28/04/11 Variational Principles ①

Finite Dimensions extremal values

$\max_x f(x) = M = f(x_M)$ $\min_x f(x) = m = f(x_m)$ Stationary points, $f'(x) = 0$

Functionals are functions on the spaces of functions.

Given a piece of string, length L, what is the maximum area we can enclose using the string as a closed curve? Describe the curve as

$t \mapsto (x(t), y(t)) \in \mathbb{R}^2$ $x(0) = x(1)$, $y(0) = y(1)$

$\text{Area} = \frac{1}{2}\int x\,dy - y\,dx = \frac{1}{2}\int (x\dot{y} - \dot{x}y)\,dt$

$\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n)\}$, $\underline{e_j} = (0, \ldots, \overset{j^{th}\text{ slot}}{1}, 0)$

Linear functions: $L: \mathbb{R}^n \to \mathbb{R}$, $L(\alpha\underline{v} + \beta\underline{w}) = \alpha L(\underline{v}) + \beta L(\underline{w})$

matrix notation

Write $\underline{x} = x_j \underline{e_j}$. $L(\underline{x}) = L(x_j \underline{e_j}) = x_j L(\underline{e_j})$

$L^T\underline{x} = \underline{L}\cdot\underline{x}$, $\underline{L} = (L_1, L_2, \ldots, L_n)$, $L_j = L(\underline{e_j})$

Definition 1.1.1 $f: \mathbb{R}^n \to \mathbb{R}$ is differentiable at $\underline{x}$ if $\exists$ a linear map $L: \mathbb{R}^n \to \mathbb{R}$ such that $f(\underline{x} + \underline{v}) - f(\underline{x}) - L(\underline{v}) = o(\|\underline{v}\|)$

Here: $\|\underline{v}\| = (\sum v_j^2)^{\frac{1}{2}}$, $q(v) = o(\|v\|)$ means $\frac{q(v)}{\|v\|} \to 0$ as $\|v\| \to 0$

This means f can be well approximated by linear functions. In full:

$\forall \varepsilon > 0, \exists \delta > 0$ such that

$\|v\| < \delta \Rightarrow \left|\frac{f(\underline{x}+\underline{v}) - f(\underline{x}) - L(\underline{v})}{\|v\|}\right| < \varepsilon$

$\Leftrightarrow |f(\underline{x}+\underline{v}) - f(\underline{x}) - L(\underline{v})| < \varepsilon\|v\|$

Exercise: Put $n=1$, and show that f is differentiable at x as just defined

$\Leftrightarrow \lim_{v\to 0}\frac{f(x+v) - f(x)}{v}$ exists, and equals $f'(x) = L$

To obtain partial derivatives, we put $\underline{v} = h\underline{e_j}$, $h \in \mathbb{R}$, then

we find that $L(\underline{e_j}) = \lim_{h\to 0}\frac{f(\underline{x}+h\underline{e_j}) - f(\underline{x})}{h} = \frac{\partial f(\underline{x})}{\partial x_j}$

We can reexpress this definition as

i.e. Taylor Expansion is valid to 1^{st} order.

$$f(\underline{x}+\underline{v}) - f(\underline{x}) - \sum_j \frac{\partial f(\underline{x})}{\partial x_j} v_j = o(\|\underline{v}\|)$$

$C^1(\mathbb{R}^n) = \{ f: \mathbb{R}^n \to \mathbb{R}$ with continuous partial derivatives $\frac{\partial f}{\partial x_j} \}$

These are differentiable at all $\underline{x} \in \mathbb{R}^n$

$C^2(\mathbb{R}^n) = \{ f: \mathbb{R}^n \to \mathbb{R}$ such that partial derivatives up to 2^{nd} order are continuous $\}$

These can be approximated by Taylor expansion up to 2^{nd} order.

$$f(\underline{x}+\underline{v}) = f(\underline{x}) + \sum_j \frac{\partial f(\underline{x})}{\partial x_j} v_j + \sum_{i,j} \frac{\partial^2 f(\underline{x})}{\partial x_i \partial x_j} v_i v_j + o(\|v\|^2)$$

Lemma 1.1.3 (1^{st} order condition for max or min). $f \in C^1$

If $f(x) \geq f(y)$ $\forall \in \mathbb{R}^n$ then $\frac{\partial f(x)}{\partial x_j} = 0$ $\forall j$
$\leq$ (also true for global minimum)

This is the same for local max and min.

Lemma 1.2.1 Let $f \in C^2(\mathbb{R}^n)$, and $\nabla f = \underline{0}$, $(\nabla f)_i = \frac{\partial f}{\partial x_i}$

i) If $\underline{x}$ is a local min, then $A_{ij} = \frac{\partial^2 f(\underline{x})}{\partial x_i \partial x_j}$ has non-negative eigenvalues.
" max " non-positive eigenvalues

ii) If all eigenvalues of A_{ij} are positive $\Rightarrow \underline{x}$ is a strict local min
" " negative $\Rightarrow \underline{x}$ is a strict local max

Notation A symmetric matrix $A = (A_{ij})$ is non-negative (or negative) if all its eigenvalues are non-negative (negative)

Equivalent formulation: A is positive $\Leftrightarrow v^T A v > 0$ $\forall \underline{v} \neq \underline{0}$

Note, same definition for $f: S \to \mathbb{R}$

1.3 Convexity

Convex functions are an important class of functions where it is relatively easy to understand max/min values and stationary points

Definition: $S \subset \mathbb{R}^n$ is convex if $\forall \underline{x}_1, \underline{x}_2 \in S$, $\{(1-t)\underline{x}_1 + t\,\underline{x}_2\}_{0 \leq t \leq 1} \in S$

2 $f: \mathbb{R}^n \to \mathbb{R}$ is convex if $\forall \underline{x}_1, \underline{x}_2$ $f((1-t)\underline{x}_1 + t\underline{x}_2) \leq (1-t)f(\underline{x}_1) + t f(\underline{x}_2)$

02/05/11

Variational Principles 2

Convexity

The epigraph $E_f \equiv \{ (\underline{x}, z) \in \mathbb{R}^{n+1} : z \geq f(\underline{x}) \}$

Exercise $E_f \subset \mathbb{R}^{n+1}$ convex $\Leftrightarrow$ f convex.

Proposition 1.3.1 If $f \in C^1(\mathbb{R}^n)$, the following are equivalent:

i) f convex ii) $f(\underline{y}) \geq f(\underline{x}) + \nabla f(\underline{x}) \cdot (\underline{y} - \underline{x}) \quad \forall \underline{x}, \underline{y}$

iii) $(\nabla f(\underline{x}) - \nabla f(\underline{y})) \cdot (\underline{x} - \underline{y}) \geq 0$

Proof (i) $\Rightarrow$ (ii) $H(t) = (1-t) f(\underline{x}) + t f(\underline{y}) - f[(1-t)\underline{x} + t\underline{y}]$

$H(t) \geq 0$ by convexity, $0 \leq t \leq 1$. $H(0) = 0$.

$\dot{H}(0) = -f(\underline{x}) + f(\underline{y}) + (\underline{y} - \underline{x}) \cdot \nabla f(\underline{x})$

$\dot{H}(0) = \lim_{t \downarrow 0} \frac{H(t) - H(0)}{t} \geq 0$, so we have (ii).

(iii) $\Rightarrow$ (ii) $\underline{v} = \underline{y} - \underline{x}$, consider $g(t) = f(\underline{x} + t\underline{v})$

$$f(\underline{y}) - f(\underline{x}) = g(1) - g(0) = \int_0^1 \frac{d}{dt} g(t)\, dt = \int_0^1 \underline{v} \cdot \nabla f(\underline{x} + t\underline{v})\, dt$$

So
$$f(\underline{y}) - f(\underline{x}) - \nabla f(\underline{x}) \cdot (\underline{y} - \underline{x}) = \int_0^1 \underline{v} \cdot \nabla f(\underline{x} + t\underline{v})\, dt - \nabla f(\underline{x}) \cdot \underline{v}$$
$$= \int_0^1 \underline{v} \cdot (\nabla f(\underline{x} + t\underline{v}) - \nabla f(\underline{x}))\, dt$$

but (iii) implies that 'integral is ≥ 0, so we recover (ii).

(ii) $\Rightarrow$ (i) $f(\underline{y}) \geq f(\underline{z}) + \nabla f(\underline{z}) \cdot (\underline{y} - \underline{z})$, $f(\underline{x}) \geq f(\underline{z}) + \nabla f(\underline{z}) \cdot (\underline{x} - \underline{z})$

$\Rightarrow (1-t) f(\underline{y}) + t f(\underline{x}) \geq (t + 1 - t) f(\underline{z}) + \nabla f(\underline{z}) \cdot [(1-t)(\underline{y} - \underline{z}) + t(\underline{x} - \underline{z})]$

$(1-t) f(\underline{y}) + t f(\underline{x}) \geq f(\underline{z}) + \nabla f(\underline{z}) \cdot [(1-t)\underline{y} + t\underline{x} - \underline{z}]$

Choose $\underline{z} = (1-t)\underline{y} + t\underline{x}$, and hence f is convex.

Geometically, ii) means that f lies above its tangent plane at any point.

Graph of $f \Leftrightarrow z - f(\underline{x}) = 0$. So the normal at $(\underline{x}, f(\underline{x}))$ is proportional to the gradient $(-\nabla f(\underline{x}), 1) \in \mathbb{R}^{n+1}$

The tangent plane at this point is given by $(\underline{y}-\underline{x}, z - f(\underline{x})) \cdot (-\nabla f(\underline{x}), 1) = 0$

$\Leftrightarrow z - f(\underline{x}) - \nabla f(\underline{x}) \cdot (\underline{y} - \underline{x}) = 0$

Also, (iii) can be interpreted as saying f is convex, if along each line, its (directional) derivative is increasing. For $n=1$, $(f'(x) - f'(y))(x-y) \geq 0$ i.e. f is non-decreasing, so we can think of convex functions as functions whose derivatives are monotonic, and non decreasing.

Definition

f is strictly convex if $f[(1-t)\underline{x} + t\underline{y}] < (1-t)f(\underline{x}) + t\underline{y}$ $\quad 0<t<1$, $\underline{x} \neq \underline{y}$

$\Leftrightarrow f(\underline{y}) > f(\underline{x}) + \nabla f(\underline{x}) \cdot (\underline{y} - \underline{x}) \quad \underline{x} \neq \underline{y}$

$\Leftrightarrow [\nabla f(\underline{x}) - \nabla f(\underline{y})] \cdot (\underline{x} - \underline{y}) > 0 \quad \underline{x} \neq \underline{y}$

Proposition 3.2

Corollary i) If f is convex, $\nabla f(\underline{x}) = \underline{0}$, then $f(\underline{x})$ is a global min by (ii)

If f is strictly convex, then $\underline{x}$ is a strict global minimum.

ii) $\nabla f(x) = \underline{b}$ for $\underline{b} \in \mathbb{R}^n$, $f \in C^1$ and strictly convex.

