CALCULUS OF VARIATIONS

1. INTRODUCTION

Functional: Mapping of some function space into real line.

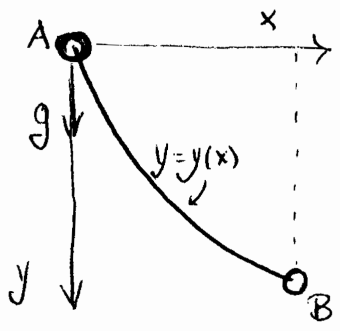
eg: $$S[y] = \int_{x_0}^{x_1} \sqrt{1+[y'(x)]^2}\, dx \quad \left(\text{arclength}\right) \tag{1}$$

$$T(f) = \max_{0 \le x \le 1} f(x) \tag{2}$$

"**Find** function for which functional is at extremum (perhaps subject to additional constraints)".

Classic problems

I Brachistochrone (John Bernoulli, 1696)



Travel time

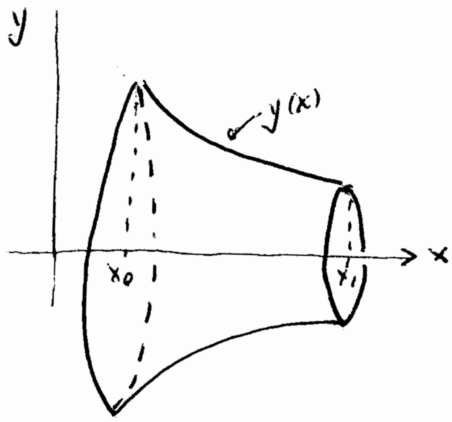
$$T = \int_A^B \frac{ds}{v}, \quad \left(\text{speed } v = \frac{ds}{dt}\right)$$

$$= \int_A^B \frac{\sqrt{1+[y'(x)]^2}}{v}\, dx \tag{3}$$

KE $\frac{1}{2}mv^2 = mgy \Rightarrow v(x) = \sqrt{2g\,y(x)}$

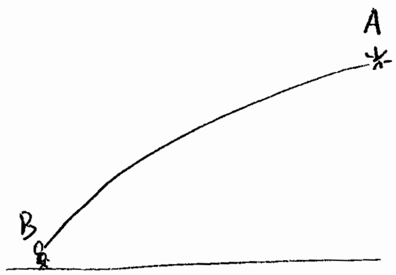
$$\sqrt{2g}\, T[y] = \int_A^B \frac{\sqrt{1+[y'(x)]^2}}{\sqrt{y(x)}}\, dx$$; "Minimize T wrt y"

II Plateau's problem (minimal surface of revolution)



$$\min \int_{x_0}^{x_1} y \sqrt{1+[y'(x)]^2}\, dx \qquad (4)$$

III. Fermat's principle (2-dim, inhomogeneous refractive index $c(x,y)$)



$$dt = \frac{1}{c}\, ds$$

$$\min t = \int \frac{ds}{c} = \int_A^B \frac{\sqrt{1+(y'(x))^2}}{c(x,y)}\, dx \qquad (5)$$

(principle of least time)

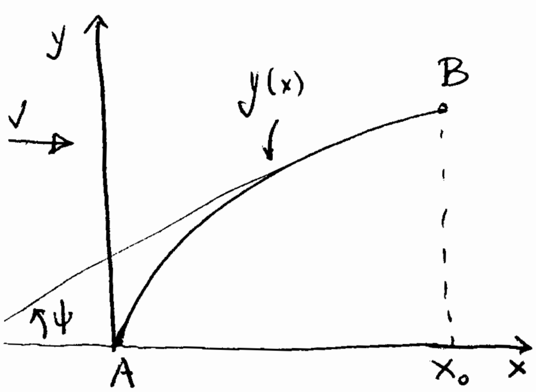
IV Newton 1686 (minimize drag on body of revolution)

Postulated: $D = 2\pi\rho V^2 \int y \sin^2\psi\, dy$

ρ: air density

V: speed

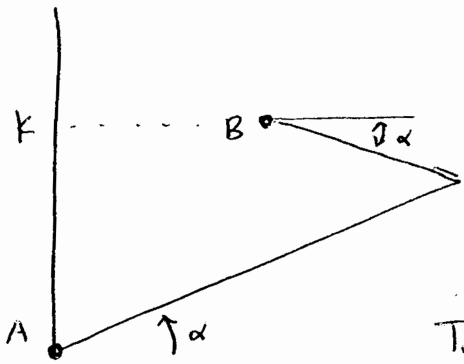
$\tan\psi = y'$ (slope)



$$\sin^2\psi\, dy = \sin^2\psi \frac{dy}{dx} dx = \sin^2\psi \tan\psi\, dx$$
$$= \tan^3\psi \cos^2\psi\, dx = \frac{\tan^3\psi}{\sec^2\psi} dx$$
$$= \frac{\tan^3\psi}{1+\tan^2\psi} dx = \frac{y'^3}{1+y'^2} dx$$

i.e.

$$D = 2\pi\rho V^2 \int_0^{x_0} \frac{y y'^3}{1+y'^2}\, dx \qquad (6)$$



Can make integral as small as we want:

$$\int_0^k y\sin^2\psi\,dx = \sin^2\alpha\int_0^k y\,dy = \frac{1}{2}k^2\sin^2\alpha$$

By making α smaller, $D \to 0$

To avoid this, add restriction $0 \leq \psi \leq \frac{\pi}{2}$.

V <u>Isoperimetric problem</u>: $\max \int y(x)\,dx$

subject to $\int \sqrt{1+[y']^2}\,dx = \text{constant}$ $\left(\int ds = \ell, \text{ length}\right)$ (7)

(Arc of fixed length $\Rightarrow$ max area for circle
Surf. of fixed area $\Rightarrow$ max volume for sphere)

VI: <u>Column</u> of fixed volume: maximize load strength.

VII <u>Geodesics</u>: shortest distance between two points on surface

VIII <u>Hamilton's Principle</u> of least action

Generally:

extremize $\int F(x,y,y')\,dx$ (8a)

subject to $\int G(x,y,y')\,dx = \text{constant}$ (8b)

Conservation laws and Differential equations

Direct $\longleftrightarrow$ Indirect

DE Extrema

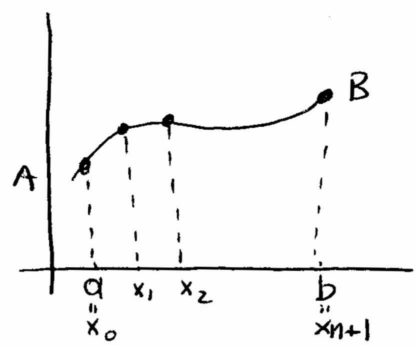
2. Numerical Methods

$$I[y] = \int_a^b F(x, y(x), y'(x))\,dx\,; \quad y(a) = A,\ y(b) = B \qquad (1)$$

Classical approaches

—— ∘ ——

1. Finite differences



$$x_{i+1} - x_i = h$$

Replace curve* $y(x)$ by polygonal approximation with vertices $(x_0, A), (x_1, y_1), \dots, (x_n, y_n = y(x_n)), (x_{n+1}, B)$ (where $y_i := y(x_i)$), with $x_0 = a$ and $x_{n+1} = b$.

