Variational Methods & Optimal Control

lecture 04

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

April 14, 2016

Fixed-end point problems

We'll start with the simplest functional maximization problem, and show how to solve by finding the first variation and deriving the Euler-Lagrange equations:

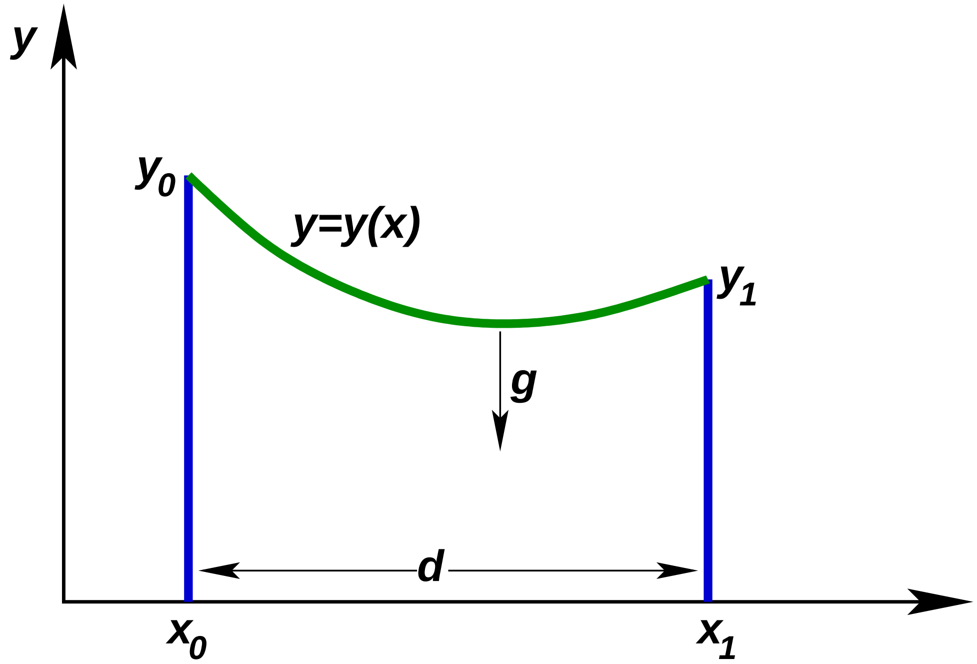
$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y} = 0$$

The Catenary

The potential energy of the cable is

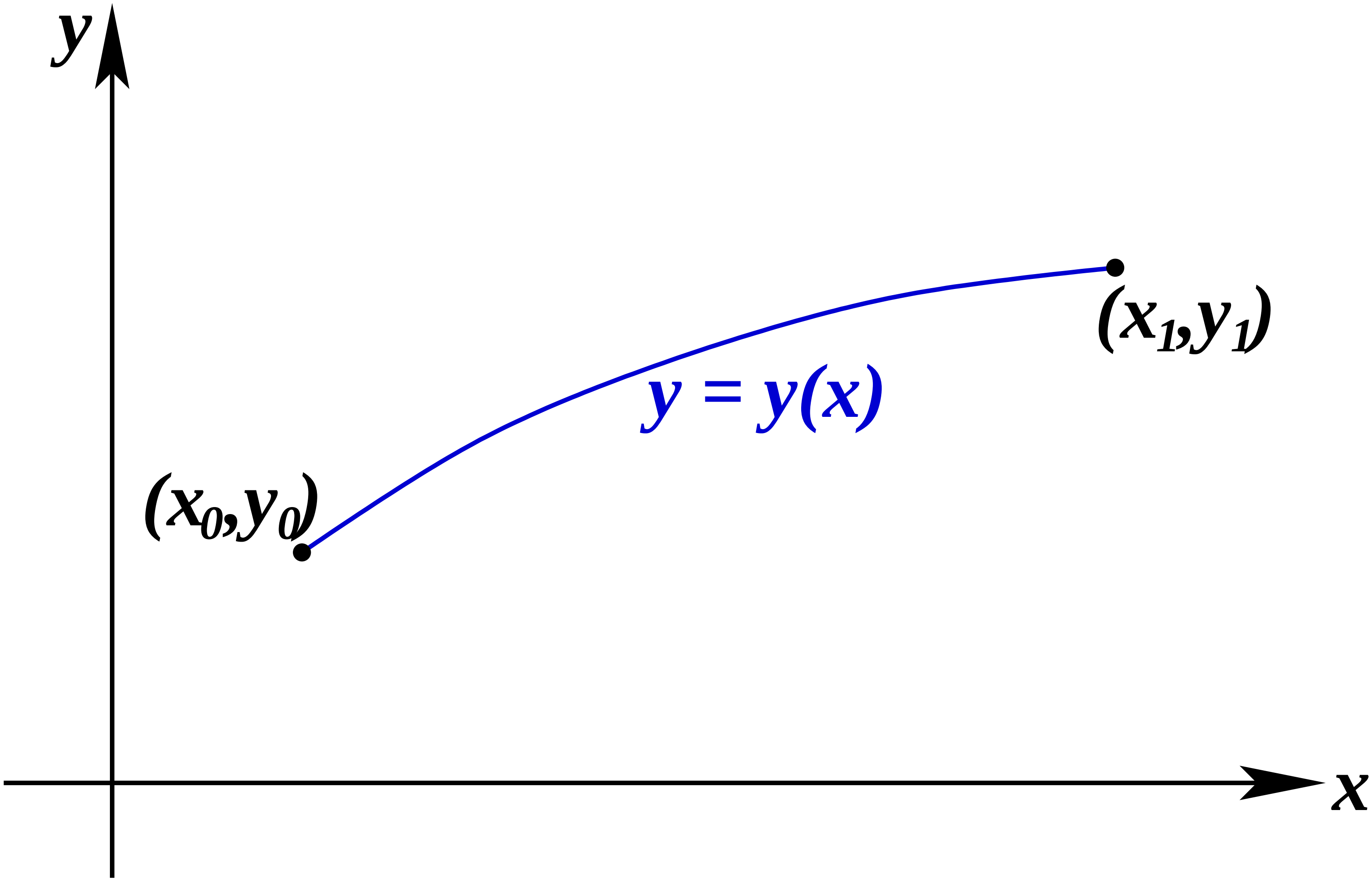
$$W_p\{y\} = \int_0^L mgy(s)ds$$

Where L is the length of the cable



Catenary problem where we have pullies on top of each pylon, and a large amount of cable. Under appropriate conditions it will reach an equilibrium shape. The critical features of this problem are that the end-points are fixed but the length L of the cable is unconstrained.

Fixed end-point variational problem



Formulation

Define the functional $F : C^2[x_0, x_1] \to I\!R$

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

where f is assumed to be function with (at least) continuous second-order partial derivatives, WRT x, y, and y'.

