3 – Calculus of Variations

Conditions for weak minima over C^1 curves

Bibliography:

Main reference: [L2012], Ch. 2, pp. 26-70

Complementary: [K1970] ch. 4, 107-183

[K1970] D. E. Kirk (1970). *Optimal Control*. Dover reprint, 1998.

[B2004] B. van Brunt (2004). *Calculus of Variations*, Springer.

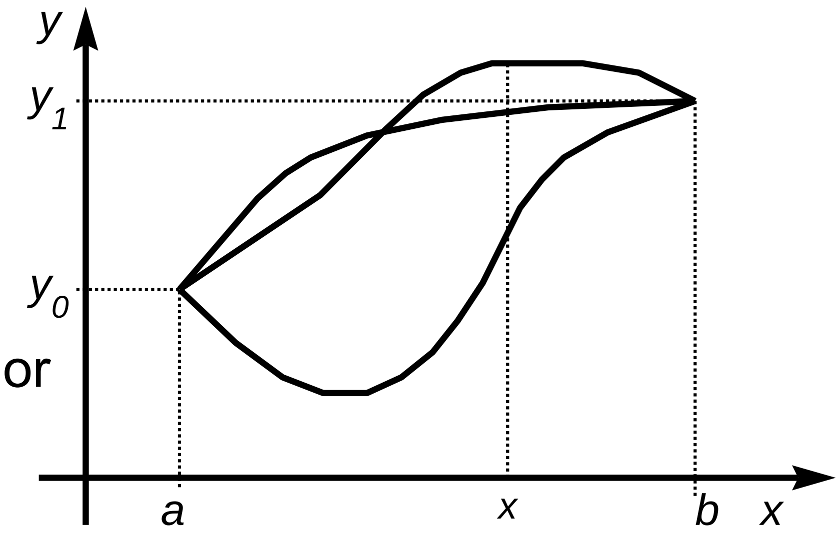
Basic Calculus of Variations Problem

Among all C^1 curves $y: [a,b] \to \mathbb{R}$ that verify the given boundary conditions

$$y(a) = y_0, \;\; y(b) = y_1,$$

Find the (local) minima of the cost function

$$J(y) \coloneqq \int_a^b L(x, y(x), y'(x))\, dx$$



$L: \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is the "lagrangian" or "runing cost".

Can be extended to $y: [a,b] \to \mathbb{R}^n$

Weak and strong extrema [L2012] p. 33-44

When considering the CV problem a norm on the functions y must be defined,

0-norm: $\|y\|_0 := \max_{a \le x \le b} |y(x)|$

1-norm: $\|y\|_1 := \max_{a \le x \le b} |y(x)| + \max_{a \le x \le b} |y'(x)|$

Strong extrema

Neighborhoods defined using the 0-norm

Weak extrema

Neighborhoods defined using the 1-norm

Exercise 1 *Weak vs. strong extrema* [L2012] p. 34

See also [SL2012] p. 7,8 for a quantitative discussion

Consider the problem of minimizing the functional

$$J(y) = \int_0^1 y(x)^2 (1 - \left(y'(x)\right)^2) dx$$

Subject to the boundary conditions

$$y(0) = y(1) = 0$$

Show that $y(x) = 0$ for $x \in [0,1]$ is a weak minimum but not a strong minimum.

<u>Solution</u>

$$J(y) = \int_0^1 y(x)^2 (1 - \left(y'(x)\right)^2) dx$$

For $y^*(x) = 0$, $J(y^*) = 0$.

If y is slightly perturbed according to the 1-norm, $|y'|$ must be small and $L = y^2(1 - (y')^2)$ is positive.

Hence, $J(y^* + \eta)$, where η is small according to the 1-norm, is positive, and thus $J(y^* + \eta) > J(y^*)$, i. e., $y^*(x) = 0$ is a minimum in the weak sense.

Consider now the 0-norm

$\|\eta\|_0$ small does not imply that $|y'|$ is small.

It is possible to find an η such that $\|\eta\|_0$ is arbitrarily small, but $|\eta'|$ is arbitrarily large. Example: $\eta(x) = \varepsilon \sin(\omega x)$, with $\varepsilon > 0$ and $\omega > 0$ constants.

In this case, $\|\eta\|_0 = \varepsilon$ and $|\eta'| = \varepsilon\omega|cos(\omega x)|$

It is thus possible to select ω sufficiently large such that $1 - (\eta')^2 < 0$ and

$$J(y = 0) > J(y = \eta)$$

Thus, $y(x) = 0$ is a weak minimum but <u>not</u> a strong minimum.

First-order necessary conditions for weak extrema

Basic CV problem:

$$\text{Minimize } J(y) := \int_a^b L(x, y(x), y'(x))\, dx$$

$$\text{s. t.} \qquad y(a) = y_0, \;\; y(b) = y_1,$$

$$y \in C^1([a, b] \to \mathbb{R})$$

Strategy to follow: Apply an admissible disturbance and impose that the first variation vanishes at an extremal.

Admissible variations

Let y a candidate to be an extremal

Perturbation $\eta: [a, b] \to \mathbb{R}$, C^1

$$y(x) + \alpha\eta(x) \qquad \alpha \in \mathbb{R}$$

For α close to 0 these perturbed curves are close to y in the 1-norm.

Since the boundary conditions must be satisfied,

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

Condition to impose (1^{st} order necessary condition):

$$\delta J|_y(\eta) = 0$$

Computing the 1st variation

Recall the definition of the 1st variation

$$J(y+\alpha\eta) = J(y) + \delta J|_y(\eta)\alpha + o(\alpha)$$

Perturbed functional $J(y+\alpha\eta) = \int_a^b L(x, y+\alpha\eta, y'+\alpha\eta')dx$

Taylor expansion with respect to α

$$J(y+\alpha\eta) = \int_a^b L\big(x,y,y') + L_y(x,y,y')\alpha\eta + L_{y'}(x,y.y')\alpha\eta' + o(\alpha)\big)dx$$

Match with the definition of the first variation

$$\delta J|_y(\eta) = \int_a^b L_y(x,y,y')\eta + L_{y'}(x,y.y')\eta' dx$$

Integration by parts

$$\frac{d}{dx}(zw) = \dot{z}w + z\dot{w}$$

$$\dot{z}w = \frac{d}{dx}(zw) - z\dot{w}$$

$$\int_a^b \dot{z}w dx = zw|_a^b - \int_a^b z\dot{w} dx$$

$$\int_a^b \dot{z} w dx = zw|_a^b - \int_a^b z\dot{w} dx$$

$$\int_a^b [L_{y'}(x,y,y')\eta'] dx$$

Eliminate the dependency on $\eta'$by integrating by parts

$$z = \eta \quad w = L_{y'}$$

$$\int_a^b [L_{y'}(x,y,y')\eta'] dx = L_{y'}(x,y,y')\eta(x)|_a^b - \int_a^b \left(\frac{d}{dx} L_{y'}(x,y,y')\right)\eta(x) dx =$$

$$= -\int_a^b \left(\frac{d}{dx} L_{y'}(x,y,y')\right)\eta(x) dx$$

Recall

$$\delta J|_y(\eta) = \int_a^b L_y(x,y,y')\eta + L_{y'}(x,y.y')\eta' dx$$

$$\int_a^b [L_{y'}(x,y,y')\eta'] dx = -\int_a^b \left(\frac{d}{dx} L_{y'}(x,y,y')\right)\eta(x)dx$$

Hence

$$\delta J|_y(\eta) = \int_a^b \left[L_y - \frac{d}{dx} L_{y\prime}\right]\eta(x)dx$$

Fundamental lemma of CV

If a continuous function $\xi: [a, b] \to \mathbb{R}$ is such that

$$\int_a^b \xi(x)\eta(x)dx = 0$$

for all C^1 functions $\eta: [a, b] \to \mathbb{R}$ with $\eta(a) = \eta(b) = 0$, then $\xi \equiv 0$.

Remark: In general if $\int_a^b \varphi(x)dx = 0$ one may not infer that $\varphi(x) = 0$ for all x.

Here, the situation is different because the integral is zero for all possible (admissible) perturbation functions η.

Proof

The proof is made by contradiction.

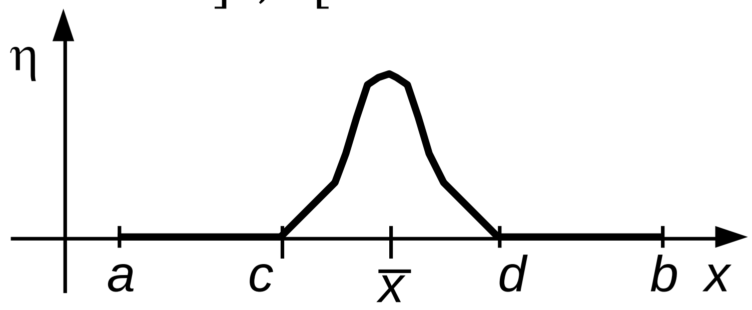
Assume that $\xi(\bar{x}) \neq 0$ for some $\bar{x} \in [a, b]$.

Since ξ is continuous, ξ is nonzero and keeps the same sign on some subinterval $[c, d]$ that contains $\bar{x}$.

Construct a function $\eta \in C^1([a, b], \mathbb{R})$ that is positive on $]c, d[$ and 0 elsewhere.

This construction yields

$$\int_a^b \xi(x)\eta(x)dx = \int_c^d \xi(x)\eta(x)dx > 0$$



And we reach a contradiction.

Hence, ξ must be always zero on $[a, b]$. q.e.d.

Back to the necessary 1st order condition

We have seen that, for any C^1 functions $\eta: [a, b] \to \mathbb{R}$ with $\eta(a) = \eta(b) = 0$

$$\delta J|_y(\eta) = \int_a^b \left[L_y - \frac{d}{dx} L_{y'}\right] \eta(x) dx$$

From the Fundamental Lemma of CV, for $\delta J|_y(\eta) = 0$ it must be

$$L_y = \frac{d}{dx} L_{y'}$$

Euler-Lagrange equation

A necessary condition for a $C^1([a,b],\mathbb{R})$ function y to be a weak extremum of the basic CV problem

$$\text{Minimize } J(y) := \int_a^b L(x, y(x), y'(x))\, dx$$

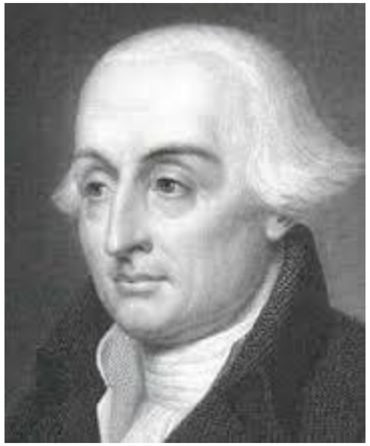
$$\text{s. t.} \qquad y(a) = y_0, \;\; y(b) = y_1$$

is that it satisfies the Euler-Lagrange equation

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

Hereafter, unless explicitly state otherwise, “extremum” refers to “weak extremum”.

Leonhard Euler (1707-1783) was born in Basel, Switzerland, a son of a Reform Church pastor. He studied with Johann Bernouilli at the University of Basel and, in 1727, moved to S. Petersburg where he stayed until 1741, when he accepted a post at the Berlin Academy, offered by the king of Prussia Frederick the Great. In Berlin he met Maupertuis and participated on the controversy that involved the President of the Berlin Academy concerning the Principle of Least Action. After becoming blind of one eye, he returned to Russia in 1760 where he stayed until his death. Euler was one of the greatest and most prolific mathematicians of all times. He discovered the formula $e^{i\pi} + 1 = 0$ and found it so elegant that he asked it to be engraved in his tomb. In 1744 he published the magnificent *The Method of Finding Plane Curves that Show some Property of Maximum or Minimum* and coined the expression *Calculus of Variations* to designate the new research area. In this 1744 book, Euler set up a general procedure to write the E-L equation and enunciated the Principle of Least Action.



Joseph-Louis de Lagrange (1736-1813), despite the French name with which he has been registered in the History of Mathematics, was actually born in Turim, Italy, under the name of Giuseppe Ludovico Lagrangia. In 1755, the young Ludovico, then with 19 years, sent to Euler a letter with an appendix in which he explained how to replace the Euler method to solve variational problems, that was based on geometrical insight, by an analytical method. Although Euler did not reply to the letter, he wrote about this work of Lagrange: *Even though the author of this* [Euler] *had meditated a long time and had revealead to friens his desire, yet the glory of first discovery was reserved to the very penetrating geometer of Turin La Grange, who having used analysis alone, has clearly attained the very same solution which the authorhad deduced by geometrical considerations.* Later, Lagrange's treaty *mécanique Analitique*, published in 1788, became a cornerstone for the development of mathematical physics.

