Problem Set 3: Calculus of Variations I

1 Geodesics (10 points)

The shortest path between two points on a manifold is called a geodesic. In general, the arc length of a path γ is given by

$$S[\gamma] = \int_\gamma ds \tag{1}$$

where ds is the infinitesimal arc length. The geodesic is the path that extremizes the arc length. To make progress, we need to pick coordinates and a parametrization of the curve γ.

a) The usual xy plane with Cartesian coordinates (x, y) has metric $ds^2 = dx^2 + dy^2$. Suppose we parametrize γ by the function $y(x)$. Then

$$S[y(x)] = \int ds = \int dx\,\sqrt{dx^2 + dy^2} = \int dx\,\sqrt{1 + y'^2}\,. \tag{2}$$

Write down the Euler-Lagrange conditions for S and show that the shortest path between two fixed points in the plane is a straight line.

SOLUTION:
We are interested in extremizing S with respect to $y(x)$, fixed at an initial and final point. At an extremum, the function $f = \sqrt{1 + y'^2}$ must therefore satisfy the Euler-Lagrange equation,

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y} = 0 \implies \frac{d}{dx}\left(\frac{y'}{\sqrt{1 + y'^2}}\right) = 0\,. \tag{3}$$

Integrating once, we find that $y' = C\sqrt{1 + y'^2}$, which you can simply rearrange to find $y' = D$ for some other real constant D. This is the equation of a straight line.

b) Now repeat the previous analysis using a general parametrization $\vec{r}(t) = (x(t), y(t))$ with t running from 0 to 1. In this case, the arc length is given by

$$S[\vec{r}(t)] = \int_0^1 dt\,\sqrt{\dot{x}^2 + \dot{y}^2}\,. \tag{4}$$

Show from the two Euler-Lagrange equations that the shortest path is a straight line.

SOLUTION:
Defining $f(\dot{x}, \dot{y}) = \sqrt{\dot{x}^2 + \dot{y}^2}$, we find that the Euler-Lagrange equation for the x variable is

$$\frac{d}{dt}\left(\frac{\partial f}{\partial \dot{x}}\right) = 0 \implies \frac{d}{dt}\left(\frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2}}\right) = 0\,, \tag{5}$$

which we can integrate once to find that $\dot{x} = C\sqrt{\dot{x}^2 + \dot{y}^2}$, where C is a real constant. By symmetry, we also know that $\dot{y} = D\sqrt{\dot{x}^2 + \dot{y}^2}$ for another constant D. From this, we find

$$\frac{\dot{y}}{\dot{x}} = \frac{dy}{dx} = \frac{D}{C}\,, \tag{6}$$

which once again is the equation of a straight line.

c) Now suppose that the path γ is restricted to lie on the surface of a sphere of radius R. Without loss of generality, we can set $R = 1$. In spherical coordinates (r, θ, ϕ), the metric is

$$ds^2 = d\theta^2 + \sin^2\theta\, d\phi^2 \,. \tag{7}$$

Write the arc length functional $S[\phi(\theta)]$ expressing the path as a function of $\phi(\theta)$. Derive the Euler-Lagrange equation and show that the geodesics between any two points on the sphere are great circles. (*Hint:* you can assume that one of the two points is at the north pole so that the curve passes through $\theta = 0$. You shouldn't be integrating something really messy!)

SOLUTION:
The arc length functional is

$$S[\phi(\theta)] = \int_0^{\theta_f} d\theta \sqrt{1 + \sin^2\theta\, \phi'^2} \,, \tag{8}$$

where we have assumed that the curve passes through the north pole with $\theta = 0$, and hence at some other polar angle θ_f finally. Defining $f(\phi', \theta) = \sqrt{1 + \sin^2\theta\, \phi'^2}$, we find that the Euler-Lagrange equations are

$$\frac{\partial f}{\partial \phi} - \frac{d}{d\theta}\left(\frac{\partial f}{\partial \phi'}\right) = 0 \implies \frac{d}{d\theta}\left(\frac{\sin^2\theta\, \phi'}{\sqrt{1 + \sin^2\theta\, \phi'^2}}\right) = 0 \,. \tag{9}$$

Integrating once we find

$$\sin^2\theta\, \phi' = C\sqrt{1 + \sin^2\theta\, \phi'^2} \implies \phi'^2 = \frac{C^2}{\sin^2\theta(\sin^2\theta - C^2)} \,. \tag{10}$$

Let's take a closer look at this equation which, importantly, is satisfied everywhere along the trajectory we are interested in. The left-hand side $\phi'^2 > 0$. In order for the right-hand side to be greater than zero, we require $C^2 < \sin^2\theta$ *for all* θ, including at the end of the trajectory where $\theta \to 0$ (remember, we are assuming that the trajectory starts at $\theta = 0$). This is only possible if $C = 0$, in which case $\phi' = 0$. These are paths of constant longitude, which are great circles.

