Math 4250 Minihomework: The Calculus of Variations

In this minihomework, we'll practice the techniques in our notes on the calculus of variations, giving proofs of some geometric facts that you've probably known for some time (but have never been able to prove!) and showing how these geometric techniques apply to some other areas of math. We will give some useful applications of these methods in physics, too. Remember that geometry is the study of shapes so there's really no area of math (or science) which is out of bounds for us as long as it involves shape.

1. (10 points) In the notes, while deriving the solution for the brachistochrone problem, we claimed that the solution to the (autonomous) differential equation

$$y'(x) = \sqrt{\frac{C - y(x)}{y(x)}}$$

with boundary conditions $y(0) = 0$ and $y(a) = b$ has the solution

$$x(t) = C\left(t - \frac{1}{2}\sin 2t\right).$$
$$y(t) = \frac{C}{2}\left(1 - \cos 2t\right).$$

where C is chosen so that the curve passes through (a, b). Prove it by solving the differential equation using our method for solving autonomous ODEs. (For less credit, check it by substituting in and differentiating.)

2. (10 points) You learned in MATH 2700 to solve ordinary differential equations by *separation of variables*. Let's refresh our memory. Suppose we have a differential equation such as

$$y'(x) = \frac{5x}{y}.$$

This is not an autonomous differential equation[1], so we can't use our previous method (from the last homework). However, we can rearrange the equation to get all y terms on the left hand side and all x terms on the right.

$$y(x)y'(x) = 5x$$

Now we can integrate both sides with respect to x,

$$\int y(x)y'(x)\,dx = \int 5x\,dx,$$

and change variables on the left hand side using $dy = y'(x)dx$,

$$\int y\,dy = \int 5x\,dx,$$

and do the integrals,

$$\frac{y^2}{2} = \frac{5}{2}x^2 + C,$$

and finally solve for $y(x)$,

$$y(x) = \pm\sqrt{5x^2 + C},$$

to get the final answer. This is the most general solution (as we've included the constant of integration). Now practice this method by finding the most general solution to

$$y'(x) = \frac{2xy}{1 + x^2}.$$

[1]because the right hand side involves x

3. (10 points) (The multivariable chain rule)

Definition. *If* $f\colon \mathbb{R}^m \to \mathbb{R}^n$*, then we can write*

$$f(\vec{x}) = (f_1(x_1, \ldots, x_m), \ldots, f_n(x_1, \ldots, x_m)).$$

The differential *or* Jacobian Df *of* f *at a point* $\vec{x} \in \mathbb{R}^n$ *is the* $n \times m$ *matrix:*

$$Df(\vec{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\vec{x}) & \cdots & \frac{\partial f_1}{\partial x_m}(\vec{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(\vec{x}) & \cdots & \frac{\partial f_n}{\partial x_m}(\vec{x}) \end{pmatrix}$$

A helpful rule for remembering the dimensions of the matrix Df is to remember that since $f\colon \mathbb{R}^m \to \mathbb{R}^n$, the differential $Df\colon \mathbb{R}^m \to \mathbb{R}^n$. We can now state:

Theorem (The multivariable chain rule)**.** *If* $f\colon \mathbb{R}^m \to \mathbb{R}^n$ *and* $g\colon \mathbb{R}^k \to \mathbb{R}^m$ *then the differential of the composition* $f(g(\vec{x}))$ *can be written*[2]

$$D(f(g(\vec{x}))) = Df(g(\vec{x}))Dg(\vec{x})$$

Here the right hand side is the $n \times k$ *matrix product of the* $n \times m$ *matrix* $Df(g(\vec{x}))$ *with the* $m \times k$ *matrix* $Dg(\vec{x})$*. The left hand side is the* $n \times k$ *matrix of partials of the composition* $f \circ g\colon \mathbb{R}^k \to \mathbb{R}^n$*.*

