function by looking at the integral

$$L[\Psi] = \frac{<\Psi|L|\Psi>}{<\Psi|\Psi>} = \frac{\int \Psi^*(x)L\Psi(x)dx}{\int \Psi^*(x)\Psi(x)w(x)dx} = \frac{\int(-\Psi^*\frac{d}{dx}(p(x)\frac{d\Psi}{dx}) - \Psi^* q(x)\Psi dx}{\int \Psi^*(x)\Psi(x)w(x)dx}$$

Out of all possible functions which satisfy the right boundary conditions, $\Psi(x) = y_0(x)$ will minimize this number and the minimal value of $L[\Psi]$will be λ_0 . Why is this so? Let's assume that any function can be written as a linear combination of orthonormal eigenfunctions

$$\Psi(x) = \sum a_n y_n(x)$$

$$Ly_n(x) == \lambda_n w(x) y$$

Then $L[\Psi]$ can be shown to be

$$L[\Psi] = \frac{\int \sum |a_n|^2 \lambda_n |y_n|^2 w(x)dx}{\int \sum |a_n|^2 |y_n|^2 w(x)dx} = \frac{\sum |a_n|^2 \lambda_n}{\sum |a_n|^2}$$

In other words, this is a weighted average of the eigenvalues of the L operator. Since $\lambda_n \geq \lambda_0$, the minimum value of this is achieved by $\Psi = y_o(x)$ as I asserted.

Here's the really good part though. Just pick some class of functions $\Psi(x; \alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_N)$ where $\alpha_i$are a bunch of parameters that we can wiggle at will. Now, it is relatively simple to computationally minimize $L[\Psi]$ with respect to the parameters (at least compared to doing an exact functional minimization)

$$\frac{\partial L[\Psi]}{\partial \alpha_i} = 0$$

It turns out that for good choices of parameters, you get extremely accurate results. I coded this in Python using the Scipy and Numpy Libraries. Scipy has built in support for interpolation by splines, so splines were used as the test function to make life easy on the programmer. Scipy also has powerful general purpose built-in optimization routines. By a little analytic work, we could perhaps exactly solve for the minimum parameter values, but gosh darn it, aren't computers supposed to take the intellectual burden off of man?

To see if this method is working, we'll use the problem

$$-\frac{d^2y}{dx^2} = \lambda y$$

with boundary conditions $y(0) = 0$, $y(\pi) = 0$ which we know has the lowest eigenfunction

$$y_0(x) = \sin(x)$$

$$\lambda_0 = 1$$

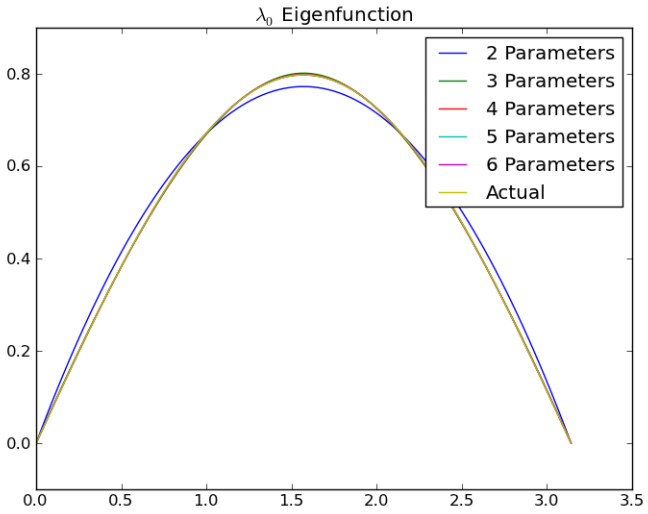

Figure 1: Plot of $\Psi_N(x)$

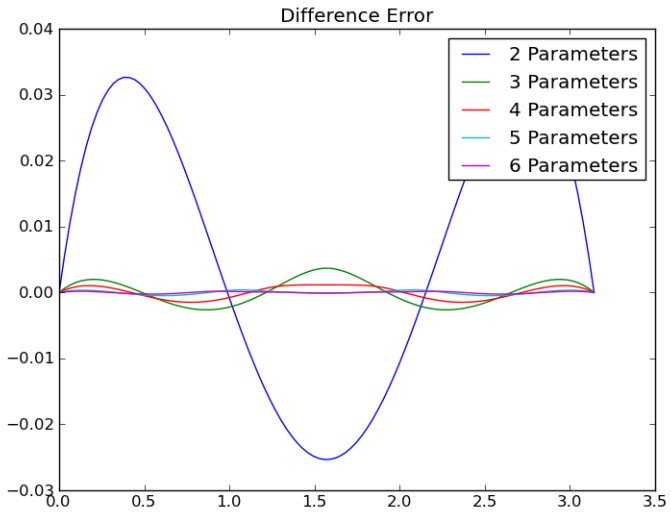

Figure 2: Error $E_N(x) = \Psi_N(x) - \sin(\pi x)$

Parameters	λ_n
2	1.013212
3	1.000261
4	1.000068
5	1.000011
6	1.000003

Table 1: Calculated Eigenvalues

Not bad! Not bad at all! With each addition of parameter, we gain about a digit of accuracy in the calculated eigenvalue and the actual function calculated is pretty darn accurate too!

This is one method (perhaps a slightly misguided one) to numerically calculate π and man is always hungry for more π. We use the identity

$$\lambda_0 = \frac{\pi^2}{L}$$

Using the boundary conditions

$$y(0) = 0$$

$$y(1) = 0$$

$$\pi = \sqrt{\lambda_0}$$

Parameters	π	Error
2	3.16227767305	0.0206850194561
3	3.14200204084	0.000409
4	3.1416992894	0.000106
5	3.14160986023	1.72×10^{-5}
6	3.14159805335	5.39×10^{-6}

Table 2: Calculating π the Rayleigh-Ritz way

For fun, let's do a function that you might not be able to guess so easily.

$$-\frac{d}{dx}(x\frac{dy}{dx}) + \frac{y}{x} = \lambda xy$$

The solution to this equation is a Bessel's function. Specifically the 1st Bessel's function $J_1(x)$ will be the minimization of the functional

$$L[x] = \frac{\int x(\frac{d\Psi}{dx})^2 + \frac{\Psi^2}{x}dx}{\int \Psi^2 x dx}$$

with the boundary conditions

$$\Psi(0) = 0$$

$$\Psi(3.8317060) = 0$$

This particular number was chosen because it is the tabulated value of the first zero of $J_1(x)$. That will make our eigenvalue λ_0 a nice pretty 1.

