Optimal Functions and Nanomechanics (OFN)

Andrew Martin

September 20, 2019

1 Intro

We will be considering similar problems to those in STD.

$$\max \ f(\mathbf{x})$$
$$A\mathbf{x} \geq \mathbf{b}$$
$$\mathbf{x} \geq 0$$

In this course we will be considering the function f as a 'functional', a function of a function. E.g.

$$f = \int_P F(\mathbf{x})$$

where P is the path taken.

For an example we might consider a thin, uniformly heavy, flexible cable suspended from the top of two poles of height y_0, y_1 respectively, spaced d distance apart. What is the shape of the cable between the two poles.

If $y(x)$ is the height of the cable at x, then we have $y(x_0) = y_0$ and $y(x_1) = y_1$

A local minimum if there is an open ball around the point x if

$$f(x) \leq f(\hat{x})$$

For all $\hat{x} \in (x - \delta, x + \delta)$ Its a global minimum if $\hat{x} \in [a, b]$ (the full set).

The maximal (minimal) set is the set of points which obtain the global maximum (minimum).

Maximum and Minimum **sufficient** condition

MVT

Taylor's Theorem

Lagrange Remainder theorem - the remainder from the first n derivatives of Taylor's theorem

The **necessary** and sufficient for maxima and minima. The $2n^{th}$ derivative has to be $f^{(2n)}(x) > 0$ given all derivatives before it are 0.

Classification of Extrema, Given $f'(x) = 0$

- local max $f''(x) < 0$

- local min $f''(x) > 0$

- turning point $f''(x) = 0, f^{(3)}(x) \neq 0$

- ...

The set of points with $f'(x) = 0$ is the set of **stationary points**

A function is piecewise continuous/differentiable if it has a finite number of discontinuities.

Del/Grad is the vector of derivative ∇x Divergence is the sum of the elements from del $\nabla \cdot x$

Curl is the cross product of the del operator $\nabla \times x$.

We probably won't do much vector stuff

A **variation** is the derivative of a functional.

A function of n variables, $\mathbf{x} = (x_1, \ldots, x_n) \in \Omega \subset \mathbb{R}^n$

Let $f : \Omega \to \mathbb{R}$

A local minima is a point $\mathbf{x}$ such that an open ball around $\mathbf{x}$ with $\hat{x} \in B(\mathbf{x}, \delta)$ we have

$$f(\hat{x}) \geq f(\mathbf{x})$$

A global minima will be over all $\hat{x} \in \Omega$ (rather than just the open ball)

Reminder of the chain rule, for 2D: if $f(x_1(t), x_2(t))$

$$\frac{df}{dt} = \frac{\partial f}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial f}{\partial x_2}\frac{dx_2}{dt}$$

Or could treat as a vector notation

$$\frac{df}{dt} = \nabla f \cdot \frac{d\mathbf{x}}{dt} = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}\right) \cdot \left(\frac{dx_1}{dt}, \frac{dx_2}{dt}\right)$$

Consider $f(x, y)$ where $x(t), y(t)$ By first principles

$$\frac{df}{dt} = \lim_{\epsilon \to 0} \frac{f(x(t+\epsilon), y(t+\epsilon)) - f(x(t), y(t))}{\epsilon}$$

Using taylor's theorem

$$x(t+\epsilon) = x + \epsilon x' + \mathcal{O}(\epsilon^2)$$

and similarly for y. Giving

$$\frac{df}{dt} = \lim_{\epsilon \to 0} \frac{f(x + \epsilon x', y + \epsilon y') - f(x(t), y(t))}{\epsilon}$$

Add and subtract $f(x, y + \epsilon y')$

$$\frac{df}{dt} = \lim_{\epsilon \to 0} \frac{f(x + \epsilon x', y + \epsilon y') - f(x, y + \epsilon y')}{\epsilon} + \lim_{\epsilon \to 0} \frac{f(x, y + \epsilon y') - f(x, y)}{\epsilon}$$

$$\frac{df}{dt} = x' \lim_{\epsilon \to 0} \frac{f(x + \epsilon x', y + \epsilon y') - f(x, y + \epsilon y')}{\epsilon x'} + y' \lim_{\epsilon \to 0} \frac{f(x, y + \epsilon y') - f(x, y)}{\epsilon y'}$$

$$= x'\frac{\partial f}{\partial x} + y'\frac{\partial f}{\partial y}$$

Taylor's Theorem in multiple dimensions

$$f(\mathbf{x} + \delta\mathbf{x}) = f(\mathbf{x}) + \delta\mathbf{x}^T \nabla f(\mathbf{x}) + \frac{1}{2}\delta\mathbf{x}^T H(\mathbf{x}) d\mathbf{x} + \mathcal{O}(\delta\mathbf{x}^3)$$

Where $H(\mathbf{x})$ is the Hessian (derivative matrix)

If a smooth function $f(\mathbf{x})$ has a local extreme at $\mathbf{x}$ then

$$\nabla f(\mathbf{x}) = \mathbf{0}$$

A sufficient condition for the extreme $\mathbf{x}$ to be a local minimum is for the quadratic form

$$Q(\delta\mathbf{x}) = \delta\mathbf{x}^T H(\mathbf{x})\delta\mathbf{x}$$

Is strictly positive definite
Where is is positive definite if

$$Q(\mathbf{x}) > 0$$

For all $\mathbf{x} \neq 0$. Also if every eigenvalue of $H(\mathbf{x}) > 0$. Can also use the principal minors.

Maxima of $f(\mathbf{x})$ are minima of $-f(\mathbf{x})$. Thus far we've assumed the functions are differentiable. If a function is continuous in an interval, it must achieve a maximum and a minimum

2 Calculus of Variations

We are not maximising the value of a function. We are maximising **functionals** - functions which take functions as arguments. We can think of it as an ∞- dimensional maximisation problem - we can choose different functions, which are not strictly constrained. This may take some effort.

A **functional** maps an element of a vector space (containing functions) to a real number. For some examples

$$F\{y(x)\} = |y(0)|$$
$$= \max_x\{y(x)\}$$
$$= \frac{dy}{dx}\Big|_{x=1}$$
$$= \sum_{n=0}^{N} a_n y(n)$$

Note that all of these map to a value - and don't really care about x.

Integral functionals are more interesting (and more useful).

$$F\{y\} = \int_a^b y(x)dx$$
$$= \int_a^b f(x)y(x)dx$$
$$= \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2}dx$$

The last one for example is the arc-length in a plane. So this could be insightful.

A simple example

$$F\{y\} = \int_{a(\epsilon)}^{b(\epsilon)} y(x, \epsilon)dx$$
$$\frac{dF}{d\epsilon} = y(b, \epsilon)\frac{db}{d\epsilon} - y(a, \epsilon)\frac{da}{d\epsilon} + \int_{a(\epsilon)}^{b(\epsilon)} \frac{\partial y(x, \epsilon)}{\partial\epsilon}dx$$

Since there is variation in the end points we have to evaluate the effect of the derivative of the end points (Leibniz rule).

If a and b are fixed then

$$\frac{da}{d\epsilon} = 0$$
$$\frac{db}{d\epsilon} = 0$$

And hence

$$\frac{dF}{d\epsilon} = \int_a^b \frac{\partial y(x, \epsilon)}{\partial \epsilon} dx$$

Consider the Brachistochrone, which involves the functional

$$F\{y\} = \int_{x_0}^{x_1} \sqrt{\frac{1 + y'^2}{y}} dx$$

Let us guess the brachistochrone takes form $y(x, \epsilon) = (1 - x)^\epsilon$. We would then use

$$\frac{dF}{d\epsilon} = 0$$

To obtain extrema.

The issue is the function may not have taken the form $(1 - x)^\epsilon$. And so this is where things get interesting.

2.1 Lagrange Multipliers

Given we need to find a minimum (or maximum) of $f(\mathbf{x})$ for $\mathbf{x} \in \mathbb{R}^\ltimes$ subject to

$$g_i(\mathbf{x}) = 0, \quad i, 1, \ldots, m < n$$

The conditions define a manifold (subset of $\mathbf{x} \in \mathbb{R}^\ltimes$). Solving this requires **Lagrange Multipliers**. Minimise (or maximise) a new function (of $m + n$ variables)

$$h(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x})$$

Where λ_i are the undetermined Lagrange multipliers.

So we take the partial derivatives of h as we would normally and get the zeros.

Consider the problem $f(\mathbf{x}) = x_1 x_2$ w $x_1 + x_2 = 1$, $x_1, x_2 > 0$, want to maximise f. So

$$h = f + \lambda g = x_1 x_2 + \lambda(x_1 + x_2 - 1)$$

$$\nabla h = \begin{pmatrix} \frac{\partial h}{\partial x_1} \\ \frac{\partial h}{\partial x_2} \\ \frac{\partial h}{\partial \lambda} \end{pmatrix} = \begin{pmatrix} x_2 + \lambda \\ x_1 + \lambda \\ x_1 + x_2 - 1 \end{pmatrix} = \mathbf{0}$$

And we get $x_1 = -\lambda, x_2 = -\lambda$, $x_1 = x_2 = \frac{1}{2}$, and $\lambda = -1/2$.

For classification we can consider the hessian (ignoring the λ component?)

$$H(h) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Which is not positive definite. But that is due to the $x_1 + x_2 = 1$ and hence the only perturbation vectors have form $(\delta x, -\delta x)^T$

$$(\delta x, -\delta x) H(h)(\delta x, -\delta x)^T = -2\delta x^2 < 0$$

Hence for all δx $f(\mathbf{x} + \delta \mathbf{x}) < f(\mathbf{x})$ and we have a local maximum.

2.2 Slack Variables

Now if we have inequalities rather than equalities, i.e

$$g_i(\mathbf{x}) \geq 0$$

Then you add slack variables, to make the constraints equalities. In this case we would have a $\mathbf{y} < 0$ so that

$$g_i(\mathbf{x}) + \mathbf{y} = 0$$

Or a $\mathbf{y} > 0$

$$g_i(\mathbf{x}) - \mathbf{y} = 0$$

So now

$$h = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i(g_i(\mathbf{x}) - y_i)$$

2.3 Vector Spaces

A vector space S is a collection of objects (vectors) defined with two operators (addition + scalar multiplication) which is closed under both operators (addition and scalar multiplication).

The operators have various properties (addition is commutive, associative, identity, inverse, distributive under scalar mult, multiplicative identity, associative scalar mult)

Some examples of vector spaces: the set of vectors $\mathbf{x} \in \mathbb{R}^n$ with vector addition and scalar multiplication

The set of all continuous functions.

The set of square integrable functions, L^2.

2.4 Normed Space

A norm on a vector space S is a real-valued function(al) whose value at $x \in S$ is denoted $||x||$ and has properties

$$||x|| \geq 0$$
$$||x|| = 0 \, iff \, x = 0$$
$$||\alpha x|| = \alpha ||x||$$
$$||x + y|| \leq ||x|| + ||y||$$

Where the last is the triangle inequality.

A vector space with a norm is a normed vector space.

For example $x \in \mathbb{R}^n$ with the Euclidean norm

$$||x||_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$$

Alternatively could use

$$||x||_1 = \sum_{i=1}^{n} |x_i|$$

The continuous function vector space $C[x_0, x_1]$ can have

$$||f||_\infty = \sup_{x \in [x_0, x_1]} |f(x)|$$

$$||f||_1 = \int_{x_0}^{x_1} |f(x)| dx$$

$$||f||_2 = \sqrt{\int_{-\infty}^{\infty} f(x)^2 dx}$$

L^2 can have the norm

$$||f||_2 = \sqrt{\int_{-\infty}^{\infty} f(x)^2 dx}$$

Let $C^n[x_0, x_1]$ be the set of functions with at least n continuous derivatives on $[x_0, x_1]$. Note

$$C^n[x_0, x_1] \subset C^{n-1}[x_0, x_1] \subset \ldots \subset C[x_0, x_1]$$

It can have $||f||_\infty, ||f||_1, ||f||_2$ among others.

Two norms $|| \cdot ||_a$ and $|| \cdot ||_b$ are equivalent if there exists positive numbers α, β such that for all $x \in S$

$$\alpha ||x||_a \leq ||x||_b \leq \beta ||x||_a$$

We will denote a normed vector space $(S, || \cdot ||)$

In finite dimensional spaces all norms are equivalent, but not in infinite dimensional spaces

Norms define distances between elements of space

Distances defines the ϵ- neighbourhood on $(S, || \cdot ||)$

2.5   Inner Product

An inner product $\langle \cdot, \cdot \rangle : S \times S \to \mathbb{R}$ i.e. mapping two elements of a vector space S to a real number, such that for any $f, g, h \in S$ and $\alpha \in \mathbb{R}$

$$\langle f, f \rangle \geq 0$$
$$\langle f, f \rangle = 0 \ iff \ f = 0$$
$$\langle f + g, h \rangle = \langle f, h \rangle + \langle g, h \rangle$$
$$\langle f, g \rangle = \langle g, f \rangle$$
$$\langle \alpha f, g \rangle = \alpha \langle f, g \rangle$$

A vector space with an inner product is an inner product space.

2.6   Functionals

Previous functionals were not interesting. integral functionals are more interesting. I.e. a function which contains the integral of another function.

Example- a hanging wire held up by points y_0, y_1. The potential energy is

$$W_p\{y\} = \int_0^L mgy(s)ds$$

Where L is the length of the cable, m is the mass per unit length and g is the gravitational constant. We want to minimise $W_p\{y\}$

Consider the brachistochrone - find a path between two points in 2D space that minimises the time taken for a rolling ball. Time taken is

$$T\{y\} = \int_0^L \frac{ds}{v(s)}$$

The energy of a body is the sum of potential and kinetic energy (assume no loss of energy)

$$E = \frac{1}{2}mv(x)^2 + mgy(x)$$

And simple conservation says E is constant, and so

$$v(x) = \sqrt{\frac{2E}{m} - 2gy(x)}$$

Example of a bent beam - two fixed end points and clamped so that

$$y(0) = 0, \ y'(0) = 0, \quad y(d) = 0, \ y'(d) = 0$$

The load on the beam at a given point x is given by $\rho(x)$

Let y denote the shape of the beam and ρ be the load per unit length on the beam (at x). For a bent elastic beam the potential energy from elastic forces is

$$V_1 = \frac{\kappa}{2} \int_0^d y''^2 dx$$

Where κ is flexural rigidity.

The potential energy is

$$V_2 = - \int_0^d \rho(x)y(x)dx$$

And hence the total potential energy is

$$V = V_1 + V_2 = \int_0^d \left(\frac{\kappa y''^2}{2} - \rho(x)y(x) \right) dx$$

3 Euler-Lagrange

3.1 Fixed endpoint problems

The simplest functional maximisation problem - we will solve by finding the **first variation** and then deriving the **Euler-Lagrange** equation:

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'} \right) - \frac{\partial f}{\partial y} = 0$$

Where f is the integrand of the functional. If we go back to the Catenary problem - a hanging wire held up by points y_0, y_1. The potential energy is

$$W_p\{y\} = \int_0^L mgy(s)ds$$

Where L is the length of the cable which is unconstrained (i.e. the cable can slack between the two points).

Define the functional $F : C^3[x_0, x_1] \to \mathbb{R}$.

$$F\{y\} = \int_{x_0}^{x_1} f(x, y, y')dx$$

Where f is assumed to be a function with (at least) continuous second-order partial derivatives with respect to x, y, y'.

We don't know how to evaluate $W_p\{y\}$ directly, so we should do a change of variables. The length of a line segment from (x, y) to $(x + \delta x, y + \delta y)$ is

$$\delta s \simeq \sqrt{\delta x^2 + \delta y^2}$$
$$= \sqrt{1 + \left(\frac{\delta y}{\delta x}\right)^2} \delta x$$
$$= \sqrt{1 + y'^2} dx$$

$$W_p = \int_0^L mgy(s)ds$$
$$= mg \int_{x_0}^{x_1} y\sqrt{1 + y'^2}dx$$
$$F\{y\} = mg \int_{x_0}^{x_1} y\sqrt{1 + y'^2}dx$$

So $f(x, y, y') = y\sqrt{1 + y'^2}$

Consider perturbations to a function $y(x)$ where η is a function change (such that it vanishes at the end points) and ϵ is a scaling factor Then we could have

$$\hat{y} = y + \epsilon\eta$$

We want y to be the maximum/minimum extremal.

We want to determine $y \in C^2[x_0, x_1]$ such that F has a local extremum.

So the space of possible curves for this problem is

$$S = \{y \in C^2[x_0, x_1] | y(x_0) = y_0, y(x_1) = y_1\}$$

And the vector space of allowable perturbations is

$$\mathcal{H} = \{\eta \in C^2[x_0, x_1] | \eta(x_0) = 0, \eta(x_1) = 0\}$$

Regard f as a function of 3 variables, x, y, y'. Take

$$\hat{y} = y + \epsilon\eta, \quad y \in S, \quad \eta \in \mathcal{H}$$
$$\hat{y}' = y' + \epsilon\eta'$$

By Taylor's theorem (2D version)

$$f(x, \hat{y}, \hat{y}') = f(x, y, y') + \epsilon\left[\eta\frac{\partial f}{\partial y} + \epsilon\frac{\partial f}{\partial y'}\right] + \mathcal{O}(\epsilon^2)$$

So

$$F\{\hat{y}\} - F\{y\} = \int_{x_0}^{x_1} f(x, \hat{y}, \hat{y}')dx - \int_{x_0}^{x_1} f(x, y, y')dx$$

$$= \int_{x_0}^{x_1} f(x, y, y') + \epsilon \left[\eta \frac{\partial f}{\partial y} + \eta' \frac{\partial f}{\partial y'} \right] + \mathcal{O}(\epsilon^2)dx - \int_{x_0}^{x_1} f(x, y, y')dx$$

$$= \epsilon \int_{x_0}^{x_1} \left[\eta \frac{\partial f}{\partial y} + \eta' \frac{\partial f}{\partial y'} \right] + \mathcal{O}(\epsilon^2)dx$$

This is the **First Variation**.

For $F\{y\}$ to be a minimum, for small ϵ, $F\{\hat{y}\} \geq F\{y\}$, so the sign of $\delta F(\eta, y)$ is determined by ϵ. And of course we can vary the sign of ϵ so for $F\{y\}$ to be a local extremum it must be true that

$$\delta F(\eta, y) = 0, \quad \forall \eta \in \mathcal{H}$$

This condition is analogous to all partial derivatives being zero. For a function of N variables to have a local extrema

$$\frac{\partial f}{\partial x_i} = 0, \quad \forall i = 1, \ldots, N$$

For a functional to be an extrema

$$\delta F(\eta, y) = \frac{d}{d\epsilon} F(y + \epsilon \eta) \Big|_{\epsilon=0} = 0$$

So now we have to minimise over an infinite dimensional space...