Exercise 2 [K1970], Example 4.2.1, p128-130.

a) Find an extremal for the functional

$$J(y) = \int_0^{\pi/2} \left[y'^2(x) - y^2(x) \right] dx$$

s. t. $\quad y(0) = 0, \quad y\left(\frac{\pi}{2}\right) = 1.$

b) Let y^* be the extremal and consider the perturbed signal

$$y(x) = y^*(x) + \eta(x)$$

where $\eta(x) = \alpha \sin{(2x)}$. Compute $J(y^*)$ and $J(y)$.

What can you say about y^* being a maximum or a minimum?

Help: EL equation $\frac{\partial L}{\partial y} = \frac{d}{dx}\left(\frac{\partial L}{\partial y\prime}\right)$

Help: To solve the equation $y + y'' = 0$, assume that the solution is of the form

$$y(x) = ke^{\gamma x}$$

for some k and γ. Replace this candidate solution in the equation and get an algebraic equation for the possible values γ_i of γ. The general solution is thus

$$y(x) = \sum_i k_i e^{\gamma_i x}$$

Finally, use the boundary conditions to find the coefficients k_i.

Other useful formulas:

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

Solution

a)

$$L = {y'}^2(x) - y^2(x)$$

$\frac{\partial L}{\partial y} = \frac{d}{dx}\left(\frac{\partial L}{\partial y\prime}\right)$ $\frac{\partial L}{\partial y} = -2y$ $\frac{\partial L}{\partial y\prime} = 2y'$ The EL equation becomes $y + y'' = 0$

Assume $y = ke^{\gamma x}$. Then $y'' = k\gamma^2 e^{\gamma x}$.

$$y + y'' = 0 \;\; \rightarrow \;\; ke^{\gamma x} + \; k\gamma^2 e^{\gamma x} = 0 \;\; \rightarrow \;\; 1 + \gamma^2 = 0 \;\; \rightarrow \;\; \gamma = \pm j$$

The general solution is thus

$$y(x) = k_1 e^{jx} + k_2 e^{-jx}$$

$$y(x) = k_1(\cos x + j \sin x) + k_2(\cos x - j \sin x)$$

$$y(x) = (k_1 + k_2) \cos x + j(k_1 - k_2) \sin x$$

$$y(x) = (k_1 + k_2)\cos x + j(k_1 - k_2)\sin x$$

Apply the boundary conditions $y(0) = 0, \quad y\left(\frac{\pi}{2}\right) = 1$

$$\begin{cases} k_1 + k_2 = 0 \\ j(k_1 - k_2) = 1 \end{cases} \quad \begin{cases} k_1 = -\frac{j}{2} \\ k_2 = \frac{j}{2} \end{cases}$$

The extremal (solution of the EL equation) is thus

$$y^*(x) = \sin x$$

b) $y^*(x) = \sin x \qquad y^{*\prime}(x) = \cos x$

$$J(y^*) = \int_0^{\pi/2} \left[{y^{*\prime}}^2(x) - {y^*}^2(x) \right] dx = \int_0^{\pi/2} [(\cos x)^2 - (-\sin x)^2] dx$$

$$= \int_0^{\pi/2} \cos(2x)\, dx = \frac{1}{2} \sin(2x) \,\Big|_0^{\pi/2} = 0$$

Furthermore

$$y(x) = \sin x + \alpha \sin(2x)$$

$$J(y) = \int_0^{\pi/2} \left[{y'}^2(x) - y^2(x) \right] dx = \frac{3}{4} \pi \alpha^2$$

Therefore

$$J(y^* + \eta) > J(y^*) \qquad \text{for } \alpha \neq 0$$

Conclusions

The extremal $y^* = \sin x$ may not be a maximum because we found an admissible perturbation that leads to a bigger value of the functional.

Although it may not be concluded that it is a minimum (since the check was done for only a particular value of η), the test is compatible with this function being a minimum.

Exercise 3

Find the extremal of the functional for the following fixed-end-points

$$J(y) = \int_0^{\pi} (y'^2 + 2y \sin x) dx$$

s. t. $y(0) = y(\pi) = 0$

Hint: EL equation

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

<u>Solution</u>

$$\frac{\partial L}{\partial y} = \frac{d}{dx}\left(\frac{\partial L}{\partial y\prime}\right) \qquad L = y'^2 + 2y\sin x$$

$$\frac{\partial L}{\partial y} = 2\sin x \qquad \frac{d}{dx}\left(\frac{\partial L}{\partial y\prime}\right) = \frac{d}{dx}2y' = 2y''$$

The EL equation reduces to $y'' = \sin x$

$y'(x) = -\cos x + \alpha$ $\qquad$ $y(x) = -\sin x + \alpha x + \beta$ $\qquad$ α and β constants

Apply the boundary conditions

$$y(0) = -\sin(0) + \alpha 0 + \beta = \beta \ \Rightarrow \ \beta = 0$$

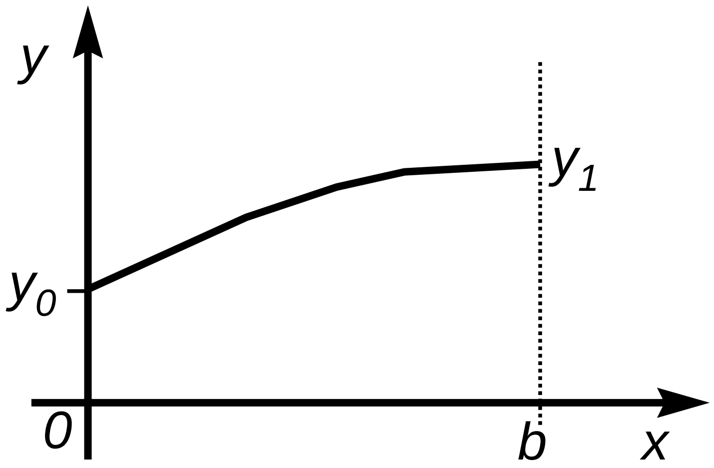
$$y(\pi) = -\sin(\pi) + \alpha\pi = \alpha\pi = 0 \Rightarrow \alpha = 0$$

Therefore, the extremal curve is given by

$$y(x) = -\sin(x)$$

Exercise [L2012] p.38, example 2.2

What is the curve with shortest length that connects the points $(0, y_0)$ and (b, y_1)?



Minimize $J(x) = \int_0^b \sqrt{1 + y'^2}dx$

s. t. $y(0) = y_0, \; y(b) = y_1$

Solution

EL equation $\frac{\partial L}{\partial y} = \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right)$ $\quad L = \sqrt{1 + y'^2} \quad L_y = 0 \quad L_{y'} = \frac{y'}{\sqrt{1+y'^2}}$

EL equation becomes $\frac{d}{dx}\left(\frac{y'}{\sqrt{1+y'^2}}\right) = 0$ or $\frac{y'}{\sqrt{1+y'^2}} = \sqrt{c_1}$

$$y'^2 = c_1(1 + y'^2) \quad \rightarrow \quad y' = \sqrt{\frac{c_1}{1-c_1}} = c_2 \quad \rightarrow \quad y(x) = c_2 x + c_3$$

Hence the solution is a straight line.

Multiple degrees of freedom

If $y = [y_1 \quad \cdots \quad y_n]^T \in \mathbb{R}^n$ the El is written componentwise

$$\frac{\partial L}{\partial y_i} = \frac{d}{dx}\left(\frac{\partial L}{\partial {y_i}'}\right)$$

Special cases of the EL equation

General case

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

Special case 1 (“no y”)

$$L = L(x, y')$$

The EL becomes

$$\frac{d}{dx} L_{y'} = 0 \qquad \Rightarrow \qquad L_{y'} = c \qquad c \text{ a constant}$$

The quantity $L_{y'}$ is called momentum.

General case

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

Special case 2 ("no x")

$$L = L(y, y')$$

Then, $L_{y'x} = 0$, and the EL becomes

$$L_y = L_{y'y} y' + L_{y'y'} y''$$

Multiply both sides by y'

$$L_{y'y}(y')^2 + L_{y'y'} y'' y' - L_y y' = 0$$

$$L_{y'y}(y')^2 + L_{y'y'}y''y' - L_y y' = 0$$

The next slide shows that this expression is equivalent to

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

Therefore

$$L_{y'}y' - L = c \qquad c \text{ a constant}$$

The quantity $L_{y'}y' - L$ is called the Hamiltonian function.

In other words: When the Lagrangian does not depend on x, the Hamiltonian is constant with respect to x.

$$\frac{d}{dx}\left(L_{y'}y' - L\right) =$$

$$\left(L_{y'y}y' + L_{y'y'}y''\right)y' + L_{y'}y'' - L_y y' - L_{y'}y'' =$$

$$= L_{y'y}(y')^2 + L_{y'y'}y''y' - L_y y'$$

Exercise 4 (The Brachistochrone problem)

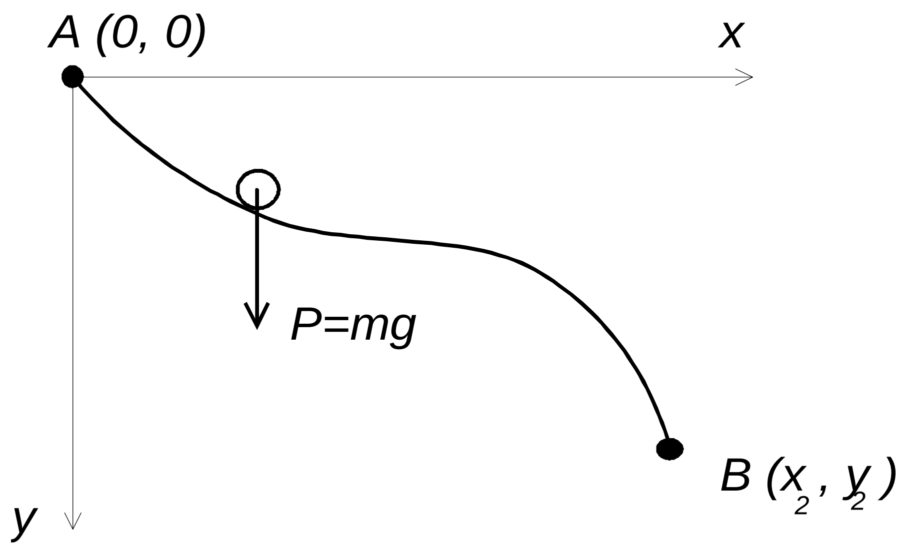
What is the shape of a curve that connects points A and B and such that a mass point slides along it, without friction, acted only by the force of gravity and going from A to B in minimum time?

Minimize

$$J(u) = \sqrt{2g}T = \int_0^{x_2} \sqrt{\frac{1+u(x)^2}{y(x)}}dx$$

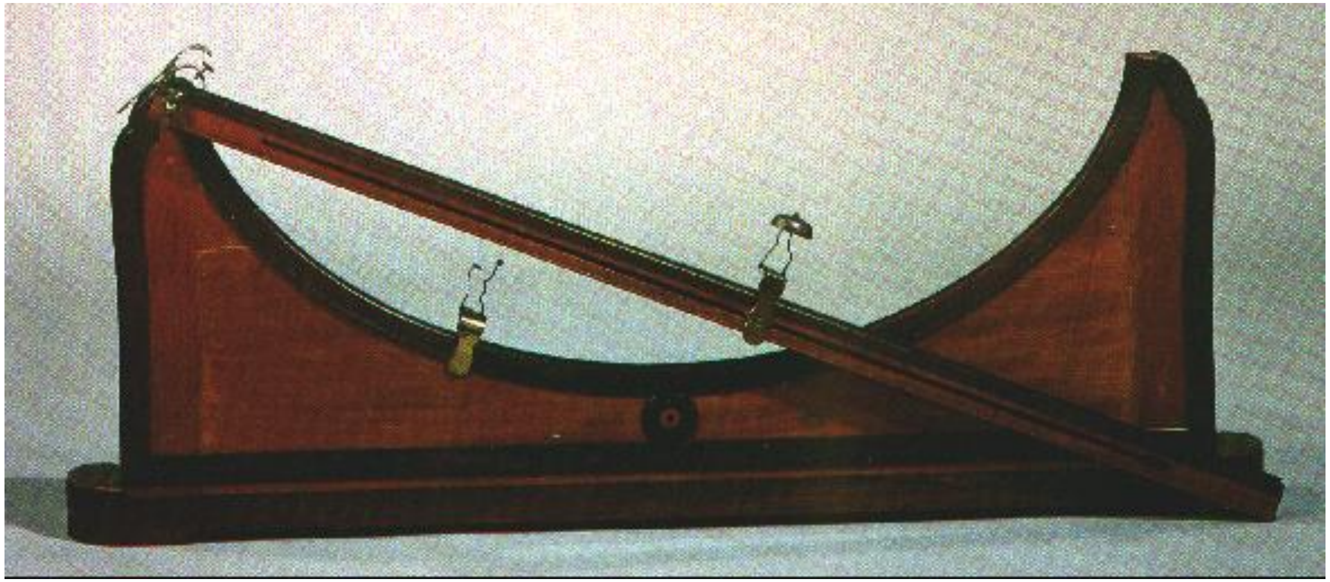
Subject to

$$y(0) = 0 \quad y(x_2) = y_2$$



A machine to demonstrate the brachistochrone

Museu Pombalino de Física da Universidade de Coimbra



If y verifies the EL equation, and L does not explicitly depends on x

$$\frac{d}{dx}\left[L-\dot{y}\frac{\partial L}{\partial \dot{y}}\right]=0$$

or

$$L-\dot{y}\frac{\partial L}{\partial \dot{y}}=const.$$

For the Brachistochrone problem

$$L(y,\dot{y}) = \sqrt{\frac{1+(\dot{y})^2}{y}} \quad \text{from which} \quad \frac{\partial L}{\partial \dot{y}} = \frac{1}{\sqrt{y}} \cdot \frac{1}{\sqrt{1+(\dot{y})^2}} \cdot \frac{1}{2} 2\dot{y}$$

The EL equation reduces thus to

$$L - \dot{y}\frac{\partial L}{\partial \dot{y}} = \frac{1}{c}$$

or

$$\sqrt{\frac{1+(\dot{y})^2}{y}} - \frac{(\dot{y})^2}{\sqrt{\left(1+(\dot{y})^2\right)y}} = \frac{1}{c}$$

$$\sqrt{\frac{1+(\dot{y})^2}{y}} - \frac{(\dot{y})^2}{\sqrt{\left(1+(\dot{y})^2\right)y}} = \frac{1}{c}$$

In order to solve this equation, start by rewriting it.