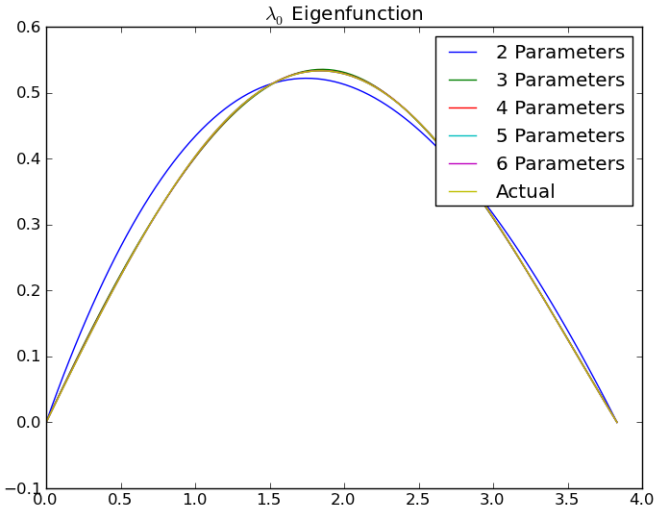

Figure 3: Plot of $\Psi_N(x)$ for $J_1(x)$

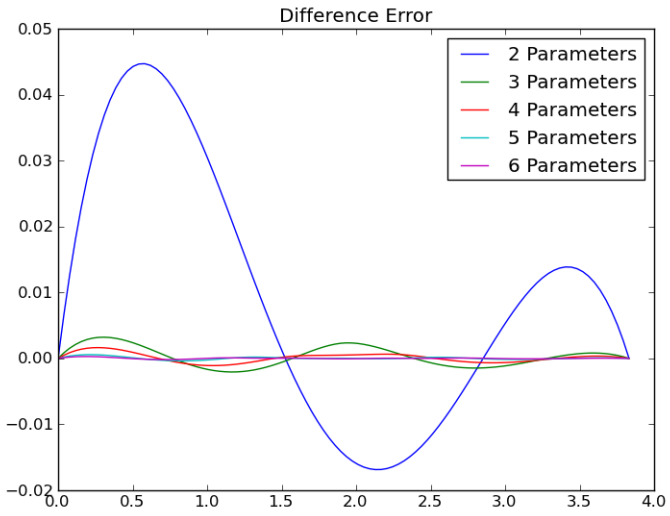

Figure 4: Error $E_N(x) = \Psi_N(x) - J_1(x)$

Parameters	λ_n
2	1.010853
3	1.000173
4	1.000030
5	1.000001
6	0.999998

Table 3: Calculated Bessel Eigenvalues

Hey it all worked! Notice that the last eigenvalue is actually below the correct lowest eigenvalue. We didn't break the pure math. I deeply suspect what we broke is the computer by reaching the limit of accuracy for the floating point numbers. Six parameters is all that took!

6 Green's Functions

The definition that is given for a Green's function is the following differential equation

$$LG(x,a) = \delta(x-a)$$

Where L is a differential operator acting on the x variable. There are two 1-D systems that provide a visual of what a Green's function looks like that I believe we should examine before we put on our fancy hats and write down differential equations.

The first is of a kicked particle. We guarantee that a particle is at $x = 0$ at time $t = 0$ and time $t = T$, which are not strange conditions if you're already comfortable with the principle of least action. We also know that it gets kicked with an impulse J at time t'. The equation of motion for this situation is

$$m\frac{d^2x}{dt^2} = J\delta(t-t')$$

We can visualize this system and realize that under no force, the particle goes in a straight line. The kick instantaneously changes the velocity of the particle in just the right way that it gets where it needs to go in the right amount of time. So we can tell that the path will be two straight lines, with just the right speeds to take the kick and end up where it needs to go

The second is the stationary string. We stretch our rubber band between two posts. The two ends struck on the posts can't move (boundary conditions). We hang a weight from a hook on this rubber band. What shape does the rubber band take? I think most would agree that it takes on this triangle shape.

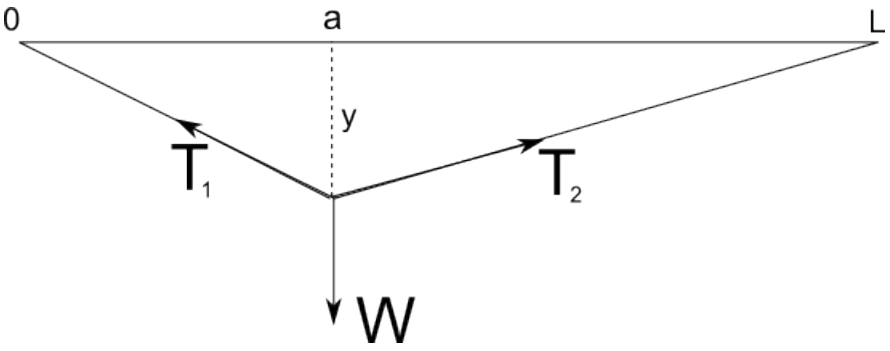

Figure 5: Force Diagram For Fixed String Ends

If you don't believe me, try it. Now, from Freshman physics, we can solve for the numerical shape of this triangle.

All the forces $(\vec{W}, \vec{T_2}, \vec{T_1})$ acting on the point must cancel each other for equilibrium to occur. Breaking this into components

$$T_{1x} = T_{2x}$$

$$T_{1y} + T_{2y} = W$$

If we assume small displacement of the rubber band from its original position, it is also safe to approximate that the x-component of the tension has the same value of the tension in the rubber band before the string was hung

$$T_{2x} \approx T \approx T_{1x}$$

Since the tensions $\vec{T_1}$and $\vec{T_2}$ lie along the rubber band, we can use similar triangle tricks to get the slopes of those two lines.

$$m_1 = \frac{-y}{a} = \frac{-T_{1y}}{T_{1x}} \approx \frac{-T_{1y}}{T}$$

$$m_2 = \frac{y}{L-a} = \frac{T_{2y}}{T_{2x}} \approx \frac{T_{2y}}{T}$$

Combining these equations, we get the relation

$$\frac{y}{a} + \frac{y}{L-a} = m_2 - m_1 = \frac{W}{T}$$

So now we have three facts about the function y(x) describing the vertical displacement of the rubber band at a point x.

1. The difference between the two slopes is $\frac{W}{T}$.

2. The rubber band has no displacement at either end of the rubber band. Thus $y(0) = 0$ and $y(L) = 0$)

3. The two lines meet at the point where the weight hangs. Thus $y^+(a) = y^-(a)$

Using these three constraints, we can write out the function that describes the y position of the rubber band at a point x with a weight hanging at a

$$y_a(x) = \begin{cases} \frac{-W}{TL}x(L-a) & x < a \\ \frac{-W}{TL}a(L-x) & x > a \end{cases}$$

Problem solved!

