Part 4: Calculus of Variations

W8C1 Lecture 12 (Feb 25)

4.1 Variational Problems

Consider the minimization of

$J[y] = \int_a^b F(x, y, y')\, dx$

among all functions $y(x)$ on $[a, b]$ with some BC $y(a) = y_a$, $y(b) = y_b$.

Here, $F(x, y, z)$ is a continuous function of three variables $x \in [a, b], y, z \in \mathbb{R}$.

📌 **Def.** A function of functions (or curves) such as $J[y]$ is called a functional.

💡 **Ex 1.**

a) $J[y] = \int_a^b (y'(x))^2\, dx \implies F(x, y, z) = z^2$

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

💡 **Ex 2.**

Let $A(a, y_a)$ and $B(b, y_b)$ be two fixed points. Find the curve joining A and B that:

a) has the shortest length

We consider the minimization $\min \int_a^b \sqrt{1 + y'(x)^2}\, dx$
The answer is a
straight line.

b) it takes the shortest time for a particle to slide down the curve under the influence of gravity.

The curve is called a brachistochrone, Greek for shortest time.

Ex 3. (Isoperimetric problem).

Among all closed curves of a given length l, the curve enclosing the greatest area is a circle.

Define $\vec{r}(s) = (x(s), y(s))$ with $0 \leq s \leq l$.

Impose $|\vec{r}'(s)| = 1$ and $\vec{r}(0) = \vec{r}(l)$.

The area is $\int_0^l \frac{1}{2}(x(s)\frac{dy}{ds} - y(s)\frac{dx}{ds})\, ds$

We want to maximize area subject to the constraints:

- $x, y : [0, l] \mapsto \mathbb{R}$
- $x'(s)^2 + y'(s)^2 = 1$ for all s.

$J = \int_0^l F(s, x, y, x', y')\, ds = \int_0^l F(s, \vec{r}, \vec{r}')\, dx$

where $\vec{r}, \vec{r}' \in \mathbb{R}^2$ and $F = \frac{1}{2}(xy' - yx')$.

4.2 Function Spaces

To talk about continuity of $J[y]$, we need a space for $y(x)$ with distance. It is usually a subset of a normed linear space $(Y, ||\cdot||)$. It is only a subset since we have boundary conditions, etc.

A norm satisfies:

1. $||y|| = 0 \iff y = 0$
2. $||\alpha y|| = |\alpha| \cdot ||y||$ for $\alpha \in \mathbb{R}, y \in Y$
3. $||x + y|| \leq ||x|| + ||y||$

Ex 4.

a) $Y = \mathbb{R}^n$, $||y|| = \begin{cases} \sqrt{y_1^2 + \cdots + y_n^2} \\ |y_1| + \cdots + |y_n| \\ \max_{1 \leq j \leq n} |y_j| \end{cases}$

b) $C^0([a, b]) = \{\text{continuous functions on } [a, b]\}$

For $k \in \mathbb{N}_0$,

$C^k([a, b]) = \{y : y, y^1, \ldots, y^{(k)} \in C^0([a, b])\}$

It has a norm:

$||y||_{C^k([a,b])} = \sum_{j=0}^k \max_{a \leq x \leq b} |y^{(j)}(x)|$

Def. The functional $J[y]$ defined in a normed space Y is said to be continuous at the point $\hat{y} \in Y$ if for any $\varepsilon > 0$, there is a $\delta > 0$ such that $|J[y] - J[\hat{y}]| < \varepsilon$ if $||y - \hat{y}|| < \delta$.

💡 **Ex 5.** The functional $J[y] = \int_0^1 (y')^2(x)\,dx$ is not continuous at $\hat{y} = 0$ in $Y = C^0([0,1])$ because $||y - \hat{y}||_{C^0} \leq \delta$ does not imply $J[y]$ is defined. It is however continuous in $y \in C^1([0,1])$:

Fix $\hat{y} \in Y$ and let $A = ||\hat{y}||_{C^1}$. If both $y, \hat{y} \in C^1([0,1])$ and $||y - \hat{y}||_{C^1} \leq \delta$, then

$||y||_{C^1} \leq ||\hat{y}||_{C^1} + ||y - \hat{y}||_{C^1} \leq A + \delta.$

$$\begin{aligned}|J[y] - J[\hat{y}]|| &= |\int_0^1 y'^2 - \hat{y}'^2(x)\,dx| \\ &\leq \int_0^1 |y' - \hat{y}| \cdot |y' + \hat{y}'|\,dx \\ &\leq \int_0^1 \delta(A + \delta + A)\,dx \\ &= (2A + \delta)\delta < \varepsilon\end{aligned}$$

if $\delta \leq 1$ and $\delta \leq \frac{\varepsilon}{2A+1}$.

W8C2 Lecture 13 (Feb 27)

4.3 The Variation of a Functional

📌 **Def.** A functional $J[y]$ on a normed linear space Y is a linear functional if

- $J[\alpha y] = \alpha Y[j], \quad \forall \alpha \in \mathbb{R}, \quad \forall y \in Y$
- $J[y_1 + y_2] = J[y_1] + J[y_2]$
- $J[y]$ is continuous

💡 **Ex 6.** Examples of linear functionals:

a) $J[y] = y(\frac{1}{2})$ for $y \in C^0([0,1])$

b) $J[y] = \int_a^b y(x)\,dx$ for $y \in C^0([a,b])$

c) $J[y] = \int_a^b \alpha(x)y(x)\,dx$ for $y \in C^0([a,b])$, where $\alpha(x)$ is fixed in $C^0([a,b])$

d) $J[y] = \int_a^b [\alpha_0(x)y(x) + \alpha_1(x)y'(x) + \cdots + \alpha_k y^{(k)}(x)]\,dx$ where $\alpha_j \in C^0([a,b])$ is a linear functional on $C^k([a,b])$, for $j = 0, \ldots, k$

📖 **Lemma 1.** Let $\alpha(x)$ be continuous in $[a,b]$.

a) If $\int_a^b \alpha(x)h(x)\,dx = 0$ for every $h(x) \in C^0([a,b])$ with $h(a) = h(b) = 0$, then $\alpha(x) = 0$ in $[a,b]$.

b) If $\int_a^b \alpha(x)h'(x)\,dx = 0$ for every $h(x) \in C^1([a,b])$ with $h(a) = h(b) = 0$, then $\alpha(x) = \mathrm{const}$ in $[a,b]$.

Proof of a):

Suppose $\alpha(x) > 0$ somewhere. Then $\alpha(x) > 0$ in some $[c, d] \subset [a, b]$.