Multiply the 1st term of the lhs by $1 = \frac{\sqrt{1+(\dot{y})^2}}{\sqrt{1+(\dot{y})^2}}$ to get

$$\frac{1+(\dot{y})^2-(\dot{y})^2}{\sqrt{\left(1+(\dot{y})^2\right)y}} = \frac{1}{c} \quad \text{or} \quad \frac{1}{\sqrt{\left(1+(\dot{y})^2\right)y}} = \frac{1}{c}$$

$$\frac{1}{\sqrt{\left(1+(\dot{y})^2\right)y}} = \frac{1}{c}$$

This equation can be written as

$$y + (\dot{y})^2 y = c^2$$

or

$$\sqrt{\frac{y}{c^2 - y}}\dot{y} = 1$$

The solution of the equation

$$\sqrt{\frac{y}{c^2 - y}}\dot{y} = 1 \qquad (*)$$

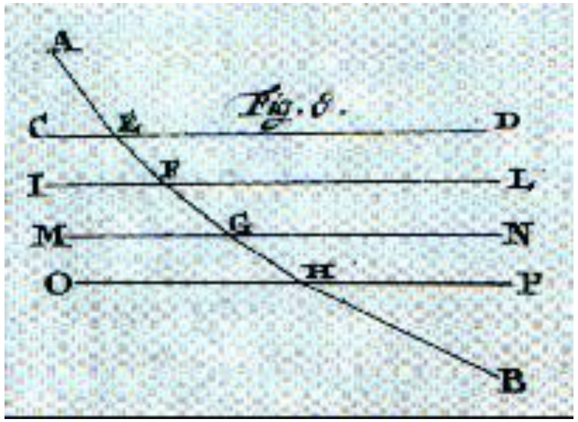
Is given by

$$x = \frac{c^2}{2}(\theta - \sin\theta) + c_1$$

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

These equations define cycloids. Constants c, c_1 and the maximum value of parameter θ are selected such as to meet the boundary constraints.

Historical remark



The equation $\sqrt{\frac{y}{c^2 - y}}\dot{y} = 1$ is found in the original work of Johann Bernouilli. This equation was obtained not from the Euler-Lagrange equation (that was unknown at the time), but instead through an ingenuous argument that uses an analogy with optics and Fermat Principle of Least Time. Bernouilli assumed that the velocity was constant along horizontal stretches and applied Snell's law to compute the angles of incidence and refraction. The sketch above is eprinted from a treaty on Physics of the XVIII that is available at the library of Universidade de Coimbra.

Details of the solution of equation(*)

Make the change of variable $\theta = \theta(x)$ such that $y = \frac{c^2}{2}(1-\cos\theta) = c^2 \sin^2\left(\frac{\theta}{2}\right)$.

In this way, $c^2 - y = c^2\left(1-\sin^2\left(\frac{\theta}{2}\right)\right) = c^2\cos^2\left(\frac{\theta}{2}\right)$ and $\dot{y} = c^2 \sin\frac{\theta}{2}\cos\frac{\theta}{2}\cdot\dot{\theta}$ where $\dot{\theta} := \frac{d\theta}{dx}$

Replacing in $\sqrt{\frac{y}{c^2-y}}\dot{y} = 1$, one gets $\sqrt{\frac{c^2\sin^2\left(\frac{\theta}{2}\right)}{c^2\cos^2\left(\frac{\theta}{2}\right)}}\cdot c^2\sin^2\frac{\theta}{2}\cos^2\frac{\theta}{2}\dot{\theta} = 1$ or,

simplifying $c^2\sin^2\frac{\theta}{2}\dot{\theta} = 1$ or else $\frac{c^2}{2}(1-\cos\theta)\frac{d\theta}{dx} = 1$

Integrating, yields $\int_0^\theta \frac{c^2}{2}(1-\cos\gamma)d\gamma = x - c_1$ from which the expression for $x(\theta)$ is obtained.

Exercise 6

Find the extremal for the following variational problem

$$J(y) = \int_1^2 \frac{y'^2}{x^3} dx$$

$$y(1) = 2, \quad y(2) = 17$$

Hint: Observe that this is special case 1, and hence the EL equation reduces to

$$\frac{\partial L}{\partial y'} = c \text{ with } c \text{ a constant}$$

Solution

$$L = \frac{y'^2}{x^3} \qquad \frac{\partial L}{\partial y'} = \frac{2y'}{x^3}$$

Since this is special case 1, the EL equation is just $\frac{\partial L}{\partial y'} = \frac{2y'}{x^3} = c$ or $y' = \frac{1}{2} c x^3$

$$y(x) - y(0) = \frac{c}{2}\int_{x_0}^{x} \sigma^3 d\sigma \quad y(x) = y(0) + \frac{c}{2}\left(\frac{x^4}{4}\right)\Big|_{x_0}^{x} \quad y(x) = y(0) + \frac{c}{8}(x^4 - x_0^4)$$

Apply the initial condition: $y(x) = 2 + \frac{c}{8}(x^4 - 1)$

Apply the final condition: $17 = 2 + \frac{c}{8} \times 15$ and $c = 8$

Solution:

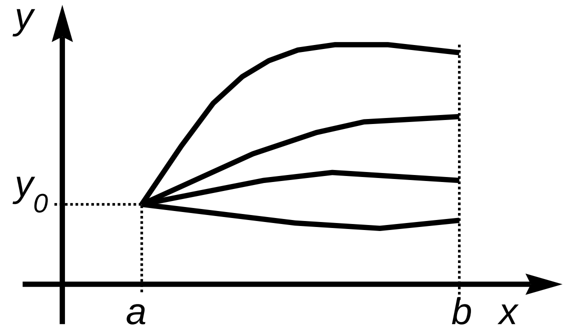
$$y(x) = x^4 + 1$$

Free end-point problems [L2012]. P.42-44, [K1970]

Find the extremals of

$$J(y) = \int_a^b L(x, y, y')dx$$

s. t. $y(a) = y_0$, $y(b)$ free



Free end-point

Admissible disturbances η satisfy

$$\eta(a) = 0, \quad \eta(b) \text{ free}$$

The first variation has now an extra term

$$\delta J|_y(\eta) = \int_a^b \left(L_y - \frac{d}{dx} L_{y'} \right) \eta(x) dx + L_{y'}(b, y(b), y'(b))\eta(b)$$

To impose $\delta J|_y(\eta) = 0$:

$$L_y - \frac{d}{dx} L_{y'} = 0 \qquad L_{y'}\big(b, y(b), y'(b)\big)\eta(b) = 0$$

Since $\eta(b)$ is arbitrary, it must be

$$L_{y'}\big(b, y(b), y'(b)\big) = 0$$

Necessary conditions for a weak minimum with free end point, fixed end time

EL equation:

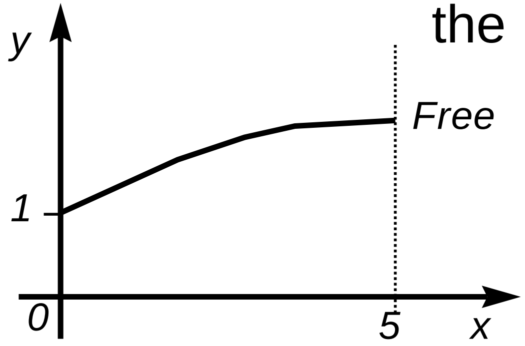
$$L_y - \frac{d}{dx} L_{y'} = 0$$

Condition that replaces the terminal condition:

$$L_{y'}\big(b, y(b), y'(b)\big) = 0$$

Exercise 8

Find the smooth curve of smallest length that connects the point $y(0) = 1$ to the vertical line in the plane (x, y) defined by $x = 5$.



Minimize $J(x) = \int_0^5 \sqrt{1 + y'^2} dx$

s. t. $y(0) = 1$

Hint: $L_y - \frac{d}{dx} L_{y'} = 0$ $L_{y'}\big(b, y(b), y'(b)\big) = 0$

Observe that this is special case 1 ("no y") of the EL equation.

Solution

$$L_y - \frac{d}{dx} L_{y'} = 0 \qquad L = \sqrt{1 + y'^2} \qquad L_y = 0 \qquad L_{y'} = \frac{y'}{\sqrt{1+y'^2}}$$

EL equation becomes $\frac{d}{dx}\left(\frac{y'}{\sqrt{1+y'^2}}\right) = 0$ or $\frac{y'}{\sqrt{1+y'^2}} = \sqrt{c_1}$

$$y'^2 = c_1(1 + y'^2) \quad \rightarrow \quad y' = \sqrt{\frac{c_1}{1-c_1}} = c_2 \quad \rightarrow \quad y(x) = c_2 x + c_3$$

The solution is a straight line! Apply now the initial condition and the condition that replaces the terminal condition to obtain c_2 and c_3

$y(x) = c_2 x + c_3$

Initial condition: $y(0) = c_3 \quad \Rightarrow \quad c_3 = 1$

Condition that replaces the terminal condition

$L_{y'}\big(b, y(b), y'(b)\big) = 0 \quad \rightarrow \quad L_{y'}|_{x=5} = \frac{y'(5)}{\sqrt{1+(y'(5))^2}} = 0 \quad \rightarrow \quad y'(5) = 0$

$y'(x) = c_2 \quad \rightarrow \quad c_2 = 0$

Conclusion: The extremal curve is the horizontal line defined by $y(x) = 1$,

$0 \leq x \leq 5$.

Exercise 10

Consider the following modification of exercise 8:

Show that the curve with shortest length that connects the point $(0, y_0)$ and the vertical line $x = b$ has a tangent at $x = b$ (when the curve touches the vertical line) that is orthogonal to the vertical line.

Minimize $J(x) = \int_0^b \sqrt{1 + y'^2} dx$

s. t. $y(0) = y_0$

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

The modified terminal condition is

$$L_{y'}|_{x=b} = 0$$

$$L_{y'}|_{x=b} = \frac{y'(b)}{\sqrt{1+(y'(b))^2}} = 0 \quad \rightarrow \quad y'(b) = 0$$

Hence, for $x = b$, the tangent to y is horizontal and hence y at $x = b$ is orthogonal to the vertical line.

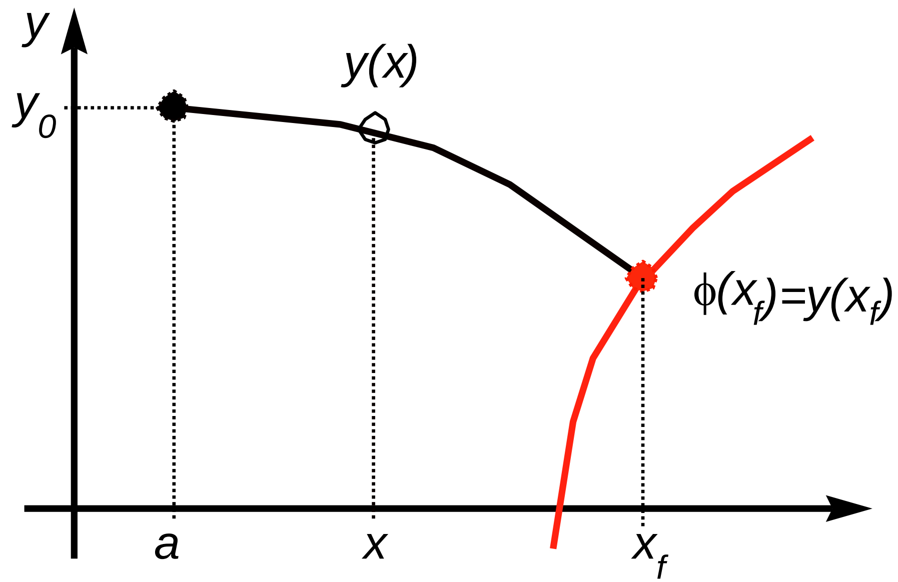
1st-order necessary conditions for free end time

Find a curve $y(x)$ that links the point (a, y_0) in the (x, y) plane with the curve $(x_f, \varphi(x_f))$, with x_f free and φ a C^1 function, $\varphi: \mathbb{R} \rightarrow \mathbb{R}$, and is an extremal of

$$J(y) = \int_a^{x_f} L(x, y, y')dx$$

a fixed

x_f unspecified



Let $y: [a, x_f] \to \mathbb{R}$ be an optimal curve.