2 (SG 1.2) Fermat's Principle and Snell's Law (5 points)

A medium is characterized optically by its refractive index n, such that the speed of light in the medium c/n. According to Fermat's principle, the path taken by a ray of light between any two points makes the travel time stationary between those points. Assume that the ray propagates in the xy-plane in a layered medium with refractive index $n(x)$. Use Fermat's principle to show that $n(x)\sin\psi = \text{constant}$ for some angle ψ related to the trajectory, by finding the equation giving the stationary paths $y(x)$ for

$$F_1[y] = \int dx\, n(x)\sqrt{1 + y'^2} \,, \tag{11}$$

where the prime denotes differentiation with respect to x. Make sure to define ψ carefully. Repeat this exercise for the case that n depends only on y and find a similar equation for the stationary paths of

$$F_2[y] = \int dx\, n(y)\sqrt{1 + y'^2} \,, \tag{12}$$

i.e. show that $n(y)\sin\theta = \text{constant}$ for some angle other angle θ related to the trajectory. Explain how these two answers are physically consistent. In the second formulation you will find it easiest to use the first integral of the Euler-Lagrange equation.

SOLUTION:
The function $y(x)$ that extremizes F_1 satisfies the Euler-Lagrange equation, which in this case is

$$\frac{d}{dx}\left(\frac{n(x)y'}{\sqrt{1+y'^2}}\right) = 0 \implies \frac{n(x)y'}{\sqrt{1+y'^2}} = C\,, \tag{13}$$

for some constant C. Given any point (x, y), let the angle made between the tangent to $y(x)$ at that point and the x-axis be ψ. We know that ψ is related to the gradient by $\tan\theta = y'$, and so $\sin\psi = y'/\sqrt{1+y'^2}$. We therefore find that

$$n(x)\sin\psi = C\,, \tag{14}$$

In the second case, instead of using the Euler-Lagrange equation, we define $f(y, y') \equiv n(y)\sqrt{1+y'^2}$ and use the first integral of the Euler-Lagrange equation, which is

$$I = f - y'\frac{\partial f}{\partial y'} = n(y)\sqrt{1+y'^2} - \frac{n(y)y'^2}{\sqrt{1+y'^2}} \tag{15}$$

The condition that $dI/dx = 0$ therefore gives

$$\frac{n(y)}{\sqrt{1+y'^2}} = C\,, \tag{16}$$

for some real constant C. Let the angle made between the tangent to $y(x)$ at a point (x, y) and the y-axis be θ. Then, we have $n(y)\sin\theta = C$.
These two answers are physically consistent, since they both show that the sine of the angle of incidence—*i.e.* the angle between the trajectory and a normal drawn perpendicular to the direction in which the refractive index is changing—is the quantity that varies inversely with n.

3 Mass on a Spring (5 points)

A mass m hangs from a hook on a spring with spring constant k subject to gravity. Show that the Euler-Lagrange equation for the Lagrangian

$$L = T - V = \frac{1}{2}m\dot{z}^2 - \frac{1}{2}kz^2 - mgz \tag{17}$$

leads to the expected equations of motion. Show that the first integral, $E = \dot{z}(\partial L/\partial\dot{z}) - L$, is the total energy of the system, and that E is conserved.

SOLUTION:
The Euler-Lagrange equation for the Lagrangian reads

$$\frac{\partial L}{\partial z} - \frac{d}{dt}\left(\frac{\partial L}{\partial\dot{z}}\right) = 0 \implies m\ddot{z} = -kz - mg\,, \tag{18}$$

which is the expected equation of motion for a mass on a spring from Newton's second law (the force of gravity always accelerates the mass downward, while the spring force opposes the displacement from equilibrium).
The first integral is

$$E = \dot{z}\frac{\partial L}{\partial\dot{z}} - L = m\dot{z}^2 - \frac{1}{2}m\dot{z}^2 + \frac{1}{2}kz^2 + mgz = \frac{1}{2}m\dot{z}^2 + \frac{1}{2}kz^2 + mgz\,, \tag{19}$$

which is the sum of the kinetic, elastic potential and gravitational potential energy, and is therefore the total energy of the system. Taking the time derivative of E, we find

$$\frac{dE}{dt} = m\dot{z}\ddot{z} + kz\dot{z} + mg\dot{z} = \dot{z}(m\ddot{z} + kz + mg) = 0\,, \tag{20}$$

since the term in parentheses is zero by the equation of motion.