Approximate $I[y]$ by

$$I_n := I(y_1, \dots, y_n) = \sum_{i=1}^{n+1} F\left(x_i, y_i, \frac{y_i - y_{i-1}}{h}\right) h$$

$$\therefore \frac{\partial I_n}{\partial y_i} = 0$$

(Sequence $\{y_i\}_1^n$ is called a "minimizing sequence")

Questions: (i) Is $\lim_{n \to \infty} I_n = \min_y I[y]$

(ii) Does polygonal line converge to the minimizer?

—— ∘ ——

2. General minimizing sequences: the Rayleigh-Ritz method.

Assume $I[y]$ possesses a glb; that is,

* the length of a curve is defined as the limit of the sequence of lengths of inscribed polygonal lines. If limit exists, curve is called rectifiable.

assume $\inf_{y \in \mathcal{A}} I[y] = d > -\infty$ for $y \in \mathcal{A}$, some class.

Then $\exists$ sequences $\phi_1, \phi_2, \ldots$ of admissible function such that $\lim_{n\to\infty} I[\phi_n] = d$ while $I[\phi_i] \geq d$ for all $\Phi_i \in \mathcal{A}$.

$\{\Phi_n\}$ are called minimizing sequences.

Questions (i) does $\lim_{n\to\infty} \phi_n$ exist?

(ii) is $\lim_{n\to\infty} I[\phi_n] = \min I[y]$?

$\{$Ritz: (to minimize $I[y]$)$\}$ Choose a complete set of admissible functions $u_1, u_2, u_3, \ldots$ Define a sequence $\{y_n\}$ ($\longrightarrow$ minimizing sequence) by

$$y_n = \sum_{i=1}^{n} a_i u_i$$

Consider $I[y_n]$; it is now a function of the $\{\alpha_i\}_1^n$;
Minimize $I[y]$ as a function of the n variables α_i, $i=1,\ldots,n$
that is $\partial I/\partial \alpha_i = 0$. Clearly $\min I[y] \leq I[y_n]$

Ex $\min I[y] = \int_0^1 (x^3 y''^2 + 100 x y^2 - 20 xy)\, dx$; $y(1) = y'(1) = 0$

Complete set: $\{x^r\}$ (Weierstrass polynomial approximation theorem: polynomials are complete)

$$y_n(x) = (x-1)^2 (a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n)$$

Take $y_1(x) = (x-1)^2 (a_0 + a_1 x)$

$y_1''(x) = 6a_1 x + 2a_0 - 4a_1$

Then $\int_0^1 [x^3(6a_1x+2a_0-4a_1)^2+10x(x-1)^4(a_0+a_1x)^2-2x(x-1)^2(\alpha_0+a_1x)]\,dx$

$$\therefore \quad I = I(a_0, a_1)$$

$$\partial I/\partial a_0 = \partial I/\partial a_1 = 0 \quad : \quad a_0 = .124, \; a_1 = .218$$

Minimizer : $I_1 = I[y_1]$: $\quad y_1 = (x-1)^2(.124+.218x)$

—·—

3. Difficulties:

(i) a minimizing sequence need not converge

(ii) even if a ms converges, its limit may be inadmissible

(iii) It is possible to have a seq. of admissible functions converging to the lub or glb of all possible values and yet have the value of the integral for the limit fn. itself be different from the lub or glb

Example: (i) $I = \iint_D (\phi_x^2+\phi_y^2)\,dx\,dy = \iint_D \left(\phi_r^2 + \frac{1}{r^2}\phi_\theta^2\right) r\,dr\,d\theta$

Set $\phi_n = \begin{cases} c_n \operatorname{Log} a_n, & 0 \le r \le a_n^2 \\ c_n \operatorname{Log}(r/a_n) & a_n^2 \le r \le a_n \\ 0 & a_n \le r \le 1 \end{cases}$

where the c_n, a_n are constant

Then $\quad I[\phi_n] = 2\pi c_n^2 \int_{a_n^2}^{a_n} \frac{1}{r^2}\, r\,dr = -2\pi c_n^2 \operatorname{Log} a_n$

Now choose $a_n = e^{-n}$ and $c_n = -n^{-2/3}$. Then

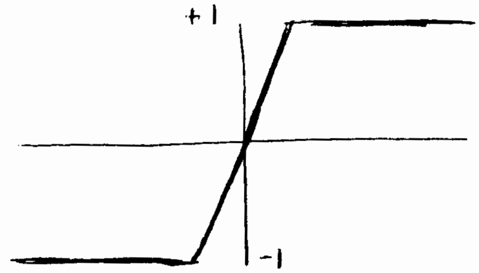
$$I[\phi_n] = 2\pi/n^{1/3} \longrightarrow 0 \quad \text{as} \quad n \to \infty$$

But $\phi_n(0) = n^{1/3} \to \infty$ as $n \to \infty$.

Example (ii) $\min \; I = \int_{-1}^{1} x^2 [u'(x)]^2 \, dx \; ; \; u(-1) = -1, \; u(1) = 1$

$I \geq 0$. Set

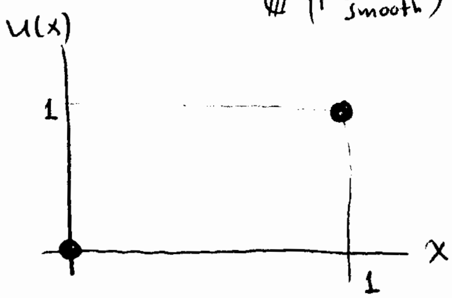
$$u_n(x) = \begin{cases} -1, & -1 \leq x \leq -\frac{1}{n} \\ nx, & -\frac{1}{n} \leq x \leq \frac{1}{n} \\ 1 & \frac{1}{n} \leq x \leq 1 \end{cases}$$



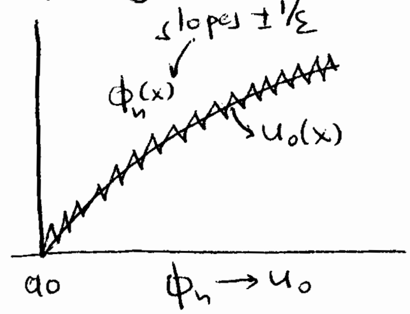
Then $I(u_n) = \int_{-1/n}^{1/n} n^2 x^2 \, dx = \frac{2}{3n} \to 0$ as $n \to \infty$

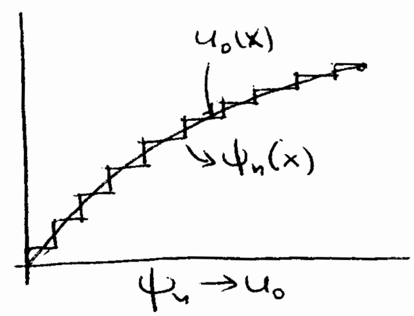
But $u_n \to u = \begin{cases} +1, & x > 0 \\ -1, & x < 0 \end{cases}$ i.e. not an admissible function (i.e. not <u>continuous</u>, piecewise smooth)

Example (iii) (piecewise smooth) $\min/\max \; I = \int_0^1 \frac{1}{1 + [u'(x)]^2} \, dx \; ; \; u(0) = 0, \; u(1) = 1$



Clearly $\text{glb}\, I = 0$, $\text{lub}\, I = 1$





$$I(\phi_n) \to 0 \neq I(u_0) \neq 1 \leftarrow I(\psi_n)$$

3. Basic Analysis definitions

1. Theorem (Weierstrass)

Every continuous function defined on a closed, bounded domain D of n-dimensional space possesses a min and max in D.