Perturbed curve

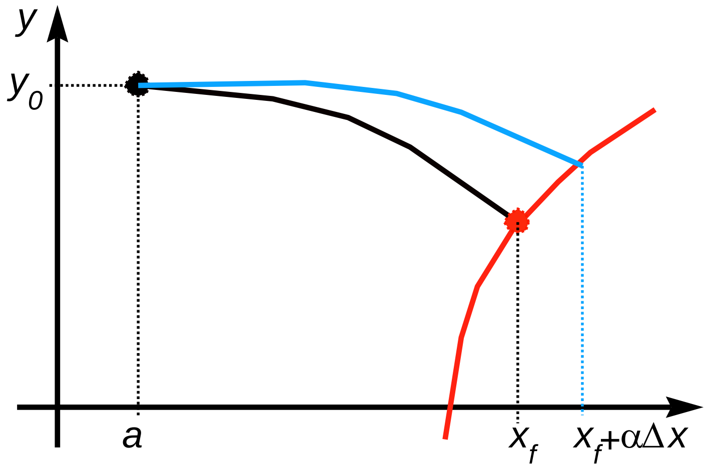
$$y + \alpha\eta$$

Since the terminal point is not fixed, let the terminal point of the perturbed curve be perturbed as well, being

$$x_f + \alpha\Delta x$$

x_f is free, but $y(x_f)$ and x_f are bound by φ.

The same applies to the perturbed curve.



Cost of the perturbed curve

$$J\left(y+\alpha\eta; x_f+\alpha\Delta x\right)=\int_a^{x_f+\alpha\Delta x} L\left(x, y(x)+\alpha\eta(x), y'(x)+\alpha\eta'(x)\right)dx$$

The first variation is the derivative with respect to α for $\alpha=0$.

Leibniz's rule for the differentiation of an integral with respect to a parameter:

$$\frac{d}{d\theta}\left(\int_{a(\theta)}^{b(\theta)} f(x,\theta)dx\right)=\int_{a(\theta)}^{b(\theta)}\frac{\partial f}{\partial\theta}dx+f(b(\theta),\theta)\cdot b'(\theta)-f(a(\theta),\theta)\cdot a'(\theta)$$

Using Leibniz's rule:

$$\delta J|_y(\eta)=\int_a^{x_f}\left(L_y\eta+L_{y'}\eta'\right)dx+L\left(x_f, y\left(x_f\right), y'\left(x_f\right)\right)\Delta x$$

$$\delta J|_y(\eta) = \int_a^{x_f} (L_y \eta + L_{y'} \eta') dx + L\left(x_f, y(x_f), y'(x_f)\right) \Delta x$$

Integrate by parts to eliminate the dependency on η':

$$\delta J|_y(\eta) = \int_a^{x_f} \left(L_y - \frac{d}{dx} L_{y'}\right) \eta dx + L_{y'} \eta|_a^{x_f} + L\left(x_f, y(x_f), y'(x_f)\right) \Delta x$$

Since perturbations η while $\Delta x = 0$ are allowed, the integral is zero and the EL equations holds.

Furthermore, because the initial condition is fixed, $\eta(a) = 0$.

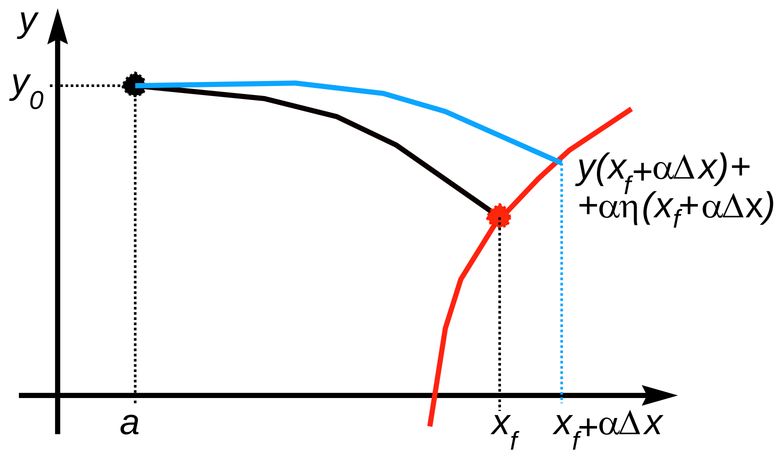
Therefore, we are left with

$$\delta J|_y(\eta) = L_{y'}\left(x_f, y(x_f), y'(x_f)\right) \eta(x_f) + L\left(x_f, y(x_f), y'(x_f)\right) \Delta x = 0$$

$$\delta J|_y(\eta) = L_{y'}\left(x_f, y(x_f), y'(x_f)\right)\eta(x_f) + L\left(x_f, y(x_f), y'(x_f)\right)\Delta x = 0$$

But $\eta(x_f)$ and Δx are related, because the terminal point of the perturbed curve must still be on the curve $y = \varphi(x_f)$:

$$y(x_f + \alpha\Delta x) + \alpha\eta(x_f + \alpha\Delta x) = \varphi(x_f + \alpha\Delta x)$$



$$y\left(x_f + \alpha\Delta x\right) + \alpha\eta\left(x_f + \alpha\Delta x\right) = \varphi\left(x_f + \alpha\Delta x\right)$$

Differentiate with respect to α

$$y'\left(x_f + \alpha\Delta x\right)\Delta x + \eta\left(x_f + \alpha\Delta x\right) + \alpha\eta'\left(x_f + \alpha\Delta x\right)\Delta x = \varphi'\left(x_f + \alpha\Delta x\right)\Delta x$$

and set $\alpha = 0$

$$y'\left(x_f\right)\Delta x + \eta\left(x_f\right) = \varphi'\left(x_f\right)\Delta x$$

$$\eta\left(x_f\right) = \left(\varphi'\left(x_f\right) - y'\left(x_f\right)\right)\Delta x$$

$$\eta(x_f) = \left(\varphi'(x_f) - y'(x_f)\right)\Delta x$$

Plug this expression in the expression for $\delta J|_y(\eta)$:

$$\delta J|_y(\eta) = L_{y'}\left(x_f, y(x_f), y'(x_f)\right)\eta(x_f) + L\left(x_f, y(x_f), y'(x_f)\right)\Delta x = 0$$

and get

$$\left[L_{y'}\left(x_f, y(x_f), y'(x_f)\right)\left(\varphi'(x_f) - y'(x_f)\right) + L\left(x_f, y(x_f), y'(x_f)\right)\right]\Delta x = 0$$

Since the equality to 0 must hold for any Δx, we conclude that the so-called transversality condition must hold:

$$L\left(x_f, y(x_f), y'(x_f)\right) + L_{y'}\left(x_f, y(x_f), y'(x_f)\right)\left(\varphi'(x_f) - y'(x_f)\right) = 0$$

1st-order necessary conditions for extremal in the free final "time" case

1) EL equation

$$L_y = \frac{d}{dx}\left(L_{y'}\right)$$

2) Initial condition

$$y(a) = y_1$$

3) Transversality condition

$$L|_{x_f, y_f, y'_f} + L_{y'}|_{x_f, y_f, y'_f}\left(\varphi'(x_f) - y'(x_f)\right) = 0$$

4) Condition that defines the arrival manifold

$$y(x_f) = \varphi(x_f)$$

Exercise 11

Write the conditions that define the curve that links the point (a, y_1) with the curve $(x_f, \varphi(x_f))$, with x_f free, and φ a C^1 function $\varphi: \mathbb{R} \to \mathbb{R}$, and such that the curve $(x, y), a \leq x \leq x_f$, is an extremal to the functional defined by its length,

$$J(y) = \int_a^{x_f} \sqrt{1 + y'(x)^2} dx$$

Hint: Write the EL equation and solve it observing that it is special case 1 to obtain the shape of the extremal. Then, apply the initial condition, the transversality condition and use the shape of φ to obtain conditions for the constants that define the extremal curve y and x_f.

Solution

$$L = \sqrt{1 + y'(x)^2} \qquad L_y = 0 \qquad L_{y\prime} = \frac{y\prime}{\sqrt{1+y\prime(x)^2}} = 0$$

EL equation: $\frac{d}{dx}\frac{y\prime}{\sqrt{1+y\prime(x)^2}} = 0$ hence $\frac{y\prime}{\sqrt{1+y\prime(x)^2}} = c$ with c a constant.

$$y' = \frac{c}{\sqrt{1 - c^2}} \coloneqq c_1$$

Hence, the extremal is a straight line, defined by

$$y(x) = c_1 x + c_2$$

c_1, c_2, and x_f are obtained from the initial condition, the transversality condition and the shape of the curve $(x_f, \varphi(x_f))$.

From the initial condition: $y_1 = c_1 a + c_2$ Furthermore: $y' = c_1$

From the transversality condition

$$L|_{x_f, y_f, y'_f} + L_{y'}|_{x_f, y_f, y'_f} \left(\varphi'(x_f) - y'(x_f) \right) = 0$$

$$\sqrt{1 + c_1^2} + \frac{c_1}{\sqrt{1 + c_1^2}} \left(\varphi'(x_f) - c_1 \right) = 0$$

Multiply by $\sqrt{1 + c_1^2}$ and simplify to get $1 + c_1 \varphi'(x_f) = 0$

From the terminal condition $y(x_f) = \varphi(x_f)$ or $c_1 x_f + c_2 = \varphi(x_f)$.

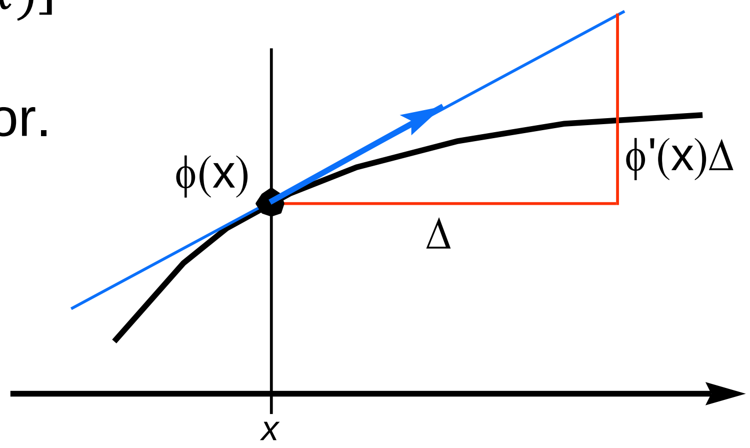
Conclusion, c_1, c_2, and x_f satisfy

$$\begin{cases} y_1 = c_1 a + c_2 \\ 1 + c_1 \varphi'(x_f) = 0 \\ c_1 x_f + c_2 = \varphi(x_f) \end{cases}$$

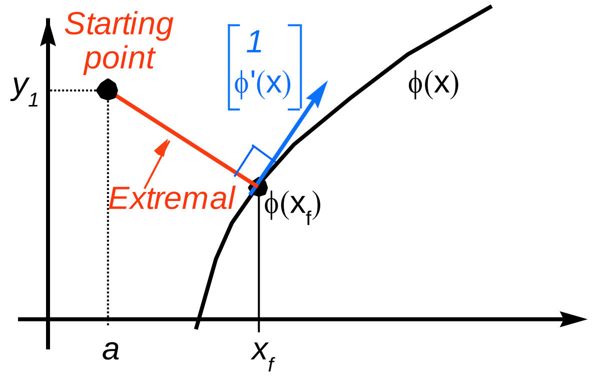
Tangent vector to a curve

Any vector $\begin{bmatrix} \Delta \\ \varphi'(x)\Delta \end{bmatrix}$ is tangent to the curve $\begin{bmatrix} x \\ \varphi(x) \end{bmatrix}$ at x.

In particular, for $\Delta = 1$, $\begin{bmatrix} 1 \\ \varphi'(x) \end{bmatrix}$ is a tangent vector.