Then, there can only be one solution ("uniqueness holds")

If $\nabla f(\underline{x}) = \underline{b}$, $\nabla f(\underline{x}) = \underline{b}$:

$0 = (\nabla f(\underline{x}) - \nabla f(\underline{y})) \cdot (\underline{x} - \underline{y})$ for $\underline{x} \neq \underline{y}$, impossible, so $\underline{x} = \underline{y}$

e.g. $f(\underline{x}) = \frac{1}{2} x^T A x = \frac{1}{2} x_i A_{ij} x_j$.

Then $\nabla f(\underline{x}) = A\underline{x}$. If all eigenvalues of $A > 0$, then f is convex, so $A\underline{x} = \underline{b}$ has a unique solution.

02/05/11

Variational Principles ②

Proposition If $f \in C^2(\mathbb{R}^n)$ then

i) f convex $\Leftrightarrow \frac{\partial^2 f(\underline{x})}{\partial x_i \partial x_j} \geq 0 \ \forall \underline{x}$

ii) If $\frac{\partial^2 f(\underline{x})}{\partial x_i \partial x_j}$ is $> 0 \ \forall \underline{x}$, then f is strictly convex.

Proof

$$(\nabla f(\underline{x}) - \nabla f(\underline{y})) = \left(\int_0^1 \frac{d}{dt}\left[\nabla f(\underline{y} + t(\underline{x}-\underline{y}))\right] dt\right) \cdot (\underline{x} - \underline{y})$$

$$= \int_0^1 \frac{\partial^2 f(\underline{y} + t(\underline{x}-\underline{y}))}{\partial x_i \, \partial x_j} (x_i - y_i)(x_j - y_j) \, dt$$

$$= \int_0^1 \frac{\partial^2 f}{\partial x_i \partial x_j} (\underline{y} + t(\underline{x}-\underline{y})) \, V_i V_j \, dt \qquad \underline{V} = \underline{x} - \underline{y}$$

If $\frac{\partial^2 f}{\partial x_i \partial x_j}$ is > 0 everywhere, this is > 0 for $\underline{V} \neq \underline{0}$, so for $\underline{x} \neq \underline{y}$, i.e. (iii) in 1.3.2 holds, so f is strictly convex, proving (ii). (i) is similar.

Moral - Convex functions are functions whose second derivative $\frac{\partial^2 f}{\partial x_i \partial x_j}$ is ≥ 0 everywhere.

04/05/11

Variational Principles ③

$f(x)$ is concave if $-f(x)$ is convex. Just reverse all inequalities.

Prove the following have a unique solution:

i) $x + x^9 = 10$ ii) $2x + 4x^3 + y = 10$, $x + 10y + 36y^3 = 20$

i) 1st Method: The Intermediate Value Theorem

$F(x) = x + x^9$ $F(0) = 0$, $F(10) = 10 + 10^9 > 10$

$\exists x_* \in (0,10)$ with $F(x_*) = 10$. F is strictly increasing so the solution is unique.

2nd Method: Consider $g(x) = \frac{x^2}{2} + \frac{x^{10}}{10} - 10x$

$\inf_x [g(x)] \le g(0) = 0$. Since $g(x) \to \infty$ as $x \to \infty$, $\exists L > 0$

such that $g(x) > 1$ if $|x| > L$.

$\inf_x [g(x)] = \inf_{x \in [-L,L]} [g(x)] = \min_{x \in [-L,L]} g(x) = \min_x g(x)$

So $\exists x_* \in (-L, L)$ with $g(x_*) = \min_x g(x)$

$\Rightarrow g'(x_*) = 0 \iff x + x^9 = 10$ is solved by x_*

To prove uniqueness, notice that g is strictly convex since $g'' = 1 + 9x^8 > 0$

Therefore $g'(x) = b$ can have at most one solution by the Corollary to 1·3·2.

ii) Define $g(\underline{x}) = x^2 + x^4 + xy + 5y^2 + 9y^4 - 10x - 20y$

Existence of solution: Remembering $xy \le \frac{x^2 + y^2}{2}$, it follows that

$|\underline{x} \cdot \underline{y}| \le |\underline{x}||\underline{y}|$

$g(x,y) \to \infty$ as $x^2 + y^2 \to \infty$. So as in i) $\exists (x_*, y_*)$ such

that $g(x_*, y_*) = \min_{(x,y) \in \mathbb{R}^2} g(x,y)$. Then, by calculus, $\frac{\partial g}{\partial x}$ and $\frac{\partial g}{\partial y}$

are equal to 0 at this point.

Uniqueness

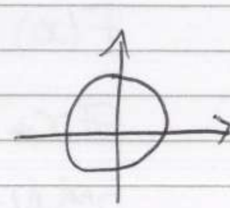
As in i), we can prove this by proving strict convexity of g.

$\begin{pmatrix} x \\ y \end{pmatrix} \begin{pmatrix} y \\ x \end{pmatrix}$ The 2^{nd} derivative matrix (Hessian) is $\begin{pmatrix} g_{xx} & g_{xy} \\ g_{yx} & g_{yy} \end{pmatrix} = \begin{pmatrix} 2+12x^2 & 1 \\ 1 & 10+108y^2 \end{pmatrix}$

Trace $= 12 + 12x^2 + 108y^2 > 0$

det $= (2+12x^2)(10+108y^2) - 1 > 0$

$\Rightarrow$ Both eigenvalues are positive. So g is strictly convex, and $\nabla g = \underline{0}$ at at most one place.



1.4 Langrange method for constrained minimisation problems

Simple Example: Maximise $f(x,y) = y^2$ on the circle

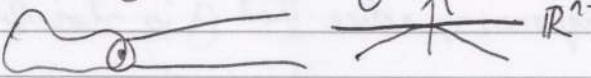
$C_e = \{ g(x,y) = x^2 + y^2 - 1 = 0 \}$

Clearly f is maximised by $(0, \pm 1)$. $\nabla f = (0, 2y) = (0, \pm 2)$

Notice at $(0, \pm 1)$, ∇f is perpendicular to C_e; this is to be expected, since you only expect a zero derivative in allowed directions, i.e tangent to C_e.

Remarks on Constraint Sets $C_e = \{ g(\underline{x}) = 0 \}$ (a hypersurface)

If $g \in C^1(\mathbb{R}^n)$ and $\nabla g(\underline{x}) \neq \underline{0}$ for all x, this set C_e is a hypersurface, i.e.



Normal space $N_x C_e = \{ \underline{v} \in \mathbb{R}^n : \underline{v} \propto \nabla g(\underline{x}) \}$

Tangent space $T_x C_e = \{$ Velocity vectors of curves passing through point $\underline{x} \in C_e \}$

$$T_x C_e = \left\{ \frac{d\underline{X}}{dt}\Big|_{t=0} \text{ where } \underset{X(0) = \underline{x}}{\underline{X}(t)} \text{ is a } C^1 \text{ curve in } C_e \right\}$$

Notice $\underline{X}(t) \in C_e \Leftrightarrow g(\underline{x}(t)) = 0 \ \forall t$

Chain rule: $\frac{d}{dt} g(\underline{x}(t)) = \nabla g(\underline{x}(t)) \cdot \dot{\underline{X}}(t) = 0$

With $t = 0 \Rightarrow \dot{\underline{X}}(0) \perp \nabla g(\underline{x})$, all tangent vectors are perpendicular to the normal

Geometrically $\mathbb{R}^n = T_x C_e \oplus N_x C_e$ (orthogonal)

To conclude, if $\underline{x}_*$ is a max or min of f on C_e, then $\nabla f(\underline{x}_*)$ is $\perp$ to C_e i.e. $\nabla f(\underline{x}_*) \in N_{x_*} C_e \Leftrightarrow \nabla f(\underline{x}_*) = \lambda \nabla g(\underline{x}_*)$. Lagrange method: to solve

06/05/

Variational Principles ④

Assume $f \in C^1$

<u>Theorem 1.4.1a</u> If $f|_{C_e}$ has a max or min at $\underline{x} \in C_e$, then $\frac{\partial}{\partial x_j}[h(\underline{x},\lambda)] = 0$, $\frac{\partial}{\partial \lambda}[h(\underline{x},\lambda)] = 0$

where $h(\underline{x},\lambda) = f(\underline{x}) - \lambda g(\underline{x})$ and λ is the Lagrange Multiplier which we must also solve for.

<u>Notice</u> $g(\underline{x}) = 0 \Leftrightarrow \frac{\partial h}{\partial \lambda} = 0$

$\frac{\partial h}{\partial x_j} = 0 \Leftrightarrow \nabla f = \lambda \nabla g$

$f|_{C_e}$ has a max/min at $\underline{x} \Leftrightarrow \phi(t) = f(\underline{x}(t))$ has a max/min at $t=0$ for all curves $t \to \underline{x}(t) \in C_e$ with $\underline{x}(0) = \underline{x}$

$\dot{\phi}(t) = \nabla f(\underline{x}(t)) \cdot \dot{\underline{x}}(t)$. $\dot{\phi}(t) = 0 \Leftrightarrow \nabla f \perp T_x C_e$

i.e. $\nabla f \propto \nabla g$. Then $\exists \lambda$ such that $\nabla f = \lambda \nabla g$ ~~for~~ at $\underline{x}$. $\square$

<u>Example</u>

Maximise $f(x,y) = y^2$ on $C_e = \{g : x^2+y^2-1 = 0\}$

<u>Solution</u> $h(x,y,\lambda) = y^2 - \lambda(x^2+y^2-1)$

$\frac{\partial h}{\partial x} = 0 \Leftrightarrow -2\lambda x = 0$ $\qquad$ $\frac{\partial h}{\partial y} = 0 \Leftrightarrow 2y - 2\lambda y = 0$

$\frac{\partial h}{\partial \lambda} = 0 \Leftrightarrow x^2+y^2=1$. Either $\lambda = 0$ <u>or</u> $x = 0$, giving

$\Rightarrow y^2 = 1$, $y = \pm 1$, maxima of f. Then we obtain $\lambda = 1$ for $y \neq 0$

If $\lambda = 0$, $y = 0$, $x^2 = 1$, i.e. $x = \pm 1$, minima of f.

<u>Theorem 1.4.1b</u> If f, g are also in C^2 and $\underline{x}$ is the max (min) of $f|_{C_e}$ then $\frac{\partial^2}{\partial x_i \partial x_j} h(\underline{x},\lambda)$ is ≤ 0 (≥ 0) (as a symmetric matrix).

<u>Why:</u> If $f|_{C_e}$ has a max at $\underline{x}$ then $\phi(t) = f(\underline{x}(t))$ has a max at $t=0$ for all curves $t \to \underline{x}(t) \in C_e$ with $\underline{x}(0) = \underline{x}$

Therefore $\ddot{\phi}(0) \leq 0$.