Now, how is this useful? While a weight hanging on a rubber band problem is extremely fascinating, what if we hung two weights, or some other cockamamie collection of weights? We'll have to go over everything all over again!

Not so!

Solving this problem for one weight essentially solves the entire problem for any configuration of weights by a superposition principle. This is perhaps not obvious by the Physics 101 arguments we've given so far, but is extremely obvious from the perspective of the differential equation or linear algebra.

Another way to go about the problem is to say, "Well, the string and mass will configure itself into its minimum energy configuration." We already derived the formula for the potential energy of the spring.

$$U = \frac{1}{2}T\int(\frac{dy}{dx})^2dx$$

Now, there is also the potential energy of the mass hanging on the rubber band

$$U = mgy(a) = Wy(a)$$

So the total potential energy (ignoring the constant from the initial stretching)

$$U = Wy(a) + \frac{1}{2}T\int_0^L(\frac{dy}{dx})^2dx$$

We can use a Dirac Delta function to put everything inside the integral

$$U = \int_0^L W\delta(x-a)y(x) + \frac{1}{2}T(\frac{dy}{dx})^2dx$$

And now use the Euler Lagrange Equations to get

$$W\delta(x-a) = T\frac{d^2y}{dx^2}$$

Our previously derived solution should fit this differential equation. When $x \neq a$, $y(x)$ is a linear function, so its second derivative vanishes. We can integrate our equation across a small area around a

$$\int_{a-\epsilon}^{a+\epsilon} dx \frac{W}{T}\delta(x-a) = \frac{W}{T} = \int_{a-\epsilon}^{a+\epsilon} dx \frac{d^2y}{dx^2} = \frac{dy}{dx}^{+} - \frac{dy}{dx}^{-} = m_2 - m_1$$

This is the same condition that we derived from Physics 101, so it all hangs together. However, now it is clear that we can represent the solution of many weights as the sum of solutions of one weight. We write this as

$$y(x) = W_1 y_{a_1}(x) + W_2 y_{a_2}(x) + W_3 y_{a_3}(x) + \ldots$$

Such that each individual $y_{a_n}(x)$satisfies the equation

$$\delta(x-a_n) = T\frac{d^2 y_{a_n}}{dx^2}$$

Then through the linearity of $\frac{d^2}{dx^2}$

$$\sum W_n \delta(x-a_n) = T\frac{d^2y}{dx^2}$$

And running the Euler Lagrange equation backwards, we see that this function minimizes

$$U = \int_0^L \sum W_n \delta(x-a_n)y(x) + \frac{1}{2}T(\frac{dy}{dx})^2 dx$$

The other main way to look at Green's functions is purely as an algebraic tool, totally abstracted from the physics of pumpkins and rubber bands.

Let's start with the absolutely simplest class of algebraic problem that I know how to construct.

$$ay = 0$$

In this equation, a is a number, like 3 or 743507683, but not 0, and y is an unknown number. This problem is so simple, that its solution may even be considered an axiom of numbers. This equation has one solution

$$y = 0$$

Great. Now suppose we add a "forcing term"

$$ay = b$$

What is the solution to this equation?

$$y = \frac{b}{a}$$

Awesome. Let's define the Green's Function (Green's Number?) for this equation

$$aG = 1$$

Which can be solved for $a \neq 0$

$$G = \frac{1}{a} = a^{-1}$$

So now we know exactly how to solve any equation with a forcing term

$$y = 0 + Gb$$

I explicitly included the 0 because that is the solution to the unforced equation. This inclusion will make more sense as we go along.

As a side track, consider the new equation with a forcing term proportional to y

$$ay = b + cy$$

There are two ways to go about this. We can define a new equation similar to the first

$$a' = a - c$$

$$a'y = b$$

Then laboriously solve for the new Green's function

$$G' = \frac{1}{a'} = \frac{1}{a - c}$$

to get our solution

$$y = G'b = \frac{b}{a - c}$$

This was awful. We had to find a whole new Green's Function! A much better approach is to find an infinite series solution to the problem by iteration. Let's assume the forcing term wasn't special and write down the solution

$$G = \frac{1}{a}$$

$$y = G(b + cy)$$

That didn't seem to get us very far. If that is a solution, you can break my legs. Our solution looks like another problem, just as hard as the original problem. However, let's write this down as a recurrence relation

$$y_{n+1} = G(b + cy_n)$$

We can dignify this as a solution scheme by saying that if this sequence converges, then it has converged to a solution. We make no guarantees as to whether it will converge however. Let's make an initial guess

$$y_0 = 0$$

Now let's iterate that sucker

$$y_1 = G(b + 0) = Gb(1)$$

$$y_2 = G(b + cG(b + 0)) = Gb + G^2bc = Gb(1 + Gc)$$

$$y_3 = G(b + cG(b + cG(b + 0))) = Gb + G^2bc + G^3bc^2 = Gb(1 + Gc + G^2c^2)$$

$$y_n = Gb\sum_{i=0}^{n-1}(Gc)^n$$

That's a geometric series! It has the solution

$$y_n = Gb\frac{1 - (Gc)^n}{1 - Gc}$$

$$\lim_{n\to\infty} y_n = \frac{Gb}{1 - Gc}, \text{ for } |Gc| < 1$$

Expanding G as $\frac{1}{a}$,

$$y = \frac{b}{a(1 - \frac{c}{a})} = \frac{b}{a - c}$$

Which is the correct solution. This iterative series that we produced is called the Neumann Series if you're a mathematician, or in physicist term, the Dyson Series or the Born Series. It is obviously a ridiculous way to go about things for our simple equation, but for complicated math things it is sometimes your only hope. Even in this example the areas of convergence are not trivial. We happened to know how it converged because the geometric series is one of the few things we're pretty sure we understand. Some very clever fellows have determined that power series of that sort converge in circles in the complex plane, but what about the equation

$$ay = b + cy^5$$

This equation is a quintic polynomial equation, which has 5 solutions. We can attempt to solve it iteratively by the recurrence relation

$$y_{n+1} = G(b + cy_n^5)$$

Computing a rough estimate of converging values of $y_0$in the complex plane using values of G, b, and c picked out of a hat gives a graph like this

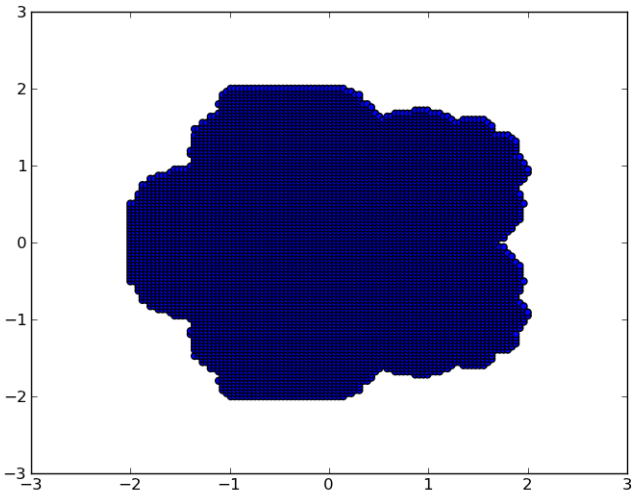

Figure 6: Convergence Region in Complex Plane

Nonlinear iterative equations of this sort are a rabbit hole that you could sink your whole life down and never fill. They are a classic generator of fractal sets, such as the Julia set and the Mandelbrot set. Definitely worth a Wikipedia search, but entirely off topic.