Let $h(x) = \begin{cases} (x-c)(d-x) & c \le x \le d \\ 0 & \text{elsewhere} \end{cases}$.

It satisfies the conditions $\int_a^b \alpha(x)h(x)\,dx = \int_c^d \alpha(x)(x-c)(d-x)\,dx > 0$.

This contradiction shows that $\alpha(x) \le 0$ for all x. Similarly, $\alpha(x) \ge 0$ for all x.

Proof of b):

Let $k = \frac{1}{b-a}\int_a^b \alpha(x)\,dx$ and $h(x) = \int_a^x [\alpha(\zeta) - k]\,d\zeta$.

Then $h(x) \in C^1([a, b])$ and $h(a) = 0 = h(b)$.

$\int_a^b (\alpha(x) - k)h'(x)\,dx = \int_a^b (\alpha(x) - k)^2\,dx \ge 0$

But also $\int_a^b (\alpha(x) - k)h'(x)\,dx = \int_a^b \alpha(x)h'(x)\,dx - k[h]_a^b = 0 - 0 = 0$

Hence we must have $\alpha(x) = k$ for all x.

> 📖 **Lemma 2.** If $\alpha(x)$ and $\beta(x)$ are continuous in $[a, b]$, and if
>
> $\int_a^b [\alpha(x)h(x) + \beta(x)h'(x)]\,dx = 0$
>
> for every $h \in C^1([a, b])$ with $h(a) = h(b) = 0$, then $\beta(x)$ is differentiable and $\beta'(x) = \alpha(x)$ for all $x \in [a, b]$.

Remarks:

i) $\beta \in C^1$ is a conclusion, not an assumption.

ii) For intuition, we can assume $\beta \in C^1$, then by IBP, $\int_a^b (\alpha(x) - \beta'(x))h(x)\,dx = 0$ for all h. By Lemma 1(a), $\alpha(x) - \beta'(x) = 0$.

Proof:

Let $A(x) = \int_a^x \alpha(\zeta)\,d\zeta$. By IBP, $\int_a^b \alpha(x)h(x)\,dx = -\int_a^b A(x)h'(x)\,dx$.

Thus, $\int_a^b [\beta(x) - A(x)]h'(x)\,dx = 0$ for all h.

By Lemma 1(b), $\beta(x) - A(x) = \text{const}$, then by the definition of A, $\beta' = \alpha$.

For a functional $J[y]$, consider its increment:

$\Delta J[y] = J[y + h] - J[y]$

corresponding to a perturbation $h(x)$ of $y(x)$.

📌 **Def.** If there is a linear functional $\varphi[y]$ such that

$\Delta J[h] = \varphi[h] + \varepsilon\|h\|$

where $\varepsilon(h) \to 0$ as $\|h\| \to 0$, we say that $J[y]$ is differentiable at y and denote $\varphi[h] = \delta J[h]$ as the principle linear part, also called the variation or differential of $J[y]$ at y.

Remarks:

i) It is unique if it exists

ii) To specify y, it is also denoted as $\delta J[y; h]$.

💡 **Ex 7.** Let $Y = \mathbb{R}^2$, $f \in C^1(\mathbb{R}^2)$ and $J[y] = f(y)$ for $y = (y_1, y_2)$.

For small $h = (h_1, h_2)$,

$\Delta J[h] = f(y+h) - f(y) = f_{y_1}(y)h_1 + f_{y_2}(y)h_2 + \text{error}$

by a Taylor expansion. Thus,

$\delta J[h] = f_{y_1}(y)h_1 + f_{y_2}(y)h_2.$

📖 **Thm 1.** A necessary condition for a differentiable function $J[y]$ to have an extrema at $y = \hat{y}$ is that its variation vanishes at $y = \hat{y}$.

$\delta J[h] = 0$ for $y = \hat{y}$ and all admissible h.

Remarks:

i) In Calc 1, if $f(x)$ attains extrema at $x = x_0$, then $f'(x_0) = 0$

ii) We may consider a minimum by considering $\tilde{J}[y] = -J[y]$ if necessary.

iii) h is not arbitrary. For example, if $Y = C^1([a, b])$ and we want y and $\hat{y}$ to have the same BC, then we need $h(a) = 0 = h(b)$. Which h is admissible depends on each problem.

Proof:

Suppose $\hat{y}$ minimizes $J[y]$. By definition, at $\hat{y}$,

$\Delta J[h] = \delta J[h] + \varepsilon\|h\|$ where $\varepsilon \to 0$ as $\|h\| \to 0$.

Suppose $\delta J[h_0] \neq 0$ for some h_0.

Then $\Delta J[\alpha h_0] = \delta J[\alpha h_0] + \varepsilon\|\alpha h_0\|$ **(☆)**

For sufficiently small $\alpha \in \mathbb{R}$, $|\alpha| < \alpha_1$, we have $|\varepsilon\|\alpha h_0\|| < \frac{1}{2}|\delta J[\alpha h_0]|$.

So the RHS of **(☆)** has the same sign as $\alpha\delta J[h_0]$ which can be + or - depending on the sign of α.

However, LHS ≥ 0 for all $\alpha \in (-\alpha_1, \alpha_1)$ since $\hat{y}$ is a minimizer.

This contradiction shows that $\delta J[h] = 0$ for all h.

W9C1 Lecture 14 (Mar 4)

4.4 Euler-Lagrange Equations

1) Consider the special case again: look for minimization of $J[y] = \int_a^b F(x, y(x), y'(x))\, dx$ among all functions $y(x)$ in the admissible class.

2) $\mathcal{A} = \{y \in C^1([a,b]), y(a) = y_a, y(b) = y_b\}$

We require $y \in C^1$ so that $J[y]$ is defined.

Consider **(T1)** for (1)-(2): we need to compute the variation δJ.

Suppose we give $y(x) \in \mathcal{A}$ an increment $h(x)$.

Since $y, y+h \in \mathcal{A}$, we must have $h \in C^1([a,b])$ and $h(a) = h(b) = 0$. Then:

$$\begin{aligned}\Delta J &= \int_a^b [F(x, y+h, y'+h') - F(x,y,y')]\, dx \\ &= \int_a^b [F_y(x,y,y') + F_z(x,y,y')h' + \text{error}]\, dx \quad = \delta J[h] + \text{error}\end{aligned}$$

where we have used a Taylor expansion for small h.

By **(T1)**, a necessary condition for $y(x)$ to be an extrema is:

$\delta J[y; h] = \int_a^b [F_y(\cdots)h + F_z(\cdots)h']\, dx$ for all admissible h.