Chain Rule $\dot{\phi}(t) = \nabla f(\underline{x}(t)) \cdot \dot{\underline{X}}(t)$

$$\ddot{\phi}(t) = \nabla f(\underline{x}(t)) \cdot \ddot{\underline{X}}(t) + \frac{\partial^2 f(\underline{x}(t))}{\partial x_i \partial x_j} X_i(t) X_j(t)$$

Now since $\underline{X}(t) \in C_c$, $g(\underline{x}(t)) = 0 \ \forall t$

$$\Rightarrow \nabla g(\underline{x}(t)) \cdot \dot{\underline{X}}(t) = \frac{d}{dt}(0) = 0$$

$$\frac{\partial g(\underline{x}(t))}{\partial x_i} \dot{X}_i(t) = 0$$

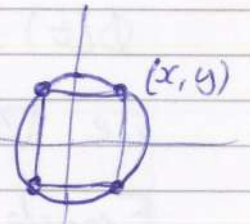
$$\Rightarrow \frac{\partial g}{\partial x_i}(\underline{x}(t)) \ddot{X}_i(t) + \frac{\partial^2 g}{\partial x_i \partial x_j} X_i(t) X_j(t) = 0$$

Put $t = 0$, $\underline{X}(0) = \underline{x}$, where $\nabla f(\underline{x}) = \lambda \nabla g(\underline{x})$

$\Rightarrow \ddot{\phi}(0)$ gives $\nabla f(\underline{x}(0)) \cdot \ddot{\underline{X}}(0) = \lambda \frac{\partial g(x)}{\partial x_i} \ddot{X}_i(0)$

$$= -\lambda \frac{\partial^2 g(x)}{\partial x_i \partial x_j} X_i(0) X_j(0)$$

Therefore $\ddot{\phi}(0) = \frac{\partial^2 h}{\partial x_i \partial x_j} \dot{X}_i(0) \dot{X}_j(0)$



Example

Find the rectangle of maximum area inscribed in a unit circle.

$g(x,y) = x^2 + y^2 - 1 = 0$

Maximise $f(x,y) = 4xy$.

$\Big\rangle$ Introduce $h(x,y,\lambda) = 4xy - \lambda(x^2+y^2-1)$

$\frac{\partial h}{\partial x} = 0 \Leftrightarrow 4y - 2\lambda x = 0$ $\quad$ $\frac{\partial h}{\partial y} = 0 \Leftrightarrow 4x - 2\lambda y = 0$

$\Rightarrow \frac{y}{x} = \frac{x}{y}$, so the maximum area is attained with a square.

Example 2 A probability distribution on $\{1, 2, \dots, n\}$ consists of $p_j \in [0,1]$ with $\sum_{j=1}^{n} p_j = 1$. Define $S(\underline{p}) = -\sum_j p_j \ln p_j$ "entropy"

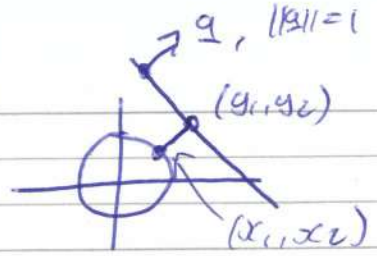
Find the distribution to maximise S. $h = S - \lambda(\sum p_j - 1)$

$$\frac{\partial h}{\partial p_j} = 0 \Leftrightarrow -(1 + \ln p_j) - \lambda = 0$$

$$\Rightarrow p_1 = p_2 = \dots = p_n \Rightarrow p_i = \frac{1}{n}. \ \left[\text{so } \lambda = -\left(1 + \ln \tfrac{1}{n}\right)\right]$$

06/05/11

Variational Principles ④



$\underline{y}\cdot\underline{a} = c$

Vector Constraints

$g_1(\underline{x}) = x_1^2 + x_2^2 - 1 = 0$

Find the closest points on the circle to the line.

$g_2(\underline{y}) = \underline{y}\cdot\underline{a} - c = 0$

Minimise $f(\underline{x}, \underline{y}) = \|\underline{x} - \underline{y}\|^2 = (x_1 - y_1)^2 + (x_2 - y_2)^2$

Assume $c > 1$, so the line is outside the circle.

Method Introduce λ_1, λ_2.

$h(\underline{x}, \underline{y}, \lambda_1, \lambda_2) = f - \lambda_1 g_1 - \lambda_2 g_2$

$\frac{\partial h}{\partial x_i} = 0$, $\frac{\partial h}{\partial y_i} = 0 \iff \begin{pmatrix} 2(\underline{x} - \underline{y}) \\ 2(\underline{y} - \underline{x}) \end{pmatrix} - \lambda_1 \begin{pmatrix} 2\underline{x} \\ 0 \end{pmatrix} - \lambda_2 \begin{pmatrix} 0 \\ \underline{a} \end{pmatrix} = 0$

Neither $\lambda_1, \lambda_2 = 0$ since this would imply $\underline{x} = \underline{y}$ but then $c = \underline{y}\cdot\underline{a} \le 1$

$$\left.\begin{array}{l} 2(1-\lambda_1)\underline{x} = 2\underline{y} \\ 2(\underline{y} - \underline{x}) = \lambda_2 \underline{a} \end{array}\right\} \Rightarrow \underline{x} = -\frac{\lambda_2}{2\lambda_1}\underline{a}$$

. Then $\underline{y} = c\underline{a}$

Then $\underline{x} = \underline{a}$ gives the min, $\underline{x} = -\underline{a}$ gives the max.

To maximise / minimise f with N constraints $g_1 = 0, g_2 = 0, \ldots, g_N = 0$

Define $h = f - \lambda_1 g_1 - \lambda_2 g_2 - \cdots - \lambda_N g_N = f - \underline{\lambda}\cdot\underline{g}$

The Lagrange augmented function.

09/05/11

Variational Principles ⑤

~~For~~ Legendre Transform and Applications

For $f \in C^1(\mathbb{R})$ define $f^*(p) = \sup(px - f(x))$

We called $f^*(p) = g(p)$ the Legendre Transform.

$\text{Domain}(f^*) = \{p : \sup(px - f(x)) < \infty\}$

Examples

1. $f(x) = ax^2,\ a > 0.\ f^*(p) = \sup_{x\in\mathbb{R}}(px - ax^2)$, attained for $x = x(p)$, $\frac{d}{dx}(px - ax^2)\big|_{x(p)} = 0$, so $x(p) = \frac{p}{2a}$
$\Rightarrow f^*(p) = p\left(\frac{p}{2a}\right) - a\left(\frac{p}{2a}\right)^2 = \frac{p^2}{4a}$ Domain$(f^*) = \mathbb{R}$.
$(f^*)^*(x) = \sup_p(xp - f^*(p)) = \frac{x^2}{4 - (\frac{1}{4a})} = ax^2 = f(x)$
2. $f(x) = ax^2,\ a < 0.\ \sup_{x\in\mathbb{R}}(px - ax^2) = +\infty\ \forall p$
Domain $(f^*) = \emptyset$
3. $f(x) = 0$, $\sup_{x\in\mathbb{R}}(px - 0) = \begin{matrix} +\infty & p \neq 0 \\ 0 & p = 0 \end{matrix} \Rightarrow \text{Domain}(f^*) = \{0\}$
4. $f(x) = ax + b$, Domain $(f^*) = \{a\}$, $f^*(a) = -b$

Proposition 1.5.1 Given $f \in C(\mathbb{R})$, Domain $(f^*) \subset \mathbb{R}$ is a convex subset, and f is a convex function.

Proof $f^*[tp_1 + (1-t)p_2] = \sup[(tp_1 + (1-t)p_2)x - f(x)]$ $0 \le t \le 1$

$tp_1x + (1-t)p_2x - f(x) = t[p_1(x) - f(x)] + (1-t)[p_2x - f(x)]$

$*$ $\le tf^*(p_1) + (1-t)f^*(p_2)$

If $p_1, p_2 \in$ Domain (f^*), then $\sup[(tp_1 + (1-t)p_2)x - f(x)] < \infty$

i.e. $tp_1 + (1-t)p_2 \in$ Domain (f^*), $0 \le t \le 1$

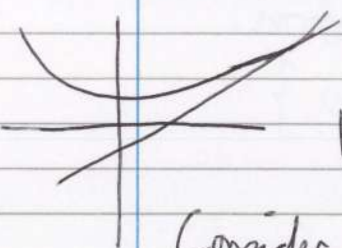
Taking the sup of $*$ gives

$f^*[tp_1 + (1-t)p_2] \le tf^*(p_1) + (1-t)f^*(p_2)$

Theorem 1.5.2

If f is convex, then $(f^*)^* = f$. This is because,

$f = (f^*)^* \Rightarrow f$ convex by the previous proposition.



Fix p, then $\sup (px - f(x)) = px(p) - f(x(p))$

We look at $f^{**}(x) = \sup_p (px - f^*(p))$

Consider the line tangent to $f(x)$ with slope p.

$$f'(x(p)) = p = \frac{y - f(x(p))}{x - x(p)}$$

$$\iff y = px - px(p) + f(x(p)) = px - f^*(x)$$

So $f^{**}(x) = \sup$ (tangent lines to the curve $f(x)$ at the point x)

f is convex, so it lies above all of it's tangent lines, so $f^{**}(x) = f(x$

In n-dimensions, we define similarly $f^*(\underline{p}) = \sup_{\underline{x} \in \mathbb{R}^n} (\underline{p} \cdot \underline{x} - f(\underline{x}))$

Example: $f(\underline{x}) = \frac{1}{2} x^T A x, \Rightarrow f^*(p) = \frac{1}{2} p^T (A^{-1}) p$ (A symmetric)

Applications

$f^*(p) = \sup (px - f(x)) \geq px - f(x), \ \forall x, p$

$\Rightarrow px \leq f(x) + f^*(p)$ (Young's Inequality)

$f(x) = ax^2, a > 0, f^*(p) = \frac{p^2}{4a} \Rightarrow px \leq ax^2 + \frac{p^2}{4a}, \forall a > 0$

$f(x) = \frac{x^\alpha}{\alpha}, 1 < \alpha < \infty, f^*(p) = \frac{p^\beta}{\beta}$, where $\frac{1}{\beta} + \frac{1}{\alpha} = 1$

$\Rightarrow xp \leq \frac{x^\alpha}{\alpha} + \frac{p^\beta}{\beta}$, e.g. $px \leq \frac{x^4}{4} + \frac{3}{4} p^{\frac{4}{3}}$

1. Mechanics: Newton's Equation $\frac{d^2 x}{dt^2} + \nabla V(\underline{x}) = 0$

The Hamiltonian Formulation: $H(\underline{p}, \underline{x}) = \frac{1}{2}|p|^2 + V(\underline{x})$

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial x_i}$$

Lagrangian Formulation: $L(\underline{x}, \dot{\underline{x}}) = \frac{1}{2}|\dot{\underline{x}}|^2 - V(\underline{x})$

09/05/11

Variational Principles 5

"Euler-Lagrange Equation" is $\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_j}\right) - \frac{\partial L}{\partial x_j} = 0$

$$H(\underline{x}, \underline{p}) = \sup_{\dot{\underline{x}}} \left(\underline{p} \cdot \dot{\underline{x}} - L(\underline{x}, \dot{\underline{x}})\right)$$

$$L(\underline{x}, \dot{\underline{x}}) = \sup_{\underline{p}} \left(\underline{p} \cdot \dot{\underline{x}} - H(\underline{x}, \underline{p})\right)$$

$\underline{x}$ is fixed here, so this is example 1 from the definition of f^*.