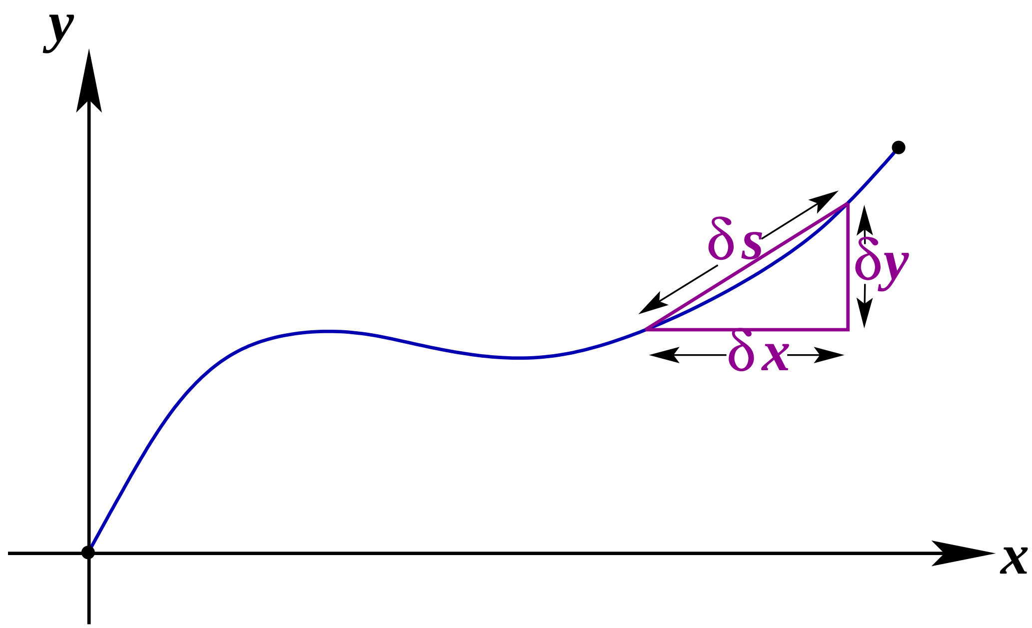
Problem: determine $y \in C^2[x_0, x_1]$ such that $y(x_0) = y_0$ and $y(x_1) = y_1$, such that F has a local extrema.

The Catenary

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

But I don't know how to evaluate this integral directly. Lets do a simple change of variables. The length of a line segment from (x,y) to $(x+\delta x, y+\delta y)$ is

$$\begin{aligned} \delta s &\simeq \sqrt{\delta x^2 + \delta y^2} \\ &= \sqrt{1 + \left(\frac{\delta y}{\delta x}\right)^2}\, \delta x \\ ds &= \sqrt{1+y'^2}dx \end{aligned}$$



The Catenary

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

Change of variables $ds = \sqrt{1+y'^2}dx$. So the functional of interest (the potential energy) is

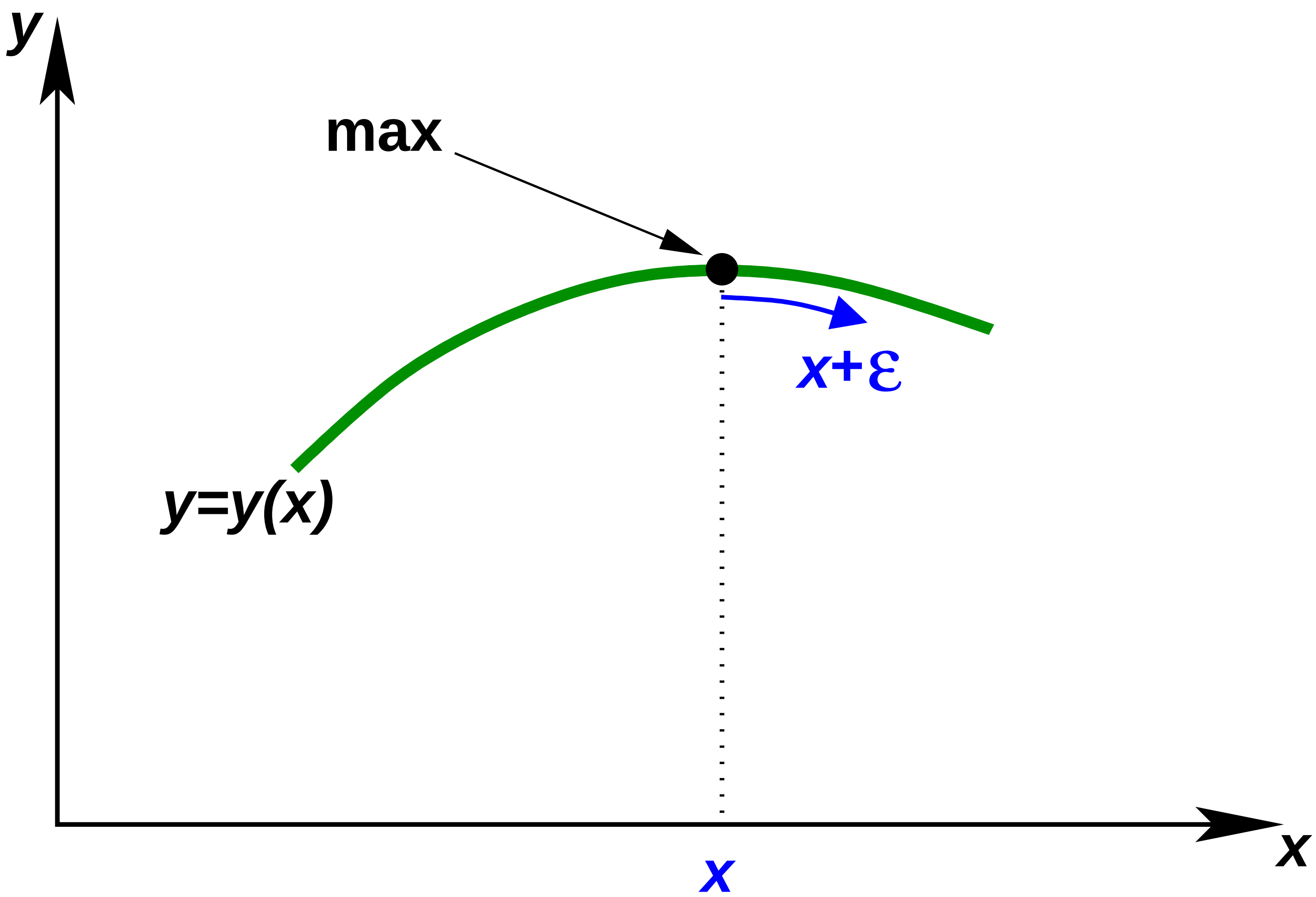
$$\begin{aligned} W_p\{y\} &= mg\int_{x_0}^{x_1} y\sqrt{1+y'^2}\,dx, \\ &= mg\int_{x_0}^{x_1} f(x,y,y')\,dx, \end{aligned}$$

where

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

How do we tackle these problems

look at small perturbations about the max/min.