◀ f continuous in $D \Rightarrow f$ bounded for all $x \in D$
Let $d = \inf_{x\in D}[f(x)]$. Then $\exists$ sequence $\{x_n\}$ of points in D such that $f(x_n) \to d$
d and D ~~closed + bdd~~ compact $\Rightarrow$ we can select a convergent subsequence $\{x_n'\}$ of seq. $\{x_n\}$

(continuity of f) &

i.e. $\lim_{n\to\infty} x_n' = x_0 \in D$, $\lim_{n\to\infty} f(x_n') = f(x_0) = d$

$\therefore$ f continuous, D compact : $f(\lim x_n') = \lim f(x_n') = d$
i.e. f achieves its minimum ▶

If we knew merely that $f(\lim x_n') \leq \lim f(x_n')$ the proof still works because

$$f(x_0) = f(\lim x_n') \leq \lim f(x_n') = d$$
$$\Rightarrow f(x_0) \leq d$$

But by definition of d we have $f(x) \geq d$ for all $x \in D$.

$$\Rightarrow \{f(x_0) \geq d \ \wedge \ f(x_0) \leq d\} \Rightarrow f(x_0) = d$$

Defn: sequence ~~$\{a_n\}$~~ $\{a_n\}$; suppose U satisfies
(i) Ultimately all terms of $\{a_n\}$ lie to the left of U (i.e. $\forall\, \varepsilon > 0\ \exists\, N > 0 \ni \{n > N \Rightarrow a_n < U + \varepsilon\}$
(ii) Infinitely many terms lie to the right of $U - \varepsilon$
(i.e. given $\varepsilon > 0$ and given $m > 0$ $\exists$ integer $n > m$

$\exists \ a_n > U - \varepsilon$

U is called limit superior

$$U = \limsup_{n\to\infty} a_n = \overline{\lim_{n\to\infty}}\, a_n$$

Ex (i) $a_n = \dfrac{(-1)^n}{1+\frac{1}{n}}$, $\underline{\lim} = -1$, $\overline{\lim} = +1$

(ii) $a_n = (-1)^n$, $\underline{\lim} = -1$, $\overline{\lim} = +1$

(iii) $a_n = (-1)^n n$, $\underline{\lim} = -\infty$, $\overline{\lim} = +\infty$

(iv) $a_n = n^2 \sin^2(\frac{1}{2} n\pi)$, $\underline{\lim} = 0$, $\overline{\lim} = +\infty$

2. Defn: A lower semi-continuous function is defined as a function $f \ni f(\lim x_n) = \underline{\lim}\, f(x_n)$ for every convergent sequence $\{x_n\}$ in the domain of f.

Theorem: Every lower (upper) semi-continuous function defined on a closed and bounded set D possesses a min (max) in D.
($f(x)$ has a max at $x = x_0$ if $f(x) < f(x_0)$ for all x in a deleted neighborhood of x_0).

* S: function space, $\underline{x} = (x_1, \ldots, x_n)$

Functional: mapping $S \to \mathbb{R}$

Linear space: $x+y = y+x$, $(\alpha+\beta)x = \alpha x + \beta x$

3. Linear space M is said to be normed if to each $x \in M$ we can assign a non-negative number $\|x\|$ called the norm of x, $\ni$:

$$\begin{cases} \|x\| = 0 \Leftrightarrow x = 0 \\ \|\alpha x\| = |\alpha| \|x\| \\ \|x+y\| \leq \|x\| + \|y\| \end{cases}$$

Ex: (i) Space $C[a,b]$ of all continuous functions on AB.
Define norm of an element $y(x) \in C[a,b]$ as

$$\|y\|_0 = \max_{a \leq x \leq b} |y(x)|$$

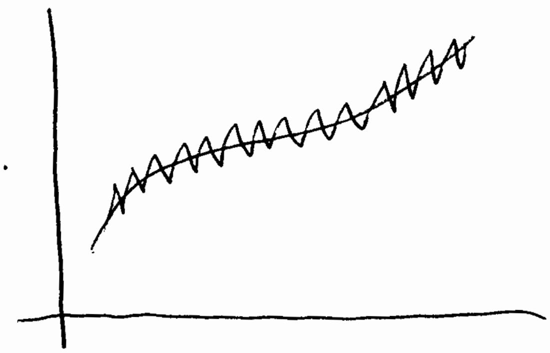
(ii) Space $C^1[a,b]$ ($\equiv \mathcal{D}_1[a,b]$ in Gelfand & Fomin) of all functions which are continuous and have continuous first derivatives on $[a,b]$.

$$\|y\|_1 = \max_{x \in [a,b]} |y(x)| + \max_{x \in [a,b]} |y'(x)|$$

(iii) $C^n[a,b]$; $\|y\|_n = \sum_{i=0}^{n} \max_{x \in [a,b]} |y^{(i)}(x)|$

$$\begin{cases} \text{I}' & C_p[a,b] \\ \text{II}' & C_p^1[a,b] \\ \text{III}' & C_p^{(n)}[a,b] \end{cases} \qquad \text{metric } \rho(u,v) = \|u-v\|_p$$

Ex.



Close in C but not in C'

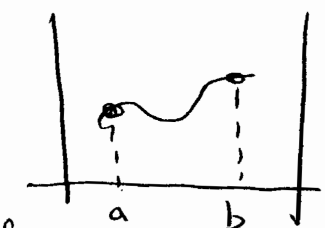
4. Functionals

1. **Definition:** weak variation ($\delta y, \delta y'$ small)
   strong " δy small

Definition: The functional is said to be continuous at the point $y_0 \in \mathcal{M}$ if $\forall \varepsilon > 0 \, \exists \delta > 0$
$\ni |I(y) - I(y_0)| < \varepsilon$ provided that $\|y - y_0\| < \delta$.

* A functional can be continuous in some space and discontinuous in another.