By **(L2)**, $F_z(x,y,y')$ is C^1 and $\frac{d}{dx}F_z(x,y,y') = F_y$.

📖 **Thm 2.** Let $J[y]$ be defined for $y \in \mathcal{A}$ as in (1) and (2). A necessary condition for $y(x)$ to be an extrema is that $y(x)$ satisfies the Euler-Lagrange equation:

$F_y(x,y,y') - \frac{d}{dx}F_z(x, y(x), y'(x)) = 0$ **(Eq. 3)**

Note that $\frac{d}{dx}$ is a total derivative, and the expanded form of **(3)** is

$F_y - F_{zx} - F_{zy}y' - F_{zz}y'' = 0$ **(Eq. 4)**

This is a 2nd order DE, linear in y'', nonlinear in y, y'.

💡 **Ex 8.** Minimize arclength in Ex 2a): $J[y] = \int_a^b \sqrt{1 + (y'(x))^2}\, dx$

$F(x,y,z) = \sqrt{1+z^2}$

$F_y = 0, F_{zx} = 0, F_{zy} = 0$

$F_z = \frac{z}{\sqrt{1+z^2}}$

$F_{zz} = \frac{1}{\sqrt{1+z^2}} + \frac{-\frac{1}{2}z}{(1+z^2)^{3/2}} \cdot 2z = \frac{1+z^2-z^2}{(1+z^2)^{3/2}} = \frac{1}{(1+z^2)^{3/2}}$

By E-L **(4)**, we have

$-\frac{1}{(1+y')^{3/2}}y'' = 0 \implies y'' = 0 \implies y(x) = mx + k$

Hence, the straight line is the only candidate for a minimizer (need to check that it is a minimizer separately) because E-L equations are a necessary condition.

Existence problem of minimizers:

(T2) says a minimizer satisfies E-L equation, with given BC. However, it may have no solution in C^2. Otherwise, y'' will be discontinuous, and the solution will satisfy E-L equation as distributions (out of scope of MATH 401).

In MATH 215/255, we studied IVPs for the same equation with IC

$$\begin{cases} y(a) = y_0 \\ y'(a) = y_1 \end{cases}$$

at same point. We have a solution $y(x)$ for $a - \delta < x < a + \delta$ for some δ.

💡 **Ex 9.** Minimize $J[y] = \int_{-1}^{1} y^2(2x - y')^2 \, dx$ among $y \in C^1, y(-1) = 0, y(1) = 1.$

Clearly, $J[y] \geq 0$, and in particular $\min J[y] = 0$ attained by:

$y(x) = \begin{cases} 0 & -1 \leq x \leq 0 \\ x^2 & 0 \leq x \leq 1 \end{cases}$ which satisfies the required conditions.

Note that $y, y' \in C^0$ but y'' does not exist at $x = 0$.

If we examine the E-L equation:

$F(x, y, z) = y^2(2x - z)^2$

$F_y = 2y(2x - z)^2$

$F_z = -2y^2(2x - z)$

both of which equal zero. So the E-L equation is empty/degenerate.

Remark: minimizer may not be in C^2. We need $F_{zz} \neq 0$, where F_{zz} is the coefficient of y'' in **(4)**.

📖 **Thm 3.** Suppose $F(x, y, z) \in C^2$, and $y(x) \in C^1$ solves the E-L equation

$F_y - \frac{d}{dx}[F_z(x, y, y')] = 0$

Then $y(x) \in C^2$ at x where $F_{zz}(x, y, y') \neq 0.$

Idea: by **(4)**,

$y'' = \frac{1}{F_z z}(F_y - F_{zx} - F_{zy} y')$

Check that the limit $\frac{\Delta y'(x)}{\Delta x}$ exists and equals the RHS. Then y'' exists and is C^0.

Special Cases:

1. $F = F(x, y'), F_y = 0$. E-L equations ⇒ $0 - \frac{d}{dx}(F_z(x, y, y')) = 0 \implies F_z(x, y, y') = c.$

   First order equation and we call it first integral (order reduction).

2. $F = F(y, y'), F_x = 0$. E-L equation ⇒ $F_y - 0 - F_{zy} y' - F_{zz} y'' = 0.$

   Multiply by y' and integrate:

$$
\begin{aligned}
F_y y' - F_{zy}(y')^2 - F_{zz}y'y'' &\iff F_y y' + F_z y'' - y'' F_z - y'(F_{zy}y' + F_{zz}y'') \\
&\iff \frac{d}{dx}(F(y,y') - y'F_z(y,y')) = 0
\end{aligned}
$$

This is also first integral, $F - y\text{'}F_z = \text{const}$. (Relates to classical mechanics, where x is time)

3. $F = F(x,y), F_z = 0$. E-L equation ⇒ $F_y(x,y) = 0$, which is an algebraic equation.
4. $F(x,y,z) = f(x,y)\sqrt{1+z^2}$.

   Let $A = \sqrt{1+z^2}, A\text{'} = \frac{z}{A}, A\text{''} = \frac{1}{A^3}$.

$$
\begin{aligned}
F_y - \frac{d}{dx}F_z &= f_y A - \frac{d}{dx}(f\frac{y'}{A}) \\
&= f_y A - f_x\frac{y'}{A} - f_y y' \cdot \frac{y'}{A} - f(\frac{y''}{A} - \frac{y'}{A^2} \cdot \underbrace{\frac{y'}{A} \cdot y''}_{\frac{dA}{dz}\,\frac{dz}{dx}}) \\
&= f_y A - f_x\frac{y'}{A} - f_y\frac{y'^2}{A} - f\frac{y''}{A^3} \\
&= \frac{1}{A}[f_y(1+y'^2) - f_x y' - f_y y'^2 - \frac{fy''}{1+y'^2}]
\end{aligned}
$$

   Hence, $f_y - f_x y\text{'} - \frac{fy\text{''}}{1+y\text{'}^2} = 0$.

   Comparing with **(4)**: still second order, linear in y'', nonlinear in y, y'. Not obviously easier.

💡 **Ex 10.** Minimize $J[y] = \int_1^2 \frac{\sqrt{1+y'^2}}{x}\,dx$, $y(1) = 0, y(2) = 1.$

This satisfies both case 1 and case 4. Case 1 is first order and easier, so use that.

$F_z = c$ for some constant.

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

$y' = \frac{cx}{\sqrt{1-c^2x^2}}.$

$y = \int \frac{cx\,dx}{\sqrt{1-c^2x^2}} = \frac{1}{c}\sqrt{1-c^2x^2} + c_1 \implies (y-c_1)^2 = \frac{1}{c^2} - x^2$

This is a circle with center $(0, c_1)$ with radius $\frac{1}{c}$.