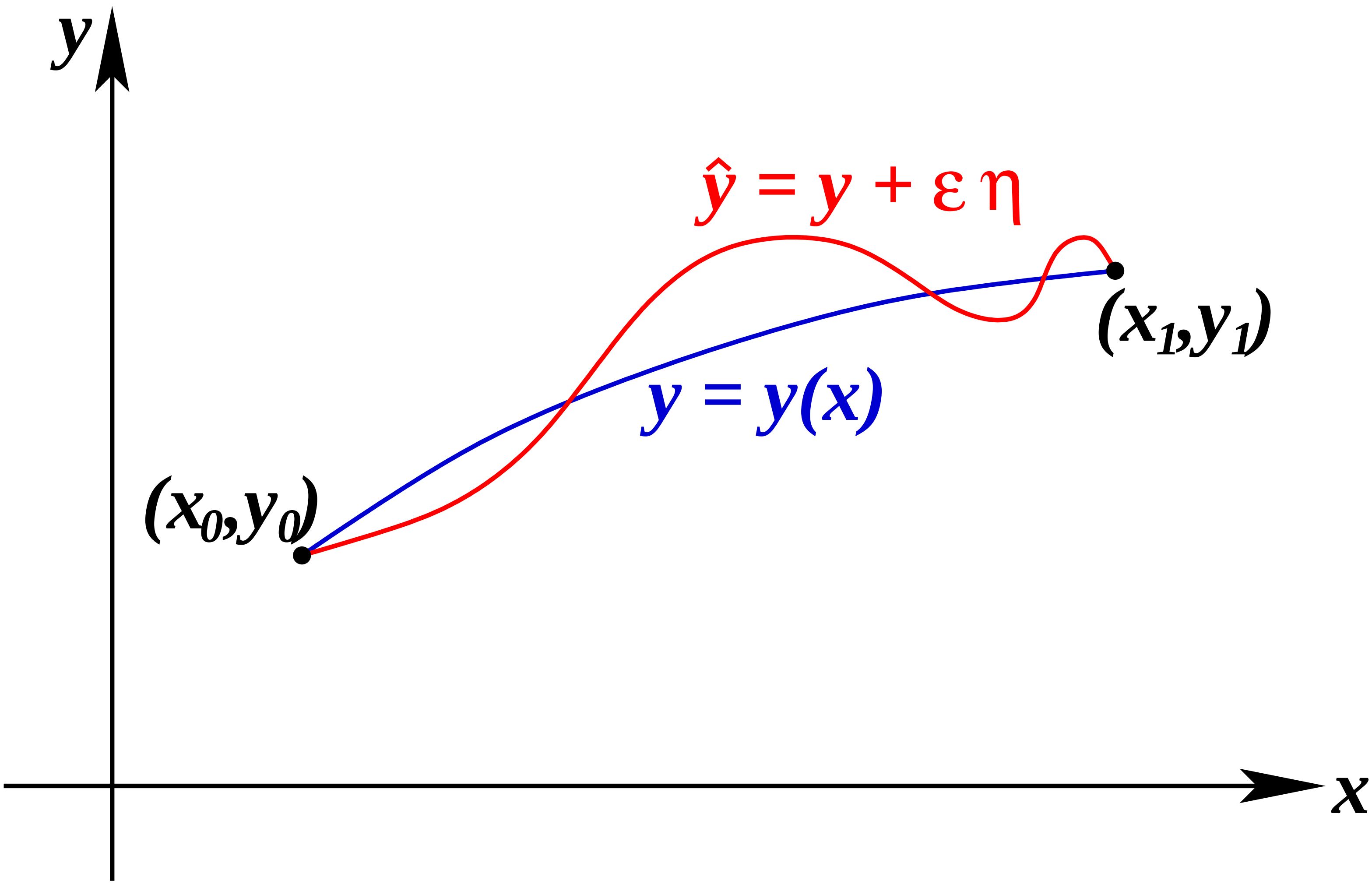


For a local maximum

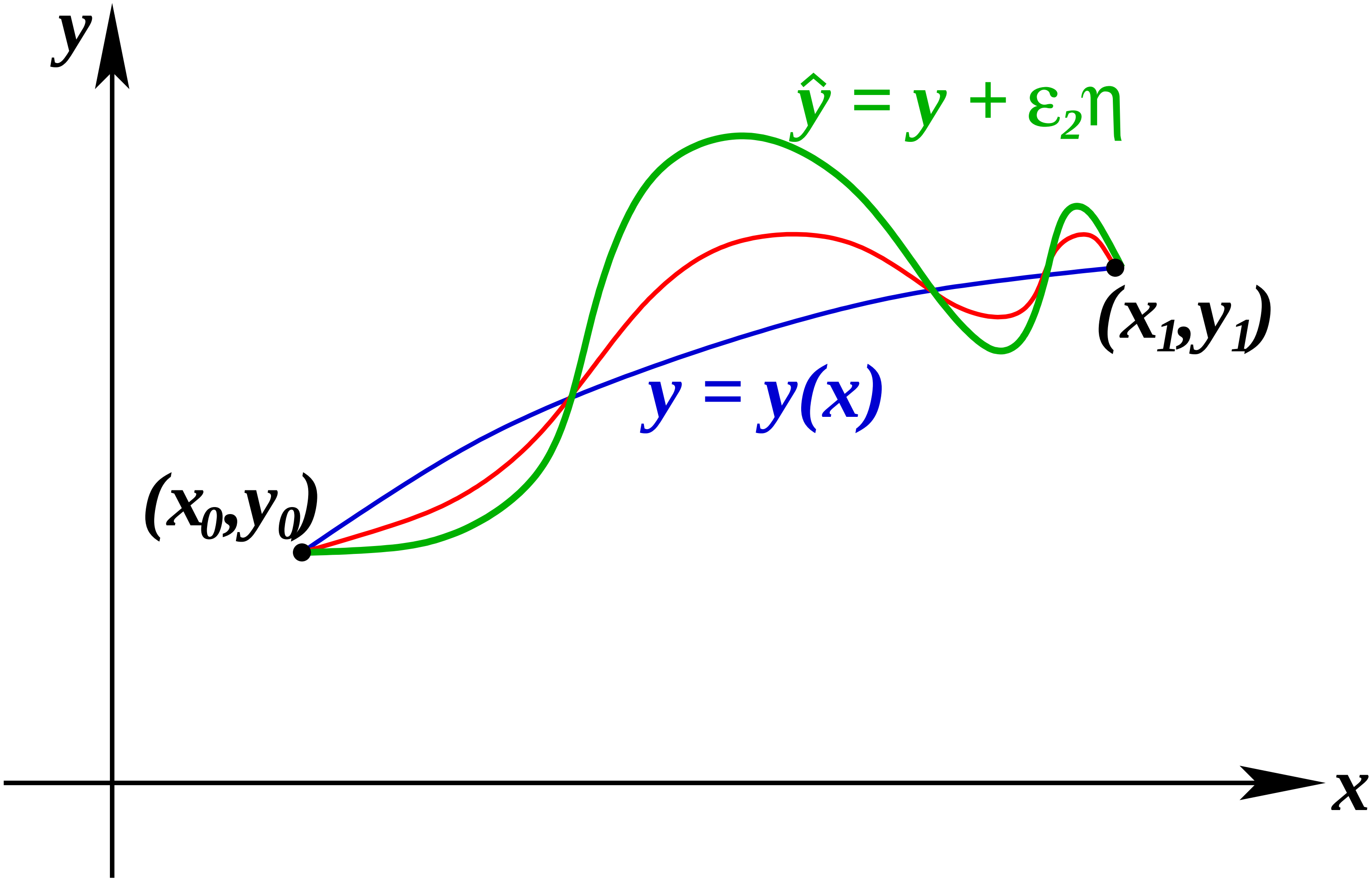
$$f(x+\varepsilon) \leq f(x)$$

$\Rightarrow$ Conditions for extremals, i.e., $f'(x) = 0$

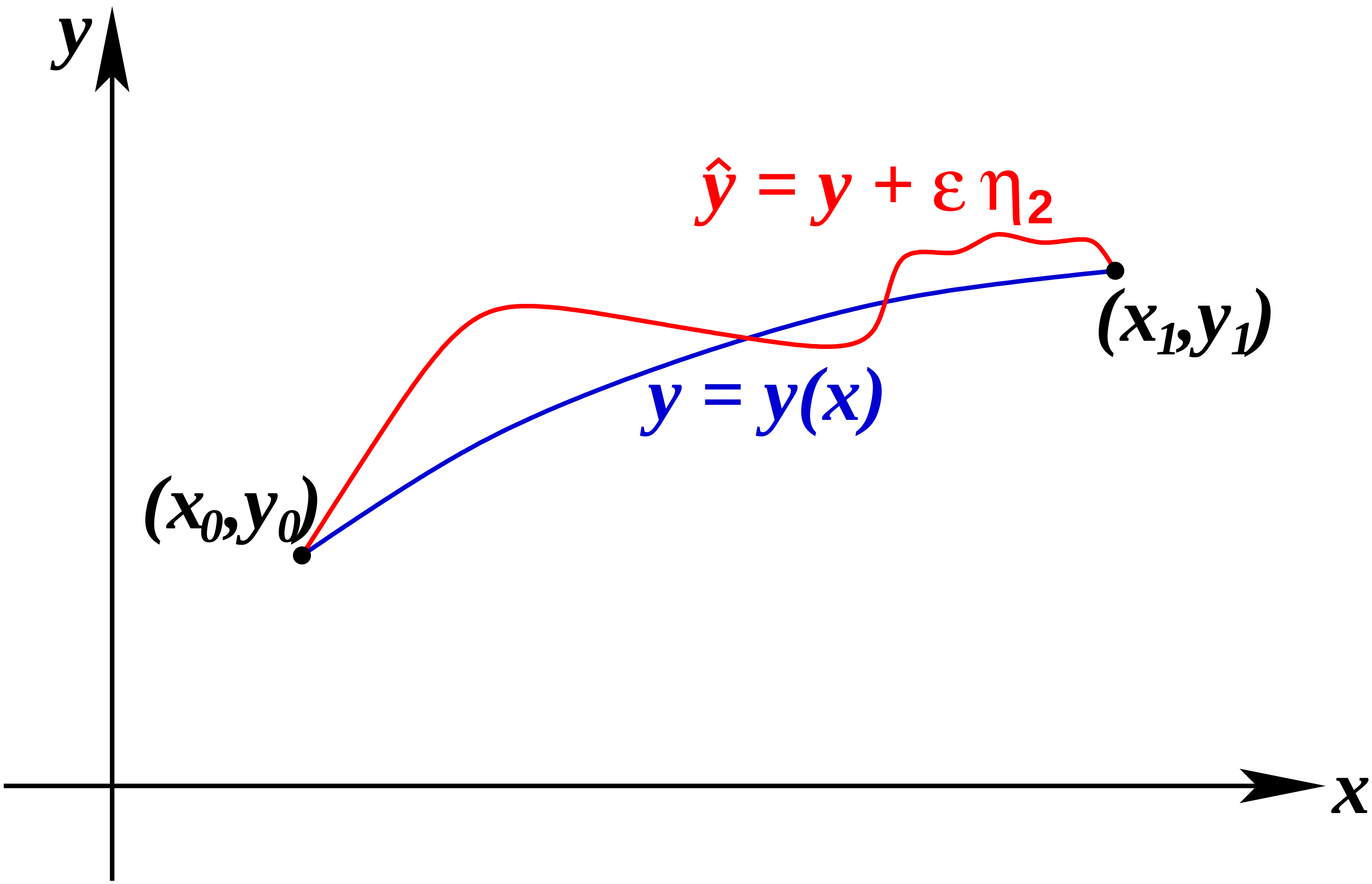
Perturbations of functions



Perturbations of functions



Perturbations of functions



The Functional of interest.

Define the functional $F : C^2[x_0, x_1] \to I\!R$

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

where f is assumed to be function with continuous second-order partial derivatives, WRT x, y, and y'.