The case where $k = n = 1$ is particularly useful for us. Here $f\colon \mathbb{R}^m \to \mathbb{R}^1$ can be written $f(\vec{x}) = f(x_1, \ldots, x_m)$ while $g\colon \mathbb{R}^1 \to \mathbb{R}^m$ can be written as $g(x) = (g_1(x), \ldots, g_m(x))$. In this case,[3]

$$\begin{aligned} \frac{d}{dx} f(g(x)) &= Df(g(x))Dg(x) \\ &= \begin{pmatrix} \frac{\partial f}{\partial x_1}(g(x)) & \cdots & \frac{\partial f}{\partial x_m}(g(x)) \end{pmatrix} \begin{pmatrix} \frac{d}{dx} g_1(x) \\ \vdots \\ \frac{d}{dx} g_m(x) \end{pmatrix} \\ &= \frac{\partial f}{\partial x_1}(g(x))g_1'(x) + \cdots + \frac{\partial f}{\partial x_m}(g(x))g_m'(x). \end{aligned} \tag{2}$$

[2]This looks a lot like the ordinary chain rule $\frac{d}{dx} f(g(x)) = f'(g(x))g'(x)$ and it reduces to this in the case where $m = n = k = 1$.

[3]Note that we used this equation in the proof of the Euler-Lagrange equation when we computed

$$\begin{aligned} \frac{d}{d\epsilon} f(x, y(x) + \epsilon v(x), y'(x) + \epsilon v'(x)) = &\left(\frac{\partial f}{\partial y}(x, y(x) + \epsilon v(x), y'(x) + \epsilon v'(x))\right) v(x) \\ &+ \left(\frac{\partial f}{\partial y'}(x, y(x) + \epsilon v(x), y'(x) + \epsilon v'(x))\right) v'(x). \end{aligned}$$

Here we were composing $f\colon \mathbb{R}^3 \to \mathbb{R}^1$ with $g(\epsilon) = (x, y(x) + \epsilon v(x), y'(x) + \epsilon v'(x))$ and computing the derivative $\frac{d}{d\epsilon} f(g(\epsilon))$. We did not include the first term $\left(\frac{\partial f}{\partial x}(x, y(x) + \epsilon v(x), y'(x) + \epsilon v'(x))\right) \frac{d}{d\epsilon} x$ on the right hand side because we knew that $\frac{d}{d\epsilon} x = 0$.

(1) (5 points) Suppose that $f\colon \mathbb{R}^3 \to \mathbb{R}^1$ is given by $f(x_1, x_2, x_3) = x_1^2 x_2 + 7x_3$ while $g\colon \mathbb{R}^1 \to \mathbb{R}^3$ is given by $g(x) = (\cos x, \sin x, x)$. Find the derivative $\frac{d}{dx} f(g(x))$ using (2).

(2) (5 points) Let $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ be $f(x_1, x_2) = (x_1, x_2, \sqrt{1 - x_1^2 - x_2^2})$ while $g\colon \mathbb{R}^2 \to \mathbb{R}^2$ be $g(r, \theta) = (r\cos\theta, r\sin\theta)$. Find the differential $D(f(g(r, \theta)))$ using the (general) multivariable chain rule.

(3) (5 points) Suppose that we have any $f\colon \mathbb{R}^3 \to \mathbb{R}^1$ and another function[4] $g\colon \mathbb{R} \to \mathbb{R}^3$ given by $g(x) = (x, y(x), y'(x))$ for some function $y\colon \mathbb{R} \to \mathbb{R}$. Find the derivative

$$\frac{d}{dx} f(g(x)) = \frac{d}{dx} f(x, y(x), y'(x))$$

using (2). Please explain your work carefully. It's ok to write the partials of f as $\frac{\partial f}{\partial x_1}$, $\frac{\partial f}{\partial x_2}$ and $\frac{\partial f}{\partial x_3}$ if you want to.

[4] Sometimes called the "1-jet" of y at x, because it carries information about y up to the 1st derivative.