Look back at δF and integrate by parts

$$\delta F = \int_{x_0}^{x_1} \left[\eta \frac{\partial f}{\partial y} + \eta' \frac{\partial f}{\partial y'} \right] dx$$

$$= \left[\eta \frac{\partial f}{\partial y'} \right]_{x_0}^{x_1} + \int_{x_0}^{x_1} \left[\eta \frac{\partial f}{\partial y} - \eta \frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right) \right] dx$$

$$= \left[\eta \frac{\partial f}{\partial y'} \right]_{x_0}^{x_1} + \int_{x_0}^{x_1} \eta \left[\frac{\partial f}{\partial y} - \frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right) \right] dx$$

$$\delta F = \int_{x_0}^{x_1} \eta \left[\frac{\partial f}{\partial y} - \frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right) \right] dx = \langle \eta, E \rangle$$

Where $E = \frac{\partial f}{\partial y} - \frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right)$ Using $\eta(x_0) = \eta(x_1) = 0$.

Since we need $\delta F = 0$, and we can pick any $\eta \in \mathcal{H}$, the inner product will only be zero if

$$E = 0 \implies \frac{\partial f}{\partial y} - \frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right) = 0$$

Which is the Euler-Lagrange first variation.

Consider the shortest distance between 2 points in 2D $(0, 0)$ and $(1, 1)$.

The length of a line segment from x to $x + \delta x$ is

$$\delta s = \sqrt{\delta x^2 + \delta y^2}$$

$$= \sqrt{1 + \left(\frac{\delta y}{\delta x} \right)^2} dx \qquad\qquad = \sqrt{1 + y'^2} dx$$

So we want to minimise

$$F\{y\} = \int_0^1 \sqrt{1 + y'^2} dx$$

So $f = \sqrt{1 + y'^2}$.

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y} = 0$$

$$\frac{d}{dx}\left(\frac{y'}{1 + y'^2}\right) - 0 = 0$$

$$\frac{y'}{\sqrt{1 + y'^2}} = const$$

$$\implies y' = c_1$$

$$\implies y = c_1 x + c_2$$

4   Special Cases

There are some special cases for which the Euler-Lagrange equation simplifies and makes our life easier. If

- f depends only on y' - EL becomes

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0$$

- f has no explicit dependence on x (autonomous)

- f has no explicit dependence on y

- $f = A(x, y)y' + B(x, y)$ (linear degenerate case)

$$\frac{\partial^2 f}{\partial y' \partial x} + \frac{\partial^2 f}{\partial y' \partial y}y' + \frac{\partial^2 f}{\partial y'^2}y'' - \frac{\partial f}{\partial y} = 0$$

4.1   Only y' dependence

So when f depends only on f' the EL equation simplifies to

$$\frac{\partial f}{\partial y'} = const$$

An example of this is calculating geodesics in the plane (which we have shown are all straight lines). We get this by

$$\frac{d}{dx}\frac{\partial f}{\partial y'} = 0$$

$$\frac{\partial^2 f(y')}{\partial y'^2}\frac{dy'}{dx} = 0$$

$$\frac{\partial^2 f(y')}{\partial y'^2}y'' = 0$$

If $f''(y') = 0$ then $f(y') = ay' + b$. if $y'' = 0$ then

$$y = c_1 x + c_2$$

So for non-degenerate problems with only y' dependence, the extremals are straight lines.

E.g.

$$F\{y\} = \int_0^3 \alpha y'^4 - \beta y'^2 dx$$

With $y(0) = 0$ and $y(1) = b$. The EL equation is

$$\left[4\alpha y'^3 - 2\beta y'\right] = 0$$

And we know from above that the solutions are of form

$$y = c_1 + c_2 x$$

And using BCs gives

$$y = bx$$

4.2 No dependence on y

Consider Fermat's principle -

> Light travels along a path between any two points such that the time taken is minimised

Take the speed of light to be dependent on the media: $c = c(x, y)$. The time taken by light along a path $y(x)$ is

$$T\{y\} = \int_{x_0}^{x_1} \frac{\sqrt{1 + y'^2}}{c(x, y)} dx$$

Fermat's principle says the path of light will be a minima for this functional.

Lets assume $c(x, y) = \frac{1}{g(x)}$. I.e. making $c(x, y) := c(x)$.

$$T\{y\} = \int_{x_0}^{x_1} g(x)\sqrt{1 + y'^2} dx$$

$$\implies f(x, y') = g(x)\sqrt{1 + y'^2} dx$$

E-L:

$$\frac{\partial f}{\partial y'} = const$$

$$g(x)\frac{y'}{\sqrt{1 + y'^2}} = c_1$$

$$\frac{y'^2}{1 + y'^2} = \frac{c_1^2}{g^2}$$

$$y'^2 = \frac{c_1^2}{g^2}(1 + y'^2)$$

$$y'^2 \left(1 - \frac{c_1^2}{g^2}\right) = \frac{c_1^2}{g^2}$$

$$y' = \sqrt{\frac{c_1^2}{g^2 - c_1^2}}$$

$$y = \int \frac{1}{\sqrt{g^2/c_1^2 - 1}} dx + c_2$$

These are not necessarily straight lines.

You can use this and the fact that in any medium the speed of light is constant to derive Snell's law for refraction.

4.3 No dependence on x

When f has no dependence on x (the problem is autonomous). Then we can replace the EL equation with

$$H(y, y') = y' \frac{\partial f}{\partial y'} - f(y, y') = const$$

H is usually a conserved quantity like energy. This is also called the Beltrami identity.

Theorem Let J be a functional of form

$$J\{y\} = \int_{x_1}^{x_2} f(y, y') dx$$

And define H by

$$H(y, y') = y' \frac{\partial f}{\partial y'} - f(y, y')$$

Then H is constant along any extremal of y.

$$\begin{aligned}
\frac{d}{dx} H(y, y') &= \frac{d}{dx}\left(y' \frac{\partial f}{\partial y'} - f(y, y') \right) \\
&= \frac{d}{dx}\left(y' \frac{\partial f}{\partial y'} - f(y, y') \right) \\
&= y'' \frac{\partial f'}{\partial y} + y' \frac{d}{dx}\frac{\partial f'}{\partial y} - y' \frac{\partial f}{\partial y} - y'' \frac{\partial f}{\partial y'} \\
&= y'\left(\frac{d}{dx}\frac{\partial f}{\partial y'} - \frac{\partial f}{\partial y} \right) \\
&= 0
\end{aligned}$$

So $H(y, y') = const$

So

$$y' \frac{\partial f}{\partial y'} - f = const$$

Is the first order ODE for extremals to J

If we go back to the Catenary example

$$W_p\{y\} = \int_0^L mgy(s) ds$$

As with the geodesics in the plane

$$ds = \sqrt{1 + y'^2} dx$$

So the functional of interest is:

$$W_p\{y\} = mg \int_{x_0}^{x_1} y\sqrt{1 + y'^2} dx$$

Which doesn't explicitly contain x so

$$H(y, y') = y' \frac{\partial f}{\partial y'} - f = const$$

Where $f = y\sqrt{1 + y'^2}$

$$c_1 = y' \frac{\partial f}{\partial y'} - f$$
$$= y'y \frac{y'}{1 + y'^2} - y\sqrt{1 + y'^2}$$
$$c_1\sqrt{1 + y'^2} = yy'^2 - y(1 + y'^2)$$
$$c_1\sqrt{1 + y'^2} = -y$$
$$y^2 = c_1^2(1 + y'^2)$$

If $c_1 = 0$ then $y = 0$. Otherwise

$$y'^2 = \left(\frac{y}{c_1}\right)^2 - 1$$

$$\implies y' = \pm\sqrt{\left(\frac{y}{c_1}\right)^2 - 1}$$

Which is separable

$$\int \frac{1}{\sqrt{\left(\frac{y}{c_1}\right)^2 - 1}} dy = \int dx$$

$$\int \frac{1}{\sqrt{u^2 - 1}} \frac{dx}{du} du = x - c_2$$

$$c_1 \cosh^{-1}(u) = x - c_2$$

$$\implies y = c_1 \cosh\left(\frac{x - c_2}{c_1}\right)$$

c_1 and c_2 are determined by the end point conditions. Note we have not specified the length of the wire L at all here.

If we set $x_0 = -1$ and $x_1 = 1$ and $y_1 = y_2$ (i.e. symmetric around 0 and same heights) then $c_2 = 0$.

$$y(x) = c_1 \cosh(x/c_1)$$

Which we can then solve numerically using $y(1) = c_1 \cosh(1/c_1) = y_1$.

But if we try for example $y_1 = 2$, then we get two possible values $c_1 = 0.47$ and $c_1 = 1.697$. The don't both have to be minima. We would have to test both to see which one is the minimum.

There are possible outcomes which give a unique solution or even none.

Once we know y we can calculate $F\{y\}$. In this case use

$$\frac{d}{dx} c_1 \cosh(x/c_1) = \sinh(x/c_1)$$

$$1 + \sinh^2(x/c_1) = \cosh^2(x/c_1)$$

$$\cosh^2(x) = (\cosh(2x) + 1)/2$$

Then we get

$$F\{y\} = \int_{-1}^{1} y\sqrt{1 + y'^2}\,dx$$
$$= \int_{-1}^{1} c_1 \cosh(x/c_1)\sqrt{1 + \sinh^2(x/c_1)}\,dx$$
$$= \int_{-1}^{1} c_1 \cosh^2(x/c_1)\,dx$$
$$= \frac{c_1}{2}\int_{-1}^{1} \cosh(\frac{2x}{c_1} + 1)\,dx$$
$$= c_1 + \frac{c_1^2}{4}[\sinh(2x/c_1)]_{-1}^{1}$$
$$= c_1 + \frac{c_1^2}{2}\sinh(\frac{2}{c_1})$$

We want to optimise this wrt c_1.

The length of the Catenary

$$L\{y\} = \int_{-1}^{1} ds$$
$$= \sqrt{1 + y'^2}\,dx$$
$$= \int_{-1}^{1} \cosh(x/c_1)\,dx$$
$$= 2c_1 \sinh(1/c_1)$$
$$\simeq 2c_1(\frac{1}{c_1} + \frac{(1/c_1)^3}{6} + \ldots)$$

We can't actually set the length, it is an output. We can constrain the length later such that it is an input to the problem.

The usual explanation for the shape of the catenary is that **forces must be balanced in equilibrium**. The tension in the cable must balance the horizontal force F_H and the downwards force F_v. Giving

$$\tan\theta = \frac{F_V}{F_H}$$
$$\frac{dy}{dx} = \frac{gms}{F_H}$$

You can do some algebra (take derivatives wrt x) giving

$$y(x) = \frac{F_H}{mg}\cosh\left(\frac{mg}{F_H}x\right)$$

If you flip a catenary upside down, you get the strongest form of an arch, since the forces at each point are balanced. This assumes that there are no forces otherwise though. So this doesn't really hold true (you actually get a parabola).

4.4   Autonomous problems contd.

Look at the EL equation for extremals

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y} = 0$$

In the autonomous case

$$\frac{\partial f}{\partial x} = 0$$

So we get

$$H(y, y') = y'\frac{\partial f}{\partial y'} - f(y, y') = const$$

Look at the Brachistochrone again. The time taken is

$$T\{y\} = \int_0^L \frac{ds}{v(s)}$$

The energy of a body is the sum of potential and kinetic energy

$$E = \frac{1}{2}mv(x)^2 + mgy(x)$$

And conservation of energy makes E constant, so

$$v(x) = \sqrt{\frac{2E}{m} - 2gy(x)}$$

$$T\{y\} = \int_0^L \frac{ds}{v(s)}$$
$$= \int_0^L \frac{\sqrt{1 + y'^2}}{\sqrt{\frac{2E}{m} - 2gy(x)}}dx$$

Define

$$w = \frac{1}{2g}\left(\frac{2E}{m} - 2gy\right) \implies w'^2 = y'^2$$

$$T\{w\} = \frac{1}{\sqrt{2g}}\int_{x_0}^{x_1} \sqrt{\frac{1 + w'^2}{w}}dx$$

Which is autonomous.

$$H = w'\frac{\partial f}{\partial w'} - f$$
$$= \frac{w'^2}{w}\left(\frac{1 + w'^2}{w}\right)^{-1/2} - \sqrt{\frac{1 + w'^2}{w}}$$
$$= \frac{w'^2}{\sqrt{w(1 + w'^2)}} - \sqrt{\frac{1 + w'^2}{w}}\frac{\sqrt{1 + w'^2}}{\sqrt{1 + w'^2}}$$
$$= \frac{-1}{\sqrt{w(1 + w'^2)}} = const$$
$$\implies w(1 + w'^2) = c_1$$

For $w' = \tan\phi$ so that $1 + w'^2 = \sec^2\phi$, Let $\kappa_1 = \frac{c_1}{2}$

$$w = \frac{c_1}{\sec^2 \phi}$$
$$= c_1 \cos^2 \phi = \kappa_1 [1 + \cos 2\phi]$$
$$\frac{dw}{d\phi} = -2\kappa_1 \sin 2\phi = -4\kappa_1 \sin \phi \cos \phi$$

So $\frac{dw}{dx} = \tan \phi \implies \frac{dx}{dw} = \cot \phi$

$$\frac{dx}{d\phi} = \frac{dx}{dw}\frac{dw}{d\phi} = -4\kappa_1 \sin \phi \cos \phi \cot \phi = -4\kappa_1 \cos^2 \phi = -2\kappa_1(1 + \cos 2\phi)$$

And hence

$$x = \int \frac{dx}{d\phi} d\phi = \kappa_2 - \kappa_1(2\phi + \sin 2\phi)$$

With

$$w = \kappa_1(1 + \cos 2\phi)$$

So we have a parametric solution $(x(\phi), w(\phi))$ - since w is a linear transformation of y.

To simplify the solution, take $\theta + \pi = 2\phi$

$$x = \kappa_2 + \kappa_1(\theta - \sin(\theta))$$
$$w = \kappa_1(1 - \cos(\theta))$$

Recall $w = \frac{1}{2g}(\frac{2E}{m} - 2gy)$ where $E = \frac{1}{2}mv^2 + mgy = const$. So

$$v(x_0) = 0$$
$$\implies E = mgy_0$$
$$y = y_0 - w$$

Where $y(x_0) = y_0$. So $w(\theta_0) = 0$, giving $\theta_0 = 0$, And $x(\theta_0) = x_0$ and hence $\kappa_2 = x_0$.

So we get

$$x = x_0 + \kappa_1(\theta - \sin \theta)$$
$$y = y_0 - \kappa_1(1 - \cos(\theta))$$

We know x_0, y_0. Just need to find κ_1 (which relates to the end point).

We have $y(x_1) = y_1$ and we need to find θ_1.

$$y_1 = y_0 - \kappa_1(1 - \cos \theta_1)$$
$$\theta_1 = \cos^{-1}\left(1 - \frac{y_0 - y_1}{\kappa_1}\right)$$
$$x_1 = x_0 + \kappa_1(\theta_1 - \sin(\theta_1))$$
$$x_1 = x_0 + \kappa_1(\cos^{-1}\left(1 - \frac{y_0 - y_1}{\kappa_1}\right) - \sin(\cos^{-1}\left(1 - \frac{y_0 - y_1}{\kappa_1}\right)))$$

And numerically solve for κ_1. Turns out there are multiple solutions, so we need to find the fastest - which we do by evaluating.

The meaning of H that we derived.

- H is a conserved quantity

- Can derive conservation laws mathematically (rather than deriving as physical laws)

- Later we will consider Noether's Theorem

4.5 Newton's Aerodynamic problem

> If in a rare medium, consisting of equal particles freely disposed at equal distances from each other, a globe and a cylinder described on equal diameter move with equal velocities in the direction of the axis of the cylinder, the resistance of the globe will be half as great as that of the cylinder... I reckon that this proposition will be not without application in the building of ships.

Consider finding the optimal shape of a rocket's nose cone in order that it creates the least resistance when passing through air.

Assumptions

- Air is thin, and composed of perfectly elastic particles

  - Particles will bounce off the nose cone with equal speed and equal angle of reflection and incidence
  - We ignore tangential friction
  - We ignore "non-Newtonian" effects such as those from the compression of the air

- As the rocket may rotate along its length, the nose cone must be circularly symmetric.

- The rocket's nose cone must have radius R at its base, and length L, and its shape should be convex.

  - its profile must be concave and non-increasing
  - $L/2R$ is called the **fineness ratio**
  - Bigger is better, though little gain for $> 5:1$.

Most realistic for high-altitude, supersonic flight (or even better - hypersonic).

It doesn't matter if the missile is moving or the medium is moving - so assume the latter for convenience. We can calculate the angle between the incident particle and the tangent to the surface by

$$\cot\theta = \tan(\pi/2 - \theta) = -y'$$

The angle of incident equals the angle of reflection. And the angle between the reflected particle and the vertical line is 2θ.

The velocity in the vertical direction after the collision is

$$s = v\cos 2\theta = v(1 - 2\sin^2\theta)$$

Use $Force = ma$ where a will be related to the change in velocity in the vertical direction:

$$a = v - s = 2v\sin^2\theta$$

Non-dimensionalise so that

$$2vm = 1$$

And hence

$$Force = \sin^2\theta = \frac{1}{1 + \cot^2\theta} = \frac{1}{1 + y'^2}$$

This is the force per particle. We need to integrate over the surface area. The surface area at radius x is

$$2\pi x\, dx$$

Scaling to remove irrelevant constants, the functional describing the resistance (Force) is

$$F\{y\} = \int_0^R \frac{x}{1 + y'^2} dx$$

With $y(0) = L$ and $y(R) = 0$, and $y' \leq 0$ and $y'' \geq 0$.

Note that we do not have any explicit dependence on y.

Plug into EL equation

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y} = 0$$

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0$$

$$\frac{d}{dx}\left(\frac{-2xy'}{(1 + y'^2)^2}\right) = 0$$

$$\frac{2xy'}{(1 + y'^2)^2} = c_1$$

$$2xy' = c_1(1 + y'^2)^2$$

$$x = \frac{c_1(1 + y'^2)^2}{2y'} \qquad x = c\left(u^3 + 2u + \frac{1}{u}\right)$$

Letting $c = -\frac{c_1}{2}$, $u = -y'$ So

$$\frac{dy}{du} = \frac{dy}{dx} \cdot \frac{dx}{du} = -uc\left(\frac{1}{u} - 3u^3 - 2u\right)$$

$$x(u) = c\left(u^3 + 2u + \frac{1}{u}\right)$$

$$y(u) = d + c(\log u - \frac{3}{4}u^4 - u^2)$$

But we can't use $x(u) = 0$. So we can't satisfy $y(0) = L$. If we look instead at

$$\frac{dx}{du} = 0$$

The minimum occurs for

$$u_0 = \frac{1}{\sqrt{3}}$$

We will use $u_0 = 1$ for reasons we will discuss in about 2 months time.

So we have $y(1) = L$ and $x(u_1) = R$, $y(u_1) = 0$. This gives a cone with a plane near the tip.

This seems counter-intuitive and so It may not be optimal, so we can try some other profiles.

A cylinder gives $F = \frac{1}{2}$. A cone gives $F = \frac{1}{4}$. A hemisphere $F = \frac{1}{4}$ which is the same as a cone.