Problem: determine $y \in C^2[x_0, x_1]$ such that $y(x_0) = y_0$ and $y(x_1) = y_1$, such that F has a local extrema.

The space of possible curves is

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

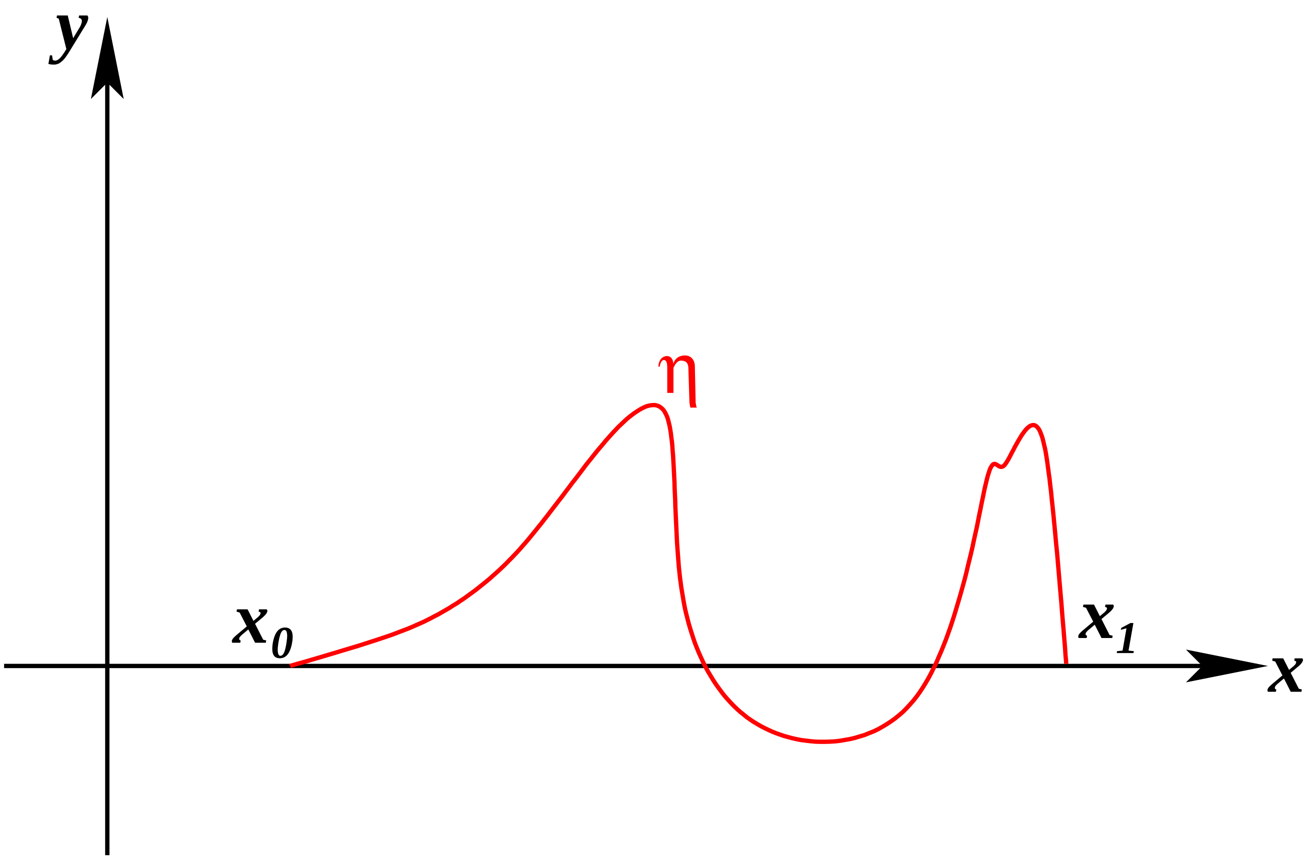
$\Rightarrow$ The vector space of allowable perturbations is

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

Perturbation functions

The vector space of allowable perturbations is

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



What to do

Regard f as a function of 3 independent variables: x, y, y'

Take $\hat{y}(x) = y(x) + \varepsilon\eta(x)$, where $y \in S$ and $\eta \in \mathcal{H}$.