4. (10 points) Fermat's principle says that a light ray travels along the path requiring the least time to join endpoints. Further, the speed of light changes with the density of the medium, slowing as the medium becomes more dense. These observations are combined into the following:

Definition. *The* optical length *of a path $\vec{\alpha}(s)$ taken by a ray of length from $\vec{\alpha}(0)$ to $\vec{\alpha}(\ell)$ is given by*

$$F(\vec{\alpha}) = \int_0^\ell n(\alpha(\vec{s}))\, ds,$$

where $n(\vec{x})$ is the index of refraction of the medium at the point $\vec{x}$ in space.

and

Theorem. *Light follows the path of least optical length between $\vec{\alpha}(0)$ and $\vec{\alpha}(\ell)$.*

(1) (5 points) Suppose that we have a parametrization $\vec{\alpha}(t)$ which is not by arclength. Rewrite the optical length functional $F(\vec{\alpha})$ as an integral with respect to t using our formula

$$s(t) = \int_0^t \|\vec{\alpha}'(x)\|\, dx$$

and the fact that $ds = s'(t)dt$.

(2) (5 points) Now suppose that $\vec{\alpha}(x) = (x, y(x))$ and $n(x, y) = \frac{1}{y}$. Over the next few parts, we will derive a formula for the path of the light ray from $(0, 1)$ to $(1, 1)$.

Write the optical length functional as an integral in the form

$$F(y) = \int_0^1 f(x, y, y')\, dx.$$

and then write down the Euler-Lagrange equation

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

Find a common denominator and simplify carefully (and a lot!). It's certainly ok to use a computer algebra system (eg. Mathematica) if you want to.

(3) (5 points) Solve the Euler-Lagrange equation.

This is a two-step process. You'll be able to write the Euler-Lagrange equation in the form $\frac{d}{dx}g(x, y(x), y'(x)) = 0$ for some function g.

You then have to solve the *first order* differential equation $g(x, y(x), y'(x)) = C$, which you can do by separation of variables.

(4) (5 points) Your solution to the Euler-Lagrange equation should involve two unknown constants.

- Find values for the constants so that $y(0) = 1$ and $y(1) = 1$.
- Write down your final solution and draw a sketch of the path of the light ray.
- An observer located at $(0, 1)$ looking in the direction of $(1, 1)$ will not see an obstacle located at $(0.5, 1)$. Why not?[5]

[5] Physicists use this idea to build "optical cloaking" devices. See `https://www.youtube.com/watch?v=oJb9RnAVDuE` for a really great demonstration.

5. (15 points) (The Hamiltonian) In the last problem, the key step in solving the second-order Euler-Lagrange equation was to reduce it to a first-order differential equation by realizing that the Euler-Lagrange equation could be written in the form $\frac{d}{dx}g(x, y, y') = 0$. We'll now see that this is a generally useful trick!

Definition. *Given a functional* $F(y) = \int f(x, y, y')\, dx$*, if* $f(x, y, y')$ *can be written as* $f(y, y')$[6] *then we define the* Hamiltonian $\mathcal{H}(y, y')$ *by*

$$\mathcal{H}(y, y') = f(y, y') - y' \frac{\partial}{\partial y'} f(y, y')$$

Prove that if $y(x)$ is a solution for the Euler-Lagrange equation

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

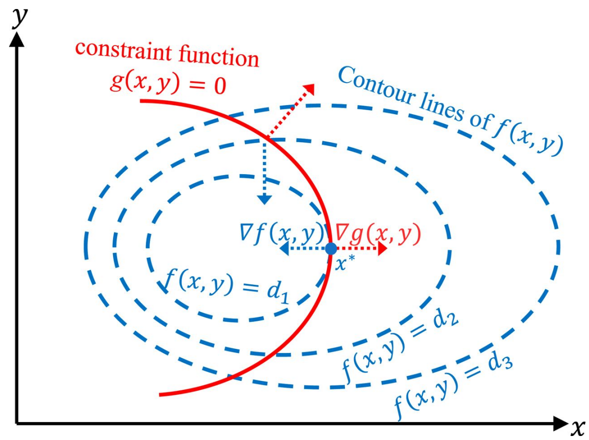
then $\frac{d}{dx}\mathcal{H}(y(x), y'(x)) = 0$ and hence $\mathcal{H}(y(x), y'(x))$ is equal to some[7] constant C for all values of x. It will help to review question 3 and in particular 3.3 when thinking about how to do the derivative.