For the frustum of a cone - i.e. what we have now. Where it is blunted for $x < a$ for some point a. I.e.

$$y' = \begin{cases} 0 & x < a \\ -L/(R - a) & x > a \end{cases}$$

And evaluating F gives $\simeq 0.191$ and we find the optimal a is

$$a \simeq 0.381966$$

If we look at the shape we obtained we get $F \simeq 0.187$ Where $x(u_0) \simeq 0.3509$

Note most of this had to be evaluated or determined numerically (e.g. in this case F is analytic but c and u_1 are not).

- Note that the frustum of a cone isn't much worse than the optimal shape

- Other shapes: ogive, Haack, ...

- In the context of bullets, a flattened end is called a **meplat**

  - Typically justified by

    * making bullets precise
    * Tips are hard to get just right
    * impact damage

  - They wouldn't do it if it wasn't working

  - High performance commercially available cartridges achieve a muzzle velocity of around 1,200 m/s (roughly Mach 3.5 at sea level)

Geodesics on non-planar objects such as the sphere are another example of a functional without explicit y dependence.

If $\frac{\partial^2 f}{\partial y'^2} \neq 0$ then we can rewrite the problem as

$$y' = g(x, c)$$

Lets look at the example as the geodesics on a sphere. Let $\phi =$ latitude and $\theta =$ longitude, Coordinates

$$x = \cos\theta \sin\phi$$
$$y = \sin\theta \sin\phi$$
$$z = \cos\phi$$

Using the chain rule(s)

$$dx = \frac{\partial x}{\partial \theta}d\theta + \frac{\partial x}{\partial \phi}d\phi = -\sin\theta \sin\phi d\theta + \cos\theta \cos\phi d\phi$$

$$dy = \frac{\partial y}{\partial \theta}d\theta + \frac{\partial y}{\partial \phi}d\phi = \cos\theta \sin\phi d\theta + \sin\theta \cos\phi d\phi$$

$$dz = \frac{\partial z}{\partial \theta}d\theta + \frac{\partial z}{\partial \phi}d\phi = -\sin\phi d\theta$$

$$ds^2 = dx^2 + dy^2 + dz^2 = \sin^2\phi d\theta^2 + d\phi^2$$

$$\int_0^1 ds = \int_{\phi_0}^{\phi_1} \left[1 + \sin^2\phi \left(\frac{d\theta}{d\phi}\right)^2\right]^{\frac{1}{2}} d\phi$$

So θ is like y, ϕ is like x, and so $\frac{d\theta}{d\phi} = \theta'$ is like y'.

Write an EL equation - Note that we have no explicit y (θ) dependence.

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y} = 0$$

$$\frac{\partial}{\partial \theta}\left(1 + \sin^2\phi\theta'^2\right)^{1/2} = c_1$$

$$\frac{\sin^2\phi\theta'}{(1 + \sin^2\phi\theta'^2)^{1/2}} = c_1$$

$$\frac{\sin^4\phi\theta'^2}{(1 + \sin^2\phi\theta'^2)} = c_1^2$$

If we note

$$\theta'^2\sin^4\phi \le \theta'^2\sin^2\phi \le 1 + \theta'^2\sin^2\phi$$

We now know that the left hand side is between $0, 1$, which means $c_1^2 \in [0, 1]$, and hence $c_1 \in [-1, 1]$. So lets use $c_1 = \sin\alpha$

$$\frac{\sin^4\phi\theta'^2}{(1 + \sin^2\phi\theta'^2)} = \sin^2\alpha$$

$$\sin^4\phi\theta'^2 = \sin^2\alpha(1 + \sin^2\phi\theta'^2)$$

$$\theta'^2\sin^2\phi(\sin^2\phi - \sin^2\alpha) = \sin^2\alpha$$

$$\theta' = \left[\frac{\sin^2\alpha}{\sin^2\phi(\sin^2\phi - \sin^2\alpha)}\right]^{1/2}$$

$$= \frac{\sin\alpha}{\sin\phi\sqrt{\sin^2\phi - \sin^2\alpha}}$$

$$\theta = \int \frac{\sin\alpha}{\sin\phi\sqrt{\sin^2\phi - \sin^2\alpha}}d\phi$$

$$= \int \frac{\sin\alpha}{\sin\phi\sqrt{\frac{\sin^2\phi\sin^2\alpha}{\sin^2\alpha} - \frac{\sin^2\alpha\sin^2\phi}{\sin^2\phi}}}d\phi$$

$$= \int \frac{\operatorname{cosec}^2\phi}{\sqrt{\operatorname{cosec}^2\alpha - \operatorname{cosec}^2\phi}}d\phi$$

Note that $\operatorname{cosec}^2 = 1 + \cot^2$

$$\theta = \int \frac{\operatorname{cosec}^2\phi}{\sqrt{\operatorname{cosec}^2\alpha - \operatorname{cosec}^2\phi}}d\phi$$

$$= \frac{1}{\cot\alpha}\int \frac{\operatorname{cosec}^2\phi}{\left(1 - \frac{\cot^2\phi}{\cot^2\alpha}\right)}d\phi$$

Let $u = \frac{\cot \phi}{\cot \alpha}$ and $du = \frac{\csc^2 \phi}{\cot \alpha} d\phi$ So

$$\theta = \int \frac{du}{(1 - u^2)^{1/2}} = \sin^{-1}\left(\frac{\cot \phi}{\cot \alpha}\right) - \beta$$

$$\sin(\theta + \beta) = \frac{\cot \phi}{\cot \alpha}$$

$$\sin \theta \cos \beta + \cos \theta \sin \beta = \frac{1}{\cot \alpha}\left(\frac{\cos \phi}{\sin \phi}\right)$$

$$\cot \alpha \sin \beta \cos \theta \sin \phi + \cot \alpha \cos \beta \sin \theta \sin \phi = \cos \phi$$

Lets call $\cot \alpha \sin \beta = A$ and $\cot \alpha \cos \beta = B$, and convert to cartesian coords.

$$Ax + By = z$$

4.6 Coordinate Transformations

Spherical coordinates in general

$$x = r \cos \theta \sin \phi$$
$$y = r \sin \theta \sin \phi$$
$$z = r \cos \phi$$

And

$$)($$

In multidimension If

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1(\mathbf{x}) \\ \vdots \\ y_n(\mathbf{x}) \end{pmatrix}$$

Then the jacobian is

$$\mathbf{J} =)($$

Sometimes we call the determinant of the Jacobian the Jacobian (good nomenclature right?). It gives the ratio of n-dimensional volume between the two coordinate systems.

The equivalent of the transformation into u in a 1D integral you would change

$$dx = \frac{dx}{du} du$$

General geodesics

$$L = \int ds = \int \sqrt{dx^2 + dy^2 + dz^2}$$

So we would change this using

$$dx = \frac{\partial x}{\partial u} du + \frac{\partial x}{\partial v} dv$$
$$dy = \frac{\partial y}{\partial u} du + \frac{\partial y}{\partial v} dv$$
$$dz = \frac{\partial z}{\partial u} du + \frac{\partial z}{\partial v} dv$$

$$dx^2 = \left(\frac{\partial x}{\partial u}\right)^2 du^2 + 2\frac{\partial x}{\partial u}\frac{\partial x}{\partial v}dudv + \left(\frac{\partial x}{\partial v}\right)^2 dv^2$$

$$dy^2 = \left(\frac{\partial y}{\partial u}\right)^2 du^2 + 2\frac{\partial y}{\partial u}\frac{\partial y}{\partial v}dudv + \left(\frac{\partial y}{\partial v}\right)^2 dv^2$$

$$dz^2 = \left(\frac{\partial z}{\partial u}\right)^2 du^2 + 2\frac{\partial z}{\partial u}\frac{\partial z}{\partial v}dudv + \left(\frac{\partial z}{\partial v}\right)^2 dv^2$$

So we could have either

$$L = \int \sqrt{P + 2Qv' + Rv'^2}du$$

Or

$$L = \int \sqrt{Pu'^2 + 2Qu' + R}dv$$

Where

$$P = \left(\frac{\partial x}{\partial u}\right)^2 + \left(\frac{\partial y}{\partial u}\right)^2 + \left(\frac{\partial z}{\partial u}\right)^2$$

$$Q = \frac{\partial x}{\partial u}\frac{\partial x}{\partial v} + \frac{\partial y}{\partial u}\frac{\partial y}{\partial v} + \frac{\partial z}{\partial u}\frac{\partial z}{\partial v}$$

$$R = \left(\frac{\partial x}{\partial v}\right)^2 + \left(\frac{\partial y}{\partial v}\right)^2 + \left(\frac{\partial z}{\partial v}\right)^2$$

4.7   Invariance of the E-L equations

Need to make sure that the extremals found using the E-L equations don't depend on the coordinate system. Obviously this is useful since coordinate transforms can drastically simplify a problem.

For instance take the coordinate transform

$$x = x(u, v)$$
$$y = y(u, v)$$

So the Jacobian for the transform can be written as

$$J =)($$

Treat u like the indep var and v like the dependent var. i.e. $v(u)$. Chain rule says

$$\frac{dx}{du} = \frac{du}{du}\frac{\partial x}{\partial u} + \frac{dv}{du}\frac{\partial x}{\partial v}$$

$$\frac{dx}{du} = \frac{\partial x}{\partial u} + \frac{\partial x}{\partial v}v'$$
$$\frac{dy}{du} = \frac{\partial y}{\partial u} + \frac{\partial y}{\partial v}v'$$

Transforming the functional we get

$$F\{y\} = \int_{x_0}^{x_1} f(x, y, y')dx$$

$$= \int_{u_0}^{u_1} f\left(x(u, v), v(u, v), \frac{\frac{\partial y}{\partial x} + \frac{\partial y}{\partial v}v'}{\frac{\partial x}{\partial x} + \frac{\partial x}{\partial v}v'}\right)(\frac{\partial x}{\partial u} + \frac{\partial x}{\partial v}v')du$$

$$= \int_{u_0}^{u_1} \hat{f}(u, v, v')$$

Relabel the functional as $\tilde{F}$ for the new transform. We want solutions of $\tilde{F}$ to be the same as those for F.

Theorem: Let $y \in S$ and $v \ in\tilde{S}$ be two functions that satisfy the smooth, non-singular transformation $x = x(u,v)$ and $y = y(u,v)$, then y is an extremal for F iff v is an extremal for $\tilde{F}$

Won't properly prove this. But basically we just show that the EL equations for both problems produce the same extremals.

As an example look at polar coordinates

$$x = r \cos \theta$$
$$y = r \sin \theta$$

With inverse

$$r = \sqrt{x^2 + y^2}$$
$$\theta = \arctan\left(\frac{y}{x}\right)$$

Find extremals of $F\{x\} = \int_{\theta_0}^{\theta_1} \sqrt{r^2 + r'^2}d\theta$

$$r_x = \frac{x}{\sqrt{x^2 + y^2}}, \quad r_y = \frac{y}{\sqrt{x^2 + y^2}}$$

Note that $\tan^{-1}(u) = \frac{1}{1+u^2}$ So

$$\theta_x = \frac{-\frac{y}{x^2}}{1 + \left(\frac{y}{x}\right)^2} = -\frac{y}{x^2 + y^2}$$

$$\theta_y = \ldots = \frac{x}{x^2 + y^2}$$

The Jacobian (determinant):

$$\mathbf{J} = \begin{vmatrix} r_x & \theta_x \\ r_y & \theta_y \end{vmatrix} = \frac{x^2 + y^2}{(x^2 + y^2)^{3/2}} = \frac{1}{\sqrt{x^2 + y^2}}$$

$$\frac{dr}{d\theta} = \frac{r_x + r_y y'}{\theta_x + \theta_y y'}$$

$$= \frac{x/\sqrt{x^2 + y^2} + yy'/\sqrt{x^2 + y^2}}{-y/(x^2 + y^2) + xy'/(x^2 + y^2)}$$

$$= \frac{x + yy'}{-y + xy'} \sqrt{x^2 + y^2}$$

Recall our functional was $\sqrt{r^2 + r'^2}$:

$$r^2 + r'^2 = x^2 + y^2 + \left(x^2 + y^2\right)\left(\frac{x + yy'}{-y + xy'}\right)^2$$

$$= \left(x^2 + y^2\right)\left(1 + \frac{x^2 + 2xyy' + y^2 y'^2}{y^2 - 2xyy' + x^2 y'^2}\right)$$

$$= \left(x^2 + y^2\right)\left(\frac{x^2 + y^2 + (x^2 + y^2)y'^2}{y^2 - 2xyy' + x^2 y'^2}\right)$$

$$= \left(x^2 + y^2\right)^2 \left(\frac{1 + y'^2}{(-y + xy')^2}\right)$$

$$\frac{d\theta}{dx} = \frac{\partial\theta}{\partial x} + \frac{\partial\theta}{\partial y}y'$$

$$= -\frac{y}{x^2+y^2} + \frac{xy'}{x^2+y^2}$$

$$\frac{\partial x}{\partial\theta} = \frac{x^2+y^2}{-y+xy'}$$

$$\implies r^2 + r'^2 = (1+y'^2)\left(\frac{dx}{d\theta}\right)^2$$

So the functional is

$$F\{r\} = \int_{\theta_0}^{\theta_1} \sqrt{r^2+r'^2}\,d\theta$$

$$= \int_{\theta_0}^{\theta_1} \sqrt{1+y'^2}\frac{dx}{d\theta}\,d\theta$$

$$\tilde{F}\{y\} = \int_{x_0}^{x_1} \sqrt{1+y'^2}\,dx$$

So that was a lot of work to derive the Cartesian form of geodesics from the polar form.

4.8   Degenerate Case

When $f = A(x,y)y' + B(x,y)$, this is the degenerate case. The EL equations reduce to

$$\frac{\partial A}{\partial x} - \frac{\partial B}{\partial y} = 0$$

Which we can't necessarily solve, and when they can be solved, the functional's value only depends on the end-points, not the actual shape of the curve.

$$F\{y\} = \int_{x_0}^{x_1} A(x,y)y' + B(x,y)dx$$

E-L gives

$$\frac{d}{dx}\frac{\partial f}{\partial y'} - \frac{\partial f}{\partial y} = 0$$

$$\frac{d}{dx}[A] - \frac{\partial A}{\partial y}y' - \frac{\partial B}{\partial y} = 0$$

$$\frac{\partial A}{\partial x} + \frac{\partial A}{\partial y}y' - \frac{\partial A}{\partial y}y' - \frac{\partial B}{\partial y} = 0$$

$$\frac{\partial A}{\partial x} - \frac{\partial B}{\partial y} = 0$$

Which is not a differential equation since we know A, B.

When there is a solution then there exists a function $\phi(x,y)$ s.t.

$$\frac{\partial\phi}{\partial y} = A$$

$$\frac{\partial\phi}{\partial x} = B$$

Thus

$$\frac{\partial A}{\partial x} = \frac{\partial^2 \phi}{\partial x \partial y} = \frac{\partial^2 \phi}{\partial y \partial x} = \frac{\partial B}{\partial y}$$

And hence $f(x, y)$ can be written as

$$f = \frac{d\phi}{dx}$$

And the functional is

$$\begin{aligned}
F\{y\} &= \int_{x_0}^{x_1} f(x, y, y')dx \\
&= \int_{x_0}^{x_1} \frac{d\phi}{dx}dx \\
&= [\phi]_{x_1} - [\phi]_{x_0}
\end{aligned}$$

As an example, let

$$f = (x^2 + 3y^2)y' + 2xy$$

Clearly this is a degen case

$$F\{y\} = \int_{x_0}^{x_1} \left[(x^2 + 3y^2)y' + 2xy\right] dx$$

E-L gives

$$\frac{\partial A}{\partial x} - \frac{\partial B}{\partial y} = 2x - 2x = 0$$

Get ϕ:

$$\phi = x^2 y + y^3 + k(x)$$
$$\frac{\partial \phi}{\partial x} = 2xy + k'(x)$$
$$\implies \phi = x^2 y + y^2 + B$$

We can choose $B = 0$.

$$\phi = x^2 y + y^3$$

So

$$\begin{aligned}
F\{y\} &= \phi(x_1, y(x_1)) - \phi(x_0, y(x_0)) \\
&= x_1^2 y_1 + y_1^3 - x_0^2 y_0 - y_0^3
\end{aligned}$$

I.e. the extremals of the function only depends on the end points.

5 Higher Derivatives and Multidimension

We have three extensions to the E-L equations so far

- f includes higher order derivatives $f = (x, y, y', , \ldots, y^{(n)})$

- Several dependent variables ($\mathbf{y}$)

- Several independent variables ($\mathbf{x}$)

5.1 Higher order derivatives

Normal E-L assumes $f(x, y, y')$. If we now include y'' Let $F : C^2[x_0, x_1] \to \mathbb{R}$

$$F\{y\} = \int_{x_0}^{x_1} f(x, y, y', y'') dx$$

Where f has continuous partial derivatives of second order wrt x, y, y', y'' and $x_0 < x_1$ are both fixed. We know that extremals are obtained for

$$\delta F$$

$$f(x, y + \epsilon\eta, y' + \epsilon\eta', y'' + \epsilon\eta'') = f(x, y, y', y'') + \epsilon(\eta\frac{\partial f}{\partial y} + \eta'\frac{\partial f}{\partial y'} + \eta''\frac{\partial f}{\partial y''}) + \mathcal{O}(\epsilon^2)$$

So

$$F\{y + \epsilon\eta\} = \int_{x_0}^{x_1} f(x, y + \epsilon\eta, y' + \epsilon\eta', y'' + \epsilon\eta'') dx$$

$$= \int_{x_0}^{x_1} f(x, y, y', y'') + \epsilon(\eta\frac{\partial f}{\partial y} + \eta'\frac{\partial f}{\partial y'} + \eta''\frac{\partial f}{\partial y''}) dx + \mathcal{O}(\epsilon^2)$$

First variation.

$$\delta F = \lim_{\epsilon \to 0} \frac{F\{y + \epsilon\eta\} - F\{y\}}{\epsilon}$$

$$= \int_{x_0}^{x_1} \eta\frac{\partial f}{\partial y} + \eta'\frac{\partial f}{\partial y'} + \eta''\frac{\partial f}{\partial y''} dx$$

$$= \left[\eta\frac{\partial f}{\partial y'} + \eta'\frac{\partial f}{\partial y''} \right]_{x_0}^{x_1} + \int_{x_0}^{x_1} \left(\eta\frac{\partial f}{\partial y} - \eta\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \eta'\frac{d}{dx}\left(\frac{\partial f}{\partial y''}\right) \right) dx$$

$$= \left[\eta\frac{\partial f}{\partial y'} + \eta'\frac{\partial f}{\partial y''} - \eta\frac{d}{dx}\left(\frac{\partial f}{\partial y''}\right) \right]_{x_0}^{x_1} + \int_{x_0}^{x_1} \eta\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) + \frac{d^2}{dx^2}\left(\frac{\partial f}{\partial y''}\right) \right) dx$$

By integrating by parts (twice)

Denote the end points

$$y(x_0) = y_0, \quad y(x_1) = y_1, \quad y'(x_0) = y_0', \quad y'(x_1) = y_1'$$

This implies

$$\eta(x_0) = \eta(x_1) = \eta'(x_0) = \eta'(x_1) = 0$$

So the square bracket term is 0, i.e.

$$\left[\eta\frac{\partial f}{\partial y'} + \eta'\frac{\partial f}{\partial y''} - \eta\frac{d}{dx}\left(\frac{\partial f}{\partial y''}\right) \right]_{x_0}^{x_1} = 0$$

And so for an extremal

$$\delta F = \int_{x_0}^{x_1} \eta\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) + \frac{d^2}{dx^2}\left(\frac{\partial f}{\partial y''}\right) \right) dx$$

$$\implies \frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) + \frac{d^2}{dx^2}\left(\frac{\partial f}{\partial y''}\right) = 0$$

Known as the 4th order Euler-Lagrange / Euler-Poisson equation.