4 (SG 1.8) The Lorentz Force Law (10 points)

The Lagrangian for a particle of charge q and mass m in an electromagnetic field is

$$L(\vec{r}, \dot{\vec{r}}) = \frac{1}{2}m\dot{\vec{r}}^2 + q\dot{\vec{r}} \cdot \vec{A}(\vec{r}, t) - q\phi(\vec{r}, t)\,, \tag{21}$$

where $\phi(\vec{r}, t)$ and $\vec{A}(\vec{r}, t)$ are the scalar and vector potentials, respectively. Show that the Euler-Lagrangian equation for the action $S[\vec{r}(t)] = \int dt\, L$ leads to the Lorentz force law

$$m\ddot{\vec{r}} = q\left(\vec{E} + \dot{\vec{r}} \times \vec{B}\right)\,, \tag{22}$$

where

$$\vec{E} = -\nabla\phi - \frac{\partial \vec{A}}{\partial t} \quad \text{and} \quad \vec{B} = \nabla \times \vec{A}\,. \tag{23}$$

Working in index notation will probably be helpful here!

SOLUTION:
First, working in Einstein notation is much better, so let's just begin by rewriting the Lagrangian in Einstein notation, where repeated indices are summed over:

$$L = \frac{1}{2}m\dot{r}_i\dot{r}^i + q\dot{r}_iA^i - q\phi\,, \tag{24}$$

keeping in mind that A^i and ϕ are themselves functions of r^i. Note that in Euclidean space, there is no distinction between upper and lower indices. We can also note the following expressions for the magnetic field:

$$\begin{aligned} B_i = \epsilon_{ijk}\frac{\partial A^k}{\partial r_j} &\implies (\vec{n} \times B)_i = \epsilon_{ilm}n^lB^m = (\epsilon_{ilm}\epsilon^{mjk})n^l\frac{\partial A_k}{\partial r^j} \\ &\implies \left(\delta_i^j\delta_l^k - \delta_i^k\delta_l^j\right)n^l\frac{\partial A_k}{\partial r^j} = n^j\left(\frac{\partial A_j}{\partial r^i} - \frac{\partial A_i}{\partial r^j}\right)\,. \end{aligned} \tag{25}$$

where ϵ^{ijk} is the 3D Levi-Civita tensor, and $\vec{n}$ is any arbitrary vector. We used the identity $\epsilon_{ilm}\epsilon^{mjk} = \delta_i^j\delta_l^k - \delta_i^k\delta_l^j$ in the second line.
Let's write down the partial derivatives of L carefully to avoid confusion:

$$\frac{\partial L}{\partial r^i} = q\dot{r}_j\frac{\partial A^j}{\partial r^i} - q\frac{\partial \phi}{\partial r^i}\,, \qquad \frac{\partial L}{\partial \dot{r}^i} = m\dot{r}_i + qA_i\,, \tag{26}$$

so that the Euler-Lagrange equation reads

$$\frac{\partial L}{\partial r^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}^i}\right) = 0 \implies q\dot{r}_j\frac{\partial A^j}{\partial r^i} - q\frac{\partial \phi}{\partial r^i} - m\ddot{r}_i - q\left(\frac{\partial A_i}{\partial t} + \frac{\partial A_i}{\partial r^j}\dot{r}^j\right) = 0\,. \tag{27}$$

Keep in mind that we are taking a *total* derivative of A^i, and not just a partial derivative. Grouping terms together nicely, we find

$$\begin{aligned} m\ddot{r}_i &= -q\left(\frac{\partial \phi}{\partial r^i} + \frac{\partial A_i}{\partial t}\right) + q\dot{r}^j\left(\frac{\partial A_j}{\partial r^i} - \frac{\partial A_i}{\partial r^j}\right) \\ &= q\vec{E}_i + q(\dot{\vec{r}} \times \vec{B})_i \,, \end{aligned} \tag{28}$$

which is precisely the Lorentz force law.

5 (SG 1.9) Euler's Equations (10 points)

Consider the action functional

$$S(\vec{\omega}, \vec{p}, \vec{r}) = \int dt\, \left[\frac{1}{2}I_1\omega_1^2 + \frac{1}{2}I_2\omega_2^2 + \frac{1}{2}I_3\omega_3^2 + \vec{p}\cdot\left(\dot{\vec{r}} + \vec{\omega}\times\vec{r}\right)\right] , \tag{29}$$

where $\vec{r}$ and $\vec{p}$ are time-dependent 3-vectors, as is $\vec{\omega} = (\omega_1, \omega_2, \omega_3)$, and I_1, I_2 and I_3 are constants. Show that the Euler-Lagrange equations for $\vec{\omega}$, $\vec{p}$ and $\vec{r}$ lead to the equations

$$\begin{aligned} I_1\dot{\omega}_1 - (I_2 - I_3)\omega_2\omega_3 = 0 \,, \\ I_2\dot{\omega}_2 - (I_3 - I_1)\omega_3\omega_1 = 0 \,, \\ I_3\dot{\omega}_3 - (I_1 - I_2)\omega_1\omega_2 = 0 \,, \end{aligned} \tag{30}$$

governing the angular velocity of a freely rotating rigid body in the rotating body frame.