Ex: Arclength $\quad S = \int_a^b \sqrt{1 + [y'(x)]^2}\, dx \equiv I[y]$



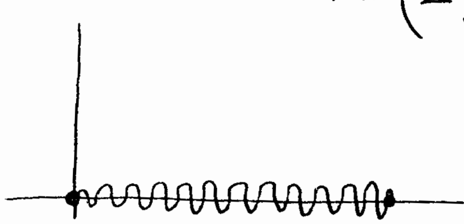
① $I[y]$ is continuous in $C^1[a,b]$, because clearly
$\|y_1 - y_2\| < \delta \Rightarrow |y_1'(x) - y_2'(x)| < \delta \Rightarrow |I[y_1] - I[y_2]| < \varepsilon$

② $I[y]$ is NOT continuous in the $\|\ \|_0$ norm, even in C^1.

Take $y(x) = a \sin(r\pi x)$ in $x \in [0, 1]$; then

$$S = \int_0^1 \sqrt{1 + r^2\pi^2 a^2 \cos^2(r\pi x)}\, dx > r\pi a \int_0^1 |\cos r\pi x|\, dx$$

$$= r\pi a \left(2r \int_0^{1/2r} \cos \pi r x\, dx \right) = r\pi a \cdot 2r \cdot \frac{1}{\pi r} = 2ar$$



Take $a = \frac{1}{n}$, $r = n^2$. Then

$I > 2n \Rightarrow I \to \infty$ as $n \to \infty$
but $|y| \le \frac{1}{n} \to 0$. $\left(y' = ar\pi \cos r\pi x = n\pi \cos n^2 \pi x \atop \|y\|_1 = n\pi + \frac{1}{n} \to \infty \right)$

$\therefore \quad I[\underset{= 0}{\lim \phi_n}] \le \underline{\lim}\, I[\phi_n]$, lower semicontinuous

<u>Theorem</u>: Every lower (upper) semicontinuous functional defined on a compact* function space possesses a min (max) in that space [proved using equicontinuity and Arzela-Ascoli: look in Courant-Friedrichs].

---

* <u>Definition</u>: A set of functions $\{f_n(x)\}$ in a space M is said to be compact if every sequence of functions in the set contains a subsequence which converges to some function in M.

$$f(x_1, \ldots, x_n) \; ; \; df = f'(x)\,dx$$

$f(x)$ $\begin{cases} \text{necessary} \\ \text{sufficient} \end{cases}$ conditions for min at $x = x_0$ are $\begin{matrix} f'(x_0) = 0, & f''(x_0) \geq 0 \\ & f''(x_0) > 0 \end{matrix}$

Ordinary functions: we say that $f(x)$ is differentiable at x,

$$f(x+h) - f(x) = \underbrace{A \cdot h}_{\text{differential}} + o(|h|) \quad \text{where} \quad A \cdot h = \sum_{i=1}^{n} A_i h_i$$

$$\lim_{\varepsilon \to 0} \left[\frac{f(\underline{x} + \varepsilon \underline{\eta}) - f(\underline{x})}{\varepsilon} \right] = A \cdot \eta \;, \quad \text{derivative: } \underline{A} = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right)$$

$$A = \nabla f = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right)$$

$$f(x + \underline{h}) - f(x) = \underbrace{A \cdot h}_{\text{linear form}} + (\text{quadratic form})$$

<u>Ex</u>: $$f(x+h, y+k) - f(x,y) = f_x(x,y)\,h + f_y(x,y)\,k + \frac{1}{2!}\left[f_{xx} h^2 + 2 f_{xy} hk + f_{yy} k^2 \right] + \frac{1}{3!}\left[f_{xxx} h^3 + 3 f_{xxy} h^2 k + 3 f_{xyy} hk^2 + f_{yyy} k^3 \right] + \ldots$$

2. Linear functionals

Defn: $I(\alpha y_1 + \beta y_2) = \alpha I(y_1) + \beta I(y_2)$: linear functional

Ex. $I(y) = \int_a^b f(x)y(x)\,dx$

$I(y) = y(x_0)$

$I(y) = \int_a^b [\alpha_0(x)y(x) + \alpha_1(x)y'(x) + \cdots + \alpha_n(x)y^{(n)}(x)]\,dx$

Defn: functional $B(x,y)$ is bilinear if it is linear in $y(x)$ for fixed x (y).

i.e. $B(\alpha x + \beta y, z) = \alpha B(x,z) + \beta B(y,z)$

$B(x, \alpha y + \beta z) = \alpha B(x,y) + \beta B(x,z)$

If we set $y=x$ we get a quadratic functional $A(x) = B(x,x)$.

If $A(x) > 0$ for $x \neq 0$, then positive definite

Ex. (i) $B(x,y) = \sum_{i,j=1}^{n} b_{ij} x_i y_j$

$A(x) = B(x,x) = \sum_{i,j=1}^{n} b_{ij} x_i x_j$

(ii) $B(x,y) = \int_a^b \alpha(t)\,x(t)\,y(t)\,dt$

$A(x) = \int_a^b \alpha(t)\,x^2(t)\,dt$

4.3 (iii) $B(x,y) = \int_a^b \int_a^b K(s,t)\,x(t)\,y(t)\,ds\,dt$

3. Defn: $J(y)$ differentiable functional

Suppose $\underbrace{J(y+h) - J(y)}_{\Delta J} = \phi(h) + \varepsilon \|h\|$ ↳ more precisely, $\phi(y,h)$

where $\varepsilon \to 0$ as $\|h\| \to 0$ and where $\phi(h)$ is a linear functional. Then the functional $J(y)$ is said to be differentiable at y and the linear functional $\phi(h)$ is

Called the first variation (or differential) of J and denoted by $\delta J[h]$

Riesz representation theorem. A linear functional $\phi(h)$ can be represented uniquely as the inner product

$$\Phi(h) = (h, u).$$

Define u as the derivative

$$\underline{\text{Defn}}: \quad \Delta J = J(y+h) - J(y) = \underset{\text{linear in } h}{\overset{\nearrow \delta J}{\Phi_1(h)}} + \underset{\text{quad in } h}{\overset{\nearrow \delta^2 J}{\Phi_2(h)}} + \overset{\text{i.e. } o(\|h\|^2)}{\varepsilon \|h\|^2}$$

$\underline{Ex}$: $J(y) = \int_a^b F(x, y(x), y'(x))\, dx$

$$\Delta J = J(y+h) - J(y) = \int_a^b F(x, y+h, y'+h')\, dx - \int_a^b F(x, y, y')\, dx$$

$$= \int_a^b \left[F_y(x,y,y')\, h + F_{y'}(x,y,y')\, h' \right] dx +$$

$$+ \frac{1}{2} \int_a^b \left[F_{yy}\, h^2 + 2F_{yy'}\, h h' + F_{y'y'}\, h'^2 \right] dx + \cdots$$

$$\delta J := \int_a^b \left(F_y h + F_{y'} h' \right) dx$$

$$\delta^2 J := \frac{1}{2} \int_a^b \left[F_{yy} h^2 + 2F_{yy'} h h' + F_{y'y'} h'^2 \right] dx$$