Let $F : C^2[x_0, x_1] \to \mathbb{R}$ be a functional of form

$$F\{y\} = \int_{x_0}^{x_1} f(x, y, y', \ldots, y^{(m)})dx$$

Where f has continuous partial derivatives of second order wrt $x, y, y', \ldots, y^{(m)}$ and $x_0 < x_1$ are both fixed.

$$\sum_{n=0}^{m} (-1)^n \frac{d^n}{dx^n}\left(\frac{\partial f}{\partial y^{(n)}}\right) = 0$$

This is sometimes called the Euler-Poisson equation

Example

$$F\{y\} = \int_0^1 (1 + y''^2)dx$$

$$y(0) = 0, \quad y(1) = 1, \quad y'(0) = 1, \quad y'(1) = 1$$

Since $F\{y\}$ does not explicitly contain x, y.

$$\frac{\partial f}{\partial y} = 0, \quad \frac{\partial f}{\partial y'} = 0 \implies \frac{d}{dx}(\frac{\partial f}{\partial y'}) = 0$$

And hence EL reduces to

$$\frac{d^2}{dx^2}\left(\frac{\partial f}{\partial y''}\right) = \frac{d^2}{dx^2}(2y'')$$

$$= 2y^{(4)}$$

$$\implies y = c_0 + c_1 x + c_2 x^2 + c_3 x^3$$

And using the end points

$$y(0) = 0 \implies c_0 = 0$$
$$y'(0) = 1 \implies c_1 = 1$$
$$y(1) = 1 \implies 1 + c_2 + c_3 = 1$$
$$y'(1) = 1 \implies 1 + 2c_2 + 3c_3 = 1$$
$$\implies c_2 = c_3 = 0$$

Hence

$$y = x$$

Example v2.

Consider an elastic beam with fixed end points and clamped s.t.

$$y(0) = y(d) = y'(0) = y'(d) = 0$$

The load per unit length on the beam is given by $\rho(x)$.

Let $y : [0, d] \to \mathbb{R}$ describe the shape of the beam, and $\rho : [0, d] \to \mathbb{R}$ be the load per unit length on the beam. For a bent elastic beam, the potential energy from elastic forces is

$$V_1 = \frac{\kappa}{2}\int_0^d y''^2 dx$$

Where κ denotes the flexural rigidity of the beam (a material property).

The potential energy is

$$V_2 = -\int_0^d \rho(x)y(x)dx$$

Hence the total potential energy is

$$V = V_1 - V_2 = \int_0^d \frac{\kappa y''^2}{2} - \rho(x)y(x)dx$$

So

$$f = \frac{\kappa y''^2}{2} - \rho(x)y(x)$$

$$\frac{\partial f}{\partial y} = -\rho(x)$$

$$\frac{\partial f'}{\partial y} = 0$$

$$\frac{\partial f''}{\partial y} = \kappa y''$$

So the EP equation gives

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) + \frac{d^2}{dx^2}\left(\frac{\partial f}{\partial y''}\right) = 0$$

$$-\rho + \kappa y^{(4)} = 0$$

$$y^{(4)} = \frac{\rho}{\kappa}$$

Note that κ is constant, but ρ may not be

$$\implies y(x) = P(x) + c_3 x^3 + c_2 x^2 + c_1 x + c_0$$

Where

$$P^{(4)}(x) = \frac{\rho(x)}{\kappa}$$

Assume $\rho(x) = \rho$ (constant) hence

$$y(x) = \frac{\rho x^4}{4!\kappa} + c_3 x^3 + c_2 x^2 + c_1 x + c_0$$

Using the boundary conditions

$$y(0) = 0 \implies c_0 = 0$$

$$y'(0) = 0 \implies c_1 = 0$$

$$y(d) = 0 \implies \frac{\rho d^4}{4!\kappa} + c_3 d^3 + c_2 d^2 = 0$$

$$y'(d) = 0 \implies \frac{\rho d^3}{3!\kappa} + 3c_3 d^2 + 2c_2 d = 0$$

Which gives

$$y(x) = \frac{\rho(d-x)^2 x^2}{24\kappa}$$

$$y'(x) = \frac{\rho x(d-x)(d-2x)}{12\kappa}$$

Which satisfies all the constraints.

5.2 Several dependent variables

Let $\mathbf{q} : [t_0, t_1] \to \mathbb{R}^n$ such that $\mathbf{q} = (q_1, \ldots, q_n)$ Where all $q_k \in C^2[t_0, t_1]$ We will use dot notation for time derivatives Define functionals for example

$$F\{\mathbf{q}\} = \int_{t_0}^{t_1} L(t, \mathbf{q}, \dot{\mathbf{q}}) dt$$

Where L has continuous 2nd order derivatives wrt t, $\mathbf{q}$, $\dot{\mathbf{q}}$. For the fixed end point problem, we look for

$$S = \{\mathbf{q} \in C^2[t_0, t_1] \mid \mathbf{q}(t_0) = \mathbf{q}_0, \quad \mathbf{q}(t_1) = \mathbf{q}_1\}$$

Note that L takes $2n + 1$ inputs.

As before look for extremals by taking perturbations of $\mathbf{q}$.

$$\tilde{\mathbf{q}} = \mathbf{q} + \epsilon \boldsymbol{\eta}$$

Where $\boldsymbol{\eta} \in \mathcal{H}^n$. Such that $\boldsymbol{\eta}(t_0) = \mathbf{0}$ and $\boldsymbol{\eta}(t_1) = \mathbf{0}$

Taylors theorem for this

$$L(t, \mathbf{q} + \epsilon\boldsymbol{\eta}, \dot{q} + \epsilon\dot{\eta}) = L(t, \mathbf{q}, \dot{\mathbf{q}}) + \epsilon \sum_{k=1}^{n} \left(\eta_k \frac{\partial L}{\partial q_k} + \dot{\eta}_k \frac{\partial L}{\partial \dot{q}_k} \right) + \mathcal{O}(\epsilon^2)$$

First variation

$$\delta F(\boldsymbol{\eta}, \mathbf{q}) = \frac{F\{\mathbf{q} + \epsilon\boldsymbol{\eta}\} - F\{\mathbf{q}\}}{\epsilon}$$

$$= \frac{1}{\epsilon} \int_{t_0}^{t_1} L(t, \mathbf{q} + \epsilon\boldsymbol{\eta}, \dot{\mathbf{q}} + \epsilon\dot{\eta}) - L(t, \mathbf{q}, \dot{\mathbf{q}})$$

$$= \int_{t_0}^{t_1} \sum_{k=1}^{n} \left(\eta_k \frac{\partial L}{\partial q_k} + \dot{\eta}_k \frac{\partial L}{\partial \dot{q}_k} \right) dt + \mathcal{O}(\epsilon)$$

$$= 0$$

For all $\eta \in \mathcal{H}^n$ as $\epsilon \to 0$. If we set $\boldsymbol{\eta}_k = (0, \ldots, \eta_k, \ldots, 0)$ For all k, we get the set of equations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_1} - \frac{\partial L}{\partial q_1} = 0$$

$$\vdots$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_n} - \frac{\partial L}{\partial q_n} = 0$$

Which is analogous to maximising a function of several variables by setting all the partial derivatives equal to 0. So we must simultaneously solve all of these equations.

Example

$$F\{\mathbf{q}\} = \int_0^1 \left(\dot{q}_1^2 + (\dot{q}_2 - 1)^2 + q_1^2 + q_1 q_2 \right) dt$$

With

$$\mathbf{q}(0) = \mathbf{q}_0, \quad \mathbf{q}(1) = \mathbf{q}_1$$

The EL equations are

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_1} - \frac{\partial L}{\partial q_1} = 0$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_2} - \frac{\partial L}{\partial q_2} = 0$$

$$\frac{\partial L}{\partial q_1} = 2q_1 + q_2 \qquad \frac{\partial L}{\partial q_2} = q_1$$
$$\frac{\partial L}{\partial \dot{q}_1} = 2\dot{q}_1 \qquad \frac{\partial L}{\partial \dot{q}_2} = 2(\dot{q}_2 - 1)$$

And hence the EL equations are

$$2\ddot{q}_1 - 2q_1 - q_2 = 0$$
$$2\ddot{q}_2 - q_1 = 0$$

Diff the second equation twice

$$2\dddot{\ddot{q}}_2 - \ddot{q}_1 = 0$$

sub into 1

$$4\dddot{\ddot{q}}_2 - 4\ddot{q}_2 - q_2 = 0$$

Characteristic equation:

$$4\lambda^4 - 4\lambda^2 - 1 = 0$$

$$\lambda^2 = \frac{1 \pm \sqrt{2}}{2}$$

Note that λ^2 could be positive or negative. Denote

$$\pm\lambda = \pm\sqrt{\frac{1 + \sqrt{2}}{2}}$$

$$\pm iu = \pm\sqrt{\frac{1 - \sqrt{2}}{2}}$$

And hence

$$q_2 = c_1 \cosh \lambda t + c_2 \sinh \lambda t + c_3 \cos ut + c_4 \sin ut$$

And

$$q_1 = 2\ddot{q}_2$$
$$= 2\lambda^2(c_1 \cosh \lambda t + c_2 \sinh \lambda t) - 2u^2 (c_3 \cos ut + c_4 \sin ut)$$

Example in 3d: the kinetic energy of a particle is

$$T = \frac{1}{2}mv^2(t) = \frac{1}{2}m \left(\dot{x}^2(t) + \dot{y}^2(t) + \dot{z}^2(t) \right)$$

Assume there exists $V(t, x, y, z)$ such that the forces acting on the particle in the direction of the cartesian axes are

$$f_x = -\frac{\partial V}{\partial x}, \quad f_y = -\frac{\partial V}{\partial y}, \quad f_z = -\frac{\partial V}{\partial z}$$

Where V is the potential energy of the particle.

The function
$$L = T - V$$
is called the Lagrangian, the path of a particle is given by $\mathbf{r}(t) = (x, y, z)$ over time $[t_0, t_1]$ We can define the action integral by

$$F\{r\} = \int_{t_0}^{t_1} L(t, \mathbf{r}, \dot{\mathbf{r}})dt$$

The path of a particle is such that F is stationary.

- Sometimes called the principle of least action

- Could be a saddle point (not just minima)

- Hamilton's principle is far more general

    - multiple particles

    - non-cartesian coordinates

    - recall changing coordinates won't change extremal curve

We could describe the mechanical system by generalised coordinates $\mathbf{q}(t)$

- The kinetic energy is given by $T(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2}\sum_{j,k=1}^{n} C_{j,k}(\mathbf{q})\dot{q}_j\dot{q}_k$

- The potential energy is given by $V(t, \mathbf{q})$

- The lagrangian is $L(t, \mathbf{q}, \dot{\mathbf{q}}) = T(\mathbf{q}, \dot{\mathbf{q}}) - V(t, \mathbf{q})$

Hamilton's principle states that the path of $\mathbf{q}(t)$ will give a stationary $F\{\mathbf{q}\}$.

Example of a simple pendulum with fixed length Kinetic energy:

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) = \frac{1}{2}ml^2\dot{\phi}^2$$

Potential energy

$$V = mg(l - y) = mgl(1 - \cos\theta)$$

The Lagrangian is

$$L(\phi, \dot{\phi}) = \frac{1}{2}ml^2\dot{\phi}^2 - mgl(1 - \cos\phi)$$

And hence the action integral is

$$F\{\phi\} = \int_{t_0}^{t_1} \left[\frac{1}{2}ml^2\dot{\phi}^2 - mgl(1 - \cos\phi)\right] dt$$

Example: Keplar's problem of planetary motion. Single planet orbiting the sun. Kinetic energy

$$T = \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2\right) = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\phi}^2)$$

Potential energy

$$V(t) = -\int f(r)dr = -\frac{GmM}{r}$$

Where (from Newton):

$$f = -\frac{dV}{dr} = -\frac{GmM}{r^2}$$

Hamilton's principle states that

$$F\{\mathbf{q}\} = \int_{t_0}^{t_1} L(t, \mathbf{q}, \dot{\mathbf{q}})dt$$

is stationary. The EL equations are

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} - \frac{\partial L}{\partial q_k} = 0$$

For all k. And so for mechanical systems, the Lagrangian will satisfy these equations.

Newton's Laws: If V depends only on location and time, and the kinetic energy depends only on the derivatives of the position. Under these circumstances, EL reduces to

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_k} + \frac{\partial V}{\partial q_k} = 0$$

Given kinetic energy of form $T(\dot{q}) = \frac{1}{2}m\sum_i \dot{q}_i^2$ then EL becomes

$$m\ddot{q}_k = -\frac{\partial V}{\partial q_k} = f_k$$

Where f_k is the force in the k direction. This is a derivation of $\mathbf{f} = m\mathbf{a}$ in more general.

If the potential does not depend on time, then we may form $H(\mathbf{q}, \dot{\mathbf{q}})$ as before

$$H(\mathbf{q}, \dot{\mathbf{q}}) = \sum_{k=1}^{n} \dot{q}_k \frac{\partial L}{\partial \dot{q}_k} - L = const$$

Given kinetic energy of the form $T(\dot{q}) = \frac{1}{2}m\sum_i \dot{q}_i^2$, then it becomes

$$H(\mathbf{q}, \dot{\mathbf{q}}) = 2T - L = T + V = const$$

If we go back to the simple pendulum example

$$F\{\phi\} = \int_{t_0}^{t_1} \left(\frac{1}{2}ml^2\dot{\phi}^2 - mgl(1 - \cos\theta)\right)dt$$

We have the appropriate form for the kinetic energy and we don't have time dependence in the potential, and hence conservation of energy holds:

$$\frac{1}{2}ml^2\dot{\phi}^2 + mgl(1 - \cos\phi) = const$$

Removing constants gives

$$\dot{\phi}^2 - \frac{2g}{l}\cos\phi = c_1$$

Diff wrt time

$$2\dot{\phi}[\ddot{\phi} + \frac{g}{l}\sin\phi] = 0$$
$$\ddot{\phi} + \frac{g}{l}\sin\phi = 0$$
$$ml^2\ddot{\phi} + gml\sin\phi = 0$$

Which is rate of change of angular momentum + torque = 0

If we pretend to be engineers and assume small ϕ

$$\ddot{\phi} + \frac{g}{l}\phi = 0$$

$$\phi = A\sin(\sqrt{\frac{g}{l}}t) + \phi_0$$

Period of oscillation is $2\pi\sqrt{g/l}$.

Example of the 3D Brachistochrone

Curve of fastest descent between the points (x_0, y_0, z_0) and (x_1, y_1, z_1) where z is height, and x, y spatial. Consider y, z to be functions of x. The time for the descent is

$$\sqrt{2g}T\{y, z\} = \int_{x_0}^{x_1} \frac{\sqrt{1 + y'^2 + z'^2}}{\sqrt{z_0 - z}}dx$$

EL equations are

$$\frac{d}{dx}\left(\frac{y'}{\sqrt{1 + y'^2 + z'^2}\sqrt{z_0 - z}}\right) = 0$$

$$\frac{d}{dx}\left(\frac{z'}{\sqrt{1 + y'^2 + z'^2}\sqrt{z_0 - z}}\right) - \frac{\sqrt{1 + y'^2 + z'^2}}{2(z_0 - z)^{3/2}} = 0$$

The first equation gives

$$\frac{y'}{\sqrt{1 + y'^2 + z'^2}} = c_1\sqrt{z_0 - z}$$

Since it is $x - indep$

$$-H = f - y'\frac{\partial f}{\partial y'} - z'\frac{\partial f}{\partial z'}$$

(we will show this later)

$$-H = \frac{\sqrt{1 + y'^2 + z'^2}}{\sqrt{z_0 - z}} - \frac{y'^2}{\sqrt{1 + y'^2 + z'^2}\sqrt{z_0 - z}} - \frac{z'^2}{\sqrt{1 + y'^2 + z'^2}\sqrt{z_0 - z}}$$

$$= \frac{1 + y'^2 + z'^2 - y'^2 - z'^2}{\sqrt{1 + y'^2 + z'^2}\sqrt{z_0 - z}} = c_2$$

And so we have

$$\frac{y'}{\sqrt{1 + y'^2 + z'^2}} = c_1\sqrt{z_0 - z}, \quad \frac{1}{\sqrt{1 + y'^2 + z'^2}} = c_2\sqrt{z_0 - z}$$

Clearly giving $y' = c_1/c_2 = const$

$$y = \frac{c_1}{c_2}(x - x_1) + y_1$$

Which is the vertical plane, so the problem reduces to the 2D Brach.

Kepler's problem of planetary motion A single planet orbiting the sun

$$L = T - V = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\phi}^2) + \frac{GmM}{r}$$

Hamilton's principle says we have to find stationary curves of the integral of L so jump into the EL equations

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

The first gives

$$mr\dot{\phi}^2 - \frac{GmM}{r} - \frac{d}{dt}(m\dot{r}) = 0$$
$$r\dot{\phi}^2 - \frac{GM}{r} - \frac{d}{dt}(\dot{r}) = 0$$
$$\ddot{r} - r\dot{\phi}^2 = -\frac{GM}{r^2}$$

The second simply gives

$$\frac{d}{dt}(r^2\dot{\phi}) = 0 \implies r^2\dot{\phi} = c$$

Use that in the first equation

$$\ddot{r} - \frac{c^2}{r^3} = -\frac{GM}{r^2}$$

Change r to a function of ϕ.

$$\dot{r} = \frac{dr}{d\phi}\frac{d\phi}{dt} = r'\dot{\phi} = \frac{cr'}{r^2}$$
$$\ddot{r} = \frac{d}{d\phi}\left(\frac{cr'}{r^2}\right)\dot{\phi} = \frac{c^2}{r^2}\left(\frac{r''}{r^2} - \frac{2r'^2}{r^3}\right)$$

Sub into the $\ddot{r}$ equation

$$\ddot{r} - r\dot{\phi}^2 = -\frac{GM}{r^2}$$
$$\frac{c^2}{r^2}\left(\frac{r''}{r^2} - \frac{2r'^2}{r^3}\right) - \frac{c^2}{r^3} = -\frac{GM}{r^2}$$
$$\frac{r''}{r^2} - \frac{2r'^2}{r^3} - \frac{1}{r} = -\frac{GM}{c^2}$$

Let $u = p/r$ so $u' = -\frac{pr'}{r^2}$ and $r'' = -\frac{pr''}{r^2} + \frac{2pr'^2}{r^3}$

$$\implies \frac{-u''}{p} - \frac{u}{p} = -\frac{GM}{c^2}$$
$$u'' + u = \frac{GMp}{c^2}$$

Which is Binet's equation. Homogeneous solution

$$u = A\cos(\phi - \omega)$$

Full solution

$$\frac{L}{r} = 1 + e\cos(\phi - \omega)$$

Where this e is the 'eccentricity' of the curve.