Plugging in boundary conditions, we get $c_1 = 2$, $\frac{1}{c^2} = 5.$

Alternatively, we can use case 4 with
$f(x, y) = \frac{1}{x}.$

$$
\begin{aligned}
f_y - f_x y' - \frac{f}{1+z^2}y'' &= 0 \\
0 + \frac{1}{x^2}y' - \frac{1}{x(1+y'^2)}y'' &= 0 \\
y'' &= \frac{1}{x}y'(1+y'^2)
\end{aligned}
$$

We have $z = y'$ and $y'' = \frac{dz}{dx}$.

$\frac{dz}{dx} = \frac{1}{x}z(1+z^2) \implies \frac{dz}{z(1+z^2)} = \frac{dx}{x}$

$\ln x = \int \frac{1}{x}\,dx = \int \frac{dz}{z(1+z^2)} = \int (\frac{1}{z} - \frac{z}{z^2+1})\,dz = \ln z - \frac{1}{2}\ln(z^2+1) + C$

Hence $cx = \frac{z}{\sqrt{z^2+1}} = \frac{y'}{\sqrt{y'^2+1}}$, and we are back to the beginning of case 1.

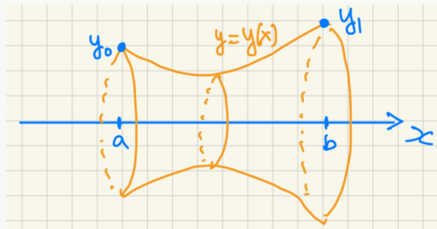
Remark: Case 4 is useful when it is not also case 1 or 2 which is 1st integral. It is the case when $F(x, y, z) = f(x, y)\sqrt{1+z^2}$ and $f(x, y)$ depends on both x and y, such as $F(x, y, z) = xy\sqrt{1+z^2}$.

W9C2 Midterm (Mar 7)

Midterm in class.

W10C1 Lecture 15 (Mar 11)

Ex 11. $J[y] = 2\pi \int_a^b y\sqrt{1+y'^2}\,dx$ with $y(a) = y_0, y(b) = y_1$ is the area of the surface of revolution by rotating $y = y(x)$ about the x-axis.



It is case 2 and 4. Using case 2:

$F_x = 0$, we have the first integral equation $F - y' F_z = c$.

$y\sqrt{1+y'^2} - y' \cdot y\frac{y'}{\sqrt{1+y'^2}} = c$

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

$$y' = \frac{\pm\sqrt{y^2-c^2}}{c}$$

We can drop the $\pm$ by changing the sign of C:

$dx = \frac{c\,dy}{\sqrt{y^2-c^2}}$

$x - c_1 = c\ln\frac{y+\sqrt{y^2-C^2}}{C}$

$y = c\cosh\frac{x-c_1}{c}$.

The resulting curve is called a catenary, and the surface is called a catenoid.

With the boundary conditions $y(a) = y_0$ and $y(b) = y_1$, we get:

$\cosh\frac{a-c_1}{c} = \frac{y_0}{c}$ and $\cosh\frac{b-c_1}{c} = \frac{y_1}{c}$.

We have two equations with two unknowns (c and c_1). There are three possibilities:

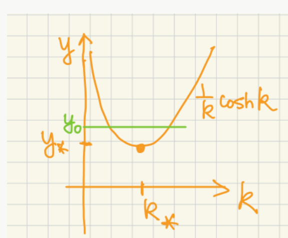
1. exactly one single solution
2. (at least) two solutions, only one of them is the area minimizer
3. no solution

Consider the symmetric case for illustration:

$a = -1, b = 1, y_0 = y_1$: then $c_1 = 0$ and $c = \frac{1}{k} > 0$.

$y = \frac{1}{k}\cosh kx$ and $y_0 = y_1 = \frac{1}{k}\cosh k$.

Let $g(k) = \frac{1}{k}\cosh k$.



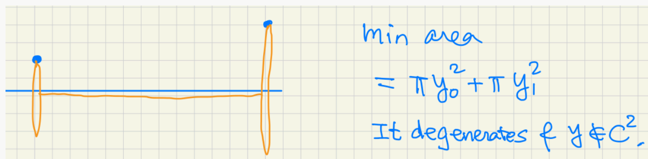
$y_* = \min g(k)$ occurs at:

$\frac{dg}{dk} = \frac{1}{k^2}(k \sinh k - \cosh k) = 0 \implies k \tanh k = 1 \implies k_* = 1.1997$

Then $y_* = 1.51$. We summarize as:

$$\begin{cases} y_0 = y_* & \text{one solution } k \\ y_0 > y_* & \text{two solutions} \\ y_0 < y_* & \text{none} \end{cases}$$

In fact, when $|y_0| + |y_1| \ll b - a$, we get a degenerate surface:



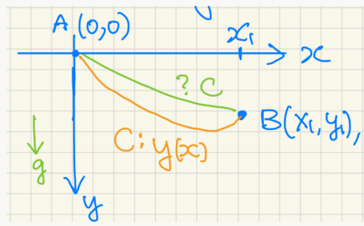
Ex 12. Consider $J[y] = \int_a^b (x-y)^2 \, dx$.

This is case 3, and we have the equation $F_y(x-y) = 0$.

We have $2(x-y) = 0 \implies y(x) = x$. However, we cannot impose boundary conditions.

💡 **Ex 13.** (Brachistochrone problem).

Find the curve that allows for the shortest time to travel from A to B under gravity with zero initial speed and no friction.



Let's choose coordinates so that A is at the origin. Point the y-axis downwards so that $y(x) \geq 0$.

Define C as the curve: $y = y(x), y(0) = 0, y(x_1) = y_1$.

We seek to minimize time $T = \int_C dt = \int_C \frac{ds}{v}$, where s is the arclength and $ds = \sqrt{1 + y'^2}\, dx$.

The speed is $v = \frac{ds}{dt}$ and has the condition $v = 0$ at A.

To find v, we use conservation of energy: total energy = KE + PE.

$E_0 = \frac{1}{2}mv^2 - mgy$ (note the negative sign on mgy because y is pointing downwards)

Hence, $v = \sqrt{2gy}$. Let's define $J[y]$ (with a factor of $\sqrt{2g}$ for convenience) as:

$J[y] = \sqrt{2g} \int_C \frac{ds}{V} = \int_C \frac{ds}{\sqrt{y}} = \int_0^b \sqrt{\frac{1+y'^2}{y}}\, dx$

So $F(x, y, z) = \sqrt{\frac{1+z^2}{y}}$ and $F_x = 0$.