SOLUTION:
Let's rewrite the Lagrangian in index notation. This is

$$L(p_i; r_i, \dot{r}_i; \omega_i) = \frac{1}{2}J_{ij}\omega^i\omega^j + p_i\dot{r}^i + \epsilon^{ijk}p_i\omega_j r_k \,. \tag{31}$$

where $J_{11} = I_1$, $J_{22} = I_2$, $J_{33} = I_3$, and $J_{ij} = 0$ for $i \neq j$. Then, the Euler-Lagrange equations for each of the variables is

$$\begin{aligned} \frac{\partial L}{\partial p_i} &= \dot{r}^i + \epsilon^{ijk}\omega_j r_k = 0 \,, \\ \frac{\partial L}{\partial r_k} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}_k}\right) &= \epsilon^{ijk}p_i\omega_j - \dot{p}^k = 0 \,, \\ \frac{\partial L}{\partial \omega_j} &= J^{ij}\omega_i + \epsilon^{ijk}p_i r_k = 0 \,. \end{aligned} \tag{32}$$

Taking the time derivative of the last equation, we have

$$J^{ij}\dot{\omega}_i + \epsilon^{ijk}\dot{p}_i r_k + \epsilon^{ijk}p_i\dot{r}_k = 0 \,. \tag{33}$$

The second term combined with the second line of Eq. (32) gives

$$\epsilon^{ijk}\dot{p}_i r_k = \epsilon^{ijk}\epsilon_{ilm}p^l\omega^m r_k = (\delta^j_l\delta^k_m - \delta^j_m\delta^k_l)p^l\omega^m r_k = p^j\omega^k r_k - p^k\omega^j r_k \,, \tag{34}$$

while the last term combined with the first line of Eq. (32) gives

$$\begin{aligned} \epsilon^{ijk}p_i\dot{r}_k &= \epsilon^{ijk}p_i\left(-\epsilon_{klm}\omega^l r^m\right) \\ &= -\left(\delta^i_l\delta^j_m - \delta^i_m\delta^j_l\right)p_i\omega^l r^m \\ &= p_i\omega^j r^i - p_i\omega^i r^j \,. \end{aligned} \tag{35}$$

Putting everything together,

$$\begin{aligned} J^{ij}\dot{\omega}_i + p^j\omega^k r_k - p^k\omega^j r_k + p_i\omega^j r^i - p_i\omega^i r^j &= 0 \\ J^{ij}\dot{\omega}_i &= -p^j\omega^k r_k + p_i\omega^i r^j \\ &= \omega^k(p_k r^j - p^j r_k)\,. \end{aligned} \tag{36}$$

However, from the third equation of Eq. (32), we see that

$$\epsilon_{jlm}J^{ij}\omega_i = -\epsilon_{jlm}\epsilon^{ijk}p_i r_k = (\delta^i_l\delta^k_m - \delta^i_m\delta^k_l)p_i r_k = p_l r_m - p_m r_l\,, \tag{37}$$

or

$$p_k r^j - p^j r_k = \epsilon_{ik}{}^j J^{li}\omega_l\,. \tag{38}$$

so that we finally have

$$J^{ij}\dot{\omega}_i = \omega^k\epsilon_{ik}{}^j J^{li}\omega_l = \epsilon^{ikj}J_{li}\omega_k\omega^l\,. \tag{39}$$

Index-by-index, we have

$$\begin{aligned} I_1\dot{\omega}_1 &= \epsilon^{ik1}J_{li}\omega_k\omega^l = J_{l2}\omega_3\omega^l - J_{l3}\omega_2\omega^l = (I_2 - I_3)\omega_2\omega_3\,, \\ I_2\dot{\omega}_2 &= \epsilon^{ik2}J_{li}\omega_k\omega^l = J_{l3}\omega_1\omega^l - J_{l1}\omega_3\omega^l = (I_3 - I_1)\omega_3\omega_1\,, \\ I_3\dot{\omega}_3 &= \epsilon^{ik3}J_{li}\omega_k\omega^l = J_{l1}\omega_2\omega^l - J_{l2}\omega_1\omega^l = (I_1 - I_2)\omega_1\omega_2\,, \end{aligned} \tag{40}$$

using the fact that J_{ij} is diagonal. This gives the desired result.