6   Special functions and Nanomechanics

Nanotechnology is engineering matter on the scale of individual atoms or small number of atoms. Includes devices, materials, structures or anything else where length scale is important.

Nanometer $1nm = 10^{-9}m$, Angstrom $1\mathring{A} = 10^{-10}m$

Nanoscale is imprecise but refers to $1\mathring{A}$ to $100nm$

Graphite is an allotrope of carbon - layers of carbon sheets covalently bonded in a hexagonal lattice. The bonding within a sheet is strong, but between sheets is weak. The interatomic separation (inside a sheet) is appx $1.42\mathring{A}$ and the interlayer separation is appx $3.35\mathring{A}$.

Graphene is a single sheet of graphitic carbon. Originally thought to be unstable (but it is stable). Graphine can be wrapped up into $0D$ fullerenes, rolled into $1D$ carbon nanotubes, or stacked into 3D graphite.

Graphene is basically the building block for a bunch of nanostructures.

Fullerene is a closed cage carbon molecule. C_{60} fullerene has 60 carbon atoms, and is the smallest stable fullerene with a radius of appx $3.55\mathring{A}$. The atom positions are simeqimately the vertices of a truncated icosahedron I.e. take an icosahedron (20 triangular faces, 12 vertices) and chop off all the vertices and you get (20 hexagonal faces, 12 pentagonal faces, 60 vertices).

There are an infinite variety of fullerenes with various symmetries and isomers. C_{70} fullerene is sphehroidal and may be imagined as C_{60} with 10 extra carbon atoms added around the equator. C_{80} has seven known isomers, one of which is basically C_{60} with 20 extra carbon atoms added around the equator

Goldberg fullerenes have icosahedral symmetry. They are 20 congruent equilateral triangles, cut from a sheet of graphene with 12 pentagons at every vertex.

One side of the fundamental equilateral triangle is characterised by two integers (n, m) and the number of atoms N making up the fullerene is given by

$$N = 20(n^2 + nm + m^2)$$

C_{60} is a Goldberg fullerene with $n = 1, m = 1$.

The vector numbers (a, b) form a basis, and can be written as

$$\mathbf{a} = \frac{3\sigma}{2}\mathbf{i} - \frac{\sqrt{3}\sigma}{2}\mathbf{j}, \quad b = \frac{3\sigma}{2}\mathbf{i} + \frac{\sqrt{3}\sigma}{2}$$

Where σ is the carbon-carbon bond length, and i, j are the cartesian basis vectors. So the vector for the triangle side C is given by

$$C = \frac{3\sigma}{2}(n + m)\mathbf{i} - \frac{\sqrt{3}\sigma}{2}(n - m)\mathbf{j}$$

The length of the side

$$s = ||C||$$

The area of an equilateral triangle is

$$A_\Delta = \sqrt{3}s^2/4$$

Each hexagon has 6 carbon atoms, where each atom participates in 3 hexagons (so each hexagon 'owns' 2 atoms).

Strictly speaking this is an icosahedron, but we will treat it as a sphere, so we want to get the radius

The belt region has 10 triangles, and therefore a length of $5s$, simeqimately equivalent to the circumference, hence $r =$

The area of the icosahedron is $A = 20A_\triangle$, and the radius of a sphere would be

$$r \simeq (A/4\pi)^{1/2}$$

Both are simeqimations which vary by simeqimately 4%.

Euler's theorem for polyhedra:

$$v - e + f = 2(1 - g)$$

Where v, e, f are the number of vertices, edges, faces respectively, and g is the genus of the polyhedron. The genus is the number of 'handles' of the solid. Fullerenes have genus $g = 0$ and nanotori have a genus $g = 1$.

So since fullerenes have genus $g = 0$

$$v - e + f = 2$$

Assuming the fullerene contains only pentagonal and hexagonal rings, denote the number of pentagons p and the number of hexagons h. Assume that every ...

Substitution into Euler's theorem gives

$$\frac{5p + 6h}{3} - \frac{5p + 6h}{2} + p + h = 2$$

And hence $p = 12$, hence there are always 12 pentagons.

This holds for any surface of genus zero.

Nanotubes can be closed cage molecules if the cylinder is capped. Alternatively it may be uncapped, exposing the inner surface for interaction. The **chirality** of a nanotube refers to the alignment of the hexagons aroudn the cylinder. The chirality angle θ is the angle between the a, b vectors defining the net as before.

The chiral vector C is defined as $C = n\mathbf{a} + m\mathbf{b}$ or written as (n, m). Known as the chiral vector numbers, where C is at a vertex. Conventionally $0 \leq m \leq n$. The angle between a and C is denoted by θ, called the Chiral angle. Perpendicular to C is the translation vector T

The rolled up model assumes that the graphene plane may be sliced and rolled into a perfect right-circular cylinder. The chiral angle is given by

$$\theta = \cos^{-1}\left(\frac{2n + m}{2\sqrt{n^2 + nm + m^2}}\right)$$

Three general cases

- $m = 0$ gives zigzag nanotubes with $\theta = 0$

- $0 < m < n$ give chiral nanotubes with $0 < \theta < \pi/6$

- $m = n$ give armchair nanotubes with $\theta = \pi/6$

The nanotube radius is given by

$$r = |C|/2\pi$$

From earlier,

$$\mathbf{a} = \frac{3\sigma}{2}\mathbf{i} - \frac{\sqrt{3}\sigma}{2}\mathbf{j}, \quad b = \frac{3\sigma}{2}\mathbf{i} + \frac{\sqrt{3}\sigma}{2}$$

Substituting gives

$$|C| = \sigma\sqrt{3(n^2 + nm + m^2)}$$

And hence the radius

$$r = \frac{\sigma\sqrt{3(n^2 + nm + m^2)}}{2\pi}$$

The translation vector T is perpendicular to the chiral vector and gives the distance along the tube axis between atoms that are equivalent and vary only by a translation in the axial direction. This is considered a unit cell of the carbon nanotube

$$T = \frac{n + 2m}{d_R}\mathbf{a} - \frac{2n + m}{d_R}\mathbf{b}$$

Where d_R is the greatest common divisor of the two numerators.

The rectangle defined by C, T gives the unit cell for a carbon nanotube. The length is given by

$$||T||$$

And so the area is

$$A = ||C|| \cdot ||T||$$

Carbon nanocones are conical instead of tubular, and are open at one end and capped at the other. they do not have chirality, but the open angle depends on the number of pentagons forming the cap.

Nanocones are formed from hexagons on a honeycombed lattice by adding fewer than the 6 needed by Euler's theorem to form the closed semi-fullerene structure. There are 5 different structures

The disclination number of pentagons on the graphene sheet gives the change with θ in form

$$\theta = \pi N_p / 3$$

It is clear that

$$\sin(\alpha/2) = r/R, \quad c = 2\pi r = 2\pi(1 - N_p/6)R$$

Therefore the relation of the cone angle and number of pentagons is

$$\sin(\alpha/2) = 1 - \frac{N_p}{6}$$

Of course this only gives 5 answers before the angle is $\alpha = 180$

7 Special Functions

Gamma function $\Gamma(z)$ defined for complex numbers defined as

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$$

And we know

$$\Gamma(z + 1) = z\Gamma(z)$$

For positive integer n

$$\Gamma(n + 1) = n!$$

For some z there are singularities (e.g. all non-positive integers), but there is a well defined integral for all z

$$\Gamma(z) = P(z) + \int_1^\infty t^{z-1} e^{-t} dt$$

Where

$$P(z) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!(z+n)}$$

Which is analytic everywhere except non-positive integers where it has simple poles with residues $(-1)^n/n!$.

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}$$

Which relates gamma to trig. We also have

$$\Gamma(2z) = 2^{2z-1}\pi^{1/2}\Gamma(z)\Gamma(z+1/2)$$

In general for multiplication

$$\prod_{l=0}^{m-1} \Gamma(z+l/m) = (2\pi)^{(m-1)/2} m^{1/2-mz}\Gamma(mz)$$

For $m = 2, 3, \ldots$.
   Get $\Gamma(1/2)$

$$\Gamma(1/2) = \int_0^{\infty} e^{-u}/\sqrt{u}\, du$$

$$\Gamma(1/2)^2 = \int\int \frac{e^{-u+v}}{\sqrt{uv}}$$

Let $u = x^2, v = y^2$ so that $du = 2x\,dx, dv = 2y\,dy$

$$\Gamma(1/2)^2 = 4\int\int e^{-(x^2+y^2)}dx\,dy$$

Take polar coords

$$\Gamma(1/2)^2 = 4\int_0^{\infty}\int_0^{\pi/2} re^{-r^2}dr\,d\theta$$

$$= 2\pi\int_0^{\infty} re^{-r^2}dr$$

$$= 2\pi\left[-\frac{1}{2}e^{-r^2}\right]_0^{\infty}$$

$$= \pi$$

$$\implies \Gamma(1/2) = \sqrt{\pi}$$

The Beta function $B(x, y)$ satisfies

$$B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$$

For non-negative integer x, y

$$B(x+1, y+1) = \frac{x!y!}{(x+y+1)!} = \left[x+y+1\binom{x+y}{x}\right]^{-1}$$

Integral form for positive real parts

$$B = \int_0^1 t^{x-1}(1-t)^{y-1}dt$$

If we take $t = \sin^2\theta$ s.t. $dt = 2\sin\theta\cos\theta d\theta$, we get

$$B(x,y) = 2\int_0^{\pi/2} \sin^{2x-1}\theta\cos^{2y-1}\theta d\theta$$

Note $B(x,y) = B(y,x)$.

he was going fuckin speeeedy on wednesday tho

Derive expression for $\int_{-\pi/2}^{\pi/2}\cos^{2n}\theta d\theta$ using factorials

$$\int_{-\pi/2}^{\pi/2}\cos^{2n}\theta d\theta = 2\int_0^{\pi/2}\cos^{2n}\theta d\theta$$
$$= B(n+\frac{1}{2},\frac{1}{2})$$
$$= \frac{\Gamma(n+\frac{1}{2})\Gamma(\frac{1}{2})}{\Gamma(n+1)}$$

Using the duplication formula,

$$\Gamma(2z) = 2^{2z-1}\pi^{1/2}\Gamma(z)\Gamma(z+\frac{1}{2})$$

$$\Gamma(n+\frac{1}{2}) = \frac{\pi^{1/2}\Gamma(2n)}{2^{2n-1}\Gamma(n)}$$

$$\int_{-\pi/2}^{\pi/2}\cos^{2n}\theta d\theta = \frac{\Gamma(n+\frac{1}{2})\Gamma(\frac{1}{2})}{\Gamma(n+1)}$$
$$= \frac{\pi\Gamma(2n)}{2^{2n-1}\Gamma(n)\Gamma(n+1)}$$
$$= \frac{\pi\Gamma(2n)}{2^{2n-1}\Gamma(n)\Gamma(n+1)}\frac{2n}{2n}$$
$$= \frac{\pi\Gamma(2n+1)}{2^{2n}[\Gamma(n+1)]^2}$$
$$= \frac{\pi(2n)!}{2^{2n}(n!)^2}$$

The Pochhammer symbol $(a)_n$ wrote

$$(a)_n = a(a+1)(a+2)\ldots(a+n-1)$$

If a is an integer

$$(a)_n = \frac{(a+n-1)!}{(a-1)!}$$

And this can be extended to include non-integer and complex vales of a by using the definition

$$(a)_n = \frac{\Gamma(a+n)}{\Gamma(a)}$$

Derive the pochhammer of a negative integer, i.e. for $(-n)_m$, where $m < n$. If $m \geq n$ it will be 0.

$$\begin{aligned}
(-n)_m &= (-n)(1-n)(2-n)\ldots(m-n-1) \\
&= (-1)^m(n-1)(n-2)\ldots(n-m+1) \\
&= (-1)^m \frac{n!}{(n-m)!}
\end{aligned}$$

So

$$(-n)_m = \begin{cases} (-1)^m \frac{n!}{(n-m)!}, & \text{if } m \leq n \\ 0 & m > n \end{cases}$$

7.1 The Hypergeometric function

Evaluation of the energy between two molecular structures having well defined shapes (e.g. cylinders, spheres, cones) generally leads to hypergeometric functions. The hypergeometric function has a series definition given by

$$F(a,b;c;z) = \sum_{n=0}^{\infty} \frac{(a)_n(b)_n}{(c)_n} \frac{z^n}{n!}$$

And integral representation

$$F(a,b;c;z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 t^{b-1}(1-t)^{c-b-1}(1-tz)^{-a}dt$$

For $Re(c) > Re(b) > 0$

Relationships to elementary functions

$$\begin{aligned}
(1+z)^a &= F(-a,b;b;-z) \\
\sin^{-1}(z) &= zF(1/2,1/2;3/2;z^2) \\
\sinh^{-1}(z) &= zF(1/2,1/2;3/2;-z^2) \\
\tan^{-1}(z) &= zF(1/2,1;3/2;-z^2) \\
\tanh^{-1}(z) &= zF(1/2,1;3/2;z^2) \\
\log(1+z) &= zF(1,1;2;-z)
\end{aligned}$$

From the series representation, the use of the pochhammer tells us

- If a or b is a negative integer, then the series terminates after finite terms

- If c is a negative integer, then the series becomes undefined after finite terms

- The ratio of successive terms approaches z as $n \to \infty$

- The series is absolutely convergent for $|z| < 1$.

Relationships to orthogonal polynomials

$$\begin{aligned}
T_n(x) &= F(-n,n;1/2;(1-x)/2) \\
U_n(x) &= (n+1)(F-n,n+2;3/2;(1-x)/2) \\
P_n(x) &= F(-n,n+1;1;(1-x)/2)
\end{aligned}$$

Where T_n and U_n are Chebyshev polynomials and $P_n(x)$ are the Legendre polynomials

Hypergeometric DEs

$$\frac{d^n}{dz^n}F(a,b;c;z) = \frac{(a)_n(b)_n}{(c)_n}F(a+n,b+n;c+n;z)$$

The hypergeometric function is a solution of a homogeneous linear second order DE with regular singularities at $z = 0, 1, \infty$. Its called the hypergeometric equation, given by:

$$z(1-z)\frac{d^2u}{dz^2}$$

Any second order Linear ODE with at most 3 regular singular points can be transformed into the hypergeometric equation. It follows that $u_1 = F(a,b;c;z)$ is a solution of this differential equation that is regular at $z = 0$. The second solution can depend on the parameter values due to certain degenerate cases. For this course we will be much more interested in the function as defined from the integral form rather than the equation.

Quadratic Transformations If any only if

$$\pm(1-c), \quad \pm(a-b), \quad \pm(a+b-c)$$

Are such that two are equal OR one equals $1/2$, then there exists higher order transformations which are known as quadratic transformations.

The transformations are revolting. Look at

https://authors.library.caltech.edu/43491/

For a bunch of them Just a useful one here:

$$F(a,b;2b;z) = (1-z)^{a/2}F\left(a, 2b-a; b+\frac{1}{2}; -\frac{(1-\sqrt{1-z})^2}{4\sqrt{1-z}}\right)$$

There are several integrals which either define or involve the hypergeometric functions and here we will only keep the most useful. Euler's integral formula is

$$F(a,b;c;z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\int_0^1 t^{b-1}(1-t)^{c-b-1}(1-tz)^{-a}dt$$

Provided $Re(c) > Re(b) > 0$. Theres the closely related

$$F(a,b;c;1-z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\int_0^1 s^{b-1}(1+s)^{a-c}(1+sz)^{-a}ds$$

Note the power swaps.

The Appell's hypergeometric equation is a formal extension of the hypergeometric equation for two variables. The most commonly used form is

$$F_1(a;b,c';c;x,y) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{(a)_{m+n}(b)_m(b')_n}{m!n!(c)_{m+n}}x^m y^n$$

Convergent for $|x| < 1$ and $|y| < 1$. The integral form:

$$F_1 = \frac{\Gamma(c)}{\Gamma(a)\Gamma(c-a)}\int_0^1 t^{a-1}(1-t)^{c-a-1}(1-tx)^{-b}(1-ty)^{-b'}dt$$

There are extensions to these which include a' and/or c'

There are a lot of gaps here, look at the notes.

If $x = 0$ or $y = 0$

$$F_1(a; b, b'; c; x, 0) = F(a, b; c; x)$$
$$F_1(a; b, b'; c; 0, y) = F(a, b'; c; y)$$

Can transform the Appell functions into other Appell functions And there is a full transformation into a normal hypergeometric function:

$$F_1(a; b, b'; c; x, y) = \sum_{m=0}^{\infty} \frac{(a)_m (b)_m}{m!(c)_m} F(a + m, b'; c + m; y) x^m$$

This representation is useful for solving Appell's function in `MATLAB`.

7.2   Elliptic Integrals

The fundamental forms of the normal elliptic integrals of the first and second kind arae

$$F(\phi, k) = \int_0^\phi \frac{d\theta}{\sqrt{1 - k^2 \sin^2 \theta}}$$
$$E(\phi, k) = \int_0^\phi \sqrt{1 - k^2 \sin^2 \theta}\, d\theta$$

Where we are using Legendre's notation and $0 \leq \phi \leq \pi/2$ is the argument and $0 \leq k \leq 1$ is the modulus.

There is a third kind but we'll ignore it completely.

When $\phi = \pi/2$ then the function is said to be **complete** and the argument is usually dropped. So if we only give one argument:

$$K(k) = F(\pi/2, k), \quad E(k) = E(\pi/2, k)$$

We define the complimentary modulus k'

$$k' = \sqrt{1 - k^2}$$

The complete elliptic integrals of first and second kind are related by

$$E(k)K(k') + E(k')K(k) - K(k)K(k') = \pi/2$$

Which is Legendre's relation. The integrals reduce to simple expressions for certain inputs:

$$F(0, k) = E(0, k) = 0$$
$$F(\phi, 0) = E(\phi, 0) = \phi$$
$$F(\phi, 1) = \sin \phi$$
$$E(\phi, 1) = \log(\tan \phi + \sec \phi)$$

For two separate molecular structures (or non-bonded), the energy E can be evaluated using either a discrete atom-atom formulation or by a continuous approach

The Discrete Method The non-bonded interaction energy may be obtained as a summation of the interaction energy between each atom pair, namely

$$E = \sum_i \sum_j \phi(\rho_{ij})$$

Where $\Phi(\rho_{ij})$ is the potential function for atoms i, j located ρ_{ij} distance apart on two distinct molecular structures.

The continuum approach assumes that atoms are uniformly distributed over the entire surface of the molecule, and the double summation is replaced by a double integral

$$E = \eta_1 \eta_2 \int_{S_2} \int_{S_1} \Phi(\rho) dA_1 dA_2$$

Where η_1, η_2 represent the mean surface density of atoms on the two interacting molecules, and ρ represents the distance between the two typical surface elements dA_1 and dA_2 located on the two interacting molecules.