2 Economics

Revenue $= S(\underline{r})$, resource r_j has price p_j

Profit $= S(\underline{r}) - \underline{p} \cdot \underline{r}$

$M(\underline{p}) = \sup_{\underline{r}} \left(S(\underline{r}) - \underline{p} \cdot \underline{r}\right)$. This is the profit as a function of the world prices of resources. Notice $M(\underline{p}) = (-S)^*(-\underline{p})$.

3. Thermodynamics

Internal Energy $U = U(S, V)$, $S =$ entropy, $V =$ volume in thermal isolation. Instead, put this in a fixed temperature "reservoir". $F(T, V)$, $T =$ temperature

$$F(T, V) = \inf_{S} \left(U(S, V) - ST\right) = -\sup_{S} \left(ST - U(S, V)\right)$$

11/05/11 Variational Principles ⑥

Functionals and Calculus of Variables

A functional $F : V \to \mathbb{R}$ is a function, and V is a set of functions

e.g. $V = C^r(\mathbb{R}^n) = \{f : \mathbb{R}^n \to \mathbb{R}$ with continuous partial derivatives up to order $r\}$

$V = C^\infty_{per}([-\pi, \pi]) = \{f \in C^\infty(\mathbb{R}) : f(x+2\pi) = f(x) \ \forall x \in \mathbb{R}\}$

$V = C([a,b]) = \{$continuous functions $f : [a,b] \to \mathbb{R}\}$

Examples of functionals

i) $\delta_{\underline{x}} : C(\mathbb{R}^n) \to \mathbb{R}$, $f \mapsto f(\underline{x})$. "Dirac functional at $\underline{x}$".

Integral functionals
ii) $I_p : C^\infty_{per}([-\pi, \pi]) \to \mathbb{R}$. $f \mapsto I_p[f] = \int_{-\pi}^{\pi} |f(x)|^p dx$, $1 \leq p \leq \infty$
iii) $L : C^\infty(\mathbb{R}) \to \mathbb{R}$. $y \mapsto L[y] = \int_a^b \sqrt{1+y'^2}\, dx$

We say a functional $F : V \to \mathbb{R}$ is continuous at $f \in F$ if $\forall \varepsilon > 0$, $\exists \delta > 0$ such that $|F(f+g) - F(f)| < \varepsilon$ if $\|g\| < \delta$

But what is $\| \ \|$? For $V = \mathbb{R}^n$ we take this to be the Euclidean Length.

Definition

A norm on a vector space V is a function $\| \ \| : V \to [0, \infty]$ such that:

i) $\|\lambda v\| = |\lambda| \|v\| \quad \forall \lambda \in \mathbb{R}, v \in V$

ii) $\|v + w\| \leq \|v\| + \|w\|$ the triangle inequality.

iii) $\|v\| = 0 \iff v = 0$

Consider for example $V = C([a,b])$. We could define

$\|f\|_\infty = \sup_{x \in [a,b]} |f(x)|$ "Uniform norm"

$\|f\|_2 = \left(\int_a^b |f(x)|^2 dx\right)^{\frac{1}{2}}$ "L2 Norm" $\quad \|f\|_p = \left(\int_a^b |f(x)|^p dx\right)^{\frac{1}{p}}$ "LP norm" $p \geq 1$

Associated with each norm is a metric. $d(V,W) = \|V - W\|$

Conditions (i) - (iii) $\Rightarrow$ d is a metric. $d_\infty(f,g) = \|f - g\|_\infty$, $d_p(f,g) = \|f-g\|_p$

From metric and topological spaces, these are equivalent. This means that the continuity of a functional depends on the choice of norm, or metric.

Before leaving this, consider the continuity of the Dirac Functional at $0 \in \mathbb{R}$. $\delta_0 : C(\mathbb{R}) \to \mathbb{R}$, $f \mapsto f(0)$.

i) Use $\|f\|_\infty = \sup_{x\in\mathbb{R}} |f(x)|$ - Choose $\delta(\varepsilon) = \varepsilon$ then

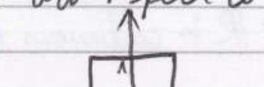
$\sup_x |g(x)| < \varepsilon \Rightarrow |\delta_0(f+g) - \delta_0(f)| = |g(0)| < \varepsilon$

δ_0 is continuous wrt $\| \ \|_\infty$

ii) Use instead $\|f\|_2 = \left(\int_{-\infty}^{\infty} |f(x)|^2 dx\right)^{\frac{1}{2}}$

δ_0 is Not continuous with respect to $\| \ \|_2$ because for

$f_n(x) = \begin{cases} n & |x| \le \frac{1}{2n^3} \\ 0 & |x| > \frac{1}{2n^3} \end{cases}$



$\|f_n\| = \left(\int_{-\infty}^{\infty} |f(x)|^2 dx\right)^{\frac{1}{2}} = \left(\frac{n^2}{2n^3}\right)^{\frac{1}{2}} \to 0$ as $n \to \infty$

$\delta_0(f_n) = f_n(0) \to \infty$

δ_0 is not continuous at $x = 0$.

To circumvent all this we will consider directional derivatives. Recall that for $f \in C^1(\mathbb{R}^n)$ its directional derivative along $\underline{v} \in \mathbb{R}^n$ is defined as

$D_{\underline{v}} f(\underline{x}) = \frac{d}{dt}\Big|_{t=0} [f(\underline{x} + t\underline{v})]$

<u>Definition</u>

A functional $I : V \to \mathbb{R}$ has directional derivative along ϕ at $u \in V$ if $\frac{d}{dt}\Big|_{t=0} I[u+t\phi]$ is defined, and we write $D_\varphi I(u) = \frac{d}{dt}\Big|_{t=0} I[u+t\varphi]$

e.g. $\delta_0 : f \mapsto f(0)$. $\frac{d}{dt}\Big|_{t=0} [\delta_0(f+t\varphi)] = \frac{d}{dt}\Big|_{t=0} [f(0) + t\varphi(0)] = \varphi(0) = \delta_0(\varphi)$

i.e. $D_\varphi \delta_0 = \delta_0(\varphi)$

11/05/11 Variational Principles ⑥

e.g. $I[f] = \int_a^b f(x)^2\, dx$

$$\frac{d}{dt}\Big|_{t=0} I[f+t\varphi] = \frac{d}{dt}\Big|_{t=0} \int_a^b (f+t\varphi)^2\, dx = 2\int_a^b f(x)\varphi(x)\, dx$$

$$D_\varphi I(f) = 2\int_a^b f(x)\varphi(x)\, dx$$

We know now what directional derivatives $D_\varphi I$ are. But also for $f \in C^1(\mathbb{R}^n)$, we have the chain rule formula

$$D_{\underline{v}} f(\underline{x}) = \nabla f(\underline{x}) \cdot \underline{v}$$

What is ∇f for a functional? What is " $\cdot$ "? The idea is to use $\langle f, g\rangle = \int f(x) g(x)\, dx$ as a replacement of the dot product. and if there exists $\frac{\delta I}{\delta f}$ such that $D_\varphi I(f) = \int \frac{\delta I}{\delta f} \varphi \, dx$, then we say $\frac{\delta I}{\delta f}$ is the "functional derivative" of I. For the Dirac Function,

$$D_\varphi(\delta_0) = \delta_0(\varphi) = \varphi(0) \neq \int (\quad)\, \varphi(x)\, dx$$

(We often see it written as $\int \delta(x)\varphi(x)\, dx$ and $\frac{\delta(\delta_0)}{\delta f} = \delta(x)$)

For $I[f] = \int_a^b f(x)^2\, dx$, we have $D_\varphi I(f) = 2\int_a^b f\varphi\, dx$, so $\frac{\delta I}{\delta f} = 2f$.

Given a functional $I: V \to \mathbb{R}$, where all directional derivatives vanish at some point $f \in V$, we say f is a <u>stationary point</u> for I.

<u>Example</u>

$$I : C^\infty_{per}([-\pi, \pi]) \to \mathbb{R}, \quad I[f] = \frac{1}{2}\int_{-\pi}^{\pi} \left(\frac{df}{dx}\right)^2 + f^2\, dx$$

Calculate $D_\varphi I(f)$ and show that it has a functional derivative.

Variational Principles (b)

e.g. $F[f] = \int_a^b f(x)^2\,dx$

$\frac{d}{d\epsilon}F[f+\epsilon\eta] = \frac{d}{d\epsilon}\Big|_{\epsilon=0}\int_a^b (f+\epsilon\eta)^2\,dx = 2\int_a^b f(x)\eta(x)\,dx$

$D_\eta F[f] = 2\int_a^b f(x)\eta(x)\,dx$

Note however, what directional derivatives $D_\eta F$ are. But also for $f \in C^1(\mathbb{R}^n)$ we have the chain rule formula

$D_v f(x) = \nabla f(x)\cdot v$

What is ∇F for a functional? What is "$\cdot$"? The idea is to use $\langle f, g\rangle = \int f(x)g(x)\,dx$ as a replacement of the dot product, and if there exists $\frac{\delta F}{\delta f}$ such that $D_\eta F[f] = \int \frac{\delta F}{\delta f}\eta\,dx$, then we say $\frac{\delta F}{\delta f}$ is the "functional derivative" of F. E.g. For the Dirac function, $D_\eta \delta_0[f] = \delta_0(\eta) = \eta(0) \neq \int (\ )\,\eta(x)\,dx$.

(We often use it anyway $\int \delta(x)\eta(x)\,dx$ and $\frac{\delta F}{\delta f} = \delta(x)$.)

For $F[f] = \int_a^b f(x)^2\,dx$, we have $D_\eta F[f] = 2\int_a^b f\eta\,dx$,

so $\frac{\delta F}{\delta f} = 2f$.

Given a functional $F: V \to \mathbb{R}$, where all directional derivatives vanish at some point $f \in V$, we say f is a stationary point for F.

Example

$F: C^\infty([-\pi,\pi]) \to \mathbb{R}$, $F[f] = \frac{1}{2}\int_{-\pi}^{\pi} (f')^2 + f^2\,dx$

Calculate $D_\eta F[f]$ and show that it has a functional derivative.