$\underline{Ex}$: $J(y) = \int_a^b F(x,y,y')$, $y(a) = A$, $y(b) = B$; $h(a) = h(b) = 0$

$$\delta J = \int_a^b \left[F_y h + F_{y'} h' \right] dx = \int_a^b \left[F_y h - \frac{d}{dx} F_{y'}\, h \right] dx \Rightarrow$$

$$\left(\int_a^b F_{y'} h'\, dx = \underset{=0 \text{ from BC. on } h}{F_{y'} h \Big|_a^b} - \int_a^b h \frac{dF_{y'}}{dx}\, dx \right)$$

$$\Rightarrow \delta J = \int_a^b \left(F_y - \frac{d}{dx} F_{y'} \right) h \, dx = \left(F_y - \frac{d}{dx} F_{y'} , h \right)$$

Functional derivative:

$$\frac{\delta J}{\delta y} := F_y - \frac{d}{dx} F_{y'}$$

Ex (1) $J(y) = y(x_0)$

$$J(y+h) - J(y) = y(x_0) + h(x_0) - y(x_0) = h(x_0) = \delta J$$

$$\Rightarrow \delta J = \int \delta(x - x_0) \, h(x) \, dx$$

i.e. $$\frac{\delta J}{\delta y} = \delta(x - x_0)$$

Ex (2) $J(y) = \int_a^b \int_a^b K(x,y) \, y(x) \, y(\xi) \, dx \, d\xi \; ; \; K(x,\xi) = K(\xi,x)$

$$\Delta J \equiv J(y+h) - J(y) = \iint K \left[y(x) h(\xi) + h(x) y(\xi) \right] dx \, d\xi + \mathcal{O}(h^2)$$

$$= 2 \iint K(x,y) \, y(\xi) \, h(x) \, d\xi \, dx \quad \text{(because } K \text{ symmetric)}$$

$$= \int_a^b \left\{ 2 \int_a^b K(x,\xi) \, y(\xi) \, d\xi \right\} h(x) \, dx$$

$$\Rightarrow \frac{\delta J}{\delta y} = 2 \int_a^b K(x,\xi) \, y(\xi) \, d\xi$$

So $\quad \Delta J = J(y+h) - J(y) = \underbrace{\Phi(h)}_{\delta J} + o(\|h\|)$

$$\delta J = \lim_{\varepsilon \to 0} \left[\frac{J(y+\varepsilon\eta) - J(y)}{\varepsilon} \right] = \left.\frac{dJ}{d\varepsilon}\right|_{\varepsilon=0} \cdot \eta$$

Ex: $J(y) = \int_a^b y^2\,dx, \quad \frac{\delta J}{\delta y} = 2y$

7.4. EXTREMA OF FUNCTIONALS (max or min)

Functional $J(y)$ has a relative extremum for $y = y_0$ if $J(y) - J(y_0)$ does not change sign in some neighborhood ($\|\cdot\|_i$) of y_0 almost always (in $C[a,b]$ or $C'[a,b]$)

1. Defn: $J(y)$ has extremum for $y = \hat{y}$ if $\exists\ \varepsilon > 0 \ni$ $J(y) - J(\hat{y})$ has the same sign for all y in domain of definition of the functional which satisfy

$\|y - \hat{y}\|_1 < \varepsilon$ (weak variation $\left.\begin{matrix} y - \hat{y} \\ y' - \hat{y}' \end{matrix}\right\}$ small)

or

$\|y - \hat{y}\|_0 < \varepsilon$ (strong variation : $y - \hat{y}$ small)

1. Theorem: A necessary condition for the differentiable functional J to have an extremum at $y = \hat{y}$ is that $\delta J = 0$ at $\hat{y}$.

◀ Suppose $\delta J \neq 0$; $\Delta J = J(y+h) - J(y) = \delta J(h) + \varepsilon\|h\|$

Take some admissible $h_0 \ni \delta J(h_0) \neq 0$. For arbitrary $\varepsilon > 0$, $J(\hat{y}+\varepsilon h_0) - J(\hat{y}) = \delta J(\varepsilon h_0) + \varepsilon \|h_0\| = \varepsilon \delta J(h_0) + \varepsilon \|h_0\|$.

$$J(\hat{y}-\varepsilon h_0) - J(\hat{y}) = \delta J(-\varepsilon h_0) = -\varepsilon \delta J(h_0) + \varepsilon \|h\|$$

by linearity of 1st variation

i.e. $J(\hat{y})$ cannot be an extremum

DEF (2) J is stationary at $\hat{y}$ if $\delta J = 0$ at $\hat{y}$.

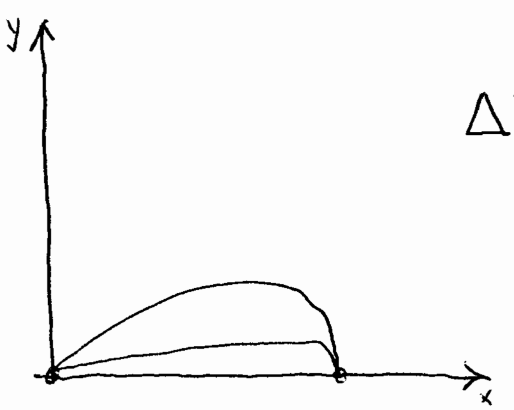
THM (2) Necessary condition for $J(y)$ to have $\underset{(\max)}{\min}$ at $\hat{y}$ is $\delta^2 J \underset{(\leq)}{\geq} 0$ at $\hat{y}$.

$$\Delta J \equiv J(y+h) - J(y) = \underbrace{\delta J(h)}_{=0} + \underbrace{\delta^2 J(h)}_{\text{quadratic (i.e. of constant sign)}} + \underbrace{\varepsilon}_{0 \text{ as } \|h\| \to 0} \|h\|^2$$

NOTE: these are NOT sufficient conditions.

EX (3) $I(y) = \int_0^1 (y'^2 + y'^3)\,dx$; $y(0) = 0$, $y(1) = 0$

$y \in C'[a,b]$



$$\Delta I = I(y+h) - I(y) = \underbrace{\int_0^1 (2y'h' + 3y'^2 h')\,dx}_{\delta I} + \underbrace{\int_0^1 (h'^2 + 3y'h'^2)\,dx}_{\delta^2 I} + \int_0^1 h'^3\,dx$$

Clearly for $y \equiv 0$ we have $\delta I = 0$ and $\delta^2 I > 0$

but consider $y_\phi(x) = \left(1/\tan\phi\right)^{3/2}\left(x\tan\phi - \tan\phi x\right)$

for which $\|y_\phi\| \to 0$ as $\phi \to \pi/2$ but $I(y_\phi) \to -\infty$.