Mean atomic surface density η is determined by dividing the number of atoms making up the molecule by the surface area of the molecule. The continuum approach is kind of like taking the average, or mean behaviour.

This is all necessary because we want to look at the **van der Waals force** between two molecules, or the intermolecular force.

The van der Waals interaction between two typical non-bonded atoms of two molecules is given by

$$\mathbf{F_{vdW}} = -\nabla E$$

So for example if the force is just in the z direction, then we get

$$F_{vdW} = k F_z$$

Where

$$F_z = -\frac{\partial E}{\partial z}$$

Since the forces in the other directions are in symmetry.

Lennard-Jones potential describes the interaction between two non-bonded atoms and is given by

$$\Phi(\rho) = -A\rho^{-m} + B\rho^{-n}$$

Where A, B are the attractive and repulsive constants (respectively) and ρ is the distance between the atoms. In many cases the values $m = 6$ and $n = 12$ are adopted (referred to as the 6-12 potential). For hydrogen bonding interactions a (10-12) potential is used.

Can alternatively be written as

$$\Phi(\rho) = 4\epsilon \left[-(\sigma/\rho)^6 + (\sigma/\rho)^{12} \right]$$

Where σ is the van der Waals distance and ϵ is the energy well depth. We get

$$\sigma = (B/A)^{1/6}, \quad \epsilon = \frac{A^2}{4B}$$

The van der Waals force is a short-range force, and so when using the Lennard-Jones potential for interaction between molecular structures, it is only necessary to include the nearest neighbour interactions.

For example, considering a molecule near the open end of a carbon nanotube, we can consider the tube to be semi-infinite in length (i.e. the rest of the tubes effects are negligible)

The Lennard-Jones potential is applied between non-bonded and non-polar atomic interactions.

Parameter values are mostly obtained empirically (sounds like we wouldn't have to obtain them ourselves).

Example: Calculate the interaction energy E and the van der Waals force F_{vdW} for a single carbon atom a perpendicular distance z from an infinite graphene sheet.

I.e. atom has point $(0, 0, z)$ and graphene sheet $(x, y, 0)$.

Get distance ρ for the atoms on the graphene plane

$$\rho = \sqrt{x^2 + y^2 + z^2}$$

$$E = \eta \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[-\frac{A}{(x^2 + y^2 + z^2)^3} + \frac{B}{(x^2 + y^2 + z^2)^6} \right] dxdy$$
$$= \eta(-AI_3 + BI_6)$$

Where
$$I_n = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{dxdy}{(x^2 + y^2 + z^2)^n}$$

Let $x = r\cos\theta$ and $y = r\sin\theta$ So $dxdy = rdrd\theta$

$$I_n = \int_0^{2\pi} \int_0^{\infty} \frac{rdrd\theta}{(r^2 + z^2)^n}$$
$$= \pi \int_0^{\infty} \frac{2rdr}{(r^2 + z^2)^n}$$
$$= \pi \left[-\frac{1}{(n-1)(r^2 + z^2)^{(n-1)}} \right]_0^{\infty}$$
$$= \frac{\pi}{(n-1)z^{2n-2}}$$

$$E = \eta(-AI_3 + BI_6)$$
$$= \pi\eta \left(-\frac{A}{2z^4} + \frac{B}{5z^{10}} \right)$$

And so the force

$$F_z = -\frac{\partial E}{\partial z}$$
$$= \pi\eta \left(-\frac{4A}{2z^5} + \frac{10B}{5z^{11}} \right)$$
$$= 2\pi\eta \left(-\frac{A}{z^5} + \frac{B}{z^{11}} \right)$$

For some reason we take $A = 17.4 eV \mathring{A}^6$ and $B = 29000 eV \mathring{A}^{12}$, $\eta = 0.3812 \mathring{A}^{-2}$

How about if we replaced the atom with another, parallel, infinite plane of graphene?

We will get an infinite interaction energy if they are truly infinite, but if we considered the interaction energy per unit area we might get something insightful. So we get

$$E_{pua} = \pi\eta^2 \left(-\frac{A}{2z^4} + \frac{B}{5z^{10}} \right)$$

With $A = 15.2$ and $B = 24100$ (apparently this is just something we choose based on the materials?). To find the equilibrium point (point where force is zero) take the derivative.

Since we are looking at per unit area, it is a pressure instead of a force, but its fundamentally the same

$$P_{vdW} = -\frac{\partial E_{pua}}{\partial z}$$

We get $z_0 = 3.41\mathring{A}$. However, earlier we said the interlayer spacing in graphite is simeqimately $3.35\mathring{A}$. Is there a reason for the discrepancy? Yes - the planes don't just interact with their nearest neighbours, but with all the sheets in the graphite. Assume an infinite number of planes all separated by the same distance z. Then the total interaction energy per unit area for each plane is given by

$$E_{pua} = 2\sum_{n=1}^{\infty} \pi\eta^2 \left(-\frac{A}{2(nz)^4} + \right)$$

Now for $m > 1$ $\sum_{n=1}^{\infty} n^{-m} = \zeta(m)$ is the Reimann zeta function. Hence

$$E_{pua} = 2\pi\eta^2 \left(-\frac{A\zeta(4)}{2z^4} + \frac{B\zeta(1)}{5z^{10}} \right)$$

Minimising the energy per unit area obtains

$$z_0 = \left(\frac{B\zeta(10)}{A\zeta(4)} \right)^{1/6}, \quad \sigma \left(\frac{\zeta(10)}{\zeta(4)} \right)^{1/6}$$

We apparently know $\zeta(4) = \pi^4/90$ and $\zeta(10) = \pi^{10}/93555$. And hence we get $z_0 = 3.37\mathring{A}$, which is closer to the experimental $3.35\mathring{A}$.

7.3 Surface integration

When integrating a scalar quantity like $\Phi(\rho)$ over a surface S we do

1. Parameterise S in terms of two variables u, v

2. Calculate the area element dA from the parameterisation

3. Evaluate the double integral

$$\int_{v_0}^{v_1} \int_{u_0}^{u_1} \Phi(\rho(u,v)) dA$$

We will be considering two-dimensional surfaces embedded in three dimensional space. Some examples:

- $x - y$ plane : $\mathbf{r} = (u, v, 0)$ with $-\infty < u, v < \infty$

- Sphere of radius a: $\mathbf{r} = (a\cos u \sin v, a\sin u \sin v, a\cos v)$ Where $-\pi < u \leq \pi$ and $0 \leq v \leq \pi$

- Cylinder of radius b : $\mathbf{r} = (b\cos u, b\sin u, v)$ Where $-\pi < u \leq \pi$ and $-\infty < v < \infty$

We get

$$dA = \left| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right| dudv$$

For cylindrical coordinates

$$\mathbf{r} = (b\cos u, b\sin u, v)$$

Tangent (derivative) vectors are

$$\mathbf{r}_u = (-b\sin u, b\cos u, 0)$$
$$\mathbf{r}_v = (0, 0, 1)$$

the cross product:

$$\mathbf{r}_u \times \mathbf{r}_v = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -b\sin u & b\cos u & 0 \\ 0 & 0 & 1 \end{vmatrix} = (b\cos u, b\sin u, 0) = b\mathbf{e}_r$$

Assuming $b > 0$ we get

$$dA = |\mathbf{r}_u \times \mathbf{r}_v| dudv = bdudv$$

Spherical coordinates:

$$\mathbf{r} = (a\cos u \sin v, a\sin u \sin v, a\cos v)$$

Tangent vectors are

$$\mathbf{r}_u = (-a\sin u \sin v, a\cos u \cos v, 0)\mathbf{r}_v = (a\cos u \cos v, a\sin u \cos v, -a\sin v)$$

Cross product

$$\mathbf{r}_u \times \mathbf{r}_v = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -a\sin u \sin v & a\cos u \cos v & 0 \\ a\cos u \cos v & a\sin u \cos v & -a\sin v \end{vmatrix}$$

$$= (-a^2\cos u \sin^2 v, -a^2\sin u \sin^2 v, -a^2\sin v \cos v)$$
$$= -a^2\sin v(\cos u \sin v, \sin u \sin v, \cos v)$$
$$= -a^2\sin v\mathbf{e}_r$$

Assuming $a > 0$ the area is

$$A = |\mathbf{r}_u \times \mathbf{r}_v| = a^2\sin v$$

$$dA = |\mathbf{r}_u \times \mathbf{r}_v| dudv = a^2\sin v dudv$$

So the surface area is

$$S = \int_0^\pi \int_{-\pi}^\pi a^2\sin v dudv$$
$$= 2\pi a^2 \int_0^\pi \sin v dv$$
$$= 2\pi a^2[-\cos v]_0^\pi$$
$$= 4\pi a^2$$

7.4 Double-walled carbon nanotubes

A double-walled carbon nanotube is simply one carbon nanotube contained within another. We use "@" to denote the encapsulation from smallest to largest. E.g. (5,5)@(10,10) denotes a (5,5) nanotube inside a (10,10).

Double walled nanotubes have similar morphology and properties as single walled, but it improves their resistance to chemicals significantly. This is especially important when *functionalisation* is required to add new properties to the carbon nanotubes. Functionalisation refers to the grafting (via covalent bonds) of chemical functions to the nanotube surface.

To model a double-walled carbon nanotube. Consider the outer nanotube with a radius b, given parametrically by $(b\cos\theta, b\sin\theta, z)$. interacting with an interior point P given by $(\delta, 0, 0)$. Thus δ is the off-zxis distance of the point, and since it is interior to the nanotube we require $\delta < b$. Assume an infinite extending nanotube, so

$$-\pi < \theta \leq \pi, \quad -\infty < z < \infty$$

So the area element for the cylinder is $dA = bd\theta dz$ (as before). And the distance from the point to the surface of the cylinder, ρ is

$$\rho^2 = (b\cos\theta - \delta)^2 + b^2\sin^2\theta - z^2$$
$$= (b - \delta)^2 + z^2 + 4b\delta\sin^2(\theta/2)$$

Used $\cos\theta = 1 - 2\sin^2\theta/2$ Thus the interaction energy between the point and the cylinder E_c is given by

$$E_c = \eta_c b(-AK_3 + BK_6)$$

Where η_c is the atomic surface density of the cylinder and

$$K_n = \int_{-\infty}^{\infty}\int_{-\pi}^{\pi} \frac{d\theta dz}{\rho^{2n}}$$
$$= \int_{-\infty}^{\infty}\int_{-\pi}^{\pi} \frac{d\theta dz}{\left[(b-\delta)^2 + z^2 + 4b\delta\sin^2(\theta/2)\right]^n}$$

Let

$$\lambda^2 = (b-\delta)^2 + 4b\delta\sin^2\theta/2$$

And let $z = \lambda\tan\phi$ with $dz = \lambda\sec^2\phi d\phi$.

$$K_n = \int_{-\pi/2}^{\pi/2}\int_{-\pi}^{\pi} \frac{\lambda\sec^2\psi d\theta d\psi}{\left[\lambda^2 + \lambda^2\tan^2\psi\right]^n}$$
$$= \int_{-\pi/2}^{\pi/2}\int_{-\pi}^{\pi} \frac{\sec^2\psi d\theta d\psi}{\lambda^{2n-1}\left[1 + \tan^2\psi\right]^n}$$
$$= \int_{-\pi/2}^{\pi/2} \cos^{2n-2}\psi d\psi \int_{-\pi}^{\pi} \lambda^{1-2n}$$

We have (apparently)

$$\int_{-\pi/2}^{\pi/2} \cos^{2n-2}\psi d\psi = \frac{\pi(2n-1)!}{2^{2n-2}[(n-1)!]^2}$$

And the lambda integral (call it I)

$$I = \int_{-\pi}^{\pi} \frac{d\theta}{\lambda^{2n-1}} = \int_{-\pi}^{\pi} \frac{d\theta}{\left[(b-d)^2 + 4b\delta\sin^2\theta/2\right]^{n-1/2}}$$
$$= 2\int_{0}^{\pi} \frac{d\theta}{\left[(b-d)^2 + 4b\delta\sin^2\theta/2\right]^{n-1/2}}$$

Let $t = \sin^2 \theta/2$ and $dt = \sin\theta/2\cos\theta/2d\theta = t^{1/2}(1-t)^{1/2}d\theta$. So $d\theta = t^{-1/2}(1-t)^{-1/2}dt$

$$\implies I = 2\int_0^1 t^{-1/2}(1-t)^{-1/2}\left[(b-\delta)^2 + 4b\delta t\right]^{1/2-n}dt$$

$$= \frac{2}{(b-\delta)^{2n-1}}\int_0^1 t^{-1/2}(1-t)^{-1/2}\left[1 - \frac{4b\delta}{(b-\delta)^2}t\right]^{1/2-n}dt$$

This is in Euler form for the hypergeometric function

$$a = n - 1/2, \quad b = 1/2, \quad c = 1, \quad z = -\frac{4b\delta}{(b-\delta)^2}$$

So

$$I = \frac{2}{(b-\delta)^{2n-1}}\frac{\Gamma(1/2)\Gamma(1/2)}{\Gamma(1)}F(n-1/2, 1/2; 1; -\frac{4b\delta}{(b-\delta)^2})$$

$$= \frac{2\pi}{(b-\delta)^{2n-1}}F(n-1/2, 1/2; 1; -\frac{4b\delta}{(b-\delta)^2})$$

So

$$K_n = \frac{\pi^2(2n-2)!}{2^{2n-2}[(n-1)!]^2(b-\delta)^{2n-1}}F(n-1/2, 1/2; 1; -\frac{4b\delta}{(b-\delta)^2})$$

Since we have the form $F(a, b; 2b; z)$ use

$$F(a, b; 2b; z) = \left[\frac{1 + (1-z)^{1/2}}{2}\right]^{-2a} \times$$

$$F\left(a, a - b + \frac{1}{2}; b + \frac{1}{2}; \left[\frac{1 - (1-z)^{1/2}}{1 + (1-z)^{1/2}}\right]^2\right)$$

For $z = -4b\delta/(b-\delta)^2$ that last term looks nice.

And we get (i'm not sure why the 2^{2n-2} became 2^{2n-3} but okay)

$$K_n = \frac{\pi^2(2n-2)!}{2^{2n-3}[(n-1)!]^2b^{2n-1}}F(n-\frac{1}{2}, n-\frac{1}{2}; 1; \frac{\delta^2}{b^2})$$

And therefore the total energy is given by

$$E_c = \frac{3\pi^2\eta_c}{4b^4}\left[-AF(5/2, 5/2; 1; \delta^2/b^2) + \frac{21B}{32b^6}F()\right]$$

Now if we consider a cylinder inside the first element, where the inner cylinder is parametrically

$$(\epsilon + b_1\cos\theta_1, b_1\sin\theta_1, z_1)$$

Where $-\pi < \theta_1 \le \pi$ and $-\infty < z_1 < \infty$. Once again we can only get the energy per unit area. So consider a single ring around the inner cylinder which gives the energy per unit length. So we can set $z_1 = 0$ for convenience, and note the line element is $dl = b_1 d\theta_1$. So the energy is given by

$$E_{cc} = \eta_c\int_{-\pi}^{\pi}D_c b_1 d\theta_1$$

$$= \frac{3\pi^2\eta_c^2 b_1}{4b_2^4}\left(-AL_5 + \frac{21B}{32b_2^6}L_{11}\right)$$

Where
$$L_n = \int_{-\pi}^{\pi} F\left(n/2, n/2; 1; \delta^2/b_2^2\right) d\theta_1$$

Use the series expansion of the hyper geometric series, and note that $0 \leq \delta^2/b_2^2 < 1$ so the expansion is absolutely convergent, so

$$L_n = \int_{-\pi}^{\pi} \sum_{k=0}^{\infty} \frac{(n/2)_k (n/2_k)}{(1)_k k!} \frac{\delta^{2k}}{b_2^{2k}} d\theta_1$$
$$= \sum_{k=0}^{\infty} \left(\frac{(n/2)_k}{k! b_2^k}\right)^2 \int_{-\pi}^{\pi} \delta^{2k} d\theta_1$$

Where

$$\delta^2 = (\epsilon + b_1 \cos \theta_1)^2 + b_1^2 \sin^2 \theta_1$$
$$= (b_1 + \epsilon)^2 - 4\epsilon b_1 \sin^2 \theta_1/2$$

So the integral expands to

$$\int_{-\pi}^{\pi} \delta^{2k} d\theta_1 = \int_{-\pi}^{\pi} \left[(b_1 + \epsilon)^2 - 4\epsilon b_1 \sin^2 \theta_1/2\right]^k d\theta_1$$
$$\vdots$$
$$= 2(b_1 + \epsilon)^2 k \int_0^1 t^{-1/2} (1-t)^{-1/2} \left(1 - \frac{4\epsilon b_1}{(b_1 + \epsilon)^2} t\right)^k dt$$

Which again is a hypergeometric function. With $a = -k, b = 1/2, c = 1, z = 4\epsilon b_1/(b_1 + \epsilon)^2$.
Hence

$$\int_{-\pi}^{\pi} \delta^{2k} d\theta_1 = 2\pi (b_1 + \epsilon)^2 k F(-k, 1/2; 1; \frac{4\epsilon b_1}{(b_1 + \epsilon)^2})$$

Using the quadratic transformation again gives

$$\int_{-\pi}^{\pi} \delta^{2k} d\theta_1 = F(-k, -k; 1; \frac{\epsilon^2}{b_1^2})$$

And hence

$$L_n = 2\pi \sum_{k=0}^{\infty} \left(\frac{(n/2)_k}{k!}\right)^2 \left(\frac{b_1}{b_2}\right)^{2k} F(-k, -k; 1; \frac{\epsilon^2}{b_1^2})$$

And note the last hypergeometric function is a terminating series since a is a negative integer.
Hence

$$F\left(-k, -k; 1; \frac{\epsilon^2}{b_1^2}\right) = \sum_{j=0}^{k} \left(\frac{(-k)_j}{j!}\right)^2 \left(\frac{\epsilon}{b_1}\right)^{2j}$$

Which can be calculated numerically with relative ease.

7.5   Nanotube@Nanotube Oscillator

A double walled nanotube starts with its inner nanotube extruded by some distance d out of the fixed, open-ended, outer nanotube. In this case, the excess van der Waals force sucks the inner nanotube back into the nanotube. As in the fullerene@nanotube oscillator, there is a reversal of direction at the opposite end of the outer nanotube which returns the oscillating inner nanotube to its starting location and the process then repeats.

We want to know for a particular double-walled nanotube:

- Will the inner nanotube be sucked in to the outer nanotube? (acceptance condition)

- How much energy will be picked up by the oscillating nanotube? (suction energy)

- What will the oscillatory frequency be?

It turns out that the DWNT (double walled nanotube) oscillator can be solved exactly without the following simplification, but the integrals are messy, and don't add much except minor 'edge effects' in most cases.

So we will assume the two nanotubes will only interact along the section of the inner nanotube that is nested. I.e. the part that is inside of the other nanotube. The rest is ignored. We will also assume it behaves like it was inside an infinitely long carbon nanotube.

So when does the acceptance condition hold? I.e. when is it energetically favourable to be inside the outer nanotube.