Geometrical interpretation of the transversality condition for the length functional



In exercise 10 it is shown that

$$1 + c_1 \varphi'(x_f) = 0$$

But in this case $y'(x_f) = c_1$ and thus

$$1 + y'(x_f)\varphi'(x_f) = \begin{bmatrix} 1 & y'(x_f) \end{bmatrix} \begin{bmatrix} 1 \\ \varphi'(x_f) \end{bmatrix} = 0$$

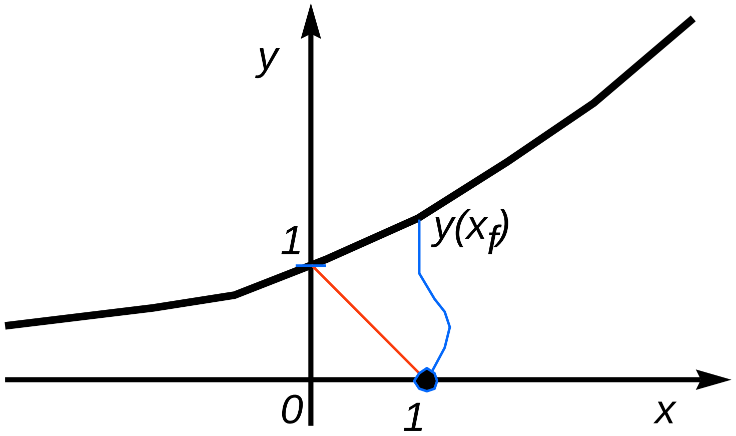
Thus, the tangent vectors to the curve $y(x)$ and $\varphi(x)$are orthogonal at x_f.

Exercise 12

Find the shortest line between the point $(1{,}0)$ and the curve $\varphi(x) = e^{x}$.

Minimize

$$J = \int_{1}^{x_f} \sqrt{1 + y'^2}\, dx$$



Compute the distance between the point $(1{,}0)$ and the curve $\varphi(x) = e^{x}$.

Hints:

$$L_y = \frac{d}{dx}\left(L_{y'}\right), \qquad L|_{x_f, y_f, y'_f} + L_{y'}|_{x_f, y_f, y'_f}\left(\varphi'(x_f) - y'(x_f)\right) = 0$$

Solution

$L = \sqrt{1 + y'^2} \quad L_y = 0 \quad L_y = \frac{d}{dx}(L_{y'})$ becomes $\frac{d}{dx}(L_{y'}) = 0$ hence $L_{y'} = c$

$$\frac{y'}{\sqrt{1+y'^2}} = c \quad \Rightarrow \quad y' = \frac{c}{\sqrt{1-c^2}} := c_1$$

Hence, the extremal is a straight line, defined by

$$y(x) = c_1 x + c_2$$

Boundary conditions

$$y(1) = 0 \ \Rightarrow \ c_1 + c_2 = 0$$

Transversality condition

$$L|_{x_f,y_f,y'_f} + L_{y'}|_{x_f,y_f,y'_f}\left(\varphi'(x_f) - y'(x_f)\right) = 0$$

$$\varphi'(x_f) = e^{x_f} \qquad y' = c_1$$

$$L|_{x_f,y_f} = \sqrt{1+c_1^2} \qquad L_{y'}|_{x_f,y_f} = \frac{c_1}{\sqrt{1+c_1^2}}$$

The transversality condition becomes thus

$$\sqrt{1+c_1^2} + \frac{c_1}{\sqrt{1+c_1^2}}(e^{x_f} - c_1)$$

Multiply by $\sqrt{1+c_1^2}$ and simplify to get $1 + c_1 e^{x_f} = 0$.

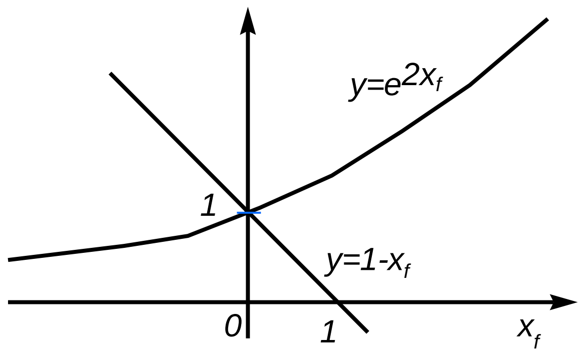
Final constraint: $y_f = e^{x_f} \quad \rightarrow \quad c_1 x_f + c_2 = e^{x_f}$

The conditions to satisfy in order to find c_1, c_2, and x_f are thus

$$\begin{cases} c_1 + c_2 = 0 \\ 1 + c_1 e^{x_f} = 0 \\ c_1 x_f + c_2 = e^{x_f} \end{cases}$$

$$\begin{cases} c_1 x_f - c_1 = e^{x_f} \\ c_1 = -\frac{1}{e^{x_f}} \end{cases} \quad \rightarrow \quad -\frac{x_f}{e^{x_f}} + \frac{1}{e^{x_f}} = e^{x_f} \quad \rightarrow \quad 1 - x_f = (e^{x_f})^2 \quad \rightarrow \quad x_f = 0$$

$c_1 = -1, \quad c_2 = 1$



The extremal solution is $y(x) = 1 - x, \quad 0 \leq x \leq 1$

The corresponding extremal “distance is

$$J^* = \int_1^0 \sqrt{2}\, dx = -\sqrt{2}$$

Hamilton's canonical equations

Hamilton's formalism (Hamilton 1835)

An alternative formulation of the results of Euler and Lagrange.

Of great significance in the context of optimal control.

[L2012] pp. 44-46.

William Rowan Hamilton (1805-1865) was born in Dublin, Ireland, and made important contributions to classical mechanics, optics, and algebra. His best known contribution to mathematical physics is the reformulation of Newtonian mechanics, now called Hamiltonian mechanics. This work has proven central to the modern study of classical field theories such as electromagnetism, as well as to the development of quantum mechanics and optimal control. In pure mathematics, he is best known as the inventor of quaternions.

Important new concepts:

Momentum

$$p \coloneqq L_{y'}(x, y, y')$$

Hamiltonian

$$H(x, y, y', p) \coloneqq p \cdot y' - L(x, y, y')$$

A general function of 4 variables.

Also a function of x alone when evaluated along a curve.

Canonical variables

y and p

Hamilton's canonical equations

y an extremal (satisfies the EL equation).

Since

$$H = py' - L,$$

it follows that (equation for y):

$$\frac{dy}{dx} = \frac{\partial H}{\partial p}$$

Furthermore, for p:

$$\frac{dp}{dx} = \frac{d}{dx}\left(L_{y'}\right) = L_y = -\frac{\partial H}{\partial y}$$

where the 1st equality follows by definition of p and the 2nd by the EL eq.

Hamilton's canonical equations

$$\frac{dy}{dx} = \frac{\partial H}{\partial p}, \qquad \frac{dp}{dx} = -\frac{\partial H}{\partial y}$$

In general, p and H need not be constant along extremals.

However, there are situations in which they are constant. The following exercise clarifies this point.

Exercise 17 [L2012], p. 45, exercise 2.7

Show that

a) When L does not depend explicitly on y, p is constant along extremals.

b) When L does not depend explicitly on x, H is constant along extremals.

Make the proof in two different ways

1. Using the EL equation
2. Using Hamilton equations

$$\frac{dy}{dx} = \frac{\partial H}{\partial p}, \qquad \frac{dp}{dx} = -\frac{\partial H}{\partial y}$$

In both cases, use the definition of p, and H:

$$p := L_{y'}(x, y, y') \qquad H = py' - L,$$

Solution

1. EL: $L_y = \frac{d}{dx} L_{y'}$

a) "No y". The EL eq. reduces to $\frac{d}{dx} L_{y'} = 0 \;\Rightarrow\; L_{y'} = c$

By definition of p, $p \coloneqq L_{y'}$, it follows that $p = c$ with c a constant.

b) "No x". By the chain rule

Then, $L_{y'x} = 0$, and the EL becomes

$$L_y = L_{y'y} y' + L_{y'y'} y''$$

Multiply both sides by y'

$$L_{y'y}(y')^2 + L_{y'y'} y'' y' - L_y y' = 0$$

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

Therefore

$$L_{y'} y' - L = c \quad \text{or} \quad p y' - L = c \quad \text{or} \quad H = c.$$

2.Proof using Hamilton equations

a) $H = py' - L \quad \frac{\partial H}{\partial y} = -\frac{\partial L}{\partial y} = 0$

Hence, since $\frac{dp}{dx} = -\frac{\partial H}{\partial y} \quad \rightarrow \quad \frac{dp}{dx} = 0$ and p is constant.

2.Proof using Hamilton equations

b) $\frac{dH}{dx} = H_y y' + H_{y'} y'' + H_p p'$

Use The definition of $H := py' - L$ and $p \coloneqq L_{y'}$ and the Hamilton equation

$\frac{dy}{dx} = \frac{\partial H}{\partial p} \qquad \frac{dp}{dx} = -\frac{\partial H}{\partial y}$ to observe that

$$H_{y'} = \frac{\partial}{\partial y'}(py' - L) = p - L_{y'}$$

$$\frac{dH}{dx} = H_y y' + H_{y'} y'' + H_p p'$$

$$\frac{dH}{dx} = -p'y' + \left(p - L_{y'}\right)y'' + y'p' = \left(p - L_{y'}\right)y'' = (p - p)y'' = 0.$$

The Hamiltonian as a stationary point [L2012] p.46

$$H(x, y, y', p) = p \cdot y' - L(x, y, y')$$

Arbitrary $x \in [a, b]$; y the corresponding $y(x)$ of the optimal curve

$p = p(x) = L_{y'}(x, y, y')$ the corresponding value of the momentum

Keep y' as a free variable denote it as z to define the function

$$H^*(z) := p \cdot z - L(x, y(x), z)$$

For $z = y'(x)$

$$\frac{dH^*}{dz}|_{z=y'(x)} = L_{y'}(x, y(x), y'(x)) - L_{y'}\big(x, y(x), y'(x)\big) = 0$$

$$\frac{dH^*}{dz}|_{z=y'(x)} = 0$$

Conclusion: Along an extremal curve, the Hamiltonian is stationary with respect to y'.

Remark: Actually, it can be shown that H is maximum with respect to y', even when H is not differentiable, or when y' takes a value in a set with a boundary and $H_{y'} \neq 0$ on this boundary.

These issues will become clearer when considering the maximum principle for optimal control.

Exercise 2 (Cont.)

Consider the problem

$$\text{Minimize} \quad J(y) = \int_0^{\pi/2} \left[y'^2(x) - y^2(x) \right] dx$$

$$\text{s. t.} \qquad y(0) = 0, \quad y\left(\frac{\pi}{2}\right) = 1.$$

In Exercise 2 we found that the extremal is $y(x) = \sin(x)$. Compute:

a) $p(x) = L_{y'}$

b) $H^*(z) \coloneqq p \cdot z - L(x, y(x), z)$

Make a sketch of this function for various values of x.

c) $H^*(z)|_{z=p'(x)}$. Is this in accordance with what you expect?

Solution $L = {y'}^2(x) - y^2(x)$

a) The extremal is $y(x) = \sin(x)$. Hence $y'^{(x)} = \cos(x)$

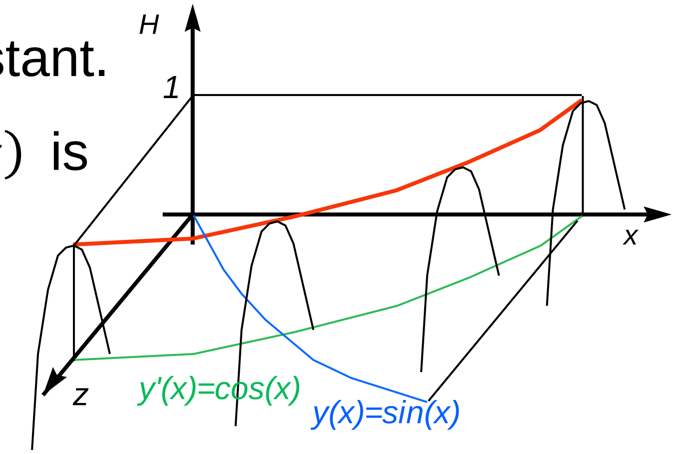
$p(x) = L_{y'} = 2y' = 2\cos\,(x)$

b) $H^*(z) := p \cdot z - L(x, y(x), z) = pz - z^2 + y^2 = 2\cos(x)z - z^2 + sin^2(x)$

c) For $z = y' = \cos\,(x)$, $H^* = 2\cos^2(x) - \cos^2(x) + 1 - \cos^2(x) = 1$, and

hence, along an optimal trajectory, H is constant.

Furthermore, for $z = y'$ (the optimal p), $H^*(z)$ is maximum with respect to z.



Legendre transform

[L2012] pp. 46-48

[H2014] pp. 93-95, 97-100.