13/05/11

Variational Principles ⑦

$V = C^2_{per}([-\pi,\pi]) = \{y \in C^2(\mathbb{R}) : y(x+2\pi) = y(x)\}$

$I[y] = \int_{-\pi}^{\pi} f(x,y,y')\,dx$

where $f = f(x,y,y')$ is a C^2 function with $f(x+2\pi,y,y') = f(x,y,y')$

Recall: I has directional derivative at y in the direction $\varphi \in V$ if

$D_\varphi I(y) = \frac{d}{dt}\Big|_{t=0} I[y+t\varphi]$

Calculate: $\frac{d}{dt}\int_{-\pi}^{\pi} f(x, y+t\varphi, y'+t\varphi')\,dx = \int_{-\pi}^{\pi}\left[\frac{d(y+t\varphi)}{dt}\frac{\partial f(x,\dots)}{\partial y} + \frac{d(y'+t\varphi')}{dt}\frac{\partial f(x,\dots)}{\partial y'}\right]$

Put $t=0$. $D_\varphi I(y) = \int_{-\pi}^{\pi}\left(\varphi\frac{\partial f}{\partial y}(x,y,y') + \varphi'\frac{\partial f}{\partial y'}(x,y,y')\right)dx$

Recall: we say a function $\frac{\delta I}{\delta y}$ is functional derivative of I if

$D_\varphi I(y) = \int_{-\pi}^{\pi}\varphi(x)\frac{\delta I}{\delta y(x)}\,dx$

Observe $D_\varphi I(y) = \int_{-\pi}^{\pi}\varphi f_y\,dx + \int_{-\pi}^{\pi} f_{y'}\,d(\varphi)$

$= \int_{-\pi}^{\pi}\varphi f_y\,dx + \underbrace{\left[f_{y'}(x,y,y')\varphi(x)\right]_{-\pi}^{\pi}}_{0 \text{ by periodicity}} - \int_{-\pi}^{\pi}\varphi\frac{d}{dx}f_{y'}\,dx$

$= \int_{-\pi}^{\pi}\varphi(x)\left(f_y - \frac{d}{dx}(f_{y'})\right)dx$

<u>Conclusion</u> I has functional derivative:

$\frac{\delta I}{\delta y} = f_y - \frac{d}{dx}(f_{y'}) = f_y - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}(x,y,y')\right)$

$= f_y - \frac{\partial^2 f}{\partial x\partial y'} - \frac{dy}{dx}\frac{\partial^2 f}{\partial y\,\partial y'} - \frac{d^2y}{dx^2}\frac{\partial^2 f}{\partial y'^2}$

Example: Given $g \in C_{per}([-\pi,\pi])$ consider $I_g[y]$, $I_g : V \to \mathbb{R}$

$I_g[y] = \int_{-\pi}^{\pi}\left(\frac{1}{2}(y'^2+y^2) - gy\right)dx$

$f(x,y,y') = \frac{1}{2}(y'^2+y^2) - g(x)\,y$

$f_y = y - g(x)$, $f_{y'} = y'$ $\qquad \frac{\delta I_g}{\delta y} = y - g(x) - \frac{d}{dx}(y') = -y'' + y - g$

1

$\forall \varphi \in V$

Assume $I : V \to \mathbb{R}$ and $y \in V$ is such that $I[y] \leq I[y+\varphi]$ i.e. y is a minimiser for I. Then also $i(t) = I[y + t\varphi]$ has minimum at $t=0$, $\forall \varphi \in V$. Therefore $\frac{d}{dt}\Big|_{t=0} i(t) = 0 \ \forall \varphi$, so $D_\varphi I(y) = 0 \ \forall \varphi$. In fact, this is equivalent to $\frac{\delta I}{\delta y} = 0$

Why? $D_\varphi I(y) = \int \frac{\delta I}{\delta y} \varphi \, dx = 0 \quad \forall \varphi \in C^2_{per}([-\pi, \pi])$

Clearly $\frac{\delta I}{\delta y} = 0 \Rightarrow D_\varphi I = 0 \ \forall \varphi$.

Lemma

If $F \in C_{per}([-\pi,\pi])$ satisfies $\int_{-\pi}^{\pi} F(x)\varphi(x)\,dx = 0$ $\forall \varphi \in C^\infty_{per}([-\pi,\pi])$ then $F \equiv 0$.

Since $\frac{\delta I}{\delta y}$ is continuous, this implies the converse statement.

This means, if $y = y(x) \in V$ is a minimiser for I, it must solve the 2nd order ODE $f_y - \frac{d}{dx}(f_{y'}) = 0$ ← Euler Lagrange equation. Also, if y solves this equation it is a stationary point for I i.e. $D_\varphi I = 0 \ \forall \varphi$.

2.2 Main Example (with boundary conditions)

Consider the same functional $I[y] = \int_a^b f(x, y, y')\,dx$

Consider y a C^2 function with $y(a) = \alpha$, $y(b) = \beta$ (fixed). Also, f is still a C^2 function but no longer required to be periodic in x.

Remark - Strictly speaking, we will consider functions $y \in C^2((a,b)) \cap C([a,b])$. We will assume functions are as smooth as necessary. For this case, we consider $D_\varphi I(y)$ with φ a smooth function such that $\varphi(a) = \varphi(b) = 0$. Then $D_\varphi I(y) = \frac{d}{dt}\Big|_{t=0} I[y + t\varphi]$

$= \int_a^b \varphi f_y + \varphi' f_{y'}\,dx$

13/05/12 Variational Principles 7

$$= \int_a^b \varphi\left(f_y - \frac{d}{dx} f_{y'}\right) dx \quad + \cancel{\left[\varphi(x) f_{y'}\right]_a^b} = 0, \ \varphi(a) = \varphi(b) = 0$$

So again I has functional derivative $\frac{\delta I}{\delta y} = f_y - \frac{d}{dx}(f_{y'})$ such that $\frac{d}{dt}\Big|_{t=0} I[y + t\varphi] = D_\varphi I(y) = \int_a^b \frac{\delta I}{\delta y} \varphi \, dx \ \forall$ smooth φ vanishing at a, b.

<u>Technical Terms:</u> Support of a function $\text{supp}(f) = \overline{\{x : f(x) \neq 0\}}$

<u>Lemma 2.2.1</u>

If $F \in C([a,b])$ is such that $\int_a^b F(x)\varphi(x)\,dx = 0$ for every smooth function φ with $\text{supp}(\varphi) \subseteq [c,d] \subset (a,b)$. Then $F \equiv 0$.

<u>Proof</u> For contradiction, assume $\exists x_0$ with $F(x_0) \neq 0$. WLOG, assume $F(x_0) = \theta > 0$. Otherwise, consider $-F$. By continuity, $\exists \delta$ such that $F(x) \geq \frac{\theta}{2}$ for $x_0 - \delta \leq x \leq x_0 + \delta$.

We need a function φ: $\varphi(x) = 0$ if $|x - x_0| \geq \delta$

$\varphi(x) > 0$ if $|x - x_0| < \delta$. If I have such a function, then $\int_a^b F(x)\varphi(x)\,dx = \int_{x_0-\delta}^{x_0+\delta} F(x)\varphi(x)\,dx \geq \frac{\theta}{2}\int \varphi\,dx > 0$, a contradiction.

$b(x) = e^{-\frac{1}{(x^2-1)^2}} \quad x^2 < 1, \quad b(x) = 0, \ x^2 \geq 1$

$\lim_{x \to \pm 1, x^2 < 1} b^{(n)}(x) = 0, \quad \lim_{u \to \infty} u^N e^{-u} = 0 \quad \forall N.$

$b_\delta(x) = b\left(\frac{x}{\delta}\right)$ looks the same on $(-\delta, \delta)$. Then we define $\varphi(x) = b\left(\frac{x - x_0}{\delta}\right)$.

16/05/11

Variational Principles (8)

<u>Problems</u> Show that $y = \frac{\sin nx}{1+n^2}$ $n \in \mathbb{Z}$ is a minimising function for $I_g[y] = \int_{-\pi}^{\pi} \left(\frac{1}{2}(y'^2 + y^2) - y(x)g(x) \right) dx$, $g(x) = \sin nx$

<u>Corollary</u>:

If $I[y] \leq I[y+\varphi]$ for all C^2 with $\varphi(a) = \varphi(b) = 0$ then $y = y(x)$ must solve $\frac{\delta I}{\delta y} = 0 \Leftrightarrow F_y - \frac{d}{dx}(F_{y'}) = 0$

<u>Why</u>: $y \in C^2$ satisfies $i(t) = I[y + t\varphi]$ satisfies $i(0) \leq i(t)$ $\forall t \in \mathbb{R}$

$\Rightarrow \frac{d}{dt} i(t) \Big|_{t=0} = 0 \Rightarrow \int_a^b \frac{\delta I}{\delta y} \varphi \, dx = 0$ $\forall$ smooth φ, $\varphi(a) = \varphi(b) = 0$

$y \in C^2 \Rightarrow \frac{\delta I}{\delta y} \in C$, $\frac{\delta I}{\delta y} = 0$

<u>Problem Solution</u>

If $y \in C^2$ minimises I_g it must satisfy $\frac{\delta I_g}{\delta y} = 0 \Leftrightarrow \frac{d}{dx}(y') + y - g = 0$

$\Leftrightarrow -y'' + y = g = \sin nx$. Look for solutions of the form $a \sin nx$

$\Rightarrow a = \frac{1}{1+n^2}$, $y = \frac{\sin nx}{1+n^2}$. We now need to check that this really does minimise our functional.

$$I_g[y+\varphi] - I_g[y] = \int_{-\pi}^{\pi} \frac{1}{2}\left[(y'+\varphi')^2 + (y+\varphi)^2 - 2g(y+\varphi)\right] - \frac{1}{2}(y'^2 + y^2 - 2gy) \, dx$$

$$= \int_{-\pi}^{\pi} (y'\varphi' + y\varphi - g\varphi) \, dx + \frac{1}{2}\int_{-\pi}^{\pi} (\varphi'^2 + \varphi^2) dx$$

$$= \int_{-\pi}^{\pi} \varphi \underbrace{(-y'' + y - g)}_{=0,\ \varphi \text{ is periodic}} dx + \underbrace{\frac{1}{2}\int_{-\pi}^{\pi} (\varphi'^2 + \varphi^2) dx}_{>0,\ \varphi \not\equiv 0}$$

Also, notice that there is only one <u>periodic</u> solution of $-y'' + y = \sin nx$

But if y_1 was another, $z = y_1 - y$ solves $-z'' + z = 0$. This has no periodic solution, because if it did we could assume $\exists x_0$ with $z(x_0) = \max_x z(x)$. Assume WLOG, $z(x_0) > 0$, then we get a contradiction as $z''(x_0) = z(x_0) > 0 \Rightarrow z \equiv 0$, uniqueness.

$\frac{\delta I}{\delta y} = 0 \iff f_y - \frac{d}{dx}(f_{y'}) = 0$, the Euler-Lagrange equation.