We addressed this when we looked at E_{cc}, i.e. the energy per unit length of offset nanotubes. If we assume the two nanotubes are coaxial then we can reuse that.

If we consider E_{cc}^* where b_1, b_2 are the radii of the nanotubes and η_c is the atomic surface density of the nanotubes

If $E_{cc}^* > 0$ then the inner tube is too large to be accepted by the other nanotube.

Given the simplifications made above, the suction energy is just the difference in energy between the extruded nanotube and the fully encapsulated nanotube. We already have E_{cc}^*, and since the extrusion length d forms a part of the ICs, the suction energy W is given by

$$W = -E_{cc}^* d$$

This is simply a statement of conservation of energy.

If we now considered finite length nanotubes, outer tube has length $2L_2$ and inner with $2L_1$, assuming $L_2 \geq L_1$.

Now arrange the coordinate system so tube axes are collinear with z axis, and the centre of the outer tube is at the origin. I.e. the end points are at $z = \pm L_2$.

Denote the centre of the oscillating inner nanotube at $z = Z$, i.e. the ends are at $z = Z \pm L_1$

There are (at most) 5 states for the inner tube

- $Z < -L_1 - L_2$, then the inner tube is fully extruded in the negative z direction

- $-L_1 - L_2 < Z < L_1 - L_2$ inner tube is partially extruded in the negative direction

- $L_1 - L_2 < Z < -L_1 + L_2$ inner tube is entirely contained

- $-L_1 + L_2 < Z < L_1 + L_2$ inner tube is partially extruded in the positive direction

- $Z > L_1 + L_2$ inner tube is fully extruded in the positive z direction.

An understanding of the four cutoff points is critical

$$Z = \pm L_1 \pm L_2$$

Taking the fully extruded state as the zero energy level, then as the tube moves toward the fully enclosed state, the total potential energy E_{cc} increases by the interaction energy per unit length E_{cc}^* times the length of the interaction, $L_1 + L_2 + Z$, i.e.

$$E_{cc} = (L_1 + L_2 + Z) E_{cc}^*$$

For state 2.

Up to a maximum when $Z = L_1 - L_2$ with $E_{cc} = 2L_1 E_{cc}^*$. Likewise when we move to the fully extruded on the right state, the potential energy is given by

$$E_{cc} = (L_1 + L_2 - Z) E_{cc}^*$$

And $E_{cc} = 0$ when it fully exits.

If we use the Heaviside unit step function $H(x)$ defined as

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

And using E_{cc}:

$$E_{cc} = E_{cc}^* [(Z + L_+)H(Z + L_+) - (Z + L_-)H(Z + L_-) - (Z - L_-)H(Z - L_-) + (Z - L_+)H(Z - L_+)]$$

Where $L_\pm = L_2 \pm L_1$

By using $F - \frac{\partial E_{cc}}{\partial z}$, we get a function for the force.

Now using Newton's second law we can look at oscillatory behaviour. If we pull out the inner tube a distance d and then released. We simeqimate the vdW force using the Heaviside unit step function and assume the inner nanotube travels through three regions

1. Region of fixed positive force

2. Region of zero net force

3. Region of fixed negative force

From newton's second law we have

$$M\frac{d^2 Z}{dt^2} = -E_{cc}^* = \frac{W}{d}$$

Where M is the mass of the oscillating nanotube. If we left it at rest at $Z = L_1 - L_2 - d$ then the velocity is

$$\frac{dZ}{dt} = \frac{W}{Md}t$$

We find the tube spends

$$t_1 = d\sqrt{\frac{2M}{W}}$$

in region one, at which time it is travelling at velocity

$$\frac{dZ}{dt}\bigg|_{t=t_1} = \sqrt{\frac{2W}{M}}$$

In region two it has no net force, and travels through $2L_2 - 2L_1$ distance, and hence the time is

$$t_2 = 2(L_2 - L_1)\sqrt{\frac{M}{2W}}$$

Hence the total period of oscillation T is given by

$$T = 4t_1 + 2t_2 = 4d\sqrt{\frac{2M}{W}} + 4(L_2 - L_1)\sqrt{\frac{M}{2W}}$$

$$= 4(L_2 - L_1 + 2d)\sqrt{\frac{M}{2W}}$$

Recall $W = -E_{cc}^* d$. Of course the frequency is the reciprocal of this.

But this is a course on variations, how would we do this variationally

Note the kinetic energy

$$K = \frac{1}{2}M\dot{Z}^2$$

And the potential energy

$$V = E_{cc}^* [(Z + L_+)H(Z + L_+) - (Z + L_-)H(Z + L_-) - (Z - L_-)H(Z - L_-) + (Z - L_+)H(Z - L_+)]$$

The Lagrangian would be $\mathcal{L} = K - V$ so the action would be given by

$$F\{Z(t)\} = \int (K - V)dt$$

And by Hamilton's principle, the path of the tube would be an extremal of the action.

Since K and V are independent of time we have

$$\dot{Z}\frac{\partial \mathcal{L}}{\partial \dot{Z}} - \mathcal{L} = const$$
$$M\dot{Z}^2 - K + V = const$$
$$K + V = const$$

I.e. sum of kinetic and potential energy is conserved. So if we take the initial energy as W (the work done to extrude the tube a distance d)

$$W = -dE_{cc}^*$$

And the kinetic energy is initially 0. We get

$$K = W + 2L_1 E_{cc}^* - V$$
$$\vdots$$
$$= k(Z)$$

(Won't write it out) And since $K = M\dot{Z}^2/2$ we derive $\dot{Z}$

And we can get the period of oscillation as

$$\frac{1}{2}T = \int \frac{dZ}{\dot{Z}}$$

(Didn't have time to put in the bounds) In each of these regions k is a simple function of Z.

This gives us a form for T.

And going through the mathematics, you get the same answer as before.

He is back on time again apparently

8 Variation with Several Independent Variables

If we have (x, y) as independent variables and we want to find an extremal for them, e.g. a surface $z(x, y)$, and where f is a function $f(x, y, z(x, y), z_x, z_y)$. Then the EL equation gives

$$\frac{\partial f}{\partial z} - \frac{\partial}{\partial x}\frac{\partial f}{\partial z_x} - \frac{\partial}{\partial y}\frac{\partial f}{\partial z_y} = 0$$

COnsider a surface minimisation problem, i.e. a surface in 3D $z(x,y)$ and x, y are indep. Problems will look like: Minimise

$$F\{z\} = \iint_\Omega (z_x^2 + z_y^2)dxdy$$

Formalisms: Ω is a simply connected bounded region of $\mathbb{R}^2$ $\delta\Omega$ is the boundary $\bar{\Omega} = \Omega \cup \delta\Omega$ is the closure of Ω.

8.1 Derivation of multi-independent variable EL

Find extremals for the functional

$$F\{z\} = \iint_\Omega f(x, y, z(x, y), z_x, z_y)dxdy$$

Analogy of fixed end points is a fixed boundary, i.e.

$$z(x, y) = z_0(x, y) \quad \forall (x, y) \in \delta\Omega$$

For some function $z_0 \in C^2(\delta\Omega)$

As before we consider perturbations, but now they are on a surface with fixed edge

$$\hat{z}(x, y) = z(x, y) + \epsilon\eta(x, y)$$

Where $\eta(x, y) = 0$ for all $(x, y) \in \delta\Omega$.

Taylor's theorem

$$f(x, y, z + \epsilon\eta, z_x + \epsilon\eta_x, z_y + \epsilon\eta_y) = f(x, y, z, z_x, z_y) + \epsilon\left[\eta\frac{\partial f}{\partial z} + \eta_x\frac{\partial f}{\partial z_x} + \eta_y\frac{\partial f}{\partial z_y}\right] + \mathcal{O}(\epsilon^2)$$

First variation, $\delta F(\eta, z) = 0$ for all η and small ϵ

$$\delta F(\eta, z) = \lim_{\epsilon \to 0} \frac{F\{z + \epsilon\eta\} - F\{z\}}{\epsilon}$$
$$= \iint_\Omega \left[\eta\frac{\partial f}{\partial z} + \eta_x\frac{\partial f}{\partial z_x} + \eta_y\frac{\partial f}{\partial z_y}\right] dxdy$$

We can't use integration by parts, so use (a form of) greens theorem

$$\iint_\Omega \left(\frac{\partial \phi}{\partial x} + \frac{\partial \psi}{\partial y}\right) dxdy = \oint_{\delta\Omega} \phi dy - \oint_{\delta\Omega} \psi dx$$

Let $\phi = \eta\frac{\partial f}{\partial z_x}$ and $\psi = \eta\frac{\partial f}{\partial z_y}$

$$\frac{\partial \phi}{\partial x} = \eta_x\frac{\partial f}{\partial z_x} + \eta\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial z_x}\right)$$

$$\frac{\partial \psi}{\partial y} = \eta_y\frac{\partial f}{\partial z_y} + \eta\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial z_y}\right)$$

So

$$\iint_\Omega \left[\eta_x \frac{\partial f}{\partial z_x} + \eta \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial z_x}\right) + \eta_y \frac{\partial f}{\partial z_y} + \eta \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial z_y}\right) \right] dxdy$$
$$= \oint_{\delta\Omega} \eta \frac{\partial f}{\partial z_x} dy - \oint_{\delta\Omega} \eta \frac{\partial f}{\partial z_y} dx$$

We know $\eta(x,y) = 0$ for all $(x,y) \in \delta\Omega$. So the RHS is 0. So

$$\iint_\Omega \left[\eta_x \frac{\partial f}{\partial z_x} + \eta \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial z_x}\right) + \eta_y \frac{\partial f}{\partial z_y} + \eta \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial z_y}\right) \right] dxdy = 0$$
$$\iint_\Omega \left[\eta_x \frac{\partial f}{\partial z_x} + \eta_y \frac{\partial f}{\partial z_y} \right] dxdy = - \iint_\Omega \left[\eta \left(\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial z_y}\right) + \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial z_x}\right) \right) \right] dxdy$$

So

$$\delta F(\eta, z) = \iint_\Omega \eta \left[\frac{\partial f}{\partial z} - \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial z_x}\right) - \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial z_y}\right) \right] dxdy$$
$$= 0$$

Using the integral lemma we get (and since η is free to take any value):

$$\frac{\partial f}{\partial z} - \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial z_x}\right) - \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial z_y}\right) = 0$$

Example:

$$F\{z\} = \iint_\Omega \left(1 + \frac{1}{2}z_x^2 + \frac{1}{2}z_y^2 \right) dxdy$$

$\frac{\partial f}{\partial z} = 0$ so we get

$$\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = 0 \implies \nabla^2 z = 0$$

Which is Laplace's equation.

We would then solve it for the boundary condition.

Example: vibrating string

- Imagine a taut string

    - flexible
    - uniform mass
    - small deflections

- Equilibrium solutions

    - String sits in a straight line
    - consider small perturbations

- Let L be the length of the string.

- Position along the string is $x \in [0, L]$.

- Constant tension τ

- string only moves up down perpendicular to x axis

- displacement at x at time t is $w(x, t) \ll L$

- no friction or damping

- only force occurs to stretch string

- constant density σ along string's length

Fixed end points so that
$$w(0, t) = w(L, t) = 0$$

Velocity of point x on the string is $w_t = \frac{\partial w}{\partial t}$ Kinetic energy of string is

$$T(t) = \frac{\sigma}{2} \int_0^L w_t^2 dx$$

Slope of the string is $\frac{\partial w}{\partial x}$ Potential energy of the string depends on how much it is stretched from its original length L Length at time t is given by

$$J(t) = \int_0^L \sqrt{1 + w_x^2} dx$$

Potential is $V = \tau(J - L)$ so

$$V(t) = \tau \int_0^L \sqrt{1 + w_x^2} - 1 dx$$

We assumed w is small so simeqimate the square root as

$$\sqrt{1 + w_x^2} \simeq 1 + \frac{1}{2} w_x^2$$

So we use

$$V(t) = \frac{\tau}{2} \int_0^L w_x^2 dx$$

So we get

$$F\{w\} = \int_{t_1}^{t_1} (T - V) \, dt = \frac{1}{2} \int_{t_1}^{t_2} \int_0^L \left(\sigma w_t^2 - \tau w_x^2 \right) dx dt$$

So the EL equation is

$$\frac{\partial f}{\partial w} - \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial w_x} \right) - \frac{\partial}{\partial t} \left(\frac{\partial f}{\partial w_t} \right) = 0$$
$$\frac{\partial}{\partial x} (\tau w_x) = \frac{\partial}{\partial t} (\sigma w_t)$$
$$\frac{\partial^2 w}{\partial t^2} = c^2 \frac{\partial^2 w}{\partial x^2}$$

O SHID ITS THE WAVE EQUATION

Example: Plateau's problem Soap film tries to minimise the surface area when stretched between a boundary. Architecture influenced by minimal surfaces

Functional of interest is

$$F\{z\} = \iint_\Omega dA$$

We have to convert this to a usable form

$$\mathbf{A} = (0, dy, z_y dy)$$
$$\mathbf{B} = (dx, 0, z_x dx)$$
$$\mathbf{A} \times \mathbf{B} = (z_x dxdy, z_y dxdy, -dxdy)$$

And

$$dA = |\mathbf{A} \times \mathbf{B}|$$
$$= \sqrt{(z_x dxdy)^2 + (z_y dxdy)^2 + (-dxdy)^2}$$
$$= \sqrt{1 + z_x^2 + z_y^2} dxdy$$

So the functional becomes

$$F\{z\} = \iint_\Omega \sqrt{1 + z_x^2 + z_y^2} dxdy$$

EL gives

$$-\frac{\partial}{\partial x}\left(\frac{z_x}{\sqrt{1 + z_x^2 + z_y^2}}\right) - \frac{\partial}{\partial y}\left(\frac{z_y}{\sqrt{1 + z_x^2 + z_y^2}}\right) = 0$$

$$\frac{\partial}{\partial x}\left(\frac{z_x}{\sqrt{1 + z_x^2 + z_y^2}}\right) = \frac{z_{xx}}{\sqrt{1 + z_x^2 + z_y^2}} - \frac{z_x(z_x z_{xx} + z_y z_{yx})}{(1 + z_x^2 + z_y^2)^{3/2}}$$
$$= \frac{z_{xx}(1 + z_x^2 + z_y^2) - z_x(z_x z_{xx} + z_y z_{yx})}{(1 + z_x^2 + z_y^2)^{3/2}}$$
$$= \frac{z_{xx}(1 + z_y^2) - z_x z_y z_{yx}}{(1 + z_x^2 + z_y^2)^{3/2}}$$

And the other gives

$$\frac{\partial}{\partial y}\left(\frac{z_x}{\sqrt{1 + z_x^2 + z_y^2}}\right) = \frac{z_{yy}(1 + z_x^2) - z_y z_x z_{xy}}{(1 + z_x^2 + z_y^2)^{3/2}}$$

And hence EL is

$$\frac{z_{xx}(1 + z_y^2) - z_x z_y z_{yx}}{(1 + z_x^2 + z_y^2)^{3/2}} + \frac{z_{yy}(1 + z_x^2) - z_y z_x z_{xy}}{(1 + z_x^2 + z_y^2)^{3/2}} = 0$$
$$\frac{z_{xx}(1 + z_y^2) + z_{yy}(1 + z_x^2) - 2z_y z_x z_{xy}}{(1 + z_x^2 + z_y^2)^{3/2}} = 0$$
$$z_{xx}(1 + z_y^2) + z_{yy}(1 + z_x^2) - 2z_y z_x z_{xy} = 0$$

Second order non-linear PDE, super scary right? Yeah. Lets drop all the square derivative terms because that seems legal. So we get

$$z_{xx} + z_{yy} = 0$$

Oh shid thats laplace's equation

I missed 5 minutes of this example but we get Laplace's equation

$$z_{xx} + z_{yy} = 0$$

With BCs

$$z(\pm a, y) = \epsilon \cos(y\pi/2/b)$$
$$z(x, \pm b) = 0$$

Solutions will look like

$$(\cosh(\lambda x) + \sinh(\lambda x))(\cos(\lambda y) + \sin(\lambda y))$$

Want even in x and want to match the zero condition at $y = \pm b$ the coeffs of $\sinh, \sin$ will be 0. Hence

$$z = A \cosh(\lambda x) \cos(\lambda y)$$
$$z = A \cosh(\frac{\pi}{2b} x) \cos(\frac{\pi}{2b} y)$$

Using the first BC we get

$$\epsilon \cos\left(\pi y/2b\right) = A \cos\left(\frac{\pi y}{2b}\right) \cosh\left(\frac{\pi a}{2b}\right)$$
$$A = \frac{\epsilon}{\cosh\left(\pi a/2b\right)}$$

And hence solution is

$$z(x, y) = \frac{\epsilon \cos(\pi y/2b) \cosh(\pi x/2b)}{\cosh\left(\pi a/2b\right)}$$

This is basically the same as the catenary example so its not surprising it looks like that

9 Numerical solutions of variational problems

Man if only we could solve these difficult PDEs without just assuming terms are 0...

The EL equations may be hard to solve, so lets use Numerical solutions. Some options:

- Numerical solution EL DE (not discussed)

- Euler's Finite Difference methods

- Ritz (Rayleigh-Ritz) and in 2D Kantorovich's Method

9.1 Finite Difference methods

Approximate the function onto a finite grid. The problem reduces to a standard multivariable max/minimisation problem, and we find the solution by setting the derivatives to zero. As the grid gets finer, this simeqimates the EL equations

- Use an arbitrary set of mesh points

$$x = a_0 < x_1 < x_2 < \ldots < x_n = b$$

- Approximate the derivative

$$y'(x_i) = \frac{y_{i+1} - y_i}{x_{i+1} - x_i} = \frac{\Delta y_i}{\Delta x_i}$$

- Use the rectangle rule for the integral

$$F\{y\} = \int_a^b f(x, y, y') dx \simeq \sum_{i=0}^{n-1} f\left(x_i, y_i, \frac{\Delta y_i}{\Delta x_i}\right) \Delta x_i = \bar{F}(\mathbf{y})$$

Where $\bar{F}(\cdot)$ is a function of the vector $\mathbf{y} = (y_1, \ldots, y_n)$

- Max/min $\bar{F}$, so require

$$\frac{\partial \bar{F}}{\partial y_i} = 0$$

For all $i = 1, 2, \ldots, n$.