Taylor's theorem (note x is kept constant below)

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

So

$$\begin{aligned} F\{\hat{y}\} - F\{y\} &= \int_{x_0}^{x_1} f(x,\hat{y},\hat{y}')dx - \int_{x_0}^{x_1} f(x,y,y')dx \\ &= \varepsilon\int_{x_0}^{x_1}\left[\eta\frac{\partial f}{\partial y} + \eta'\frac{\partial f}{\partial y'}\right]dx + O(\varepsilon^2) \end{aligned}$$

The first variation

For small ε the quantity

$$\delta F(\eta, y) = \lim_{\varepsilon \to 0} \frac{F\{y + \varepsilon\eta\} - F\{y\}}{\varepsilon} = \int_{x_0}^{x_1} \left[\eta \frac{\partial f}{\partial y} + \eta' \frac{\partial f}{\partial y'} \right] dx$$

is called **the First Variation**.

For $F\{y\}$ to be a minimum, for small ε, $F\{\hat{y}\} \geq F\{y\}$, so the sign of $\delta F(\eta, y)$ is determined by ε.

As before, we can vary the sign of ε, so for $F\{y\}$ to be a local minima it must be the case that

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

Analogy to functions

This condition on the first variation is analogous to all partial derivatives being zero!

For a function of N variables to have a local extrema

$$\frac{\partial f}{\partial x_i} = 0, \qquad \forall i = 1, \ldots, n$$

For a functional to be an extrema

$$\delta F(\eta, y) = \left. \frac{d}{d\varepsilon} F(y + \varepsilon\eta) \right|_{\varepsilon=0} = 0, \qquad \forall \eta \in \mathcal{H}$$

Note now that we have to minimize over an infinite dimensional space $\mathcal{H}$, instead of $\mathbb{R}^n$.

Simplification

Integrate the second term by parts

$$\begin{aligned}\delta F(\eta, y) &= \int_{x_0}^{x_1} \left[\eta \frac{\partial f}{\partial y} + \eta' \frac{\partial f}{\partial y'}\right] dx \\ &= \left[\eta \frac{\partial f}{\partial y'}\right]_{x_0}^{x_1} + \int_{x_0}^{x_1} \eta \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right] dx\end{aligned}$$

But note that by the problem definition $\eta \in \mathcal{H}$, and so $\eta(x_0) = \eta(x_1) = 0$, and so the first term is zero.

The function inside the integral exists, and is continuous by our assumption that f has two continuous derivatives, so for $E(x) = \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right]$

$$\delta F(\eta, y) = \int_{x_0}^{x_1} \eta(x) E(x) dx = \langle \eta, E \rangle^2 = 0$$

Euler-Lagrange equation

Theorem 2.2.1: Let $F : C^2[x_0, x_1] \to \mathbb{R}$ be a functional of the form

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

where f has continuous partial derivatives of second order with respect to x, y, and y', and $x_0 < x_1$. Let

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

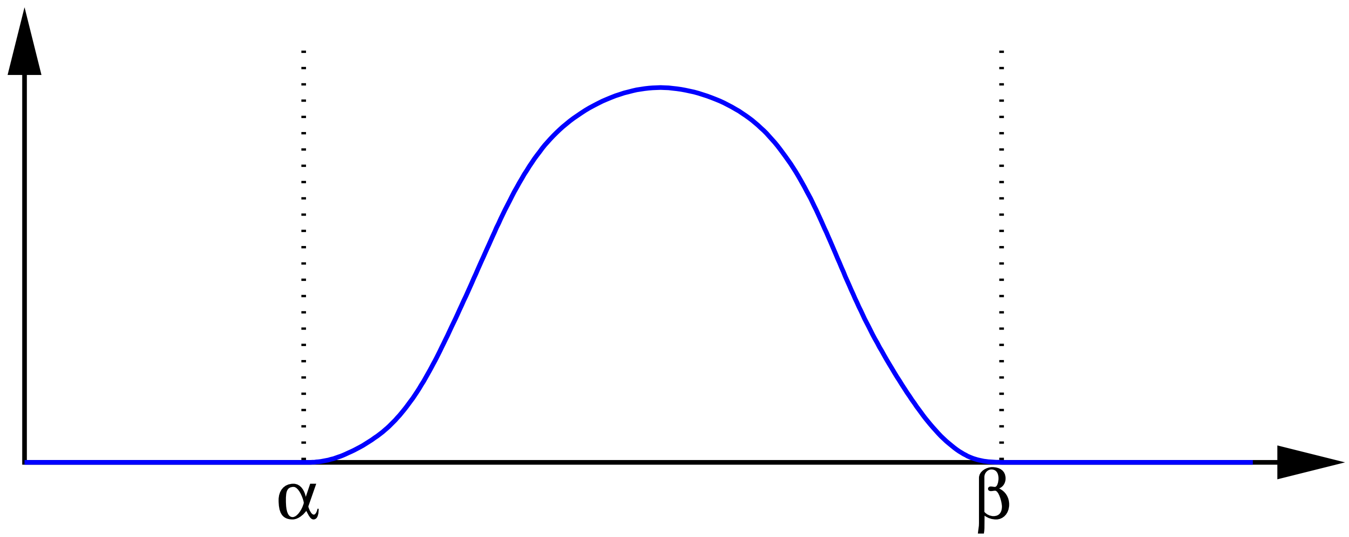
where y_0 and y_1 are real numbers. If $y \in S$ is an extremal for F, then for all $x \in [x_0, x_1]$

$$\boxed{\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y} \;=\; 0} \Leftarrow \text{the Euler-Lagrange equation}$$

A useful lemma

Lemma 2.2.1: Let $\alpha, \beta \in \mathbb{R}$, such that $\alpha < \beta$. Then there is a function $\nu \in C^2(\mathbb{R})$, such that $\nu(x) > 0$ for all $x \in (\alpha, \beta)$ and $\nu(x) = 0$ otherwise.
Proof: by example

$$\nu(x) = \begin{cases} (x-\alpha)^3(\beta - x)^3, & \text{if } x \in (\alpha, \beta) \\ 0, & \text{otherwise.} \end{cases}$$



A second useful lemma

Lemma 2.2.2: Suppose $\langle \eta, g \rangle = 0$ for all $\eta \in \mathcal{H}$. If $g : [x_0, x_1] \to I\!R$ is a continuous function then $g(x) = 0$ for all $x \in [x_0, x_1]$.
Proof: Suppose $g(x) > 0$ for $x \in [\alpha, \beta]$. Choose ν as in Lemma 2.2.1.

$$\langle \nu(x), g(x) \rangle^2 = \int_{x_1}^{x_2} \nu(x) g(x) dx = \int_{\alpha}^{\beta} \nu(x) g(x) dx > 0$$

Hence a contradiction.
Similar proof for $g(x) < 0$.