This is case 2 and 4. Using the first integral equation for case 2:

$$F - y' F_z = c$$

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

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

$$\sqrt{y}\sqrt{1+y'^2} = \tfrac{1}{c} > 0$$

$$y(1 + y'^2) = k \text{ where } k = \frac{1}{c^2} > 0$$

Hence, $y'^2 = \frac{k}{y} - 1 = \frac{k-y}{y} \implies \frac{dy}{dx} = \pm\sqrt{\frac{k-y}{y}} \implies \pm dx = \sqrt{\frac{y}{k-y}}\, dy$

Let $k - y = kw^2$, so $y = k - kw^2 = k(1 - w^2)$ and $dy = -2kw\, dw$.

$\pm dx = \sqrt{\frac{k(1-w^2)}{kw^2}}(-2kw\, dw) = -2k\sqrt{1 - w^2}\, dw$.

Let $w = \cos t$, so $dw = -\sin t\, dt$.

$\pm dx = 2k \sin t \cdot \sin t\, dt = k(1 - \cos 2t)\, dt$

Integrating, we get $\pm x = k(t - \frac{1}{2}\sin 2t)$. Now let $2t = \theta$ and $\frac{k}{2} = R$.

$+x = R(\theta - \sin\theta) + x_0$, where we have chosen the positive sign so that $\frac{dx}{d\theta} = y \geq 0$.

From previously, we have:

$y = k - kw^2 = k - k\cos^2\frac{\theta}{2} = k - \frac{k}{2}(\cos\theta + 1) = R(1 - \cos\theta)$

Hence, $(x, y) = \vec{r}(\theta) = \big(R(\theta - \sin\theta) + x_0, R(1 - \cos\theta)\big)$

For the boundary conditions, we have, for some $\theta_0 < \theta_1$:

$A = (x(\theta_0), y(\theta_0)) = (0, 0)$

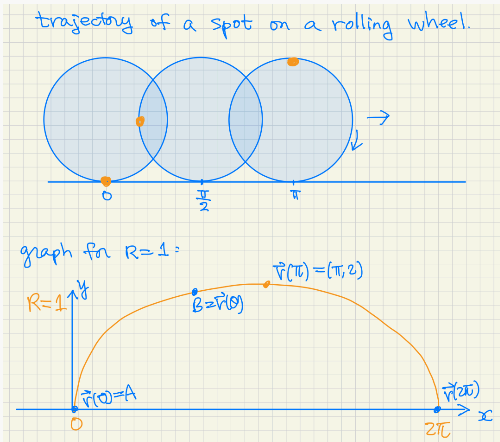
$B = (x(\theta_1), y(\theta_1)) = (x_1, y_1)$

Hence $y(\theta_0) = 0 \implies 1 - \cos\theta_0 = 0 \implies \theta_0 = 0$

We also have $x_0 = x(0) = 0$.

The trajectory becomes $\vec{r}(\theta) = R\big((\theta, 1) - (\sin\theta, \cos\theta)\big)$.

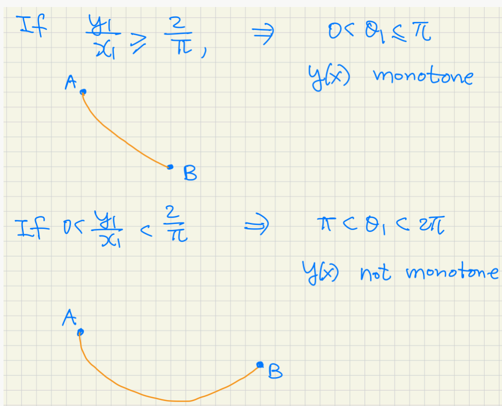
This is the trajectory of a spot on a rolling wheel.



AB has slope $\frac{y_1}{x_1} = m(\theta) = \frac{y(\theta)}{x(\theta)} \in (0, \infty)$.

For each $m \in (0, \infty)$, there is a unique θ_1 such that $\frac{y_1}{x_1} = m(\theta_1)$.

After θ_1 is found, we can solve for R.



W10C2 Lecture 16 (Mar 13)

4.5 Functions of Several Variables

Let $\Omega \subset \mathbb{R}^n$ be a bounded domain, $n = 2, 3$. Consider the minimization problem of:

$J[y] = \int_\Omega F(x, u, \nabla u)\, dx$, among $\{u : \Omega \to \mathbb{R}, u \in C^2, u|_{\partial\Omega} = g\}$

where the Lagrangian is:

$F(x, y, p) : \Omega \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$, and $x = (x_1, \ldots x_n) \in \Omega, p = (p_1, \ldots, p_n)$

💡 **Ex 14.** Let $\Omega = \mathbb{R}^3$, and consider $J[y] = \int_\Omega (\frac{1}{2}|\nabla u|^2 + q(x)u(x))\, dx$, where $u \in C^1(\overline{\Omega})$ and $u|_{\partial\Omega} = g$.

We have $F(x, u, p) = \frac{1}{2}|p|^2 + q(x)u$.

The necessary condition **Thm 1** that $\delta J[u; h] = 0\, \forall\, h$ is still valid, but **Thm 2** (Euler-Lagrange equation) only applies to $x \in \mathbb{R}$ and needs modification.

📖 **Lemma 3.** Let $x \in \Omega \subset \mathbb{R}^n$. If $\alpha(x) \in C(\overline{\Omega})$ and $\int_\Omega \alpha(x)h(x)\, dx = 0$ for all $h \in C^2(\overline{\Omega})$ with $h|_{\partial\Omega} = 0$, then $\alpha(x) = 0$ in Ω.

Proof:

Suppose $\alpha(x_0) > 0$ for some $x_0 \in \Omega$. Then $\alpha(x) > 0$ in some ball with radius $\varepsilon > 0$

$B_\varepsilon(x_0) : |x - x_0| < \varepsilon \subset \Omega$.

Let $h(x) = \begin{cases} (\varepsilon^2 - |x - x_0|^2)^3 & |x - x_0| < \varepsilon \\ 0 & \text{else} \end{cases}$

Then $\int_\Omega \alpha h\, dx = \int_{B_\varepsilon(x_0)} \alpha h\, dx > 0$.

This contradiction shows that $\alpha(x) \leq 0$. Likewise, we can show that $\alpha(x) \geq 0$, so $\alpha(x) = 0$.

Derivation of Euler-Lagrange Equation

To compute the variation $\delta J[u; h]$, let $h(x) \in C^2(\overline{\Omega})$ with $h|_{\partial\Omega} = 0$.