Typically use a uniform grid so that $\Delta x_i = \Delta x = (b - a)/n$

Example: extremals for

$$F\{y\} = \int_0^1 \left[\frac{1}{2} y'^2 + \frac{1}{2} y^2 - y\right] dx$$

With $y(0) = 0$ and $y(1) = 0$. EL equation gives $y'' - y = -1$

The solution gives

$$y(x) = A \cosh x + B \sinh x + 1$$

We get $A = 1$ and $B = \ldots$

But lets do it numerically

- Take the grid $x_i = i/n$ for $i = 0, \ldots, n$

- end points $y_0 = 0$ and $y_n = 0$

- Even spacing so $\Delta x = 1/n$

- $\Delta y_i = y_{i+1} - y_i$

- $y_i' = \Delta y_i / \Delta x = n(y_{i+1} - y_i)$

$$y_i'^2 = n^2(y_i^2 - 2y_i y_{i+1} + y_{i+1}^2)$$

- Approximate $F\{y\}$ with

$$\bar{F}(\mathbf{y}) = \sum_{i=0}^{n-1} f(x_i, y_i, y_i')\Delta x$$

$$= \sum_{i=0}^{n-1} \frac{1}{2}n^2(y_i^2 - 2y_iy_{i+1} + y_{i+1}^2)\Delta x + (y_i^2/2 - y_i)\Delta x$$

$$= \sum_{i=0}^{n-1} \frac{1}{2}n(y_i^2 - 2y_iy_{i+1} + y_{i+1}^2) + \frac{y_i^2/2 - y_i}{n}$$

With $y_0 = y_n = 0$

- Include the BCs in the objective function using Lagrange multipliers

$$\bar{H}(\mathbf{y}) = \sum_{i=0}^{n-1} \frac{1}{2}n(y_i^2 - 2y_iy_{i+1} + y_{i+1}^2) + \frac{y_i^2/2 - y_i}{n} + \lambda_0 y_0 + \lambda_n y_n$$

- Hence the system:

$$\frac{\partial \bar{H}(\mathbf{y})}{\partial y_i} = \begin{cases} n(y_0 - y_1) + \frac{y_0 - 1}{n} + \lambda_0, & \text{for } i = 0 \\ n(2y_i - y_{i+1} - y_{i-1}) + \frac{y_i - 1}{n} & \text{for } i = 1, \ldots, n-1 \\ n(y_n - y_{n-1}) + \lambda_n & \text{for } i = n \end{cases}$$

- This system has $n + 3$ equations and $n + 3$ variables. So we can write it as

$$A\mathbf{z} = \mathbf{b}$$

Where the first $n + 1$ terms of $\mathbf{z}$ are $\mathbf{y}$ and the last two are the Lagrange multipliers.

There was a convoluted proof of convergence to the EL equations.
There are a few ways to improve the finite difference method

- Use a better method of 'numerical quadrature' (integration)

  - trapezoidal rule
  - Simpson's rule
  - Romberg's method

- Use a non-uniform grid - make it finer where there is more variation

- Or just use a different approach

9.2 Ritz's method

Approximate the functions (the extremal in particular) using a family of simple functions. Reduce the problem into a multivariable maximisation problem, but now we seek coefficients for our simeqimation.

- Approximate $y(x)$ with

$$y(x) = \phi_0(x) + c_1\phi_1(x) + \ldots c_n\phi_n(x)$$

Where $\phi_j(x)$ are a convenient set of functions, and find the values of c_j which produce an extremal.

- For the fixed end problem:

  - $\phi_0(x)$ satisfies the end conditions
  - $\phi_j(x_0) = \phi_j(x_1) = 0$ for $j \neq 0$.

- The ϕ can be chosen from any standard set of linearly independent functions.

The method

- Select $\{\phi_j\}_{j=0}^n$

- Approximate $y_n(x) = \phi_0(x) + c_1\phi_1(x) + \ldots + c_n\phi_n(x)$

- Approximate $F\{y\} \simeq F\{y_n\} = \int_{x_0}^{x_1} f(x, y_n, y_n')dx$

- Integrate to get $F\{y_n\} = F_n(c_1, \ldots, c_n)$

- F_n is a known function of n variables, so we can maximise (or minimise) using

$$\frac{\partial F_n}{\partial c_i} = 0$$

For all i

Assume the extremal of interest is a minimum. Then

$$F\{y\} < F\{\hat{y}\}$$

For all $\hat{y}$ within the neighbourhood of y. Assume y_n is close enough to be in that neighbourhood, so

$$F\{y\} \leq F\{y_n\} = F_n(\mathbf{c})$$

And so the simeqimation provides an **upper bound** on the minimum $F\{y\}$.

Same example as before

$$F\{y\} = \int_0^1 \left[\frac{1}{2}y'^2 + \frac{1}{2}y^2 - y \right] dx$$

With $y(0) = 0$ and $y(1) = 0$.

$$y_n(x) = \phi_0(x) + \sum_{i=1}^n c_i\phi_i(x)$$

Take $\phi_0(x) = 0$ and $\phi_i(x) = x^i(1-x)^i$. Collect powers of c_1 since thats what we take the derivative with respect to.

$$F_1(c_1) = \int_0^1 \left[\frac{c_1^2}{2}(1-2x)^2 + \frac{c_1^2}{2}x^2(1-x)^2 - c_1x(1-x) \right] dx$$

$$= \frac{c_1^2}{2}\left[x - 2x^2 + 5x^3/3 - x^4/2 + x^5/5 \right]_0^1 + c_1\left[-x^2/2 + x^3/3 \right]_0^1$$

$$= \frac{c_1^2}{2}\frac{11}{30} - \frac{c_1}{6}$$

$$\frac{dF_1}{dc_1} = \frac{11}{30}c_1 - \frac{1}{6} = 0$$

And hence $c_1 = 5/11$. So our simeqimate extremal is

$$y_1(x) = \frac{5}{11}x(1-x)$$

Giving the simeqimated functional:

$$F_1(5/11) = \frac{c_1^2}{2}\frac{11}{30} - \frac{c_1}{6} = -\frac{5}{132}$$

Which will be an upper bound on the true value of the functional on the extremal.

Alternatively try choosing $\phi_1(x) = \sin(\pi x)$.

$$F_1(c_1) = F\{c_1\phi_1\} = \int_0^1 \left[\frac{1}{2}c_1^2\phi_1'^2 + c_1^2\frac{1}{2}\phi_1^2 - c_1\phi_1\right]dx$$

$$= \frac{c_1^2}{2}\frac{1}{2}\left[\pi^2 + 1\right] - \frac{2}{\pi}c_1$$

As another example, consider the catenary

$$W_p\{y\} = mg\int_{x_0}^{x_1} y\sqrt{1 + y'^2}dx$$

Take symmetric with fixed end points, $x_0 = -1$, $x_1 = 1$ and $y(-1) = y(1) = a$.

Approximate with a polynomial

$$y(x) = \sum_{i=0}^{\infty} a_i x^i$$

Note that symmetry makes y even so

$$y(x) = \sum_{i=0}^{\infty} a_{2i} x^{2i}$$

So to second order

$$y(x) \simeq a_0 + a_2 x^2$$

And since $y(1) = y_1$ we simplify to

$$y(x) \simeq a_0 + (y_1 - a_0)x^2$$

This actually comes out to a really revolting expression in a_0. Its actually easy to plot and solve numerically though.

We see that there is a max and a min that correspond to the two solutions we initially found to the catenary problem, and we can use this to find out which was a min!

The beauty of this method is it gives us insight into the nature of the extremals we obtain. If

- approximations are near to the actual extrema

- there are no other extrema close by

- the functional is smooth (no jumps)

Then the type of extrema we get using Ritz will be the same as the analytic extrema.

9.2.1 Ritz in more than one independent variable

$$z(x,y) \simeq z_n(x,y) = \phi_0(x,y) + \sum_{i=1}^{n} c_i \phi_i(x,y)$$

Where ϕ_0 satisfies the BCs, and the $\phi_i = 0$ on the boundary. And solve the problem as before, by solving the system of equations

$$\frac{\partial F_n}{\partial c_i} = 0$$

9.3 Kantorovich's method

Approximate with

$$z(x,y) \simeq z_n(x,y) = \phi_0(x,y) + \sum_{i=1}^{n} c_i(x)\phi_i(x,y)$$

The only difference being that the c_i are now functions of one independent variable, x. Hence we get a larger class of functions to be used.

$$F\{z_n\} = \iint_{\Omega} z_n(x,y)dxdy = \sum_{i=0}^{n} \int c_i(x) \left[\int_{y_0(x)}^{y_1(x)} \phi_i(x,y) \right] dx$$

Integrating the inner integral gives

$$F\{z_n\} = \sum_{i=0}^{n} c_i(x)\Phi_i(x)dx$$

Which is just a function of x, and we can apply the E-L machinery. The method approximately separates x and y.

Ex.

$$F\{z(x,y)\} = \int_{-b}^{b} \int_{-a}^{a} (z_x^2 + z_y^2 - 2z)dxdy$$

With $z = 0$ on the boundary.

EL gives the poisson equation:

$$\frac{d}{dx}\frac{\partial f}{\partial z_x} + \frac{d}{dy}\frac{\partial f}{\partial z_y} = \frac{\partial f}{\partial z}$$

$$\frac{d}{dx}2z_x + \frac{d}{dy}2z_y = -2$$

$$\nabla^2 z(x,y) = -1$$

Try

$$z_1 = c(x)(b^2 - y^2)$$

$$\frac{\partial z_1}{\partial x} = c'(b^2 - y^2)$$

$$\left(\frac{\partial z_1}{\partial x}\right)^2 = c'^2(b^2 - y^2)^2$$

$$\left(\frac{\partial z_1}{\partial y}\right)^2 = 4c^2y^2$$

Subbing in:

$$F\{z_1(x,y)\} = \int_{-b}^{b}\int_{-a}^{a} \left(z_x^2 + z_y^2 - 2z\right) dxdy$$

$$= \int_{-a}^{a}\left[\int_{-b}^{b} \left[c'^2(b^2-y^2)^2 + 4c^2y^2 - 2c(b^2-y^2)\right] dy\right] dx$$

$$= \int_{-a}^{a}\left[c'^2(b^4y - 2b^2y^3/3 + y^5/5) + 4c^2y^3/3 - 2c(b^2y - y^3/3)\right]_{-b}^{b} dx$$

$$= \int_{-a}^{a}\left[\frac{16}{15}b^5c' + \frac{8}{3}b^3c^2 - \frac{8}{3}b^3c\right] dx$$

And use EL on this system

$$F\{z\} = \int_{-a}^{a} f(x,c,c')dx$$

$$f(x,c,c') = \frac{16}{15}b^5c' + \frac{8}{3}b^3c^2 - \frac{8}{3}b^3c$$
$$\frac{\partial f}{\partial c} = \frac{16}{3}b^3c - \frac{8}{3}b^3$$
$$\frac{\partial f}{\partial c'} = \frac{32}{15}b^5c'$$
$$\frac{d}{dx}\frac{\partial f}{\partial c'} = \frac{32}{15}b^5c''$$

$$\frac{d}{dx}\frac{\partial f}{\partial c'} - \frac{\partial f}{\partial c} = 0$$
$$\frac{32}{15}b^5c'' - \frac{16}{3}b^3c + \frac{8}{3}b^3 = 0$$
$$c'' - \frac{5}{2b^2}c = -\frac{5}{4b^2}$$

Gives

$$c = k_1\cosh\left(\sqrt{\frac{5}{2}}\frac{x}{b}\right) + k_2\sinh\left(\sqrt{\frac{5}{2}}\frac{x}{b}\right) + \frac{1}{2}$$

And the BCs $c(\pm a) = 0$ gives $k_2 = 0$ and

$$k_1 = -\frac{1}{2\cosh\left(\sqrt{\frac{5}{2}}\frac{a}{b}\right)}$$

So our solution to the problem is

$$z_1 = \frac{1}{2}(b^2-y^2)\left[1 - \frac{\cosh\left(\sqrt{\frac{5}{2}}\frac{x}{b}\right)}{\cosh\left(\sqrt{\frac{5}{2}}\frac{a}{b}\right)}\right]$$

For a better solution, try z_2

$$z_2 = c_1(x)(b^2-y^2) + c_2(x)(b^2-y^2)^2$$

Some notes

- Obviously the quality of solutions depends on:

  - The family of functions chosen

  - The number of terms used, n.

- Could test convergence by increasing n, and seeing the difference in $|F\{y_{n+1}\} - F\{y_n\}|$, but thats not guaranteed to show it

- We would instead show convergence by using a lower bound

$$\text{lower bound} \leq F\{y\} \leq \text{upper bound}$$

- We would do this by using the **complementary variation principle** - but this is complicated and we won't cover it

10 Integral constraints and non-fixed endpoints

We now can include additional constraints into the problems E.g.

- Integral constraints of form
$$\int g(x, y, y')dx = const$$

- Holonomic constraints, e.g. $g(x, y) = 0$

- Non-holonomic constraints e.g. $g(x, y, y') = 0$

- We won't consider inequality constraints

10.1 Integral Constraints

They are of the form
$$\int g(x, y, y')dx = const$$

The standard example is Dido's problem. We refer to these constraints as **isoperimetric**.

Dido was a Carthaginian queen, she fled to North Africa where a chief offered her as much land as an oxhide could contain. Cut the oxhide into thin strips and then used them to surround a patch of ground (to found Carthage). Maximise the contained area, given a fixed length of oxhide.

This falls under the set of **isoperimetric** problems, iso meaning the same, and perimetric - relating to perimeter. In general they have a spatial constraint

We can write isoperimetric problems as the problems of finding extremals of the functional $F : C^2[x_0, x_1] \to \mathbb{R}$ given by
$$F\{y\} = \int_{x_0}^{x_1} f(x, y, y')dx$$

With all the normal conditions, plus the constraint

$$G\{y\} = \int_{x_0}^{x_1} g(x, y, y')dx = L$$

Simplified form of Dido's problem - fixed end points along the coast to maximise area given a fixed length L. Maximise the area

$$F\{y\} = \int_{x_0}^{x_1} y dx$$

Encompassed by y, such that

$$G\{y\} = \int_{x_0}^{x_1} \sqrt{1 + y'^2} dx = L$$

With $y(-a) = 0$ and $y(a) = 0$. For simplicity take $2a < L \leq \pi a$.

Approach:

- perturb the curve and consider the first variation

- Cannot perturb by $\epsilon \eta$ since $G\{y + \epsilon \eta\} = L$ might be violated!!

- Use Lagrange multipliers

Minimise/maximise $f(\mathbf{x})$ for $\mathbf{x} \in \mathbb{R}^n$ With

$$g_i(\mathbf{x}) = 0$$

Lagrange Multipliers, minimise (or maximise)

$$h(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x})$$

Where λ_i are theh LM With the functional we want to find extremals of

$$H\{y\} = \int_{x_0}^{x_1} h(x, y, y') dx = \int_{x_0}^{x_1} f(x, y, y') + \lambda g(x, y, y') dx$$

The Euler-Lagrange equations become

$$\frac{d}{dx}\left(\frac{\partial h}{\partial y'}\right) - \frac{\partial h}{\partial y} = 0$$

Where $h = f + \lambda g$

For this problem:

$$H\{y\} = \int_{x_0}^{x_1} \left(y + \lambda\sqrt{1 + y'^2}\right) dx$$

For EL

$$\frac{\partial h}{\partial y} = 1, \quad \frac{\partial h}{\partial y'} = \frac{\lambda y'}{\sqrt{1 + y'^2}}$$

So EL gives

$$\frac{d}{dx}\left(\frac{\lambda y'}{\sqrt{1+y'^2}}\right) = 1$$

$$\frac{\lambda y'}{\sqrt{1+y'^2}} = x + c_1$$

$$\frac{y'}{\sqrt{1+y'^2}} = \frac{x + c_1}{\lambda} = \tilde{x}$$

$$y' = \tilde{x}\sqrt{1+y'^2}$$

$$(1 - \tilde{x}^2)y'^2 = \tilde{x}^2$$

$$y'^2 = \frac{\tilde{x}^2}{1 - \tilde{x}^2}$$

$$y = \int \frac{\tilde{x}}{\sqrt{1 - \tilde{x}^2}}dx$$

$$\tilde{x} = \frac{x + c_1}{\lambda} = \sin\theta, \quad dx = \lambda\cos\theta d\theta$$

$$y = \int \frac{\sin\theta}{\sqrt{1 - \sin^2\theta}}\lambda\cos\theta d\theta$$

$$= \lambda \int \sin\theta d\theta = -\lambda\cos\theta + c_2$$

$$x = \lambda\sin\theta - c_1$$

$$y = -\lambda\cos\theta + c_2$$

The arc-length will be $L = 2\theta_1 r$ and $x_1 = a$ gives $r = -\lambda = a/\sin\theta_1$

And determine θ_1 from

$$\sin\theta_1 = \frac{2a}{L}\theta_1$$

And compute

$$r = a/\sin\theta_1$$

And can see that

$$c_2 = -\cos(\theta_1)$$

Using $y_1 = 0$.

What if we had a realistic coast? Let the coast be defined as $c(x)$.

$$\text{Area} = \int_{x_0}^{x_1} y - c\,dx$$

But note that c doesn't depend on y, y'. so the EL equations are unchanged provided $c(x) < y(x)$ for the extremal.

New catenary problem. Fix the length of the wire.

Minimise the potential energy

$$W_p\{y\} = mg\int_{x_0}^{x_1} y\sqrt{1+y'^2}dx$$

But constrain with the length L

$$G\{y\} = \int_{x_0}^{x_1} \sqrt{1 + y'^2} dx = L$$

So seek extremals of

$$H\{y\} = \int_{x_0}^{x_1} (y + \lambda)\sqrt{1 + y'^2} dx$$

So we compute

$$H(y, y') = y'\frac{\partial h}{\partial y'} - h = const$$

Where this H is different bc of notation reuse

$$H(y, y') = \frac{(y + \lambda)y'^2}{\sqrt{1 + y'^2}} - (y + \lambda)\sqrt{1 + y'^2} = const$$

Let $u = y + \lambda$ and $u' = y'$ so

$$\frac{uu'^2}{\sqrt{1 + u'^2}} - u\sqrt{1 + u'^2} = c_1$$

$$\frac{u^2}{1 + u'^2} = c_1^2$$

$$u = c_1 \cosh \frac{x - c_2}{c_1}$$

$$y = c_1 \cosh \frac{x - c_2}{c_1} - \lambda$$

Two end points, $y(x_0) = y_0$ and $y(x_1) = y_1$ where $x_0 = -1$ and $x_1 = 1$, and length constraint

$$L\{y\} = \int_{-1}^{1} \sqrt{1 + y'^2} dx$$

$$= \int_{-1}^{1} \cosh\left(\frac{x}{c_1}\right) dx$$

$$= c_1 \sinh\left(\frac{x}{c_1}\right) \Big|_{-1}^{1}$$

$$= 2c_1 \sinh\left(\frac{1}{c_1}\right) = L$$

And use that to calculate c_1 given L, and then we calculate λ to satisfy the end heights $y_0 = y_1$.

The functional becomes

$$F\{y\} = \dots$$

$$= c_1 + \frac{c_1^2}{2}\sinh(2/c_1) - 2\lambda c_1 \sinh(1/c_1)$$

This assumes $y < 0$ is possible.

If $y < 0$ isn't possible, the wire will lie on the ground sometimes. Also L has to be greater than $x_1 - x_0$ to obtain physically possible solutions, and if L much greater than y_1 the wire will drag on the ground.

10.2   Rigid Extremals

Extremals that cannot be perturbed, but still satisfy the constraint. E.g.

$$G\{y\} = \int_0^1 \sqrt{1 + y'^2}dx = \sqrt{2}$$

With BCs $y(0) = 0$ and $y(1) = 1$. This constraint will only be satisfied with $y = x$, so we can't perturb around this curve to find conditions for viable extremals.

11   Formulations, conservation laws and classification