Problem

Show $y(x) = \alpha + \frac{x-a}{b-a}(\beta - \alpha)$ minimises $I[y] = \int_a^b \sqrt{1+y'^2}\,dx$

$y(a)=\alpha, y(b) = \beta$

Solution Assume y is a C^2 function satisfying $I[y] \le I[y+\varphi]$ for all smooth φ with $\varphi(a) = \varphi(b) = 0$. As before, this means that y solves $\frac{\delta I}{\delta y} = 0$, $-\frac{d}{dx}\left(\frac{y'}{\sqrt{1+y'^2}}\right) + 0 = 0$, $\frac{y'}{\sqrt{1+y'^2}} = c$

$\Rightarrow y' = \text{constant}$. $y = A + Bx$, $y(a) = \alpha$, $y(b) = \beta$

$\Rightarrow y(x) = \alpha + \frac{x-a}{b-\alpha}(\beta - \alpha)$

To prove y really minimises, use convexity. Any stationary point of a strictly convex function is a strict global minimum. For any $\varphi \in C^2$ with $\varphi(a) = \varphi(b) = 0$, consider $i(t) = I[y+t\varphi]$

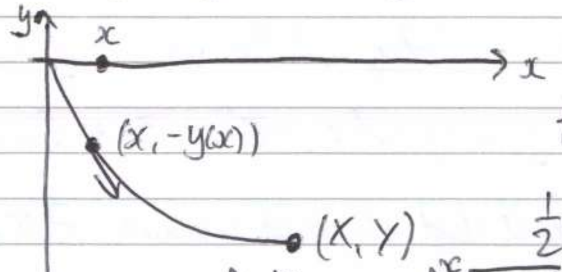
Exercise $\frac{d^2 i}{dt^2} = \int_a^b \frac{dx}{\sqrt{1+(y'+t\varphi')^2}} > 0$.

So i is a strictly convex function of $t \Rightarrow i(0) < i(t)\ \forall t \neq 0, \varphi \not\equiv 0$.

$t=0$ is a strict minimum $\Rightarrow$ a straight line minimises length.

Problem (Brachistochrone)

Find $y = y(x)$ which minimises the time of descent for a bead, moving under gravity, without friction, on a wire with shape $y = -y(x)$.



After descending to depth $|y(x)| = y(x)$ the speed v is given by $\frac{1}{2}mv^2$,

$\frac{1}{2}mv^2 = mgy$, $v = \sqrt{2gy}$

$$T = \int \frac{ds}{v} = \int_0^x \frac{\sqrt{1+y'^2}}{\sqrt{2gy}}\,dx$$

16/05/11 Variational Principles (8)

$$T = \frac{1}{\sqrt{2g}} \int \sqrt{\frac{1+y'^2}{y}}\, dx$$

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

<u>More general Euler-Lagrange equations</u>

1. Replace $y(x)$ with $\underline{X}(t) \in \mathbb{R}^N$ is a vector valued curve.

$I[X] = \int_{t_0}^{t_1} L(t, \underline{X}, \dot{\underline{X}})\, dt$, $L : \mathbb{R} \times \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ "Lagrangian"

$\frac{\delta I}{\delta x_j} = 0$, $j \in \{1, \ldots, N\}$, a system of Euler-Lagrange equations

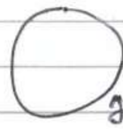
$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_j}\right) - \frac{\partial L}{\partial x_j} = 0$. e.g. $L = \frac{1}{2} m \|\dot{\underline{x}}\|^2 - V(\underline{x})$

$\frac{\partial L}{\partial \dot{x}_j} = m\dot{x}_j$, $\frac{\partial L}{\partial x_j} = -\frac{\partial V}{\partial x_j}$, $m\ddot{x}_j + \frac{\partial V}{\partial x_j} = 0$, Newton's Law

18/05/11

Variational Principles ⑨

<u>Generalisation: Multidimensional Integrals</u>

 $B = \{x \mid \|x\| < R\}$, $\partial B = \{x \mid \|x\| = R\}$

$u : \mathbb{R}^n \to \mathbb{R}$

$$I[u] = \int_B \left(\tfrac{1}{2} \|\nabla u\|^2 - \rho(x)\, u \right) d^n x$$

Assume u, φ are C^2.

Consider $I[u + \varepsilon\varphi] = i(\varepsilon)$

$$D_\varphi I(u) = \frac{di}{d\varepsilon}\Big|_{\varepsilon=0} = \frac{d}{d\varepsilon} \int_B \left(\tfrac{1}{2} \|\nabla u + \varepsilon \nabla\varphi\|^2 - \rho(x)(u + \varepsilon\varphi) \right) d^n x$$

$$= \int \left(\frac{d}{d\varepsilon} (\nabla u + \varepsilon\nabla\varphi)\cdot(\nabla u + \varepsilon\nabla\varphi)\Big|_{\varepsilon=0} - \rho\varphi \right) d^n x$$

$$= \int_B (\nabla\varphi\cdot\nabla u - \rho\varphi)\, d^n x$$

We need to write this is

$$= \int_B \frac{\delta I}{\delta u(x)} \varphi(x)\, d^n x$$

$\frac{\partial u}{\partial n} = \underline{n}\cdot\nabla u$

From Vector Calculus, recall: $\int_B \nabla u \cdot \nabla\varphi\, d\underline{x} = \int_{\partial B} \frac{\partial u}{\partial n} \varphi\, dS - \int_B (\varphi \Delta u)\, d^n x$

Now let φ vanish on ∂B, then

$$D_\varphi I(u) = \int_B (\nabla\varphi\cdot\nabla u - \rho\varphi)\, d^n x = \int_B (-\Delta u - \rho)\varphi\, d^n x$$

<u>Conclusion</u> $\frac{\delta I}{\delta u} = -\Delta u - \rho$

A generalisation of the 1D case ~~B~~ states that if $u \in C^2$ satisfies $I[u] \leq I[u+\varphi]$ for all smooth (C^1) φ with $\varphi|_{\partial B} = 0$, then $\frac{\delta I}{\delta u} = 0$ i.e. $\nabla^2 u = -\rho$, the Poisson Equation.

For a functional of the form $I[u] = \int_B f(x, u, \nabla u)\, dx$, the functional derivative is $\frac{\delta I}{\delta u} = \frac{\partial f}{\partial u} - \nabla_j \left(\frac{\partial f}{\partial(\nabla_j u)} \right)$

If $f = \tfrac{1}{2}\|\nabla u\|^2 - \rho u = \tfrac{1}{2}\sum(\nabla_j u)^2 - \rho u \Rightarrow \frac{\partial f}{\partial(\nabla_j u)} = \nabla_j u$, $\frac{\partial f}{\partial u} = -\rho$.

$$\frac{\partial f}{\partial u} - \nabla_j\left(\frac{\partial f}{\partial(\nabla_j u)}\right) = -\rho - \nabla_j(\nabla_j u)$$

$\int_A \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA = \int P\, dx + Q\, dy$

Conservation Laws

$I = \int_a^b f(x, y, y')\, dx \qquad y(a) = \alpha, \ y(b) = \beta$

Assume $y = y(x)$ is a C^2 solution of the Euler-Lagrange equation.

$f_y - \frac{d}{dx}(f_{y'}) = 0$

i) If f is independent of y, $\frac{d}{dx}(f_{y'}) = 0$, $f_{y'} = c$, constant.

In the Vector-Valued case, $L = L(t, \underline{x}, \underline{\dot{x}})$, this gives

$\frac{\partial L}{\partial \dot{x}_j}$ = constant if L is independent of x_j. $L = \frac{1}{2} m \|\underline{\dot{x}}\|^2 - V(\underline{x})$

V independent of x_j then $p_j = m \frac{\partial L}{\partial \dot{x}_j} = m\dot{x}_j$ = constant.

ii) If f is independent of x, then $y' f_{y'} - f$ = constant

To prove this, take $\frac{d}{dx}(y' f_{y'} - f) = y'' f_{y'} + y' \frac{df_{y'}}{dx} - \frac{df}{dx}$

$= y'' f_{y'} + y' f_y - (y' f_y + y'' f_{y'}) = 0$

In the vector valued case, $L = \frac{1}{2} m \|\underline{\dot{x}}\|^2 - V(\underline{x})$. The corresponding conservation law is $\dot{x}_j \frac{\partial L}{\partial \dot{x}_j} - L$ = constant $= \underline{\dot{x}} \cdot \frac{\partial L}{\partial \underline{\dot{x}}} - L$

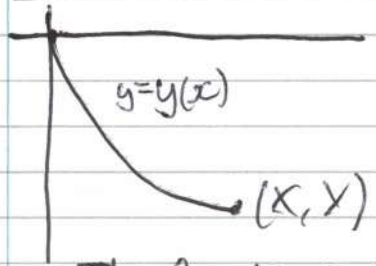
<u>Exercise</u> $\frac{\partial L}{\partial \dot{x}_j} = m\dot{x}_j = p_j$, so

$\sum \dot{x}_j \frac{\partial L}{\partial \dot{x}_j} - L = \sum_j \dot{x}_j m(\dot{x}_j) - \left(\frac{1}{2} m \sum \dot{x}_j^2 - V(\underline{x})\right) = \underbrace{\frac{1}{2} m \|\underline{\dot{x}}\|^2 + V(\underline{x})}_{= \text{energy}}$

<u>Moral</u> If we have symmetry (e.g. f independent of x, symmetric under translation $x \to x + a$) $\Rightarrow$ conservation law, "Noether Principle".

18/05/11 Variational Principles (9)

<u>Brachistochrone Problem</u>



$$T = \int_0^X \frac{\sqrt{1+y'^2}}{\sqrt{y}}\,dx \qquad y(X) = Y.$$

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

The function in the integrand is independent of X. Then if $y \in C^2$ solves our equation, then $y'F_{y'} - F = \dfrac{y'^2}{\sqrt{y(1+y'^2)}} - \dfrac{\sqrt{1+y'^2}}{\sqrt{y}} = C$

Exercise: Solve for y'.

$$y' = \frac{(1-c^2y)^{\frac{1}{2}}}{cy^{\frac{1}{2}}} \qquad \int dx = \int \frac{c\sqrt{y}}{\sqrt{1-c^2y}}\,dy$$

$$y = \frac{1}{c^2}\sin^2\frac{\theta}{2}, \quad dy = \frac{1}{c^2}\sin\frac{\theta}{2}\cos\frac{\theta}{2}\,d\theta$$

$$\int \frac{\sin\frac{\theta}{2} \times \frac{1}{c^2}\sin\frac{\theta}{2}\cancel{\cos\frac{\theta}{2}}}{\cancel{\cos\frac{\theta}{2}}}\,d\theta = \frac{1}{c^2}\int \sin^2\frac{\theta}{2}\,d\theta = \frac{1}{2c^2}\int 1-\cos\theta\,d\theta$$

$$= \frac{1}{2c^2}(\theta - \sin\theta) = x \qquad (\theta = 0 \text{ when } x = 0)$$

$$y = \frac{1}{c^2}\sin^2\frac{\theta}{2} = \frac{1}{2c^2}(1-\cos\theta)$$

This gives the curve in parametrised form, a cycloid.