$\Delta J = J[u + h] - J[u] = \int_\Omega [F(x, u + h, \nabla u + \nabla h) - F(x, u, \nabla u)]\, dx$.

By a Taylor expansion, we have:

$\Delta J = \int_\Omega [F_u h + F_{p_1}\partial_1 h + \cdots + F_{p_n}\partial_n h]\, dx + \int_\Omega (\text{higher order terms})\, dx$

Hence, $\delta J[h] = \int_\Omega [F_u h + \sum_{k=1}^n F_{p_k}\partial_k h]\, dx$.

To remove the ∂_k on h, perform IBP by using the divergence theorem:

$$\begin{aligned} \delta J[y] &= \int_\Omega [F_u h + \textstyle\sum_{k=1}^n \partial_{x_k}(F_{p_k} h) - \sum_{k=1}^n (\partial_{x_k} F_{p_k})h]\, dx \\ &= \int_\Omega [F_u - \textstyle\sum_{k=1}^n \partial_{x_k} F_{p_k}]h\, dx + \cancel{\int_{\partial\Omega} \hat{n} \cdot F_{p_k} h\, dS} \end{aligned}$$

where $\hat{n}$ is the unit outer normal and the second integral vanishes because $h = 0$ on the boundary.

We recover the Euler-Lagrange equation for higher dimensions:

$F_u - \sum_{k=1}^n \partial_{x_k}(F_{p_k}(x, u, \nabla u)) = 0$.

Ex 14. (Revisit) Let $\Omega = \mathbb{R}^3$, and consider $J[y] = \int_\Omega (\frac{1}{2}|\nabla u|^2 + q(x)u(x))\,dx$, where $u \in C^1(\overline{\Omega})$ and $u|_{\partial\Omega} = g$.

We have $F(x, u, p) = \frac{1}{2}|p|^2 + q(x)u$.

We have $F_u = q(x)$ And $F_{p_k} = p_k$.

The Euler-Lagrange equation becomes:

$0 = q(x) - \sum_{k=1}^{3} \partial_k(\partial_k u) = q(x) - \Delta u$ (we recover Laplace's equation)

Ex 15. (minimal surface problem)

The surface area of a membrane $u(x, y) : \Omega \to \mathbb{R}, \quad \Omega \subset \mathbb{R}^2$ is given by:

$J[u] = \int_\Omega \sqrt{1 + u_x^2 + u_y^2}\,dx\,dy.$

We want to minimize $J[y]$ subject to the given boundary height: $u|_{\partial\Omega} = g$.

The E-L equation gives:

$F_u - \partial_x F_{p_1} - \partial_y F_{p_2} = 0.$

$$
\begin{aligned}
0 &= -0 + \partial_x\left(\frac{u_x}{\sqrt{1+u_x^2+u_y^2}}\right) + \partial_y\left(\frac{u_y}{\sqrt{1+u_x^2+u_y^2}}\right) \\
&= \frac{u_{xx}}{\sqrt{}} - \frac{u_x}{2\sqrt{}^3} \cdot 2(u_x u_{xx} + u_y u_{yx}) + \frac{u_{yy}}{\sqrt{}} - \frac{u_y}{2\sqrt{}^3} \cdot 2(u_x u_{xy} + u_y u_{yy}) \\
&= \frac{1}{\sqrt{}^3}\begin{pmatrix} u_{xx}(1 + u_x^2 + u_y^2 - u_x^2) \\ +u_{xy}(-u_x u_y - u_y u_x) \\ +u_{yy}(1 + u_x^2 + u_y^2 - u_y^2) \end{pmatrix}
\end{aligned}
$$

Our resulting PDE is:

$(1 + u_y^2)u_{xx} - 2u_x u_y u_{xy} + (1 + u_x^2)u_{yy} = 0.$

Geometric meaning (MATH 424):

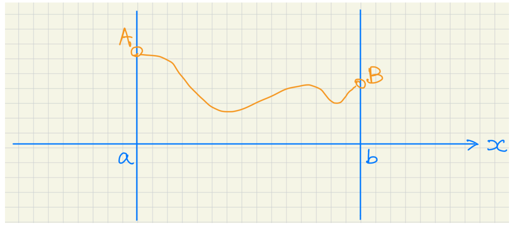
$H = \frac{1}{2}(\kappa_1 + \kappa_2) = \frac{(1+u_y^2)u_x x - 2u_x u_y u_{xy} + (1+u_x^2)u_{yy}}{\sqrt{1+u_x^2+u_y^2}}$ is the mean curvature.

For the surface, the intersection curve with any normal plane has curvature $\kappa_1 = \max \kappa$ and $\kappa_2 = \min \kappa$.

A surface with zero mean curvature $H = 0$ is called a minimal surface.

4.6 Variable Endpoint Problems

We want to minimize $J[y] = \int_a^b F(x, y, y')\,dx$ among $y \in C^1([a, b])$ without imposing boundary conditions.



In other words, the end points are allowed to slide up and down, and we now have a larger admissible set of $y(x)$.

An admissible increment h satisfies: $h \in C^1([a, b])$, and there are no boundary conditions for h.

As before, the principle linear part for ΔJ is:

$\delta J[h] = \int_a^b (F_y h + F_z h') \, dx.$

However, we gain boundary terms by IBP:

$\delta J[h] = \int_a^b (F_y - \frac{d}{dx} F_z) h(x) \, dx + [F_z h]_{x=a}^b$

If y is an extrema, then $\delta J = 0$. In particular, $\delta J[h] = 0$ for all $h \in C^1([a, b])$ with $h(a) = h(b) = 0$ (a subset of admissible h).

Hence, $y(x)$ still satisfies the Euler-Lagrange equation: $F_y - \frac{d}{dx} F_z(x, y, y') = 0.$

Hence for all $h \in C^1([a, b])$, we have $0 = [F_z h]_{x=a}^b = F_z h|_{x=b} - F_z h|_{x=a}$

Since $h(a)$ and $h(b)$ are arbitrary, we get $F_z|_{x=a} = 0$ and $F_z|_{x=b} = 0$. These are the new boundary conditions for $y(x)$.

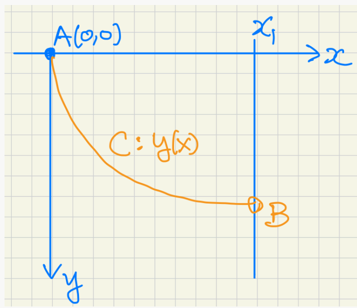
Mixed Case Boundary Conditions

We can also consider a mixed case: where one end is fixed (say at $x = a$) and the other end is variable.