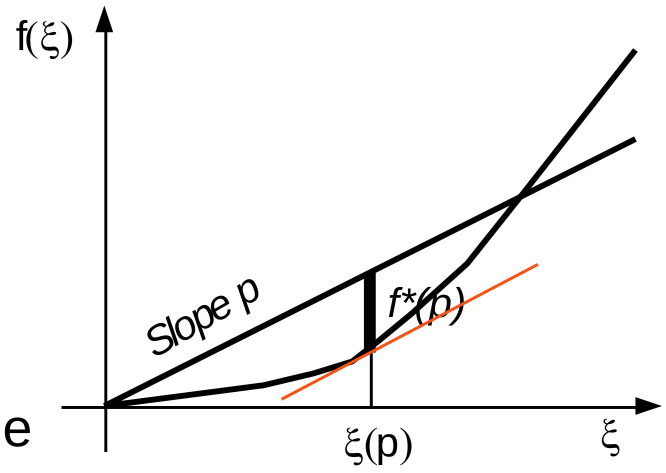
[H2014] P. Hamill (2014). *A student's guide to Lagrangians and Hamiltonians*. Cambridge University Press.

A function $f: \mathbb{R} \to \mathbb{R}$ with argument ξ

$$\downarrow$$

A function f^* with argument $p \in \mathbb{R}$

Defining the Legendre Transform



For a given p:

Draw a line of slope p through the origin.

$$\xi(p) \coloneqq \arg\max_{\xi}\{p\xi - f(\xi)\}$$

$\xi(p)$ corresponds to the point at which the vertical line distance from f to this line is maximized:

$$f^*(p) \coloneqq p\xi(p) - f\big(\xi(p)\big) = \max_{\xi}\{p\xi - f(\xi)\}$$

Furthermore

$$f^*(p) + f\big(\xi(p)\big) = p\xi$$

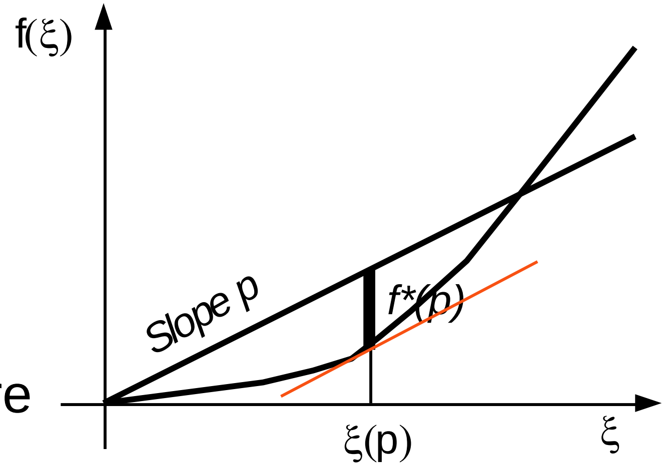
$$f^*(p) = \max_{\xi}\{p\xi - f(\xi)\}$$

When f is differentiable

$$p - f'(\xi) = 0$$

Geometrically: the tangent line must be parallel to the straight line of slope p.

For convex functions, both definitions are equivalent.



Properties of the Legendre Transform [L2012], p. 47

f^* is a convex function even if f is not convex.

The Legendre transform is involutive: If f is convex, $f^{**} = f$.

The Legendre transform and the Hamiltonian

The Hamiltonian H can be obtained by applying the Legendre transform to the Lagrangian L

$$L \rightarrow H$$

Consider $L(x, y, y')$ as a function of $\xi = y'$

The relation

$$p - f'\big(\xi(p)\big) = 0$$

Becomes

$$p - L_{y'}\big(x, y, y'(p)\big) = 0$$

that corresponds to the definition of the momentum p.

Consider now the general definition of the Legendre transform

$$f^*(p) = p\,\xi(p) - f(\xi(p))$$

Apply it to L, to yield

$$L^*(x,y,p) = py'(p) - L\big(x,y,y'(p)\big)$$

that is the definition of the Hamiltonian.

Variational problems with constraints

Integral constraints

[L2012] pp. 52-55

Heuristic discussion

$$C(y) = \int_a^b M(x, y, y')dx = C_0$$

y an extremal

Perturbed curves $y + \alpha\eta$

To be admissible, η must preserve the constraint

$$C(y + \alpha\eta) = C_0 \quad \forall_\alpha \text{ close to } 0$$

$$\delta C|_y(\eta) = 0$$

Consequence (1st-order necessary condition for constrained optimality)

$\exists \lambda^*$ (Lagrange multiplier) and λ_0 such that

$$\lambda_0 \left(L_y - \frac{d}{dx} L_{y'} \right) + \lambda^* \left(M_y - \frac{d}{dx} M_{y'} \right) = 0$$

Rearranging terms

$$(\lambda_0 L + \lambda^* M)_y = \frac{d}{dx} (\lambda_0 L + \lambda^* M)_{y'}$$

That amount to say that the Euler Lagrange equation holds for the augmented Lagrangian

$$\lambda_0 L + \lambda^* M$$

λ_0 is usually 1 and is called the abnormal multiplier.

From constrained to unconstrained optimization

Constrained problem:

$$\min_y J(y) = \int_a^b L dx$$

$$\text{s.t. } C(y) = \int_a^b M dx = C_0$$

$$\downarrow$$

Unconstrained problem

$$\min_y \max_\lambda \int_a^b L dx + \lambda \left(\int_a^b M dx - C_0 \right)$$

For curves that satisfy the constraint, the values of the two functionals coincide.

Procedure to solve problems with integral constraints

1. Solve the Euler-Lagrange equation considering the augmented Lagrangian

$$\lambda_0 L + \lambda M$$

2. Obtain a solution $y(x, \lambda)$ that depends on λ.

3. Plug the solution $y(x, \lambda)$ in the integral constraint to obtain an equation on λ, and solve it to obtain λ.

4. Eliminate λ from $y(x, \lambda)$.

Exercise 20

Minimize $\int_0^1 y'^2 dx$ s. t. $y(0) = 2$, $y(1) = 4$ and the integral constraint

$$\int_0^1 y dx = 1$$

Hints:

$$(L + \lambda^* M)_y = \frac{d}{dx}(\lambda_0 L + \lambda^* M)_{y'}$$

Solve EL with constraint as a function of λ.

Use the integral constraint to find λ^*.

Solution

$L = y'^2 \qquad M = y$

$\mathcal{L} = L + \lambda M = y'^2 + \lambda y \qquad \mathcal{L}_y = \frac{d}{dx}\mathcal{L}_{y'}$

$\mathcal{L}_y = \lambda \qquad \mathcal{L}_{y'} = 2y'$

The EL for $\mathcal{L}$ becomes $\lambda = 2y'' \quad \rightarrow \quad y(x) = \frac{\lambda}{4}x^2 + c_1 x + c_2$

$y(0) = 2 \quad \rightarrow \quad c_2 = 2 \qquad y(1) = 4 \quad \rightarrow \quad c_1 = 2 - \frac{\lambda}{4}$

$\rightarrow \quad y(x) = \frac{\lambda}{4}x^2 + \left(2 - \frac{\lambda}{4}\right)x + 2$

Use now the integral constraint to find λ:

$$\int_0^1 \left(\frac{\lambda}{4}x^2 + \left(2 - \frac{\lambda}{4}\right)x + 2\right)dx = 1$$

$$\int_0^1 \left(\frac{\lambda}{4}x^2 + \left(2 - \frac{\lambda}{4}\right)x + 2\right) dx = 1$$

$$\left[\frac{\lambda}{12}x^3 + \left(1 - \frac{\lambda}{8}\right)x^2 + 2x\right]_0^1 = 1 \quad \rightarrow \quad \frac{\lambda}{12} + \left(1 - \frac{\lambda}{8}\right) + 2 = 1 \quad \rightarrow \quad \lambda^* = 48$$

$$y(x) = \frac{\lambda}{4}x^2 + \left(2 - \frac{\lambda}{4}\right)x + 2$$

$$y(x) = 12x^2 - 10x + 2$$

Exercise 22

Find the extremals of $J(y)=\int_0^\pi y'^2 dx$ s. t. $y(0)=y(\pi)=0$ and

$$C(y)=\int_0^\pi y^2 dx=\frac{\pi}{2}$$

Show that there is an infinite set of extremals. Evaluate the functional on a typical extremal.

Hints:

1. Look for solutions of the EL equation with $\lambda<0$.
2. Useful trigonometric formulas

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

Solution

$L = y'^2 \qquad M = y^2$

$\mathcal{L} = L + \lambda M = y'^2 + \lambda y^2 \qquad \mathcal{L}_y = \frac{d}{dx}\mathcal{L}_{y'}$

$\mathcal{L}_y = 2\lambda y \qquad \mathcal{L}_{y'} = 2y'$

The EL becomes $\lambda y = y''$

Assume $y(x) = e^{\gamma x}$

The EL yields the characteristic equation: $\gamma^2 = \lambda \quad \rightarrow \quad \gamma = \pm j\sqrt{-\lambda}$

$$y(x) = k_1 e^{j\sqrt{-\lambda}x} + k_2 e^{-j\sqrt{-\lambda}x}$$

$$y(x) = k_1\left(\cos\left(\sqrt{-\lambda}x\right) + jsin\left(\sqrt{-\lambda}x\right)\right) + k_2\left(\cos\left(\sqrt{-\lambda}x\right) - jsin\left(\sqrt{-\lambda}x\right)\right)$$

$$y(x) = k_1\left(\cos\left(\sqrt{-\lambda}x\right) + jsin\left(\sqrt{-\lambda}x\right)\right) + k_2\left(\cos\left(\sqrt{-\lambda}x\right) - jsin\left(\sqrt{-\lambda}x\right)\right)$$

Boundary conditions

$$y(o) = 0 \quad \rightarrow \quad k_1 + k_2 = 0 \quad \rightarrow \quad k_2 = -k_1 \quad \rightarrow \quad y(x) = 2jk_1\, sin\left(\sqrt{-\lambda}x\right)$$

Consider now $y(\pi) = 0$

We seek a solution for $k_1 \neq 0$. Thus, $\sqrt{-\lambda} = n, \quad n \in \mathbb{Z}$

$$y(x) = 2jk_1\, sin(nx)$$

Use now the integral constraint to compute k_1:

$$\int_0^\pi y^2 dx = \frac{\pi}{2} \quad \rightarrow \quad -4k_1^2 \int_0^\pi sin^2(nx)dx = \frac{\pi}{2}$$

$$k_1^2 = -\frac{\pi}{8} \cdot \frac{1}{\int_0^\pi sin^2(nx)dx}$$

$$\int_0^{\pi} sin^2(nx)dx = \frac{1}{2}\int_0^{\pi}[1-\cos(2nx)]dx = \frac{1}{2}\left(\pi - \frac{1}{2n}\sin(2n\pi)\right) = \frac{\pi}{2}$$

$$k_1^2 = -\frac{\pi}{8}\cdot\frac{1}{\int_0^{\pi} sin^2(nx)dx} = \frac{-\frac{\pi}{8}}{\frac{\pi}{2}} = -\frac{1}{4} \qquad \rightarrow \qquad k_1 = \pm\frac{1}{2}j$$

Therefore, the solution for the extremals is

$$y^*(x) = \sin(nx)$$

Value of the functional:

$$J(y) = \int_0^{\pi} y'^2 dx = \int_0^{\pi} n^2 cos^2(nx)dx = \frac{n^2}{2}\int_0^{\pi}(1+\cos(2nx))dx =$$

$$= \frac{n^2}{2}\left(\pi + \frac{1}{2n}\sin(2nx)|_0^{\pi}\right) = \frac{n^2\pi}{2} \qquad J(y) = \frac{n^2\pi}{2}$$

Exercise 24 (Dido's isoperimetric problem) [L2012] p. 55. Ex. 2.10 a))

Let $y: [a, b] \subset \mathbb{R} \to \mathbb{R}$ be a C^1 function with argument x.

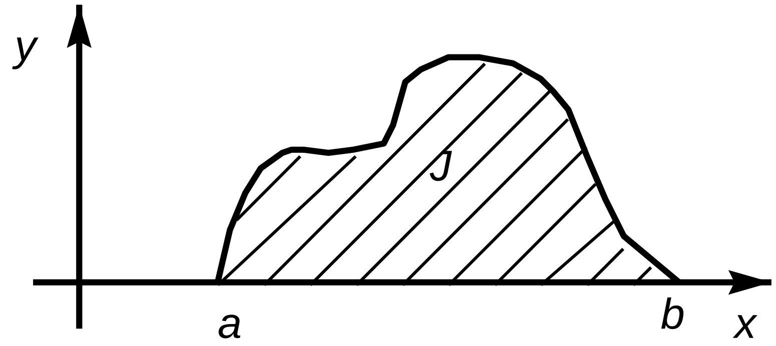
Show that the function that solves the variational problem

Minimize $J = \int_a^b y(x)dx$

Subject to

$$\int_a^b \sqrt{1 + (y'(x))^2}dx = C_0,$$

with $C_0 \in \mathbb{R}^+$ a constant, is an arc of circle.



Hints:

1) After writing the EL equation with constraint, integrated both sides with respect to x.