Proof of Euler-Lagrange equation

As noted earlier, at an extremal the first variation

$$\delta F(\eta, y) = \langle \eta(x), E(x) \rangle^2 = \int_{x_0}^{x_1} \eta(x) E(x) dx = 0$$

for all $\eta(x) \in \mathcal{H}$. From Lemma 2.2.2, we can therefore state that

$$E(x) = \left[\frac{\partial f}{\partial y} - \frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right) \right] = 0,$$

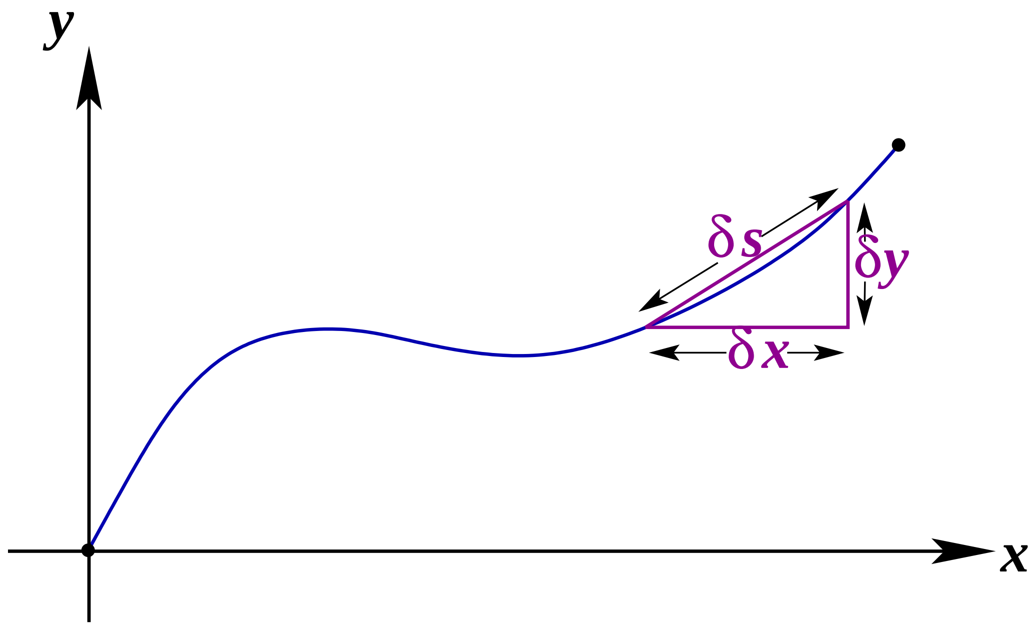
the Euler-Lagrange equation.$\Box$

Example: geodesics in a plane

Let $(x_0, y_0) = (0,0)$ and $(x_1, y_1) = (1,1)$, find the shortest path between these two points.

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

$$\begin{aligned} \delta s &= \sqrt{\delta x^2 + \delta y^2} \\ &= \sqrt{1 + \left(\frac{\delta y}{\delta x}\right)^2}\,\delta x \\ ds &= \sqrt{1 + y'^2}\,dx \end{aligned}$$



So the total path length is $F\{y\} = \int_{x=0}^{x=1} ds = \int_0^1 \sqrt{1 + y'^2}\,dx$

Example: geodesics in a plane

The arclength of a curve described by $y(x)$ will be

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

Then

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

So $\frac{y'}{\sqrt{1+y'^2}}$ is a constant, implying $y' = const$. Hence $y(x) = c_1 x + c_2$, the equation of a straight line.

- Q: how do I know this is a minimum?

Special cases

Now that we know the Euler-Lagrange (E-L) equations, we can use them directly, but there are some special cases for which the equations simplify, and make our life easier:

- f depends only on y'
- f has no explicit dependence on x (autonomous case)
- f has no explicit dependence on y
- $f = A(x,y)y' + B(x,y)$ (degenerate case)

Special case 1

When f depends only on y' the E-L equations simplify to

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

An example of this is calculating geodesics in the plane (which we all know are straight lines).

f depends only on y'

Geodesics in the plane are a special case of $f = f(y')$, with no explicit dependence on y. Apply the chain rule to the E-L equation and we get

$$
\begin{aligned}
\frac{d}{dx}\frac{\partial f}{\partial y'} &= 0 \\
\frac{d^2 f(y')}{dy'^2}\frac{dy'}{dx} &= 0 \\
\frac{d^2 f(y')}{dy'^2} y'' &= 0
\end{aligned}
$$

so one of the two following must be true

$$
\begin{aligned}
f''(y') &= 0 \\
y'' &= 0
\end{aligned}
$$

f depends only on y'

- If $f''(y') = 0$, then $f(y') = ay' + b$. We will later see that problems in this form are “degenerate”, and solutions don’t depend on the curve’s shape.
- If $y'' = 0$, then
$$y = c_1 x + c_2.$$

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

- e.g. geodesics in the plane

Example f depends only on y'

Consider finding the extremals of

$$F\{y\} = \int_0^1 \alpha y'^4 - \beta y'^2 , dx$$

such that $y(0) = 0$ and $y(1) = b$.
The Euler-Lagrange equation is

$$\frac{d}{dx}\left[4\alpha y'^3 - 2\beta y'^2\right] = 0$$

We could play around with this for a while to solve, but we already know the solutions are straight lines, so the extremal will be

$$y = bx$$

Fermat's principle

Fermat's principle of geometrical optics:

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

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

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

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

Speed of light

The speed of light (EM radiation) is only constant in a vacuum

medium	speed (km/s)	refractive index
vacuum	300,000	1.0
water	231,000	~ 1.3
glass	200,000	~ 1.5
diamond	125,000	~ 2.4
silicon	75,000	~ 4.0

Refractive index = c/v

Example

Consider $c(x,y) = 1/g(x)$

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

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

f has no explicit dependence on y so

$$\begin{aligned}
\frac{\partial f}{\partial y'} &= const \\
g(x)\frac{y'}{\sqrt{1+y'^2}} &= const
\end{aligned}$$

Example (ii)

$$
\begin{aligned}
g(x)\frac{y'}{\sqrt{1+y'^2}} &= c_1 \\
\frac{y'^2}{1+y'^2} &= \frac{c_1^2}{g(x)^2} \qquad \text{implies } c_1^2 \leq g(x)^2 \\
y'^2 &= \frac{c_1^2}{g(x)^2}(1+y'^2) \\
y'^2\left(1-\frac{c_1^2}{g(x)^2}\right) &= \frac{c_1^2}{g(x)^2} \\
y' &= \sqrt{\frac{c_1^2}{g(x)^2-c_1^2}}
\end{aligned}
$$

Example (iii)