$\therefore$ $y \equiv 0$ is <u>not</u> a strong relative min $\left(I(0)=0\right)$

But $y \equiv 0$ does provide a weak relative minimum because

$$\|y\|_1 < \varepsilon \Rightarrow |y'| < \varepsilon \Rightarrow I \leq \int_0^1 (\epsilon^2 + \varepsilon^3)\,dx < 2\varepsilon^2$$

% Let $\int_0^1 u'^2\,dx = \varepsilon^2 c > 0$, $(0 < c < 1)$

then $\int_0^1 |u'^3|\,dx \leq \varepsilon^3 c$ and

$$I = \int_0^1 (u'^2 + u'^3)\,dx \geq \underbrace{\int_0^1 u'^2\,dx}_{\text{strictly positive}} - \int_0^1 |u'|^3\,dx \geq \varepsilon^2 c - \varepsilon^3 c = \varepsilon^2 c(1-\varepsilon) > 0$$

i.e. $0 < \varepsilon^2 c(1-\varepsilon) \leq I(y_\varepsilon) \leq 2\varepsilon^2$, i.e. $I(y_\varepsilon) \xrightarrow{\varepsilon \to 0^+} 0^+$.

or $I(y) - I(0) > 0$ for all $y_\varepsilon \ni \|y_\varepsilon - 0\|_1 < \varepsilon$.

5. FUNDAMENTAL LEMMA OF THE CALCULUS OF VARIATIONS

5.1. Preliminaries

<u>Lemma I</u>: If $\alpha(x) \in C[a,b]$ and if $\int_a^b \alpha(x) h(x)\, dx = 0$ for every function $h(x) \in C[a,b] \ni h(a) = h(b) = 0$ then $\alpha(x) \equiv 0$ in $[a,b]$.

◀ Suppose $\alpha(x_0) > 0$ for some $x_0 \in [a,b]$. Then by continuity $\alpha(x) > 0$ for $x \in [x_0-\delta, x_0+\delta]$ for δ suff. small, >0.

Take

$$h(x) = \begin{cases} (x - x_0 + \delta)(x - x_0 - \delta), & \text{for } x \in [x_0-\delta, x_0+\delta] \\ 0 & \text{otherwise} \end{cases}$$

for which $\int_a^b \alpha(x) h(x)\, dx > 0$, QED ▶
(by contradiction)

<u>Corollary</u>: same result if we replace C by $C^{(n)}$.

<u>Lemma II</u>: if $\alpha(x) \in C[a,b]$ and if $\int_a^b \alpha(x) h'(x)\, dx = 0$ $\forall\, h \in C'[a,b]$ with $h(a) = h(b) = 0$, then $\alpha(x) \equiv$ constant, x in $[a,b]$

◀ $\left(\int_a^b \alpha h' = \alpha h \big|_a^b - \int_a^b \alpha' h = 0 \ \forall h \Rightarrow \alpha' = 0 \text{ by I} \Rightarrow \alpha = \text{const} \right.$
does not work because don't know α' exists$\left. \right)$

Choose $h(x) = \int_a^x [\alpha(\xi) - c]\, d\xi$ where $c = \frac{1}{b-a} \int_a^b \alpha(\xi)\, d\xi$

Then:

$$0 = \int_a^b \alpha(x) h'(x)\, dx = \int_a^b \alpha(x) h'(x)\, dx - c\,[\underbrace{h(b) - h(a)}_{0}]$$

$$= \int_a^b [\alpha(x) - c]\, h'(x)\, dx = \int_a^b [\alpha(x) - c]^2\, dx$$

since $h' = \alpha(x) - c$

$\Rightarrow \alpha(x) \equiv c$ throughout $(\alpha \in C[a,b])$. ▶

Lemma III: if $\alpha(x) \in C[a,b]$ and if $\int_a^b \alpha(x) h''(x)\,dx = 0$ for every $h \in C^{(2)}[a,b]$; $h(a) = h'(a) = h(b) = h'(b) = 0$
then $\alpha(x) = C_0 + C_1 x$ $\quad$ (C_0, C_1 constants)

◀ (similar, see G & F) ▶

Lemma IV If $\alpha(x), \beta(x) \in C[a,b]$ and if

$$\int_a^b [\alpha(x) h(x) + \beta(x) h'(x)]\,dx = 0$$

for every function $h(x) \in C'[a,b]$ s.t. $h(a) = h(b) = 0$
then $\beta(x)$ is differentiable and $\beta'(x) = \alpha(x)$, for $x \in [a,b]$
◀ (formal integration by parts gets the answer by assuming β differentiable : no good!)

Let $A(x) = \int_a^x \alpha(\xi)\,d\xi \Rightarrow \int_a^b \alpha h\,dx = \int_a^b A' h\,dx = \underbrace{Ah\Big|_a^b}_{0} - \int_a^b A h'\,dx$

$$\Rightarrow \int_a^b (\alpha h + \beta h')\,dx = \int_a^b (-A + \beta) h'\,dx = 0$$

$$\Rightarrow \text{(Lemma II)} \quad -A + \beta = \text{constant} \Rightarrow \beta(x) = \text{constant} + \int_a^x \alpha(\xi)\,d\xi$$

$$\Rightarrow \beta' = \alpha(x)$$ ▶

5.2 Legendre's condition for weak extremum

Let $I(y) = \int_a^b F(x, y, y')\,dx$; $y(a) = A$, $y(b) = B$

($F \in C^2$ in all arguments)

$\delta I = \int_a^b (F_y h + F_{y'} h')\,dx$ $\quad$ i.e. $\frac{d}{dx} F_{y'} = F_y$ or

(Euler equation) $\quad \frac{d}{dx} F_{y'} - F_y = 0 \qquad (+\text{b.c.}) \qquad (1)$

is a necessary condition for extremum. All solutions are called extremals

Alternatively: $\quad F_{y'y'}\, y'' + F_{yy'}\, y' + F_{y'x} - F_y = 0 \qquad (2)$

(as a result y'' continuous where $F_{y'y'} \neq 0$).

<u>Theorem(1)(du Bois-Raymond):</u> Suppose $y \in C'[a,b]$ and it satisfies the Euler equations with $F \in C^2$ in all arguments, then y has continuous 2nd derivative wherever $F_{y'y'}(x,y,y') \neq 0$.

<u>2nd variation:</u> $(h(a) = h(b) = 0)$

$$\delta^2 I = \frac{1}{2}\int_a^b \left(F_{yy} h^2 + \underbrace{2F_{yy'} hh'} + F_{y'y'} h'^2\right) dx$$

Assumes C^3. (*) too much! (*)

$$\int_a^b 2F_{yy'} hh' dx = \underbrace{\cancel{h^2} F_{y'y}\Big|_a^b}_{=0} - \int_a^b \left(\frac{d}{dx} F_{yy'}\right) h^2 dx$$

$$\Rightarrow \quad \delta^2 I = \int_a^b \left(P h'^2 + Q h^2\right) dx$$

where $\quad P = \frac{1}{2} F_{y'y'} \qquad (3)$

$\qquad\quad Q = \frac{1}{2}\left(F_{yy} - \frac{d}{dx} F_{yy'}\right) \qquad (4)$

Apply theorem 4.4.2 which states that a necessary condition for $I(y)$ to have a minimum at $\hat{y}$ is that $\delta^2 I \geq 0$ at $\hat{y}$.