2) Solve the resulting equation with respect to y.

3) Show that the primitive (indefinite integral) of

$$\frac{x+c}{\sqrt{\lambda^2-(x+c)^2}} \qquad \lambda,\, c \text{ constants}$$

is

$$-\sqrt{\lambda^2-(x+c)^2}+d \qquad d \text{ a constant}$$

Solution (Dido's isoperimetric problem)

$$\mathcal{L}(y,\lambda) = y + \lambda\sqrt{1 + (y'(x))^2} \qquad \frac{\partial \mathcal{L}}{\partial y} = 1 \qquad \frac{\partial \mathcal{L}}{\partial y'} = \lambda \frac{1}{2} \cdot \frac{2y'}{\sqrt{1+(y'(x))^2}}$$

EL equation

$$\frac{\partial \mathcal{L}}{\partial y} = \frac{d}{dx} \cdot \frac{\partial \mathcal{L}}{\partial y'} \qquad \rightarrow \qquad 1 = \frac{d}{dx}\left(\lambda \frac{y'}{\sqrt{1+(y'(x))^2}}\right)$$

Integrate both sides with respect to y:

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

c a constant

Solve with respect to y':

$$\frac{\lambda^2 (y')^2}{1+(y'(x))^2} = (x+c)^2 \quad \rightarrow \quad \lambda^2 (y')^2 = (x+c)^2 (1 + (y'(x))^2)$$

$$(\lambda^2 - (x+c)^2)(y'(x))^2 = (x+c)^2$$

$$(\lambda^2 - (x+c)^2)(y'(x))^2 = (x+c)^2$$

$$y' = \pm \frac{x+c}{\sqrt{\lambda^2 - (x+c)^2}} \qquad (*)$$

The primitive is

$$y = \pm\sqrt{\lambda^2 - (x+c)^2} + d \qquad (**)$$

This expression for the primitive of (*) can be readily checked by differentiating (**) with respect to x to obtain (*).

Eq. (**) can be written as

$$(y-d)^2 + (x+c)^2 = \lambda^2$$

That shows that the solution of the EL equation is a segment of circle with center at $(-c, d)$ and radius λ.

Exercise 25 (The catenary)

Consider the problem of finding $y \in C^1([a,b] \to \mathbb{R})$ such that $y(a) = y(b) = y_0$ and that

Minimize

$$J = \int_a^b y(x)\sqrt{1 + (y'(x))^2}\,dx$$

Subject to

$$\int_a^b \sqrt{1 + (y'(x))^2}\,dx = C_0,$$

with $C_0 \in \mathbb{R}^+$ a constant.

Show that the solution is given by $y(x) = cosh\left(\frac{x}{c}\right)$, $c > 0$, modulo a translation along x.

Hints

Use the special case of the EL equation in which the Lagrangian does not depend on x, that amounts to state that the Hamiltonian is constant, *i.e.*

$$\frac{\partial \mathcal{L}}{\partial y'} \cdot y' - \mathcal{L} = c$$

with $\mathcal{L}$ the lagrangian augmented with the Lagrange multiplier and c a constant. Solve this condition with respect to $y' = \frac{dy}{dx}$.

Use the fact that

$$\int \frac{1}{\sqrt{(y+\lambda)^2}} dy = cosh^{-1}\left(\frac{y+\lambda}{c}\right)$$

Solution

$$\frac{\partial \mathcal{L}}{\partial y'} \cdot y' - \mathcal{L} = c$$

$$\mathcal{L} = y(x)\sqrt{1 + \left(y'(x)\right)^2} + \lambda\sqrt{1 + (y'(x))^2}$$

$$\frac{\partial \mathcal{L}}{\partial y'} = \frac{1}{2} \cdot \frac{2yy'}{\sqrt{1 + (y')^2}} + \frac{1}{2} \cdot \frac{2\lambda y'}{\sqrt{1 + (y')^2}}$$

The EL becomes

$$(y + \lambda)\frac{(y')^2}{\sqrt{1 + (y')^2}} - (y + \lambda)\sqrt{1 + (y')^2} = c$$

$$(y+\lambda)\frac{(y')^2}{\sqrt{1+(y')^2}}-(y+\lambda)\sqrt{1+(y')^2}=c$$

$$(y+\lambda)\frac{(y')^2-\left(1+(y')^2\right)}{\sqrt{1+(y')^2}}=c \qquad \rightarrow \qquad (y+\lambda)\frac{1}{\sqrt{1+(y')^2}}=c$$

$$(y+\lambda)=c\sqrt{1+(y')^2} \qquad (y')^2=\frac{(y+\lambda)^2}{c^2}-1 \qquad \frac{dy}{dx}=\frac{\sqrt{(y+\lambda)^2-c^2}}{c}$$

$$\int\frac{1}{\sqrt{(y+\lambda)^2-c^2}}dy=\int\frac{1}{c}dx$$

$$cosh^{-1}\left(\frac{y+\lambda}{c}\right)=\frac{1}{c}(x+d)$$ d a constant

$$y=c\cdot cosh\left(\frac{x+d}{c}\right)-\lambda$$

<u>Exercise 26</u> (The Principle of Maximum Entropy)

Consider a gas-filled box in the idealized situation in which the box is unidimensional and lays along the x coordinate, between $x = 0$ and $x = L$.

Let $p(x)$ be the probability density function of the number of molecules of gas. That is to say, the number of molecules between x and $x + dx$, with dx small is approximately $p(x)dx$ multiplied by the total number of molecules inside the box. Clearly, $p(x) = 0$ for $x < 0$ and $x > L$. *What is $p(x), 0 \leq x \leq L$ in the situation of maximal disorder*?

According to the Principle of Maximum Entropy, the answer is given by the solution of the following variational problem.

Solve the following variational problem

$$\min_p J(p) = \int_{-\infty}^{\infty} p(x)\log(p(x))\,dx$$

Subject to $\int_{-\infty}^{\infty} p(x)dx = 1$

Solution

$$\mathcal{L}(p,\lambda) = p\log p + \lambda p$$

$$\mathcal{L}_p = \frac{d}{dx}\mathcal{L}_{p'} \qquad \mathcal{L}_{p'} = 0 \quad \rightarrow \quad \mathcal{L}_p = 0$$

$$\mathcal{L}_p = \log p + p\cdot\frac{1}{p} + \lambda = \log p + 1 + \lambda = 0 \quad \rightarrow \quad p(x) = c \qquad c \text{ a constant}$$

From the constraint, it follows that

$$p(x) = \frac{1}{L} \qquad\qquad 0 \le x \le L$$

2nd order conditions

[L2012] pp. 26-36. [B2004] pp. 221-253, Chap. 10.

Analysis based on the 2nd order expansion

$$J(y+\alpha\eta) = J(y) + \delta J|_y(\eta)\alpha + \delta^2 J|_y(\eta)\alpha^2 + o(\alpha^2)$$

that defines the quadratic form $\delta^2 J|_y(\eta)$ called the second variation.

Issues:

- 2nd order necessary condition: $\delta^2 J|_y(\eta) \geq 0$
- Sufficient conditions for a weak minimum. It is not enough to ask

$$\delta^2 J|_y(\eta) > 0$$

This condition must be strengthened to ensure that the 2nd order term dominates the higher order terms $o(\alpha^2)$.

Legendre's necessary condition for a weak minimum

[L2012] p. 59-62

<u>Compute</u> $\delta^2 J|_y(\eta)$

Perturbed functional

$$J(y+\alpha\eta)=\int_a^b L(x,y+\alpha\eta,y'+\alpha\eta')dx$$

2nd order Taylor expansion with respect to α

$$J(y+\alpha\eta)=\int_a^b L(x,y,y')dx+\alpha\int_a^b\left[L_y(x,y,y')\eta+L_{y'}(x,y,y')\eta'\right]dx+$$

$$+\frac{\alpha^2}{2}\int_a^b\left[L_{yy}\eta^2+2L_{yy'}\eta\eta'+L_{y'y'}\eta'^2\right]dx$$

$$J(y+\alpha\eta)=\int_a^b L(x,y,y')dx+\alpha\int_a^b \left[L_y(x,y,y')\eta+L_{y'}(x,y,y')\eta'\right]dx+$$

$$+\frac{\alpha^2}{2}\int_a^b \left[L_{yy}\eta^2+2L_{yy'}\eta\eta'+L_{y'y'}\eta'^2\right]dx$$

Compare with

$$J(y+\alpha\eta)=J(y)+\delta J|_y(\eta)\alpha+\delta^2 J|_y(\eta)\alpha^2+o(\alpha^2)$$

2nd variation

$$\delta^2 J|_y(\eta)=\frac{1}{2}\int_a^b \left[L_{yy}\eta^2+2L_{yy'}\eta\eta'+L_{y'y'}\eta'^2\right]dx$$

The integrand is evaluated along $(x,y(x),y'(x))$.

Eliminate the term on $\eta\eta'$ using integration by parts

$$\int_a^b 2L_{yy'}\eta\eta' dx = \int_a^b L_{yy'}\frac{d}{dx}(\eta^2)dx = L_{yy'}\eta^2|_a^b - \int_a^b \frac{d}{dx}\left(L_{yy'}\right)\eta^2 dx$$

Since $L_{yy'}\eta^2|_a^b = 0$ by the boundary conditions, it follows that

$$\delta^2 J|_y(\eta) = \frac{1}{2}\int_a^b L_{y'y'}\eta'^2 + \left[L_{yy} - \frac{d}{dx}\left(L_{yy'}\right)\right]\eta^2 dx$$

$$\delta^2 J|_y(\eta) = \frac{1}{2}\int_a^b L_{y'y'}\eta'^2 + \left[L_{yy} - \frac{d}{dx}\left(L_{yy'}\right)\right]\eta^2 dx$$

Define

$$P(x) := \frac{1}{2}L_{y'y'} \qquad\qquad Q(x) := \frac{1}{2}\left[L_{yy} - \frac{d}{dx}\left(L_{yy'}\right)\right]$$

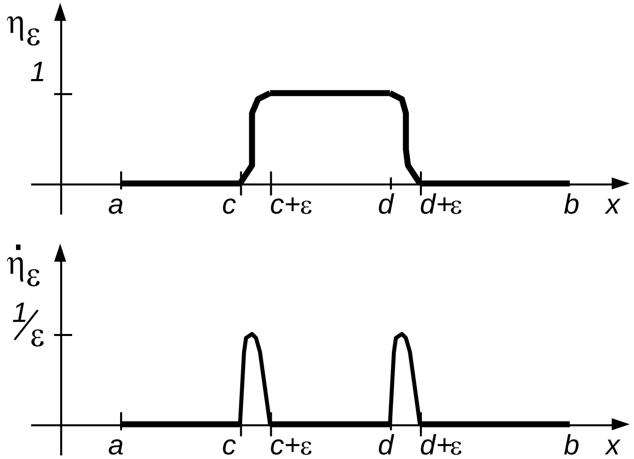
The 2nd variation is written

$$\delta^2 J|_y(\eta) = \int_a^b \left[P(x)\eta'^2 + Q(x)\eta^2\right]dx$$

We know that if y is a minimum for all C^1 perturbations η, $\eta(a) = \eta(b) = 0$, then $\delta^2 J|_y(\eta) \geq 0$ or $\int_a^b \left[P(x)\eta'^2 + Q(x)\eta^2\right]dx \geq 0$.

Now, we want to restate this condition in terms of P and Q only.

Consider a family of perturbations parameterized by small $\varepsilon > 0$



Bound on Q

$$\left|\int_a^b Q(x)\eta_\varepsilon^2(x)dx\right| \le \int_a^b |Q(x)|dx$$

This bound is uniform over ε.

Bound on P

Because of the "peaks" of order $\frac{1}{\varepsilon}$,

$$\int_a^b P(x)\eta_\varepsilon'^2(x)dx$$

Does not stay bounded when $\varepsilon \to 0$.

If $P(\bar{x}) < -\delta$ for some $\bar{x}$, $\delta > 0$, then $P(x) < 0$ in some interval around $\bar{x}$

$$\Rightarrow$$

It is possible to build a disturbance η_ε such that

$$\int_a^b P(x)\eta_\varepsilon'^2(x)dx \le -\frac{\gamma}{\varepsilon}$$ with $\gamma > 0$ a constant.

$$\int_a^b P(x)\eta_\varepsilon'^2(x)dx \leq -\frac{\gamma}{\varepsilon}$$

By making $\varepsilon \longrightarrow 0$,

$$\int_a^b P(x)\eta_\varepsilon'^2(x)dx$$

Can be made arbitrarily negative and dominate

$$\int_a^b Q(x)\eta_\varepsilon^2(x)dx$$

Conclusion

If $P(x)$ takes a negative value for some x, we can chose a perturbation η such that

$$\delta^2 J|_y(\eta) = \int_a^b \left[P(x)\eta'^2 + Q(x)\eta^2\right] dx < 0$$

and y cannot be a minimum.