Then the boundary condition is $y(a) = y_a, F_z(b, y(b), y'(b)) = 0$>

Ex 16. (Brachistochrone variant): we fix the point A but allow B on a line $x = x_1$.

We want to find B and the curve with least time: $J[y] = \int_0^{x_1} \frac{\sqrt{1+y'^2}}{\sqrt{y}}\, dx$.



Since $y(x)$ satisfies the E-L equation, from **Ex 13**, we have

$\vec{r}(\theta) = (x(\theta), y(\theta)) = (R(\theta - \sin\theta) + x_0, R(1 - \cos\theta))$.

$A = \vec{r}(\theta_0) \implies \theta_0 = 0, x_0 = 0$.

To find $B = \vec{r}(\theta_1)$, we need:

$0 = F_z|_{x=x_1} = \frac{y'}{\sqrt{y(1+y'^2)}}|_{x=x_1}$.

Hence, $y'(\theta_1) = 0$, so $\theta_1 = \pi$.

$\vec{r}(\theta_1) = (R(\pi - 0), R(1 - (-1)) = (x_1, y_1)$.

Thus, $R = \frac{x_1}{\pi}$ and $y_1 = 2R = \frac{2x_1}{\pi}$.

The trajectory is $\vec{r}(\theta) = \frac{x_1}{\pi}(\theta - \sin\theta, 1 - \cos\theta), 0 \le \theta \le \pi$.

W11C1 Lecture 17 (Mar 18)

Remark on Extensions:

a. Higher dimensions: $J[y] = \int_\Omega F(x, y, \nabla u)\, dx$
   - Variable BC: no BC specified
   - Boundary integral $J[y] = \int_\Omega F(x, y\nabla u)\, dx + \int_{\partial\Omega} G(x, y, \nabla u)\, dx$ (HW7 Q2)

b. Higher order derivatives (skipped):
   - $J[y] = \int_a^b F(x, y, y', y'')\, dx$
   - $J[u] = \int_\Omega |\Delta u|^2 + q(x)u\, dx$ among $u \in C^4$, $u = \frac{\partial u}{\partial n} = 0$ on $\partial\Omega$.

c. Vector valued functions (skipped):

   $u = (u_1, \dots, u_m) : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$

   $J[u] = \int_\Omega F(\underbrace{x}_{\mathbb{R}^m}, \underbrace{u}_{\mathbb{R}^m}, \underbrace{\nabla u}_{\in\mathbb{R}^{n\times m}})\, dx$

   Need for geodesics: minimal length curve on a surface between 2 points (MATH 424). See Gelfand-Fomin Ex 2, p. 49.

4.7 Variational Problems with Constraints

We studied the minimization of $J[u] = \int_\Omega F(x, u, \nabla u)\, dx$ in the admissible class $\mathcal{A} = \{u \in C^1(\overline{\Omega}), u|_{\partial\Omega} = g\}$.

In addition to boundary conditions, we may add other conditions (constraints) to $\mathcal{A}$. For example, we may further impose $M[u] = \int_\Omega G(x, u, \nabla u)\, dx = m_0$.

💡 **Ex 17.** (Isoperimetric problem)

a. Among all closed curves of a given length ℓ, find the curve enclosing the greatest area (Ex 3 of Lecture 12)
b. Maximize $\int_a^b y(x)\, dx$ subject to $y(a) = y_0, y(b) = y_1$, and $\int_a^b \sqrt{1 + y'^2(x)}\, dx = \ell$.

To state the theorem for constrained minimization, we introduce for intuition and convenience the variational derivative.

Recall the principle linear part of ΔJ is

$\delta J[y; h] = \int_a^b (F_y h + F_z h')\, dx = \int_a^b (F_y - \frac{d}{dx} F_z) h\, dx = \langle \frac{\delta J}{\delta y}, h \rangle$

📌 **Def.**

$\frac{\delta J}{\delta y} = J'[y] = F_y - \frac{d}{dx} F_z$ is the variational derivative of J at y. It is the part of δJ without h.

Remarks:

1. In Calc 1, $\Delta f = f(x+h) - f(x) = f'(x)h + \text{h.o.t.}$

For higher dimensions,

$\delta J[u;h] = \int_\Omega (F_u h + \sum_k F_{p_k} \partial_{x_k} h)\, dx = \int_\Omega \frac{\delta J}{\delta u} h\, dx$

where $\frac{\delta J}{\delta u} = F_u - \sum_k \partial_{x_k} F_{p_k}$

2. We computed $\delta J[u;h]$ by Taylor's expansion. It can also be computed as a directional derivative (HW6 Q3) $\delta J[u;h] = \lim_{t\to 0} \frac{1}{t}\{J[u+th] - J[u]\}$, which is weaker

📖 **Thm 4.** If $u^*(x)$ is an extremal of $J[u] = \int_\Omega F(x, u, \nabla u)\, dx$ in

$\mathcal{A} = \{u \in C^2(\overline{\Omega}), u|_{\partial\Omega} = g, M[u] = \int_\Omega G(x, y, \nabla u)\, dx = m_0\}$

and $\frac{\delta M}{\delta u}(u^*) \neq 0$. Then there exists a constant λ such that

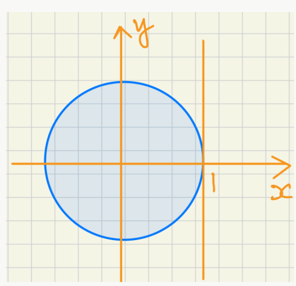
$\frac{\delta J}{\delta u}(u^*) + \lambda \frac{\delta M}{\delta u}(u^*) = 0.$

λ is called a Lagrange multiplier.

💡 **Ex 18.**

In Calc 3, if $p = (x_0, y_0)$ minimizes $f(x,y)$ on the curve $g(x,y) = m$ and $\nabla g(p) \neq 0$, then $\nabla f(p) + \lambda \nabla g(p) = 0$ for some $\lambda \in \mathbb{R}$.

a. $\min_{x-1=0} x^2 + y^2 = 1$ occurs at $p = (1,0)$, with $\nabla f = (2,0), \nabla g = (1,0), \nabla f + (-1)\nabla g = 0$



b. $\min_{(x-1)^4=0} x^2 + y^2 = 1$ occurs at $p = (1,0)$, but $(x-1)^4 = 0$. $\nabla f = (2,0), \nabla g = (0,0)$ and λ does not exist. We need the condition $\nabla g(p) \neq 0$.