Let's look at the next level of complication, a matrix equation

$$A\underline{x} = 0$$

$$A\underline{x} = \underline{b}$$

Now A is a matrix like this

$$A = \begin{pmatrix} 678 & 678 & 5 & 4 \\ 3 & 6 & 4 & 6 \\ 8 & 5 & 7 & 1 \\ 67 & 4 & 6 & 457 \end{pmatrix}$$

and $\underline{x}$ and $\underline{b}$ are vector such as

$$\underline{x} = \begin{pmatrix} 6 \\ 7 \\ 5 \\ 20394 \end{pmatrix}$$

A linear differential equation has a roughly equivalent matrix form. We can split up the interval L into N chunks of length $\frac{L}{N}$. This is the essence of the discrete time step method for numerically solving differential equations. We can write a vector that looks a lot lot like our function by filling it with the values of the function at the edges of our chunks. For example, if we broke the interval L into 4 chunks,

$$\underline{f_4} = \begin{pmatrix} f(0) \\ f(\frac{L}{4}) \\ f(\frac{2L}{4}) \\ f(\frac{3L}{4}) \\ f(L) \end{pmatrix}$$

As we break the interval into smaller and smaller chunks, the vector will represent the function better and better. We can write a first derivative matrix

$$f'(x) \approx \frac{f(x+\Delta x) - f(x)}{\Delta x},\ \Delta x = \frac{L}{N}$$

For the interval separated into 4 chunks this can be written out in matrix form

$$\Delta = \frac{4}{L} \begin{pmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}$$

As we let $N \to \infty$, this matrix will produce a vector that is closer and closer to the first derivative of the function $f(x)$. Notice that this matrix is not a square matrix, and thus not invertible. The derivative throws away a constant from the function, so a little bit of information is lost. This matrix does the same thing, reducing the dimension of our function vector by one. Similarly, we can construct a second derivative matrix

$$\Delta^2 = \frac{16}{L^2} \begin{pmatrix} 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \end{pmatrix}$$

We lost another dimension! How in the world can we solve the matrix equation $\Delta^2 \underline{f} = \underline{g}$ when Δ^2isn't a square matrix with the same dimensions as $\underline{f}$? The answer is that we have to incorporate the boundary conditions. The separation between a differential equation and boundary conditions is a

somewhat arbitrary one. One should really never be considered without the other.

Let's pad out our matrix with zeroes, to make it square.

$$\Delta^2 = \frac{16}{L^2}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Now we create a boundary value matrix

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

This is the appropriate matrix for specifying values at the endpoints of the interval. Now the matrix $B+\Delta^2$ is invertible. Notice that the matrix B doesn't scale with the number of chunks, so as the number of chunks gets large, it becomes more natural to think of B as separate from Δ^2until it ends up like this:

$$\frac{d^2y}{dx^2} = F(x),\ y(0) = 0,\ y(L) = 0$$

The following matrix is a good representation of this differential equation

$$K = \frac{16}{L^2}\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$K\underline{y} = \frac{L^2}{16}\begin{pmatrix} 0 \\ \underline{F} \\ 0 \end{pmatrix}$$

The first and last row of the matrix K guarantee that the componentsy_0 and y_L both equal zero, which is the same as our boundary conditions in the differential equation. The middle section is the second derivative matrix with some ones cut out of it. Why did I cut out these ones? First of all, I was able to do it because those ones didn't matter. Since we knowy_0 and y_L both equal zero, the ones that have been removed added nothing to the equation. Secondly, it keeps our matrix pleasingly symmetric, a property that is well worth seeking out by any means necessary.

The matrix K possesses an inverse G such that

$$GK = I$$

With this G in hand, we can easily solve our matrix equation

$$\underline{y} = G\frac{L^2}{16}\begin{pmatrix} 0 \\ \underline{F} \\ 0 \end{pmatrix}$$

To sate your curiosity, here is G for the K we've been considering

$$G = \frac{1}{4}\begin{pmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & -3 & -2 & -1 & 0 \\ 0 & -2 & -4 & -2 & 0 \\ 0 & -1 & -2 & -3 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix}$$

Not so interesting perhaps. Let's look at the plots of the rows of G

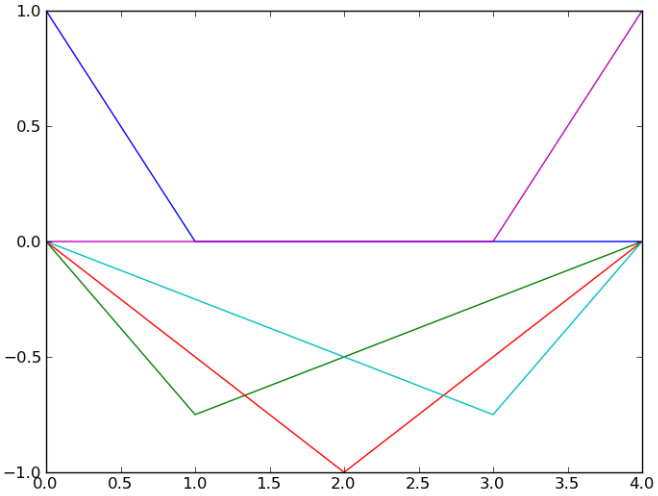

Figure 7: Impulse Responses

Hmm, Interesting. Now let's take a look at the green's function of an interval separated into 50 chunks.

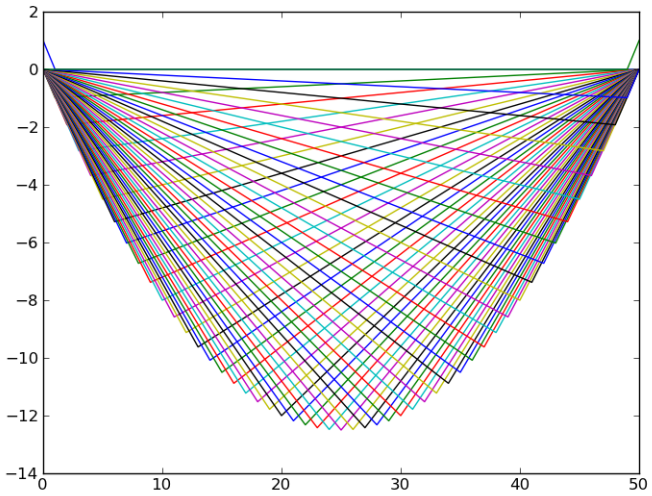

Figure 8: A Whole Bunch of Green's Functions

We see that the rows of this matrix seem to form piecewise linear functions, with an envelope that appears to be parabolic. That's exactly what we got for our green's functions of strings earlier!