Thus, for y to be a minimum, it must be

$$P(x) \geq 0 \qquad \forall x \in [a, b]$$

This is Legendre's condition, stated in the next slide.

Recall that $P(x) = L_{y'y'}$

Legendre's condition – 2nd order necessary condition for optimality

Legendre, 1786.

$$L_{y'y'}(x, y(x), y'(x)) \geq 0 \qquad \forall x \in [a, b]$$

This result also holds in the multivariable case, where $L_{y'y'}$ is a symmetric matrix that must be positive semidefinite for all x.

Legendre's condition and the Hamiltonian

Hamiltonian:

$$H(x, y, y', p) = py' - L(x, y, y')$$

The function H^* (the Hamiltonian computed along an optimal curve), when taken as a function of $z = y'$, has a stationary point for $z = y^{*\prime}$ (the derivative of the optimal curve).

Furthermore,

$$H_{y'y'} = -L_{y'y'}$$

From the Legendre condition, along an optimal trajectory

$$H_{y'y'} \leq 0$$

Adrien-Marie Legendre (1752-1833)

French mathematician who did important contributions to statistics, number theory and mathematical analysis, and has left is name connected to topics like the Legendre polynomials, the Legendre transformation (already mentioned in this course), and the Legendre necessary condition in the Calculus of Variations. He was the first to publish the method of least squares to estimate parameters in linear models, a fact that caused a controversy with Gauss who conceived this method many years before, but did not publish it.

Sufficient conditions for a weak minimum

Conjecture (not enough!)

$$\delta^2 J|_y(\eta) > 0 \quad \Rightarrow \quad optimality$$

Actually, in addition, we need an extra global condition to dominate $o(\alpha^2)$.

Implication of $\delta^2 J|_y(\eta) > 0$ on P

For every differentiable function $w = w(x)$

$$0 = w\eta^2|_a^b = \int_a^b \frac{d}{dx}(w\eta^2)dx = \int_a^b [w'\eta^2 + 2w\eta\eta']dx$$

Use this formula to rewrite the second variation as

$$\delta^2 J|_y(\eta) = \int_a^b \left[P(x)\eta'^2 + Q(x)\eta^2\right]dx =$$

$$= \int_a^b \left[P(x)\eta'^2 + Q(x)\eta^2\right]dx + \int_a^b [w'\eta^2 + 2w\eta\eta']dx =$$

$$= \int_a^b \left[P(x)\eta'^2 + 2w\eta\eta' + (Q + w')\eta^2\right]dx$$

$$\delta^2 J|_y(\eta) = \int_a^b \left[P(x)\eta'^2 + 2w\eta\eta' + (Q + w')\eta^2\right] dx$$

Objective: Find w that makes the integrand a perfect square.

Select w to satisfy the Riccati type differential equation

$$P(Q + w') = w^2$$

$$P(Q+w')=w^2$$

Assume $P>0$ and that w exists.

Since

$$Q+w'=\frac{w^2}{P}$$

it follows that

$$\delta^2 J|_y(\eta)=\int_a^b\left[P(x)\eta'^2+2w\eta\eta'+(Q+w')\eta^2\right]dx=$$

$$=\int_a^b P\left[\eta'^2+2\eta'\frac{w\eta}{P}+\left(\frac{w\eta}{P}\right)^2\right]dx=\int_a^b P\left(\eta'+\frac{w\eta}{P}\right)^2 dx$$

$$\delta^2 J|_y(\eta) = \int_a^b P\left(\eta' + \frac{w\eta}{P}\right)^2 dx$$

We now prove that, for an arbitrary η, this integral is strictly positive.

If the integral is 0, then

$$\eta' + \frac{w\eta}{P} = 0$$

This is a linear ODE that, together with the initial condition

$$\eta(a) = 0$$

Implies that $\eta(x) = 0, \forall x \in [a, b]$.

(See the detour in the next page)

A detour: Homogeneous linear ODE's

The solution of

$$y' + p(x)y = 0, \quad y(0) \text{ given}$$

is given by

$$y(x) = e^{-\int_0^x p(\sigma)d\sigma} y(0)$$

Proof: $$y(0) = e^{-\int_0^0 p(\sigma)d\sigma} y(0) = e^0 y(0) = y(0)$$

And hence it verifies the initial condition. Furthermore

$$y'(x) = -p(x)e^{-\int_0^x p(\sigma)d\sigma} y(0) = -p(x)y(x)$$

And hence the proposed solution verifies the ODE.

END DETOUR

Conclusion

If the solution of the Riccati equation

$$P(Q + w') = w^2$$

exists, then

$$P = L_{y'y'} > 0$$

and

$$\delta^2 J|_y(\eta) > 0$$

For all admissible perturbations.

Caveat:

The Riccati ODE may have a finite escape time, i. e., the solution may not exist on the whole interval $[a, b]$.

This issue was raised by Lagrange in 1797.

Conjugate points [L2012], pp. 64,65

Issue to address: Existence of solution to the Riccati equation on the entire interval $[a, b]$.

Jacobi, 1837.

Idea: Make a change of variable to reduce the quadratic Riccati equation

$$P(Q + w') = w^2$$

To a linear equation on v, using

$$w = -\frac{Pv'}{v}, \qquad v \neq 0$$

Riccati equation

$$P\left(Q-\frac{\frac{d}{dx}(Pv')v-Pv'^2}{v^2}\right)=\frac{P^2v'^2}{v^2}$$

Multiply by v ($v \neq 0$) and divide by P ($P > 0$)

$$Qv-\frac{d}{dx}(Pv')+P\frac{v'^2}{v}=P\frac{v'^2}{v}$$

$$Qv=\frac{d}{dx}(Pv')$$

Jacobi equation

$$Qv = \frac{d}{dx}(Pv')$$

Since it is of 2nd order, the solution of this ODE is defined by the initial conditions $v(a)$ and $v'(a)$.

Since if v is a solution, then γv is also a solution, we may assume without loss of generality

$$v(a) = 0,\ v'(a) = 1$$

A point c is conjugate to a if $v(c) = 0$.

See [SL2012] pp. 37-48 for properties of conjugate points.

The initial condition $v(a) = 0$ is inadequate because it yields $w = \infty$.

We may however take $v(a) = \varepsilon$ with ε vanishingly small.

(see [L2012] p.65, Exercise 2.13, on this issue)

In the interval $[a, c]$, with c conjugate to a, we have shown that

$$\delta^2 J|_y(\eta) > 0$$

for any η not identically zero.

It can be proved ([L2012] p. 67) that $\delta^2 J|_y(\eta)$ actually dominates the higher-order term $o(\alpha^2)$.

2^{nd} order sufficient conditions for optimality

An extremal $y(.)$ is a strict local minimum in the weak sense if

$$L_{y'y'}(x,y,y') > 0 \quad \forall x \in [a,b]$$

and the interval $[a,b]$ does not contain any point conjugate to a.

Exercise 27 – Sufficient conditions of minimum

[P1993] p. 39 ex. 3.1 and p. 46

Consider the problem of finding the curve that minimizes

$$J(y) = \int_1^2 y'^2 \, x^3 dx$$

with $y(1) = 0$, $y(2) = 3$.

a) Find the extremal y^*.

b) Compute $\Delta J = J(y) - J(y^*)$, where $y = y^* + \eta$, with η C^1 curve such that $\eta(1) = \eta(2) = 0$, and conclude that y^* is actually a minimum.

c) Show that $L_{y'y'} > 0$ along y and that there are no conjugate points in the interval $[1,2]$. For this sake write the Jacobi equation.

Solution

a) $L = y'^2 x^3$ From which $L_y = 0,\quad L_{y'} = 2y'x^3$

The EL equation $L_y = \frac{d}{dx} L_{y'}$ becomes $\frac{d}{dx}(2y'x^3) = 0$ or $2y'^{x^3} = c_1$

$y' = \frac{c_1}{2x^3} \quad \Rightarrow \quad y(x) = -\frac{c_1}{4x^2} + c_2$

$y(1) = 0 \quad \Rightarrow \quad c_2 = \frac{c_1}{4}$

$y(x) = \frac{c_1}{4}(1 - \frac{1}{x^2}) \qquad y(2) = 3 \quad \Rightarrow \quad \frac{c_1}{4}\left(1 - \frac{1}{4}\right) = 3 \quad \Rightarrow \quad c_1 = 16$

$y(x) = 4(1 - \frac{1}{x^2})$

b) $y' = \frac{8}{x^3}$

$$\Delta J = \int_1^2 (\frac{8}{x^3} + \eta')^2 x^3 dx - \int_1^2 \left(\frac{8}{x^3}\right)^2 x^3 dx =$$

$$= \int_1^2 16\eta' dx + \int_1^2 \eta' x^3 dx = 16\eta(x)|_1^2 + \int_1^2 \eta' x^3 dx =$$

$$= \int_1^2 \eta' x^3 dx > 0 \quad \text{for} \quad \eta(x) \neq 0 \text{ in }]1,2[$$

Hence, $J(y) > J(y^*)$ for any admissible η (C^1, satisfying the boundary conditions $\eta(1) = \eta(2) = 0$), and y^* is a minimum.

In general, the direct computational computation of ΔJ is impossible to perform due to its complexity.

c) $L_{y'} = 2y'x^3$, $\quad L_{y'y'} = 2x^3$ and hence, for $x \in [1,2]$, $L_{y'y'} > 0$.

Study of the conjugate points:

$$P = \frac{1}{2} L_{y'y'} = x^3 \qquad Q = \frac{1}{2}\left(L_{yy} - \frac{d}{dx} L_{yy'}\right) = 0$$

Jacobi equation

$Qv = \frac{d}{dx}(Pv')$ Reduces to $\frac{d}{dx}(Pv') = 0$ or $Pv' = c_1$, $\quad v' = \frac{c_1}{x^3}$, $\quad v = \frac{c_1}{2x^2} + c_2$

Initial conditions: $v(1) = 0$, $\quad v'(1) = 1 \quad \Rightarrow \quad c_1 = -1$, $\quad c_2 = \frac{1}{2}$

The solution of the Jacobi equation is thus

$$v(x) = \frac{1}{2}\left(1 - \frac{1}{x^2}\right)$$

$$v(x) = \frac{1}{2}(1 - \frac{1}{x^2})$$

For $x > 1,\ 1 - \frac{1}{x^2} > 0\ \ \forall x$ and hence there are no conjugate points (zeros of the Jacobi equation) for $x \in [1,2]$.

This fact together with $L_{y'y'} > 0$ means that the sufficient condition for a weak local minimum is verified by the extremal obtained in a).

Hamilton's Principle

Also known as the Principle of Least Action.

Action $S = \int_0^{t_f} L(t, y, y')dt$ $L = T - U$ $T =$ kinetic energy, U= Potential energy

The action is an extremal.

Exercise 28

Consider a particle of mass m that moves in a straight line along a coordinate y, with kinetic energy $T = \frac{1}{2}my'^2$ and potential energy $U = \frac{1}{2}ky^2$ (a mass connected to a spring). Justify the term "Principle of Least Action" by showing that the extremals of the action integral are actually minima on sufficiently small time intervals.1

$$L = \frac{1}{2}my'^2 - \frac{1}{2}ky^2$$

$$L = \frac{1}{2} m y'^2 - \frac{1}{2} k y^2 \qquad L_{y\prime} = my \qquad L_{y\prime y\prime} = m > 0$$

Study of the conjugate points:

$$P = \frac{1}{2} L_{y\prime y\prime} = \frac{1}{2} m \qquad Q = \frac{1}{2}\left(L_{yy} - \frac{d}{dx} L_{yy'}\right) = -\frac{1}{2} k$$

Jacobi equation

$$Qv = \frac{d}{dx}(Pv') \quad \rightarrow \quad -\frac{1}{2} kv = \frac{1}{2} m \frac{d}{dx} v' \quad \rightarrow \quad -kv = mv'' \quad \rightarrow \quad v'' + \frac{k}{m} v = 0$$

Assume solutions of the form $v(x) = e^{\lambda x}$.

Characteristic equation: $\lambda^2 + \frac{k}{m} = 0 \quad \rightarrow \quad \lambda_{1,2} = \pm j\sqrt{\frac{k}{m}} \qquad \omega = \sqrt{\frac{k}{m}}$

$$v(x) = k_1 e^{j\omega x} + k_2 e^{-j\omega x}$$

$$v(x) = k_1 e^{j\omega x} + k_2 e^{-j\omega x}$$

$$v(x) = (k_1 + k_2)\cos(\omega x) + j\,(k_1 - k_2)\sin(\omega x)$$

Initial conditions $v(0) = 0$, $v'(0) = 1$:

$k_1 + k_2 = 0$, $\quad j(k_1 - k_2)\omega = 1 \qquad k_1 = -k_2 = -j\frac{1}{\omega}$

$$v(x) = \frac{1}{\omega}\sin\omega x$$

There is a conjugate point at $\frac{\pi}{\omega}$.