[6] That is, x does not appear in the integrand for $F(y)$.

[7] There may be many solutions $y(x)$ to the Euler-Lagrange equation, since it is a differential equation and we haven't specified any initial conditions or boundary conditions. The value of C depends on which solution we pick.

6. (10 points) (The method of Lagrange Multipliers) At some point in your multivariable calculus course, you should have learned the method of Lagrange multipliers for solving optimization problems with constraints with many variables. We're now going to do a quick review of the topic in order to help you remember how it works.

Suppose you have a function $f(x, y)$ that you want to maximize along the level curve $g(x, y) = 0$. That is, you want to find the largest d so that the level curve $f(x, y) = d$ intersects the level curve $g(x, y)$. As the picture below[8] shows, this ought to happen when the two level curves are tangent to one another and ∇f is a scalar multiple of ∇g.



Theorem. *(Lagrange Multipliers theorem) If $f(x_0, y_0)$ is a maximum or minimum for $f(x, y)$ subject to the constraint $g(x, y) = 0$ and $\nabla g(x_0, y_0) \neq 0$, then there exists some λ_0 so that*

$$\nabla(f(x, y) + \lambda_0 g(x, y))|_{(x_0,y_0)} = 0, \quad \textit{or} \quad \nabla f(x_0, y_0) = \lambda_0 \nabla g(x_0, y_0).$$

The scalar λ_0 is called the "Lagrange multiplier" and the function $\mathcal{L}(x, y) = f(x, y) + \lambda g(x, y)$ is called the "Lagrangian".

(1) (5 points) Suppose that $f(x, y) = x + y$ and $g(x, y) = x^2 + y^2 - 1$. Write down a formula for the Lagrangian $\mathcal{L}(x, y)$ and find the gradient[9] $\nabla \mathcal{L}$.

[8] Copied from Vadlamani, S.K., Xiao, T.P. and Yablonovitch, E. (2020). PNAS, 117(43), 26639-26650.

[9] Remember, the gradient of a scalar function $f(x_1, \ldots, x_n)$ is the vector $(\partial/\partial x_1 f, \ldots, \partial/\partial x_n f)$.

(2) (5 points) Now find both solutions to the system of equations

$$\nabla\mathcal{L}(x_0, y_0) = (0,0) \quad \text{and} \quad g(x,y) = 0,$$

for the three unknowns x_0, y_0 and λ_0 and find the points on the circle $g(x,y) = 0$ which maximize and minimize the function $f(x,y)$.

7. (10 points) (Strictly convex and concave functions) If a function $f(x)$ has $f''(x) > 0$, we say that $f(x)$ is strictly convex (or strictly concave up). If $f(x)$ has $f''(x) < 0$, we say that $f(x)$ is strictly concave (or strictly concave down).

   Suppose that $f(x)$ is smooth and strictly convex or strictly concave. Prove that there do not exist three distinct points $x_1 < x_2 < x_3$ so that

   $$f(x_1) = f(x_2) = f(x_3)$$

   Hint: Rolle's theorem.

8. (15 points) (The hanging chain) Ever since I was a little kid, I've been fascinated by high-tension towers.[10] One of the things that I always admired were the graceful curves the cables make when stretched from tower to tower. This kind of problem is exactly what we study in differential geometry. What shape is that curve? And why?

In the right hand picture, you can see the problem more clearly. A cable or chain is suspended from two fixed anchors, which may as well be located at $(-1, 0)$ and $(1, 0)$. What shape does the chain make? In this question, we'll use the calculus of variations (together with Lagrange multipliers) to answer this question.

We'll need two ideas from physics:

1. The gravitational potential energy of a mass m is given by mgh, where h is the height of the mass and g is the (positive) acceleration of gravity.
2. The hanging chain is minimizing gravitational potential energy, subject to the constraint of fixed length.