There is another good set of homogenous boundary conditions given by

$$y'(0) = 0$$
$$y(L) = 0$$

When we set the derivative of the function to zero at the boundaries, these are called Neumann boundary conditions, whereas before the function being zero at the boundaries is called Dirichlet boundary conditions

We use the names Dirichlet, Neumann, and Robin Boundary (sort of a mix of the first two) conditions because we prefer to give credit than give them self-explanatory titles. I'm not saying that those guys didn't earn their keep, but I predict that in 1000 years this will change. We don't call addition the "Cryntok the Babylonian Operation", even thought Cryntok was really smart.

Physically, this corresponds to a stretched rubber band where one end is free to move up and down a slick pole. In other words, the left hand side will not support any tension force in the vertical direction but will supply any horizontal force asked of it. If we hang a weight at point a, we still have three forces in equilibrium, but the torque connecting to the sliding end must be horizontal and all of the vertical weight must be held by the fixed end of the string. Our

physical intuition says that the Green's function for these boundary conditions ought to look like the following

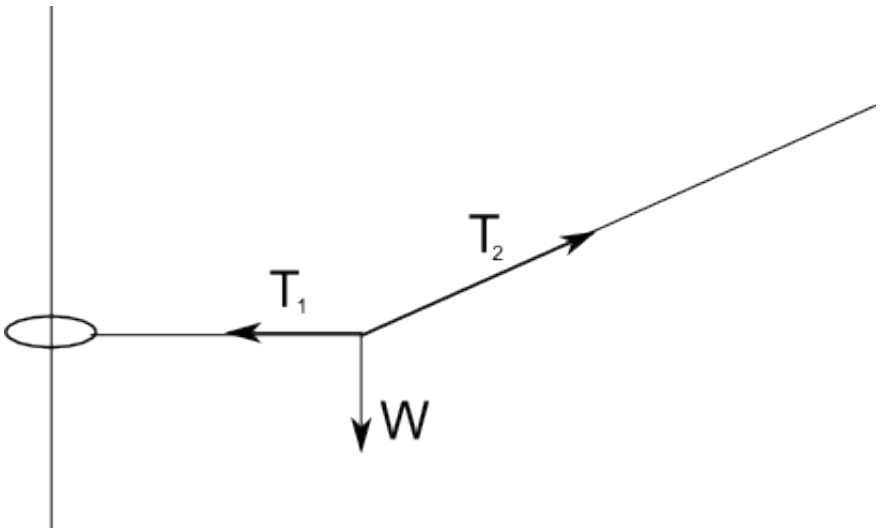

Figure 9: Force Diagram without Fixed Ends

The nice symmetric matrix that corresponds to this differential equation and boundary conditions is

$$K = \frac{16}{L^2} \begin{pmatrix} -1 & 1 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Let's make our computer plot the inverse of this matrix and see how it compares to our intuition

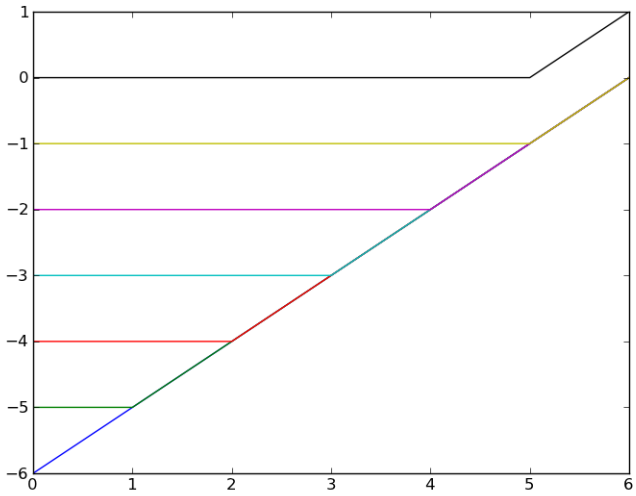

Figure 10: Rock On, Physical Intuition.

7 Mass on a String

Just for kicks, let's put a bead on our string. We can write down the action fairly easily by adding in the kinetic energy of the bead to our Lagrangian for the string

$$S = \int \left[\int \left[\frac{1}{2}\sigma(\frac{dy}{dt})^2 - \frac{1}{2}\tau(\frac{dy}{dx})^2 \right] dx + \frac{1}{2}m\dot{y}(a)^2 + \frac{1}{2}m\dot{a}^2 \right] dt$$

Where $\dot{y}(a)$ is the total time derivative given by

$$\dot{y}(a) \equiv \frac{\partial y(a)}{\partial t} + \dot{a}\frac{\partial y(a)}{\partial x} \equiv y_t + \dot{a}y_x$$

We will focus on the dynamics of the bead. What follows is some finicky derivative work, but the equations of motion of the string itself are even worse. Let's start by using the Euler-Lagrange equations

$$\frac{\partial L}{\partial a} - \frac{d}{dt}\frac{\partial L}{\partial \dot{a}} = 0$$

$$m\dot{y}\frac{\partial \dot{y}}{\partial x} - \frac{d}{dt}(my_x\dot{y} + m\dot{a}) = 0$$

Luckily $\frac{\partial}{\partial x}$ commutes with the $\frac{d}{dt}$operator, allowing some term cancellation. Also we divide by m because it is worthless.

$$y_x \ddot{y} + \ddot{a} = 0$$

We expand out those total derivatives.

$$y_x \frac{d}{dt}(y_t + \dot{a} y_x) + \ddot{a} = y_x(y_{tt} + 2\dot{a} y_{tx} + \dot{a}^2 y_{xx}) + \ddot{a}(1 + y_x^2) = 0$$

Let's divide everything by $\sqrt{1 + y_x^2}$to make things pretty (trust me on this)

$$\frac{y_x(y_{tt} + 2\dot{a} y_{tx} + \dot{a}^2 y_{xx})}{\sqrt{1 + y_x^2}} + \ddot{a}\sqrt{1 + y_x^2} = 0$$

What a bunch of gobbledy-gook! Turns out that each part of this equation has a fairly simple interpretation. Check out this triangle