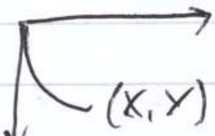
So far, <u>if</u> there is a solution given ~~by~~ $y(x)$, $y \in C^2$, it must be given by a cycloid. We still need to show it minimises the time of descent, and that there exists a unique cycloid joining $(0,0)$ to (X,Y).

20/05/11

Variational Principles (10)

Brachistochrone

$$I = \int_0^X \frac{\sqrt{1+y'^2}}{\sqrt{y}}\,dx$$



$$x = \frac{1}{2c^2}(\theta - \sin\theta) \qquad y = \frac{1}{2c^2}(1 - \cos\theta)$$

i) Given (X, Y), we must be able to find a unique c, Θ such that

$$X = \frac{1}{2c^2}(\Theta - \sin\Theta), \quad X = \frac{1}{2c^2}(1 - \sin\Theta)$$

Restrict $0 < \Theta < 2\pi$

q/12, Sheet II : $\exists$ a unique cusp free cycloid arc joining $(0,0)$ to $(X, Y)_{X<0}$

ii) q/18, Sheet I : $y(x) = [\phi(x)]^2$

$$J[\phi] = I[\phi^2] = \int_0^X \left(\phi^{-2} + 4(\phi')^2\right)^{\frac{1}{2}} dx$$

Exercise : The function $L(u,v) = (u^{-2} + 4v^2)^{\frac{1}{2}}$, defined on $\{(u,v) \in \mathbb{R}^2 : u > 0\}$ is strictly convex. From this we can show that the arc of a cycloid really minimises I among C^2 functions $y = y(x)$.

iii) In calculus of variations, some minimising functions for some functions aren't smooth.

q/6, Sheet II, $I[y] = \int_{-1}^{1} (1 - y_x^{\,2})^2\,dx$, $y(\pm 1) = 1$

minimises, $\min I[y] = 0$ amongst piecewise C^1 functions.

Note : This could <u>not</u> happen for the Brachistochrone problem.

Also note that we used the <u>indirect</u> method. <u>If</u> the solution exists, it solves the Euler - Lagrange equation. Then solve this equation, and then show at the end that your solution minimises our functional.

<u>Direct Approach</u> $\inf I[y] = m$. Does $\exists y$ such that $I[y] = m$?

By definition of $\inf I[y]$, $\exists$ a sequence of functions y_n such that $I[y_n] \to m$.

This requires understanding compactness properties of sets in spaces of functions, this is very complex.

Variational Problems with Constraints

1. (Dido) Find $y = y(x)$ with $y(\pm a) = 0$ such that $I[y] = \int_{-a}^{a} y\, dx$ for given $\int_{-a}^{a} \sqrt{1+y'^2}\, dx = L$, fixed.
As in finite-dimensional problems we introduce a multiplier λ for each constraint, so here $J[y] = \int_{-a}^{a} \sqrt{1+y'^2}\, dx - L = 0$

$$\Phi[y, \lambda] = I[y] + \lambda J[y]$$

If y minimises, and is C^2, then $\frac{\delta \Phi}{\delta y} = 0 \Leftrightarrow \frac{\delta I}{\delta y} + \lambda \frac{\delta J}{\delta y} = 0$

$$\Phi = \int_{-a}^{a} y + \lambda\sqrt{1+y'^2}\, dx$$

$$\frac{\delta \Phi}{\delta y} = 0 \Leftrightarrow f_y - \frac{d}{dx}(f_{y'}) = 0 \Leftrightarrow 1 - \lambda \frac{d}{dx}\left(\frac{y'}{\sqrt{1+y'^2}}\right) = 0$$

$$\frac{y'}{\sqrt{1+y'^2}} = \frac{x}{\lambda} + C$$

$$\int dy = \pm \int \frac{\frac{1}{\lambda}(x+c)}{\left(1 - \left(\frac{x+c}{\lambda}\right)^2\right)^{\frac{1}{2}}}\, dx \qquad \frac{x+c}{\lambda} = \sin\theta$$

$$\Rightarrow y = y_0 \pm \lambda \cos\theta, \quad x + C = \lambda \sin\theta$$

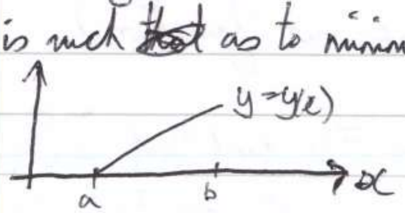
$(y - y_0)^2 + (x+c)^2 = \lambda^2$. Find (c, y_0) and λ to match $y(\pm a) = 0$ and $J[y] = 0$.

If there are N constraints $J_\alpha = 0$ consider $\Phi = I[y] + \sum_{\alpha=1}^{N} \lambda_\alpha J_\alpha$

But sometimes we have contraints which are themselves functions.
In this case, we need a Lagrange multiplier function.

Applications of Euler-Lagrange Equations

1. Fermat's Principle of Least Time :
Light moves in medium with speed c (non constant). It's path is such ~~that~~ as to minimise the time $T = \int \frac{ds}{c}$



$c = c(x)$, $y > 0$ $\qquad T = \int \frac{\sqrt{1+y'^2}}{c(x)}\, dx$

The path $y(x)$ must satisfy $f_y - \frac{d}{dx}(f_{y'}) = 0$. Recall from last time that since f is independent of y, the 1st conservation law gives $f_{y'} = \text{constant}$
i.e.

$$\frac{y'}{c(x)\sqrt{1+y'^2}} = A \qquad y(a) = 0, \quad y(b) = \beta$$

20/05/11

Variational Principles (10)

e.g. $[a,b]=[1,2]$, $c(x)=x$

$\frac{y'}{\sqrt{1+y'^2}}=Ax$, $y'=\frac{Ax}{(1-A^2x^2)^{\frac{1}{2}}} \Rightarrow \int_0^y y'\,dx = y(x) = \int_1^x \frac{Ax}{\sqrt{1-A^2x^2}}\,dx$

Choose A to match $y(2)=\beta$. If $c=c(y)$, use the 2^{nd} conservation law.

2. <u>Lagrangian - Mechanics</u> $L(t,\underline{X},\dot{\underline{X}}) = \frac{1}{2}m\|\dot{\underline{X}}\|^2 - V(\underline{X})$

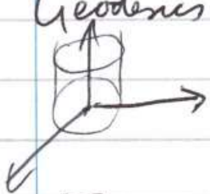
$S[\underline{X}] = \int L(t,\underline{X},\dot{\underline{X}})\,dt \qquad \frac{\delta S}{\delta X_\alpha}=0 \Leftrightarrow m\ddot{X}_\alpha + \frac{\partial V}{\partial X_\alpha} = 0$

3 <u>Geodesics</u>

In general, a geodesic is a length minimising curve

In the plane: $y=y(x)$, geodesics minimise $\int\sqrt{1+y'^2}\,dx$

Geodesics on the Euclidean Plane are straight lines. What about a cylinder?



$x_1^2+x_2^2=1$

x_3 arbitrary

$\|\dot{\underline{X}}\|^2 = \sum_{j=1}^3 \dot{x}_j^2$

The length of a curve in parametrised form is $I[\underline{X}] = \int \|\dot{\underline{X}}\|\,dt$

Constraint: $x_1(t)^2+x_2(t)^2=1 \quad \forall t$

Introduce multiplier function $\lambda=\lambda(t)$

$\Phi[\underline{X},\lambda] = \int \|\dot{\underline{X}}\| + \lambda(x_1^2+x_2^2-1)\,dt$

$\frac{\delta\Phi}{\delta x_j}=0$

23/05/11

Variational Principles (11)

$$\frac{\delta \Phi}{\delta x_i} = \frac{\partial \Phi}{\partial x_i} - \frac{d}{dt}\left(\frac{\partial \Phi}{\partial \dot{x}_i}\right) = \begin{pmatrix} 2\lambda x \\ 2\lambda y \\ 0 \end{pmatrix} - \frac{1}{\|\dot{x}\|}\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} = 0$$ For a geodesic curve

$\Rightarrow$ These would be easy without $\|\dot{x}\|$, but we can choose a parametisation proportional to arc length to make this constant.

Explicitly $\underline{x} = \underline{x}(t)$. Consider instead $\tau = \tau(t)$, $\tau \in C^1$, $\frac{d\tau}{dt} > 0$

then as a function of τ we have $\frac{d\underline{x}}{d\tau} = \frac{d\underline{x}}{dt}\frac{dt}{d\tau} = \frac{1}{\dot{\tau}}\frac{d\underline{x}}{dt}$

Then, choose $\dot{\tau}$ to be $\|\frac{d\underline{x}}{dt}(t)\|$, then $\frac{\dot{x}}{\|\dot{x}\|} = \frac{dx}{d\tau}$ etc

So the Euler - Lagrange equation becomes $2\lambda x - \frac{d\tau}{dt}\frac{d}{d\tau}\left(\frac{dx}{d\tau}\right) = 0$

and similarly for y. $\quad -\frac{d\tau}{dt}\frac{d}{d\tau}\left(\frac{dz}{d\tau}\right) = 0$.

Define $\mu(\tau) = \frac{\lambda}{\dot{\tau}}$ and we have:

$$\begin{pmatrix} -\frac{d^2x}{d\tau^2} + 2\mu x \\ -\frac{d^2y}{d\tau^2} + 2\mu y \\ -\frac{d^2z}{d\tau^2} \end{pmatrix} = \underline{0}$$

To deal with the Lagrange multiplier μ, we use the constraint $x^2+y^2=1 \ \forall \tau$.

$$-x\frac{d^2x}{d\tau^2} - y\frac{d^2y}{d\tau^2} + 2\mu(\underbrace{x^2+y^2}_{=1}) = 0$$

$$2\mu = x\frac{d^2x}{d\tau^2} + y\frac{d^2y}{d\tau^2}$$

But $x^2+y^2 = 1 \quad \Rightarrow x\frac{dx}{d\tau} + y\frac{dy}{d\tau} = 0$

$$\Rightarrow x\frac{d^2x}{d\tau^2} + y\frac{d^2y}{d\tau^2} + \left(\frac{dx}{d\tau}\right)^2 + \left(\frac{dy}{d\tau}\right)^2 = 0$$

So $2\mu = -\left(\left(\frac{dx}{d\tau}\right)^2 + \left(\frac{dy}{d\tau}\right)^2\right) = \left(\frac{dz}{d\tau}\right)^2 - \left\|\frac{d\underline{x}}{d\tau}\right\|^2 = \text{constant} = -\omega^2$

Note: $\left\|\frac{d\underline{x}}{d\tau}\right\|^2 = \frac{1}{\dot{\tau}^2}\left\|\frac{d\underline{x}}{dt}\right\|^2 = \frac{\|\frac{d\underline{x}}{dt}\|^2}{\|\frac{d\underline{x}}{dt}\|^2} = 1$

$\frac{d^2z}{d\tau^2} = 0 \Rightarrow \frac{dz}{d\tau} = \text{constant}$

So $\frac{d^2x}{d\tau^2} + \omega^2 x = 0$, $\frac{d^2y}{d\tau^2} + \omega^2 y = 0$, $\frac{d^2z}{d\tau^2} = 0$

$x(\tau) = \cos(\omega\tau + \alpha)$, $y(\tau) = -\sin(\omega t + \alpha)$, $z(t) = A + B\tau$

So geodesics on C with arc length parametisation are the superposition of uniform motion along the axis of C with transverse circular motion.