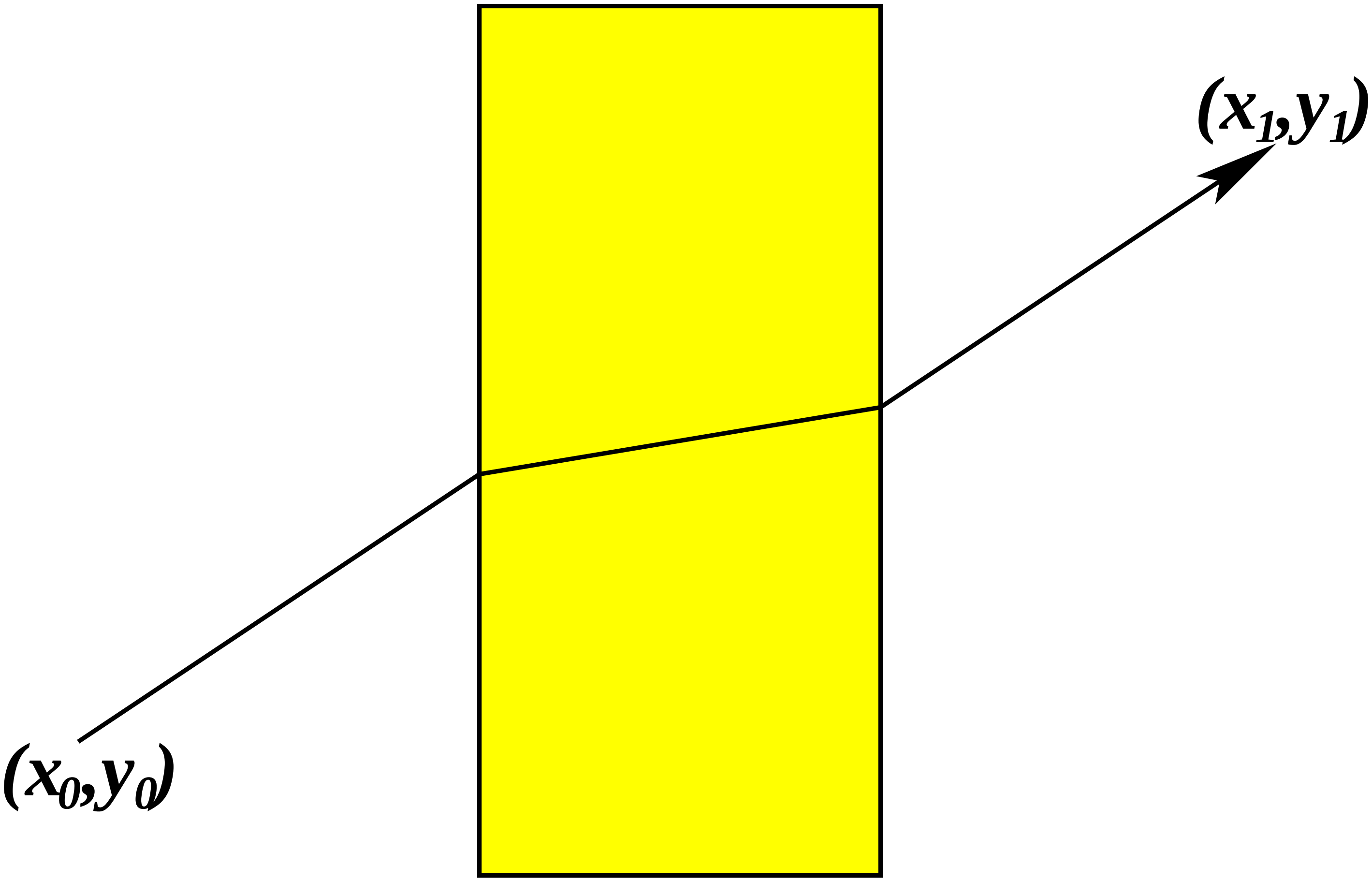
$$
\begin{aligned}
y' &= \sqrt{\frac{c_1^2}{g(x)^2 - c_1^2}} \\
y &= c_1 \int \frac{1}{\sqrt{g(x)^2/c_1^2 - 1}} dx + c_2
\end{aligned}
$$

The constants, c_1 and c_2 are determined by the fixed end points.

- so not all extremals are straight lines
- we had to include an x term here to make it more interesting

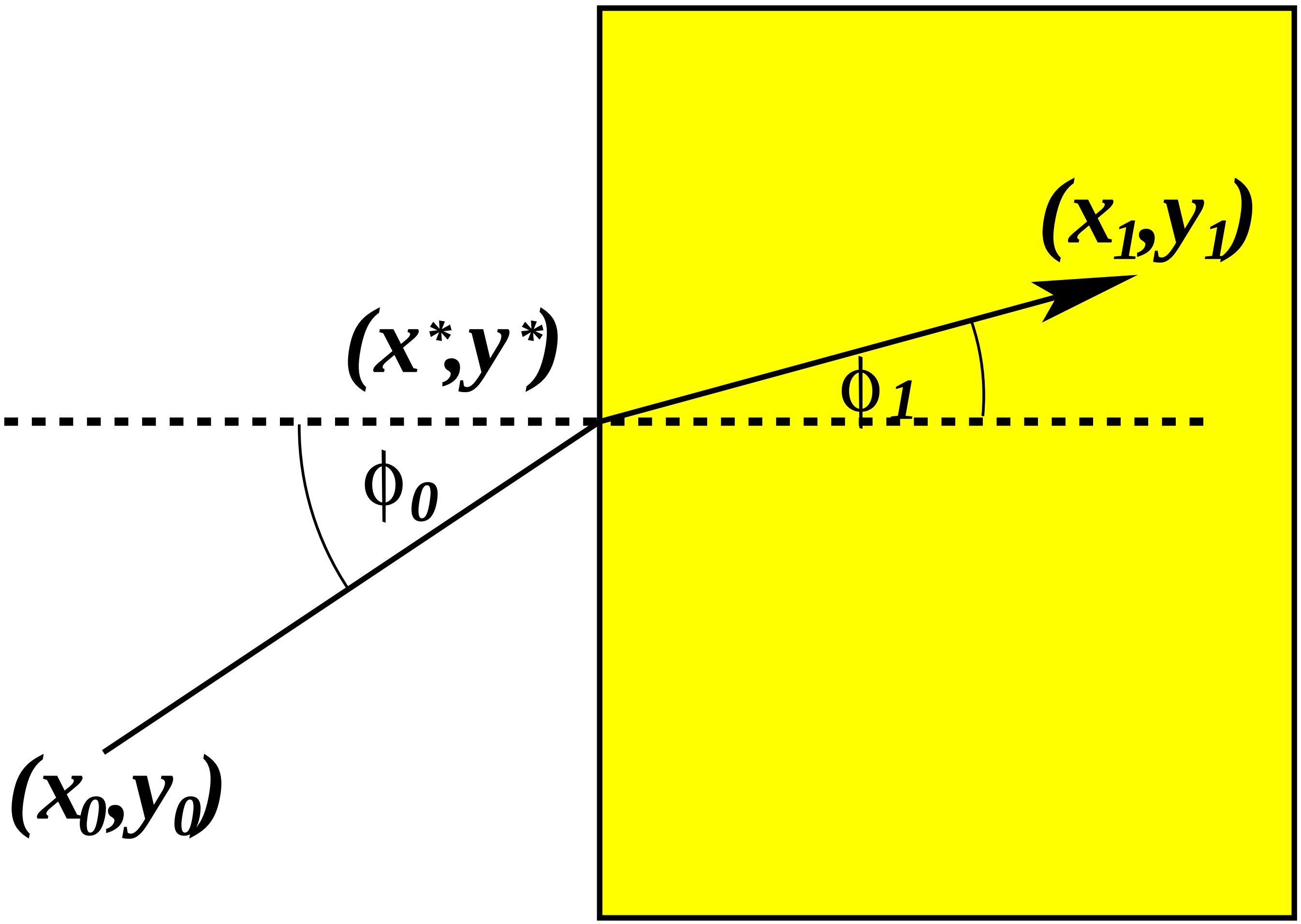
What we can't do (yet)

member, f must have at least two continuous derivatives. If the speed of light $x,y)$ has discontinuities, then we are in trouble.