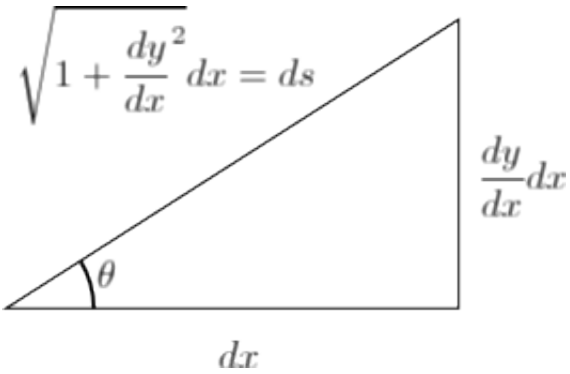

Figure 11: Infinitesimal Arc length Triangle

From this we get some trigonometric identities

$$\frac{y_x}{\sqrt{1 + y_x^2}} = \sin(\theta)$$

$$\sqrt{1 + y_x^2} = \sec(\theta)$$

Where θ is the angle the curve makes with the horizontal x-axis. This means we can write the component of velocity along the curve as

$$\frac{ds}{dt} = \dot{s} = \dot{a}\sqrt{1 + y_x^2} = \frac{\dot{a}}{\cos(\theta)}$$

Likewise, the acceleration along the curve can be written

$$\frac{d^2s}{dt^2} = \ddot{s} = \ddot{a}\sqrt{1+y_x^2} = \frac{\ddot{a}}{\cos(\theta)}$$

First suppose that the time derivative $\dot{a}$ is zero. The only terms remaining in the equation are

$$\frac{y_x y_{tt}}{\sqrt{1+y_x^2}} + \ddot{a}\sqrt{1+y_x^2} = 0$$

Replacing the ugliness with our new trigonometric quantities

$$\ddot{s} = -\sin(\theta) y_{tt}$$

We've seen this one! That is the inclined plane equation, with g replaced by y_{tt}. Since the effects of gravity are equivalent to those of an accelerating reference frame, this makes sense. If the string is accelerating, the bead moving with it experiences an incline plane type force.

What about those $\dot{a}$ terms. Consider a string that is stationary, like a wire. Then all time derivative of $y(a)$ drop out.

$$\frac{y_x \dot{a}^2 y_{xx}}{\sqrt{1+y_x^2}} + \ddot{a}\sqrt{1+y_x^2} = 0$$

Now let's pull a trick and multiply the first term by 1

$$\frac{y_x(\dot{a}\sqrt{1+y_x^2})^2 y_{xx}}{(\sqrt{1+y_x^2})^3} + \ddot{a}\sqrt{1+y_x^2} = 0$$

Oh, but here's an interesting formula for the radius of curvature that I pull from thin air

$$\frac{y_{xx}}{(\sqrt{1+y_x^2})^3} = \frac{1}{R}$$

What this means is that locally, the curve y looks like a circle of radius R. Making all these substitutions

$$\sin(\theta)\frac{\dot{s}^2}{R} + \ddot{a} = 0$$

We've seen this one too! For an unmoving wire, the spatial curvature creates centripetal force. and this is just the x-component of that force!

The last term that we haven't discussed is

$$\frac{y_x 2\dot{a} y_{tx}}{\sqrt{1+y_x^2}}$$

I do not believe this term reduces down to anything extremely familiar, but we can examine it by imaging a curve $y(x,t)$ that is a line with constantly increasing slope

$$y(x,t) = cxt + d$$

The second derivatives y_{xx} and y_{tt} both vanish for this curve, removing the other force terms from the equations of motion, isolating the effects of this term. It is similar in nature to the other two. The changing slope combined with the bead's velocity gives rise to an effective curvature as far as the bead is concerned.

8 Simulating Partial Differential Equations

The number one way to simulate ordinary or partial differential equations is that you do the dumbest thing you can do. The definition of the derivative is

$$\frac{dy}{dx} = \lim_{dx \to 0} \frac{y(x+dx) - y(x)}{dx}$$

Well, just cut off that limit, set dx equal to some fairly small number, cross your fingers, and have your computer crank out the solution. This is called the Euler method. You can get a little fancier and use the Taylor expansion

$$f(x + \Delta x) = f(x) + f'(x)\Delta x + \frac{1}{2!}f''(x)\Delta x^2 + \dots$$

Then using some clever choices of Δx and taking some combinations, you can remove the error proportional to small powers of Δx. This is the essence of the Runge-Kutta Method and other Linear Multistep Methods.

All this works for higher order differential equations as well

$$\frac{d^2y}{dx^2} = \lim_{dx \to 0} \frac{y(x+dx) - 2y(x) - y(x+\Delta dx)}{dx}$$

When we cut off the limit, we get the matrix approximation of the second derivative that we were discussing earlier.

We can use this to simulate the Heat equation

$$\frac{\partial y}{\partial t} = \kappa \frac{\partial^2 y}{\partial x^2}$$

$$y(x, t + \Delta t) = \frac{\kappa \Delta t}{\Delta x^2}(y(x + \Delta x, t) - 2y(x,t) + y(x - \Delta x, t)) + y(x,t)$$

With this equation, given $y(x, 0)$,we can find the function at any future time we please. This was done in python, using the Numpy library for linear algebra to speed things along a little.

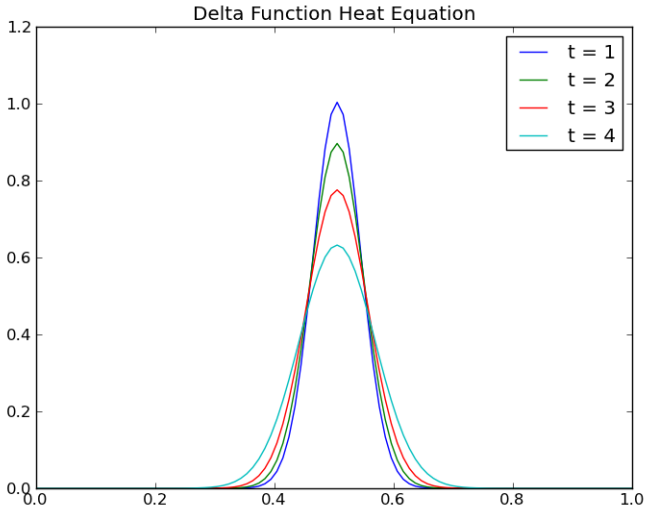

Figure 12: Simulated Delta Function Initial Condition Heat Equation

Looks like Gaussians! Just like it should!