Moral - Always easiest to work with arc length parametisation.

In fact, geodesics with arc-length parametisation are obtained from the functional
$$\tilde{I}[\underline{x}] = \int_{t_0}^{t_1} \|\dot{\underline{x}}\|^2 dt$$
The Euler-Lagrange equations for $\tilde{I}$ give the geodesic in arc-length parametisation.

The Second Variation

If $f \in C^2(\mathbb{R}^n)$, $f(\underline{x}+\underline{y}) = f(\underline{x}) + \underline{\nabla} f(\underline{x}) \cdot \underline{y} + \frac{1}{2}\frac{\partial^2 f(\underline{x})}{\partial x_i \partial x_j} y_i y_j + o(\|\underline{y}\|^2)$
We will try to generalise this for a functional.

Consider $I_g : C^2_{per}([-\pi,\pi]) \to \mathbb{R}$, $I_g[y] = \int_{-\pi}^{\pi} \left(\frac{1}{2}(y'^2+y^2) - g(x)y\right)dx$
Let's consider a 1-parameter family of variations $y \to y + \varepsilon\varphi$.

$$I_g[y+\varepsilon\varphi] = \int_{-\pi}^{\pi} \left[\frac{1}{2}(y'+\varepsilon\varphi')^2 + \frac{1}{2}(y+\varepsilon\varphi)^2 - g(y+\varepsilon\varphi)\right]dx$$
$$= I[y] + \underbrace{\varepsilon\int_{-\pi}^{\pi}(y'\varphi' + y\varphi - g\varphi)\,dx}_{\text{first variation } \delta I_g} + \underbrace{\varepsilon^2\int_{-\pi}^{\pi}\frac{1}{2}(\varphi'^2+\varphi^2)\,dx}_{\text{second variation}}$$

So far we have studied δI_g, using
$$\delta I_g = \int_{-\pi}^{\pi}(y'\varphi' + y\varphi - g\varphi)dx = D_\varphi I_g(y) = \int_{-\pi}^{\pi}\frac{\delta I_g}{\delta y}\varphi\, dx$$

where $\frac{\delta I_g}{\delta y} = -y'' + y - g$.

$\frac{\delta I_g}{\delta y} = 0 \iff y$ solves the Euler-Lagrange equation $\iff y$ is stationary
i.e. $\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} I_g[y+\varepsilon\varphi] = 0 \;\forall \varphi \in C^2_{per}([-\pi,\pi])$.

$\delta^2 I_g$ is analogous to $\frac{1}{2}\frac{\partial^2 f}{\partial x_i \partial x_j}$ for stationary functions $f \in C^2(\mathbb{R}^n)$

<u>Theorem 5.1</u> If $D_\varphi I_g[y] = 0$ and $\delta^2 I \geq c\int_{-\pi}^{\pi}(\varphi'^2+\varphi^2)dx$ for some $c>0$, then y is a weak local minimum, i.e. $I[y] < I[y+\varphi]$ for all φ, smooth, with $\max(|\varphi|+|\varphi'|)$ sufficiently small, $\varphi \equiv 0$ not allowed.

25/05/17

Variational Principles (12)

2nd Variation $\delta^2 I = \frac{1}{2} D^2 I[y](\varphi) = \frac{1}{2}\int_{-\pi}^{\pi} (\varphi'^2 + \varphi^2)\, dx$

Definition $I[y] = \int_a^b f(x, y, y')\, dx \quad y(a) = \alpha,\ y(b) = \beta$
has a weak local minimum at y if $I[y+\varphi] \geq I[y]$
for C^1 φ with $\max_{[a,b]}(|\varphi| + |\varphi'|)$ sufficiently small, and $\varphi(a) = 0 = \varphi(b)$

This generalises to the periodic case in the obvious way.

Expression for $\delta^2 I$

$$I[y+\varepsilon\varphi] = \int f(x, y+\varepsilon\varphi, y'+\varepsilon\varphi')\, dx$$
$$= \int_a^b f(x, y, y')\, dx + \varepsilon\int_a^b \varphi f_y(x,y,y') + \varphi' f_{y'}(x,y,y')\, dx$$
$$+ \frac{\varepsilon^2}{2}\int_a^b \varphi^2 f_{yy} + 2\varphi\varphi' f_{yy'} + \varphi'^2 f_{y'y'}\, dx + o(\varepsilon^2)$$

$$\delta^2 I = \int_a^b \varphi^2 f_{yy} + \frac{d}{dx}(\varphi^2) f_{yy'} + \varphi'^2 f_{y'y'}\, dx$$
$$= \int_a^b \varphi^2 \left(f_{yy} - \frac{d}{dx}(f_{yy'})\right) + (\varphi')^2 f_{y'y'}\, dx$$
$$= \int_a^b P(x)\varphi'^2 + Q(x)\varphi^2\, dx \quad \leftarrow \text{Sturm - Liouville Functional}$$

a) **Theorem 5.1** Let $y \in C^2$ be a weak local minimum for I. Then
$$\frac{\delta I}{\delta y} = f_y - \frac{d}{dx}(f_{y'}) = 0$$
$$\delta^2 I = \int_a^b (P\varphi'^2 + Q\varphi^2)\, dx \geq 0 \quad \forall C^1 \varphi \text{ with } \varphi(a) = \varphi(b) = 0$$

b) If $y \in C^2$ satisfies $\frac{\delta I}{\delta y} = 0$, and $\delta^2 I = \int_a^b (P\varphi'^2 + Q\varphi^2)\, dx$,
$\delta^2 I \geq c\int_a^b (\varphi'^2 + \varphi^2)\, dx$ for some $c > 0$, and all $C^1 \varphi$ with
$\varphi(a) = \varphi(b) = 0$, then y is a weak local minimum.

$g(x) = \sin nx$

Example #1 $I_g[y] = \int_{-\pi}^{\pi} \frac{1}{2}(y'^2 + y^2) - g(x)\, y \, dx$

The Euler-Lagrange equation is $-y'' + y = g = \sin nx$ $(n \in \mathbb{Z})$

Solution $y = \frac{1}{1+n^2} \sin nx \in C^\infty_{per}([-\pi, \pi])$. To apply Theorem 5·1, we look at the second variation.

$\delta^2 I = \frac{1}{2}\int_{-\pi}^{\pi} (\varphi'^2 + \varphi^2)\, dx$, choose $c = \frac{1}{2}$, so y is a weak local min.

Example #2 $I[y] = \int_0^1 \frac{1}{2}y'^2 + \frac{\lambda}{4}y^4 \, dx$ $\quad y(0) = y(1) = 0$

Find a range of λ for which $y(x) \equiv 0$ is a weak local min.

Solution The Euler-Lagrange equation is $-y'' + \lambda y^3 = 0$ has solution $y(x) \equiv 0$.

$\delta^2 I = \frac{1}{2}\int_0^1 \varphi'^2 \, dx$

Does $\exists c > 0$ such that $\int_0^1 \varphi'^2 \, dx \geq c \int_0^1 (\varphi'^2 + \varphi^2)\, dx$ for C^1 functions with $\varphi(a) = \varphi(b) = 0$.

Method 1 $\varphi(x) = \int_0^x \varphi'(t)\, dt$ since $\varphi(0) = 0$

Therefore $\forall x$, $|\varphi(x)| \leq \int_0^x |\varphi'(t)|\, dt \leq \int_0^1 |\varphi'(t)|\, dt \leq \left(\int_0^1 |\varphi'(t)|^2 dt\right)^{\frac{1}{2}}$

(Holder's inequality : $\int_a^b f(t) g(t)\, dt \leq \left(\int_a^b f^2 dt\right)^{\frac{1}{2}} \left(\int_a^b g^2 dt\right)^{\frac{1}{2}}$

So $|\varphi(x)|^2 \leq \int_0^1 \varphi'(t)^2 \, dt \quad \forall x \in [0, 1]$

Therefore $\int_0^1 \varphi(x)^2 \, dx \leq \int_0^1 \varphi'(t)^2 \, dt$ and so

$\int (\varphi(x)^2 + \varphi'(x)^2)\, dx \leq 2\int_0^1 \varphi'(x)^2 \, dx$

So we can choose $c = \frac{1}{2}$, so by Theorem 5·1, $y(x) \equiv 0$ is a weak local min for any value of λ.

In general, to decide if $\int_a^b P\varphi'^2 + Q\varphi^2 \, dx \geq c \int_0^1 \varphi'^2 + \varphi^2 \, dx$ we need to consider eigenvalues / eigenfunctions of the Sturm-Liouville problem: $-\frac{d}{dx}(P\varphi') + Q\varphi = \lambda\varphi$, where λ is the eigenvalue and φ is the eigenfunction. $L\varphi = \lambda\varphi$. We need to check that all the eigenvalues are positive.

25/05/11

Variational Principles

Example #3

For which values of μ are the eigenvalues of $\int_{\pi}^{\pi} \varphi'^2 + \mu \varphi^2 \, dx$ all positive? $\varphi \in C^\infty_{per}([-\pi, \pi])$

We have to solve $-\varphi'' + \mu\varphi = \lambda\varphi$, $\varphi(\pi) = \varphi(-\pi)$

$\varphi'' = (\mu - \lambda)\varphi$. For $\mu - \lambda = -\omega^2 < 0$, we have solutions $e^{\pm i\omega x}$. These are 2π periodic only if $e^{i\omega 2\pi} = 1$ i.e if $\omega = n \in \mathbb{Z}$.
i.e. we only get a periodic solution for $\lambda = \mu + n^2$.
For $\mu - \lambda > 0$, we have no periodic solutions.
For $\mu = \lambda$, φ = constant is the only periodic solution.
So the eigenvalues are $(\mu + n^2)_{n \geq 0}$, so for $\mu > 0$, the functional $\int_{\pi}^{\pi} (\varphi'^2 + \varphi^2) dx$ is strictly positive.

A modification of this is $\int_0^1 \varphi'^2 + \mu\varphi^2 \, dx$ with $\varphi(0) = 0 = \varphi(1)$.
$-\varphi'' + \mu\varphi = \lambda\varphi$
For $\lambda = \mu + (n\pi)^2$ there is a solution $\sin n\pi x$, so that the smallest eigenvalue is $\mu + \pi^2$. e.g. $\int_0^1 \varphi'^2 - \frac{\pi^2}{2}\varphi^2 \, dx$ is strictly positive.