How we might solve

Break into two problems, with a boundary point (x^*, y^*), which has a fixed value of x^* (the location of the boundary), but a movable value for y^*.



The functional

$$F\{y\} = \int_{x_0}^{x^*} \frac{\sqrt{1+y'^2}}{c_0}\,dx + \int_{x^*}^{x_1} \frac{\sqrt{1+y'^2}}{c_1}\,dx$$

Separate into two problems, as if we knew (x^*, y^*). Each is a geodesic in the plane problem. So the solutions are straight lines

$$y(x) = \begin{cases} (x-x_0)\frac{y^*-y_0}{x^*-x_0} + y_0 & x \leq x^* \\ (x-x^*)\frac{y_1-y^*}{x_1-x^*} + y^* & x \geq x^* \end{cases}$$

Now we can explicitly compute $F\{y\}$ as a function of x, by differentiating y, and then we can treat it as a minimization problem in one variable y^*.

The total time taken

We can simplify the integrals by noting from Pythagoras that the lengths of the two lines are

$$\sqrt{(x^* - x_0)^2 + (y^* - y_0)^2} \quad \text{and} \quad \sqrt{(x^* - x_1)^2 + (y^* - y_1)^2}$$

and that the time take to traverse the pair of line segments will be

$$T\{y\} = \frac{\sqrt{(x^* - x_0)^2 + (y^* - y_0)^2}}{c_0} + \frac{\sqrt{(x^* - x_1)^2 + (y^* - y_1)^2}}{c_1}$$

$$\frac{dT}{dy^*} = \frac{(y^* - y_0)}{c_0 \left[(x^* - x_0)^2 + (y^* - y_0)^2\right]^{1/2}} - \frac{(y_1 - y^*)}{c_1 \left[(x^* - x_1)^2 + (y^* - y_1)^2\right]^{1/2}}$$

The result

$$
\begin{aligned}
\frac{dT}{dy^*} &= \frac{(y^* - y_0)}{c_0\left[(x^* - x_0)^2 + (y^* - y_0)^2\right]^{1/2}} - \frac{(y_1 - y^*)}{c_1\left[(x^* - x_1)^2 + (y^* - y_1)^2\right]^{1/2}} \\
&= \frac{\sin\phi_0}{c_0} - \frac{\sin\phi_1}{c_1}
\end{aligned}
$$

which we require to be zero to find the minimum. Hence

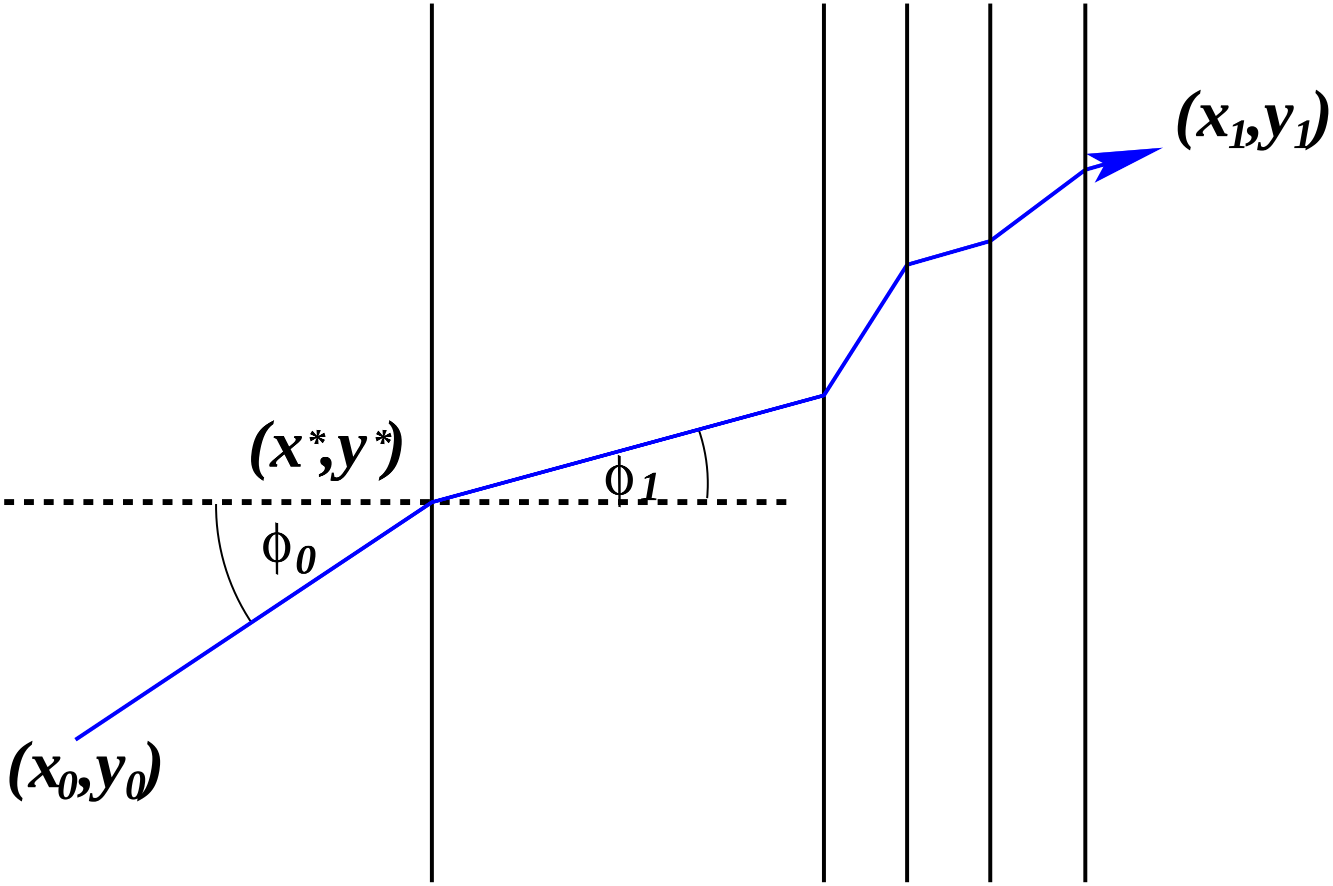
$$
\frac{\sin\phi_0}{c_0} = \frac{\sin\phi_1}{c_1} \quad \Leftarrow \textbf{ Snell's law for refraction}
$$

Hence there are often ways around discontinuities, though it may involve some pain

(e.g. what about internal reflection)

More than one boundary

Snell's law applies at each boundary



Dealing with “kinks”

- We’ll spend a fair bit of time later on dealing with “kinks” in curves
- Underlying point
  - The integral can still be well defined even if extremal isn’t “smooth”
  - But the Euler-Lagrange equations don’t work at the kinks
  - Use the Euler-Lagrange equations everywhere except the kinks
  - Do something else at the kinks