Lemma (1): A necessary condition for

$$\delta^2 I = \int_a^b (Ph'^2 + Qh^2)\,dx, \quad h(a)=h(b)=0,$$

$h(x) \in C'[a,b]$ to be non-negative is that $P \geq 0$ on $[a,b]$.

◀ Suppose $P \geq 0$ doesn't hold. Then $P = -2\beta$ $(\beta > 0)$ at some point x_0. Then $\exists\, \delta > 0 \ni P(x) < -\beta$ in

$$a \leq x_0 - \delta < x < x_0 + \delta \leq b$$

Consider an admissible $h(x)$ given by

$$h(x) = \begin{cases} \sin^2 \frac{\pi(x-x_0)}{\delta} & \text{for } x_0 - \delta \leq x \leq x_0 + \delta \\ 0 & \text{otherwise} \end{cases}$$

Then

$$\int_a^b (Ph'^2 + Qh^2)\,dx = \int_{x_0-\delta}^{x_0+\delta} P \frac{\pi^2}{\delta^2} \sin^2 \frac{2\pi(x-x_0)}{\delta}\,dx + \int_{x_0-\delta}^{x_0+\delta} Q \sin^4 \frac{\pi(x-x_0)}{\delta}\,dx$$

$$< -\frac{\beta\pi^2}{\delta} + 2M\delta, \quad \text{with } M = \max|Q|$$

which is negative for δ sufficiently small. ▶

THEOREM 1 (LEGENDRE): A necessary condition for the functional

$$I(y) = \int_a^b F(x,y,y')\,dx, \; y(a) = A, \; y(b) = B$$

to have a weak relative minimum at $y = \hat{y}(x)$ is that the inequality

$$F_{y'y'} \geq 0 \qquad \text{(Legendre's condition)} \qquad (5)$$

be satisfied at every point of the curve $\hat{y}(x)$.
(Same for maximum $\rightarrow$ invert inequality).

From theorem 4.4.2 we must show that $\delta^2 J \geq 0 \rightarrow F_{y'y'} \geq 0$

Now $\delta^2 J = \varepsilon^2 \int_a^b (P\eta^2 + 2Q\eta\eta' + R\eta'^2)\,dx$ (6) $\begin{cases} P = F_{yy}(x, \hat{y}, \hat{y}') \\ Q = F_{yy'} \quad " \\ R = F_{y'y'} \quad " \end{cases}$

where we assumed $\hat{y}$ is an extremal

$\eta(a) = \eta(b) = 0$

and we are considering variations $\hat{y} \rightarrow \hat{y} + \varepsilon\eta$, $\eta \in C^1[a,b]$. (piecewise smooth variations are also OK – see alternate proof).

Proof (1) (Legendre – with a flaw, fixed by Weierstrass): let $w \in C^1$, then

$w(a), w(b)$ arbitrary

(7) $$\varepsilon^2 \int_a^b (2\eta\eta' w + \eta^2 w')\,dx \equiv 0 \quad \left(= \varepsilon^2 \eta^2 w \Big|_a^b\right)$$

and adding to (6):

(8) $$\delta^2 J = \varepsilon^2 \int_a^b \left((P + w')\eta^2 + 2(Q + w)\eta\eta' + R\eta'^2\right) dx$$

The discriminant of this quadratic form in (η, η') is

(9) $$D = (Q + w)^2 - R(P + w') = 0$$

Setting it equal to zero for a proper choice of w, (8) is rewritten as

(10) $$\delta^2 J = \varepsilon^2 \int_a^b R\left[\eta' + \frac{Q + w}{R}\eta\right]^2 dx$$

i.e. if $R \neq 0$ in $[a,b]$ then $\delta^2 J$ has the same sign as R.

However this assumes w can be chosen always to satisfy

(11) $$\frac{dw}{dx} = -P + \frac{(Q + w)^2}{R}$$

To overcome this, we resort to the theory of 1^{st}-order ODE: let $R(x_0) < 0$ for $a < x_0 < b$. Then $\exists\ \delta > 0$:

<u>Lemma:</u> $\begin{cases} R(x) < 0 \text{ for } x \in [x_0-\delta, x_0+\delta) \\ \text{There exists a solution of (11), } w(x) \in C^1(x_0-\delta, x_0+\delta) \end{cases}$

◀ Indeed, since $R(x)$ is continuous, $\exists\ \delta_1 : R < 0,\ x \in (x_0-\delta_1, x_0+\delta_1)$

Hence (11) has rhs that is continuous and has a continuous partial derivative wrt w in the neighborhood of (x_0, w_0) (w_0 being arbitrary).

[Here we have assumed that $\hat{y}(x)$ is continuous], ~~piecewise~~ smooth ~~continuous con stx~~

From elementary existence theory, $\exists\ \delta_2 : x \in (x_0-\delta_2, x_0+\delta_2)$
$\Rightarrow$ (11) has a ~~continuous 1^{st} i~~ C^1 1^{st} integral.

Take $\delta = \min(\delta_1, \delta_2)$. ▶ (<u>Note</u>: no guarantee we can find such integral in $[a,b]$)

So, choose $\eta = \begin{cases} (x-x_0-\delta)(x-x_0+\delta), & x \in (x_0-\delta, x_0+\delta) \\ 0 & \text{otherwise} \end{cases}$

Then $\delta^2 J = \varepsilon^2 \int_{x_0-\delta}^{x_0+\delta} R\left(\eta' + \frac{\varphi + w}{R}\eta\right)^2 dx.$

So if $\hat{y}$ was a minimizer, and $R(x_0) < 0$ we are led to a contradiction, i.e.

$$\hat{y} \text{ minimizer} \Rightarrow R(x) \geq 0, \quad x \in [a, b]$$

(Note: as we have seen, $F_{y'y'} \neq 0 \Rightarrow y''$ continuous in $[a,b]$

<u>Note</u>: we still do not know that $F_{y'y'} > 0$ in $[a,b]$ $\Rightarrow \delta^2 J > 0$

Ex.(1): $I(y) = \int_0^1 \underbrace{(y'^2 + 12xy)}_{F(x,y,y')}\, dx \quad ; \quad y(0) = 0, \ y(1) = 1$

$$F_y = 12x \ , \quad F_{y'} = 2y' \ ; \quad \frac{d}{dx} F_{y'} = 2y''$$

$$F_{y'y'} = 2 > 0$$

$$\frac{d}{dx} F_{y'} - F_y = 0 \quad \Rightarrow \quad \left.\begin{array}{l} 2y'' - 12x = 0 \\ y(0) = 0 \\ y(1) = 1 \end{array}\right\} \Rightarrow \ y = x^3$$