There is another slightly different way to go about this however. We could choose to evaluate our spatial derivative at the future time step and then solve a linear equation to get $y(x, t + \Delta t)$

$$y(x,t+\Delta t) - \frac{\kappa \Delta t}{\Delta x^2}(y(x+\Delta x, t+\Delta t) - 2y(x,t+\Delta t) + y(x-\Delta x, t+\Delta t)) = y(x,t)$$

This is known as the implicit Euler method, and strangely enough, it is superior in many ways to the previous method. In a very uninformative analogy, the implicit Euler is more stable the same way dragging a long rod through the mud is more stable than pushing it.

It's time to reveal that I tricked you a little bit. The explicit method that I used first had to be massaged carefully to keep the solution stable. For bad values of grid spacing the solution of the first method was awful.

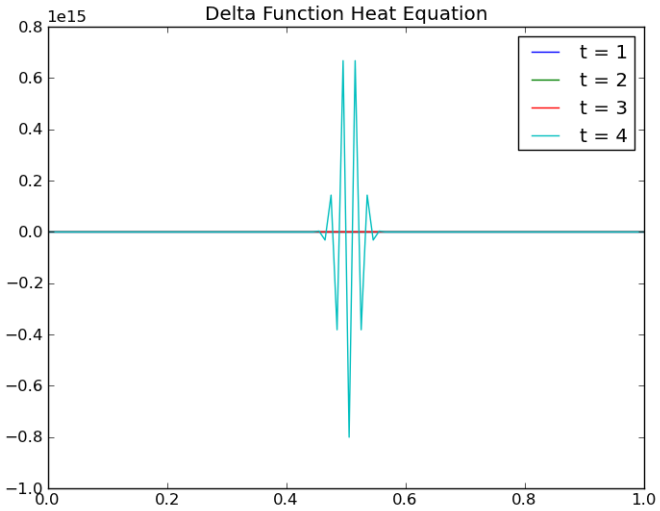

Figure 13: This is Terrible

That is a 10^{15} on the y axis. That's not so good.
With our wave equation

$$\frac{1}{c^2}\frac{\partial^2 y}{\partial t^2} - \frac{\partial^2 y}{\partial x^2} = 0$$

We can pull some similar jive as with the heat equation.

$$y(x,t+\Delta t) = \frac{\kappa \Delta t^2}{\Delta x^2}(y(x+\Delta x,t)-2y(x,t)+y(x-\Delta x,t))-y(x,t-\Delta t)+2y(x,t)$$

Let's look at some simple test functions. First of all, let's try a traveling pulse

$$y(x,0) = e^{-\frac{(x-b)^2}{\sigma^2}}$$

$$y_t(x,0) = cy_x(x,0)$$

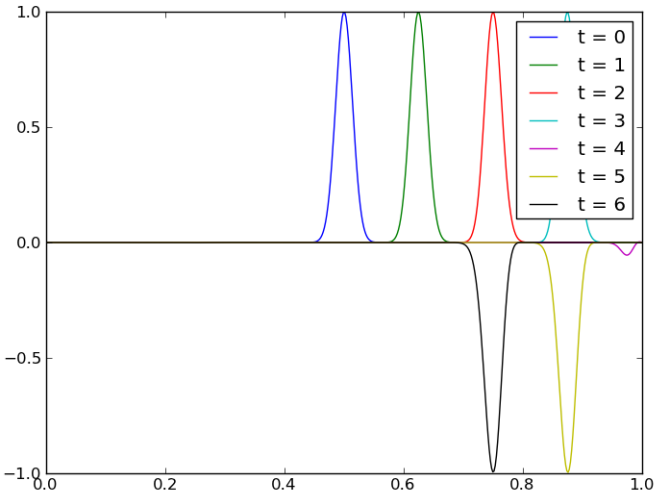

Figure 14: Traveling Pulse

That's nice looking. Exactly what we expected. Good so far. Now let's try

$$y(x,0) = e^{-\frac{(x-b)^2}{\sigma^2}}$$

$$y_t(x,0) = 0$$

D'Alembert has a nice solution to the wave equation

$$y(x,t) = \frac{1}{2}(y_0(x-ct) + y_0(x+ct)) + \frac{1}{2c}\int_{x-ct}^{x+ct} y_{t0}(\eta)d\eta$$

From the D'Alembert solution, we'll expect the solution to be of the form

$$y(x,t) = \frac{1}{2}(y_0(x-ct) + y_0(x+ct))$$

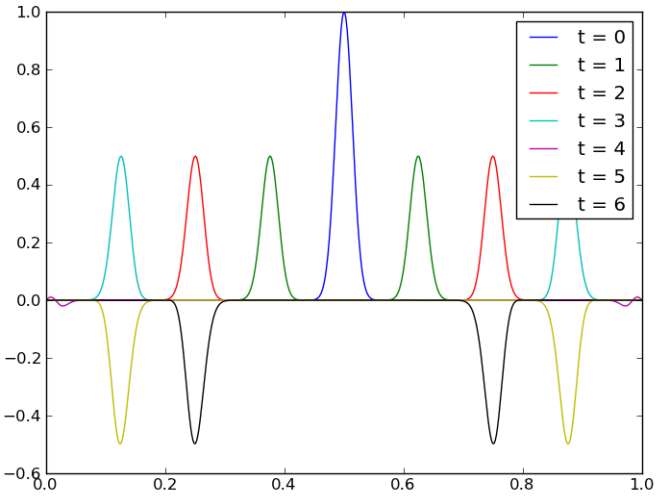

Figure 15: Initially Unmoving Pulse

Alright! Still looking good. How about we look at a normal mode

$$y(x,0) = \sin(\frac{n\pi x}{L})$$

$$y_t(x,0) = \frac{n\pi}{L}\cos(\frac{n\pi x}{L})$$

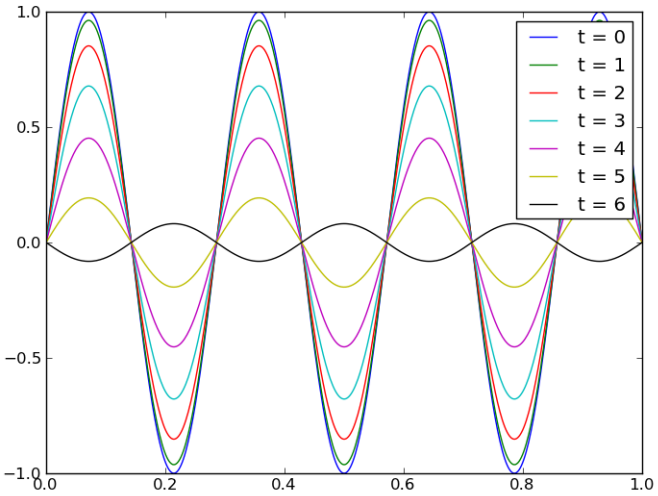

Figure 16: Normal Mode ($n = 7$)

Now we're feeling extremely confidant. Everything has worked so far. Let's go for a $\delta(x)$.