Proof of Thm 4:

Let u^* be an extremal. Let $\{u_\varepsilon(x)\}_{-\varepsilon_1 < \varepsilon < \varepsilon_1} \subset \mathcal{A}$ be a one-parameter family of functions in $\mathcal{A}$, with $u_0 = u^*$. We have $u_\varepsilon|_{\partial\Omega} = g, M[u_\varepsilon] = m_0$ for all ε. u_ε is a perturbation of u^*.

Let $\xi(x) = \frac{d}{d\varepsilon}|_{\varepsilon=0} u_\varepsilon(x)$.

Since u^* is an extremal in $\mathcal{A}$,

$0 = \frac{d}{d\varepsilon}|_{\varepsilon=0} J[u_\varepsilon] = \frac{d}{d\varepsilon}|_{\varepsilon=0} \int_\Omega F(x, u_\varepsilon, \nabla u_\varepsilon)\, dx = \int_\Omega \frac{\delta J}{\delta y}[u^*]\xi\, dx.$

If ξ were arbitrary, then $\frac{\delta J}{\delta y} = 0$. But ξ is not arbitrary: it needs to ensure the existence of $\{u_\varepsilon\}_\varepsilon$ such that $M[u_\varepsilon] = m_0$. A similar calculation gives

$0 = \frac{d}{d\varepsilon}|_{\varepsilon=0} M[u_\varepsilon] = \int_\Omega \frac{\delta M}{\delta u}[u^*]\xi\, dx.$

Denote $(u,v) = \int_\Omega u(x)v(x)\, dx$, and let $f = \frac{\delta J}{\delta u}[u^*]$ and $g = \frac{\delta M}{\delta u}[u^*]$.

The last condition is $(g,\xi) = 0, g \perp \xi$ (ξ is tangent to the level set $M[u] = m_0$, and g is normal).

Claim 1: when $(g,\xi) = 0$, there is a family $\{u_\varepsilon\}_\varepsilon \subset \mathcal{A}$ such that $M[u_\varepsilon] = m_0$.

Proof of Claim 1:

Since $g \neq 0$ by assumption, there is $B_\varepsilon(x_0) \subset \Omega$ such that $g(x) \neq 0$ for $x \in B_\varepsilon(x_0)$. Hence $g(x)$ has constant sign (always positive or always negative).

Fix $\zeta(x) = \begin{cases} (\varepsilon^2 - |x-x_0|^2)^3 & B_\varepsilon(x_0) \\ 0 & \text{else} \end{cases}$

Hence $(g,\zeta) \neq 0$.

Try the correction $u_\varepsilon = u^* + \varepsilon\xi + \delta\zeta$ for $|\delta| \ll |\varepsilon| \ll 1$.

$M[u_\varepsilon] - M[u^*] = \int_\Omega g(\varepsilon\xi + \delta\zeta) + O(\varepsilon^2 + \delta^2)\, dx = 0 + \delta\underbrace{(g,\xi)}_{\neq 0} + \int_\Omega O(\varepsilon^2+\delta^2)\, dx$

Hence $\delta = O(\varepsilon^2)$ can be solved by the mean value theorem, which proves the claim.

Assume Claim 1. If $(g,\xi) = 0$, then u_ε exists, so $(f,\xi) = 0$.

For any $\eta \in C^2(\Omega), \eta|_{\partial\Omega} = 0$, decompose $\eta = ag + \tilde{\eta}$, where $a = \frac{(\eta,g)}{(g,g)}$. Then $\tilde{\eta} \perp g$.

By the previous conclusion, $0 = \int_\Omega f(x)\tilde{\eta}(x)\, dx = (f,\eta) - (f,ag) = (f - \frac{(f,g)}{(g,g)}g, \eta)$

Let $\lambda = -\frac{(f,g)}{(g,g)}$, then $(f + \lambda g, \eta) = 0$ for all η. By Lemma 1, $f + \lambda g = 0$, which completes the proof.

💡 **Ex 17(b).** (again). To maximize $J[y] = \int_a^b y\, dx$ among $y(a) = y_0, y(b) = y_1, M[y] = \int_a^b \sqrt{1+y'^2}\, dx = \ell$, we have the E-L equation $J'[y] + \lambda L'[y] = 0$.

$1 + \lambda(-\frac{d}{dx}(\frac{y'}{\sqrt{1+y'^2}})) = 0 \implies \frac{y'}{\sqrt{1+y'^2}} = cx + d$ where $c = \frac{1}{\lambda}$.

$$\begin{aligned} (y')^2 &= (cx+d)^2(1+y'^2) \\ (1-(cx+d)^2)y'^2 &= (cx+d)^2 \\ y' &= \frac{cx+d}{\sqrt{1-(cx+d)^2}} \\ y &= \int \frac{cx+d}{\sqrt{1-(cx+d)^2}}\, dx \\ &= -\frac{1}{c}\sqrt{1-(cx+d)^2} + c_1 \end{aligned}$$

We get $(y-c_1)^2 + (x + \frac{d}{c})^2 = \frac{1}{c^2}$, a circle. We can solve for c, d, c_1 using the three conditions. The special case $y_0 = y_1 = 0$ is similar to **Ex 19**.

W11C2 Lecture 18 (Mar 20)

💡 **Ex 19.** (Minimizing surface area under a fixed volume). Let $\Omega \subset \mathbb{R}^2$ bounded, and $u : \Omega \to \mathbb{R}, u \geq 0$
.

Minimize $J[y] = \int_\Omega \sqrt{1 + |\nabla u|^2}\, dx$ among $u|_{\partial\Omega} = 0, M[u] = \int_\Omega u\, dx = m_0$.

The E-L equation for some $\lambda \in \mathbb{R}$ is:

$$\begin{aligned} \tfrac{\delta J}{\delta u} + \lambda \tfrac{\delta M}{\delta u} &= 0 \\ F_u - \sum_i \partial_{x_i} F_{p_i} + \lambda(G_u - \sum_i \partial_{x_i} G_{p_i}) &= 0 \\ 0 - \sum \partial x_i \tfrac{\partial_i u}{\sqrt{1+|\nabla u|^2}} + \lambda(1-0) &= 0 \\ \operatorname{div} \tfrac{\nabla u}{\sqrt{1+|\nabla u|^2}} &= \lambda \end{aligned}$$

We consider the special case $\Omega = B_r(0)$.

We claim that $u(x) = \sqrt{R^2 - |x|^2} - \sqrt{R^2 - r^2}, R \geq r$ is a solution to the E-L equation.

The choice of R makes $M[u] = m_0$ possible if m_0 is not too large.

If $m_0 \sim 0$ then take $R \sim \infty$ and we get $\min J[u] \sim \pi r^2$

If $m_0 \uparrow$ then take $R = r$.