Ex.(2) $\quad I(y) = \int_a^b F(y, y')\, dx$

$$\frac{d}{dx} F_{y'} - F_y = 0$$

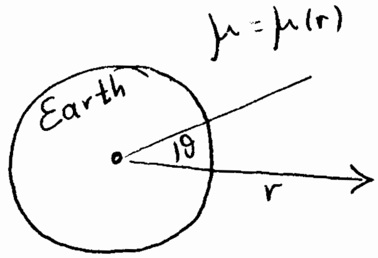
$$\hookrightarrow y'\left(F_{y'y'}\, y'' + F_{y'y}\, y' - F_y = 0\right) \qquad (\text{since } F_{y'x} \equiv 0)$$

$$\Rightarrow \frac{d}{dx}\left(y' F_{y'} - F\right) = 0$$

$$\left(= \left(\underbrace{\cancel{\partial_x}}_{\equiv 0} + y'\partial_y + y''\partial_{y'}\right)\left(y'F_{y'} - F\right) = y'^2 F_{yy'} - y' F_y + \cancel{y'' F_{y'}} + y'' y' F_{y'y'} - \cancel{y'' F_{y'}} \right)$$

$$\Rightarrow \ y' F_{y'} - F = \text{constant}$$

Ex.(3) : Fermat's principle



$$\frac{ds}{dt} = c \Rightarrow dt = \frac{ds}{c} = \mu\, ds$$

minimize $\displaystyle\int \mu(r)\,ds = \int \mu(r)\sqrt{r'^2 + r^2}\,d\theta \quad (*)_1$

$$\left.\begin{array}{c}\displaystyle\int F(\theta, r, r')\,d\theta \\ \Downarrow \\ \dfrac{d}{d\theta}F_{r'} - F_r = 0\end{array}\right\}$$

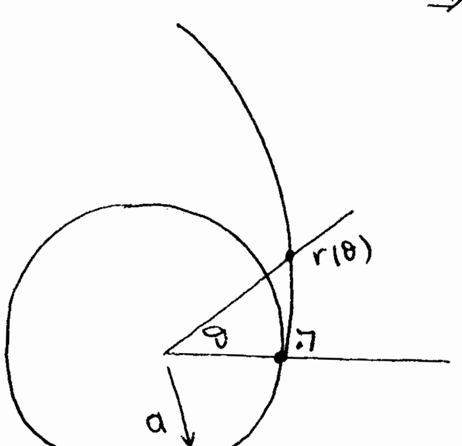
(assuming fixed BC - but leave unspecified)

$\Rightarrow$ (Ex.2): $\quad r' F_{r'} - F = \text{const. } K$ (indep. of θ)

where $F = \mu(r)\sqrt{r'^2 + r^2}$

$$\therefore \quad \frac{r'^2 \mu(r)}{\sqrt{r'^2 + r^2}} - \mu(r)\sqrt{r'^2 + r^2} = k$$

$$\Rightarrow \boxed{\frac{\mu(r)\, r^2}{\sqrt{r'^2 + r^2}} = k} = a\,\mu(a)$$



If the light ray is at right angles to the radius vector at $\theta = 0$, $r = a$, we have $r' = 0$ at $r = a$ so

$$\boxed{k = a\,\mu(a)}$$

E.g. $\mu(r) = \sqrt{1 + \frac{\sigma^2}{r^2}}$ $\qquad$ Let $\mu(a) = \sqrt{1 + \frac{\sigma^2}{a^2}} = \mu_0 > 1$

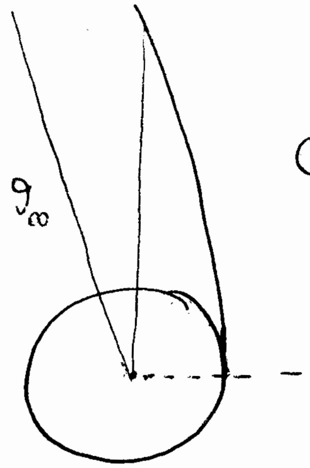
then, equ. for r' becomes

$$r'^2 = \frac{r^2(r^2 - a^2)}{a^2 \mu_0^2} \Longrightarrow r(\theta) = a \sec(\theta/\mu_0)$$

$(*)_1$ $\quad ds^2 = dx^2 + dy^2 = \left\{\left(\frac{dx}{d\theta}\right)^2 + \left(\frac{dy}{d\theta}\right)^2\right\} d\theta^2 = \{r'^2 + r^2\}\, d\theta^2$

$$\left.\begin{array}{l} x = r(\theta)\cos\theta \\ y = r(\theta)\sin\theta \end{array}\right\} \quad \frac{dx}{d\theta} = r'\cos\theta - r\underbrace{\theta'}_{=1}\sin\theta$$

$$\frac{dy}{d\theta} = r'\sin\theta + r\cos\theta$$



Clearly, $r = a$ when $\vartheta = 0$ while

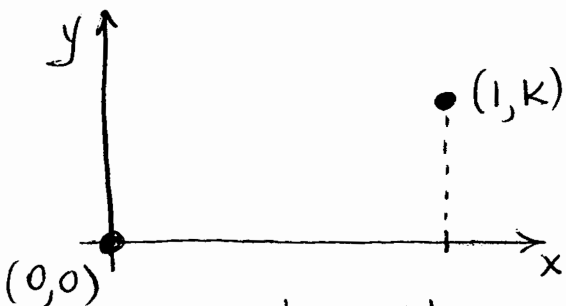
$$r \to \infty \quad \text{as} \quad \vartheta \to \left(\frac{\mu_0 \pi}{2}\right)^-$$

refractive index of air, $\lambda = 5.184 \times 10^{-5}$ cm

Note: for $\mu_0 = \frac{c_\infty}{c_\alpha} \sim 1.000294$ (Born & Wolf) p. 101, 7^{th} ed.

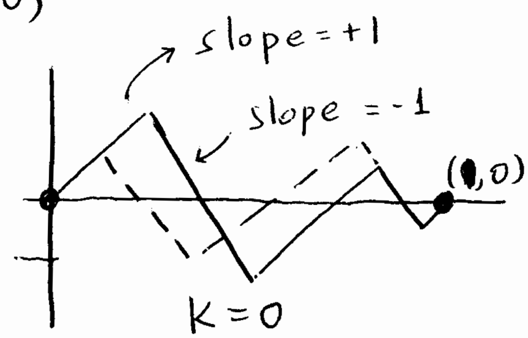
we find $\mu_0 \frac{\pi}{2} \simeq \frac{\pi}{2} + 1'35''$

Ex. (4) Minimize $\displaystyle\int_0^1 (y'^2 - 1)^2\, dx$; $y(0) = 0$, $y(1) = k$



let $k = \tan\phi$, $0 \le \phi \le \frac{\pi}{2}$

Now, $I(y) \geq 0$; lower bound (i.e. 0) is not attained for any function $y \in C'[0,1]$ (except for $\phi = \pi/4$, $k = 1$, see below)



However, can construct infinitely many curves that have only piecewise continuous derivatives

In general: $F = (y'^2 - 1)^2$, $F_{y'} = 4y'(y'^2 - 1)$

$$F_{y'y'} = 12\left(y'^2 - \tfrac{1}{3}\right)$$

Euler equation: $\displaystyle\frac{d}{dx} F_{y'} = 0 \Rightarrow F_{y'y'}\, y'' = 0 \Rightarrow \begin{cases} F_{y'y'} = 0 \\ y'' = 0 \end{cases}$