The curve that minimizes mgh will also minimize mh, so we can ignore g.[11] Now the mass is the mass of the cable or chain itself, so it's proportional to length along the curve. This means that the functional to minimize is given by:

$$\operatorname{gp}(y) = \int_0^L y(s)\, ds = \int_{-1}^{1} y(x)\sqrt{1 + (y'(x))^2}\, dx$$

and there is a constraint functional

$$\operatorname{len}(y) = \int_{-1}^{1} \sqrt{1 + (y'(x))^2}\, dx = L.$$

We will now define a Lagrangian:

$$\mathcal{L}(y) = \operatorname{gp}(y) + \lambda \operatorname{len}(y)$$

[10]There are so many different designs and structures for the towers themselves: why didn't one "best" design ever win out over all the others?

[11]The other thing that mathermaticians tend to say in situations like this is "choose units so that $g = 1$". Which also works, but I never really liked it– after all, it assumes that there *are* such units. This is true, but unjustified. And anyway, nobody would want to do *that* unit conversion.

In light of Problem 5, because our Lagrangian $\mathcal{L}(y) = \mathrm{gp}(y) + \lambda\,\mathrm{len}(y)$ can be written in the form

$$\mathcal{L}(y) = \int y(x)\sqrt{1+y'(x)^2} + \lambda\sqrt{1+y'(x)^2}\,dx = \int f(y,y')\,dx.$$

instead of solving the second order Euler-Lagrange differential equation for $y(x)$, we can instead solve the first-order differential equation $\mathcal{H}(y,y') = c$.

(1) (5 points) Write down $\mathcal{H}(y,y') = f(y,y') - y'\frac{\partial}{\partial y'}f(y,y')$ and solve the equation

$$\mathcal{H}(y,y') = c$$

for $y'(x)$ to find a differential equation for the shape of the cable.

Note: You may assume that $c \neq 0$. (The case $c = 0$ is a little tricky, but ultimately doesn't contribute a meaningful solution.)

(2) (5 points) Now solve your differential equation for $y(x)$. It will be helpful to remember that $c \neq 0$ and so we have the integration formula

$$\int \frac{|c|}{\sqrt{x^2 - c^2}}\, dx = c \operatorname{arccosh}\left(\frac{x}{c}\right) + D.$$

(3) (5 points) Your solution to the last part should involve three unknown constants: L, C, and a new constant of integration D. Solve for these constants using the boundary and constraint equations:

$$y(-1) = 0, \quad y(1) = 0, \quad \int_{-1}^{1} \sqrt{1 + y'(x)^2}\, dx = L$$

Prove that no solutions exist with $L \leq 2$.

Notes: It's a lot easier to apply the first two equations to simplify $y(x)$ as much as you can before integrating to apply the last one. Remember that $c \neq 0$. It will help to use the result of 7.

9. (20 points) (The isoperimetric inequality, Extra Credit Problem[12]) Here's another classical variation problem. Among all curves joining $(-1,0)$ and $(0,1)$ enclosing area $\pi/2$ with the x-axis, which is the shortest? You can probably guess "the semicircle". Let's prove it! In this problem, we have a functional to minimize:

$$\mathrm{len}(y) = \int_{-1}^{1} \sqrt{1 + (y'(x))^2}\, dx$$

and a constraint functional

$$\mathrm{area}(y) = \int_{-1}^{1} y(x)\, dx = \frac{\pi}{2}.$$

Just as we did above, we'll define a Lagrangian:

$$\mathcal{L}(y) = \mathrm{len}(y) + \lambda\, \mathrm{area}(y).$$

(1) (10 points) Write down the Euler-Lagrange equation

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

and simplify as much as possible.

[12] This problem is optional, though very cool, and will be graded for bonus points on Gradescope.

(2) (10 points) The Euler-Lagrange equation involves only $y''(x)$ and $y'(x)$, so we can solve for $y'(x)$ either as an autonomous ODE or by separation of variables. Choose a method and solve, giving us (as usual) another first order ODE for $y'(x)$. Solve that, too, to get a solution involving λ and two constants of integration C and D. Last, apply the boundary conditions and the condition that the area under the curve is $\pi/2$ to get a final solution for $y(x)$.