$$y(x, 0) = 0$$

$$y_t(x, 0) = \delta(x)$$

The solution to this can also be found using D'Alembert's formula.

$$y(x, t) = \frac{1}{2c}(H(x + ct) - H(x - ct))$$

Where $H(x)$ is the Heaviside Function

$$H(x) = \begin{cases} 1 & x > 0 \\ 0 & x < 0 \end{cases}$$

Let's take a look at what our computer has to say

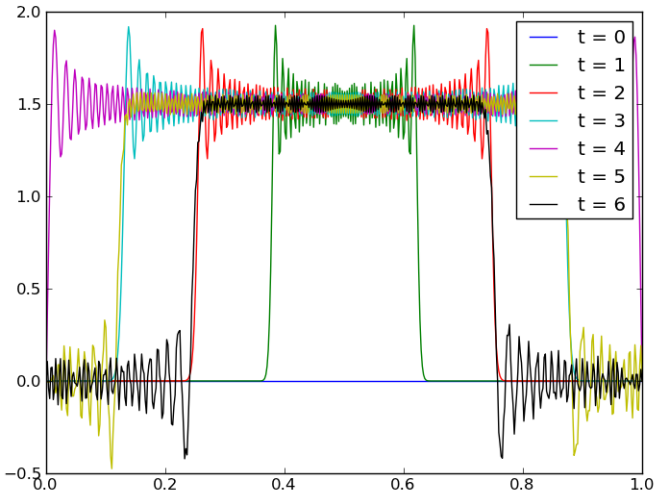

Figure 17: Oy Vey!

Here is our first sign of trouble. We've got something similar to a Gibbs phenomenon occurring from the looks of it. This isn't surprising that this simulation can't handle $\delta(x)$ very well. Anything that is under the scale of the lattice won't simulate well.

Let's try something a little smoother

$$y(x,0) = 0$$

$$y_t(x,0) = e^{-\frac{(x-b)^2}{\sigma^2}}$$

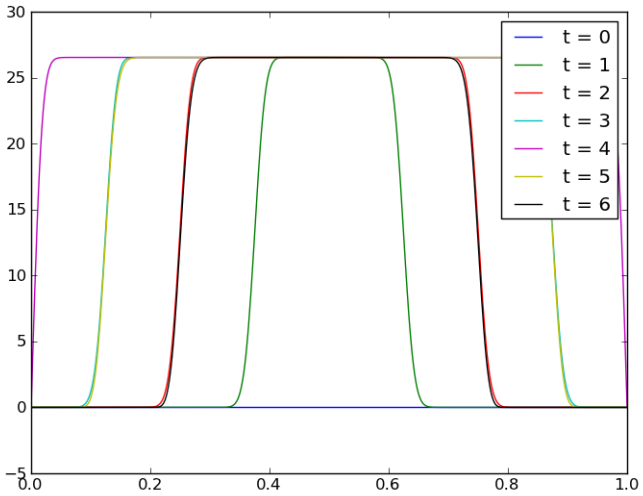

Figure 18: Much Better

A very powerful method, especially for strange domains and boundary conditions, is the Finite Element Method. Given some differential equation, you can multiply it by an unknown function v and integrate it

$$Lu = f \rightarrow \int vLudx = \int vfdx$$

Usually, we'll integrate by parts the second derivative stuff in L so that the left hand side has differential operations on both v and u. This new form is very similar to the form of the action in classical mechanics and the variational function in the Rayleigh-Ritz Variational method, except now there is a test function v instead of a second factor of u. If you can find the $u(x)$ such that this equality holds for any function v, then it seems like it ought to solve the original problem, and it almost always does, as this new condition is only slightly weaker than the original differential equation.

The last main ingredient of the Finite Element Method is the test function that are used. By restricting the possible functions to a set of parametrized functions (again, very similar to the variational method) the problem becomes computable. Often, the test functions are a bunch of piecewise linear functions localized around the grid points, making the integrals very easy to compute. The grid on which they're based can be chosen almost completely arbitrarily.

I did not build my own finite element solver, but instead used one freely available for Python called FEniCS. Here are some pretty pictures that I made.

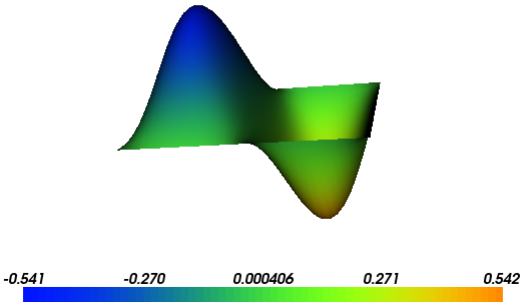

Figure 19: 2-1 Normal Mode in Square Domain

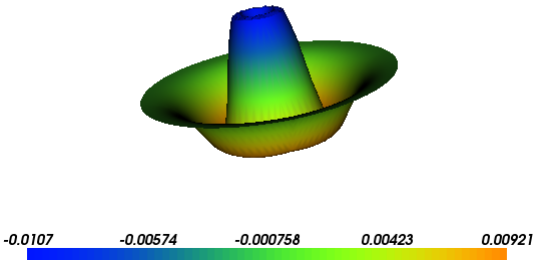

Figure 20: Gaussian Propagated in Circular Membrane

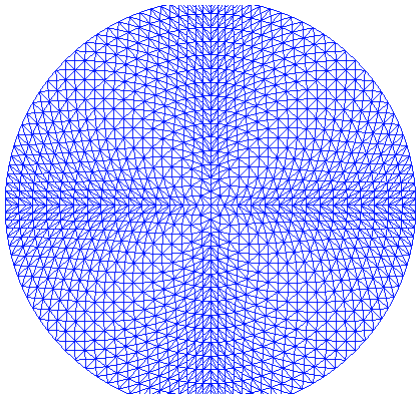

Figure 21: Circular Domain Mesh

References

[1] Richard Haberman, "Elementary Partial Differential Equations." Prentice-Hall (1987)

[2] A. Fetter and D Walecka "Theoretical Mechanics of articles and Continua." Dover

[3] Philip Morse and Herman Feshbach "Methods of Theoretical Physics" Mcgraw-Hill 1953.

[4] G. Arfken "Mathematical Methods for Physicists" Academic Press. 1985

[5] R. Courant and D. Hilbert "Methods of Mathematical Physics Vol I." Interscience Publishers. 1961

[6] T. Pang "An Introduction to Computational Physics" University Press. 2006

[7] U.H. Gerlach "LINEAR MATHEMATICS IN INFINITE DIMENSIONS Signals Boundary Value Problems and Special Functions" http://www.math.osu.edu/~gerlach/math/BVtypset/node77.html

[8] Hans Petter Langtangen "The FEniCS Tutorial" http://fenicsproject.org/