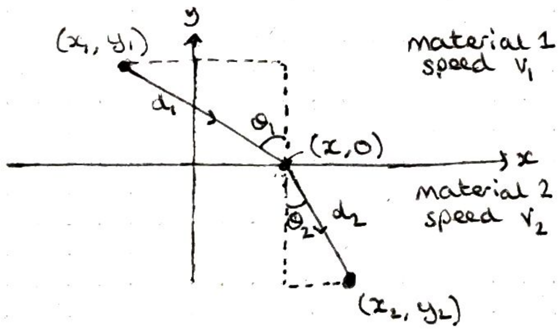
Least Action - Optics

Newton's three laws of dynamics excellently describe the motion of objects. However, we can arrive at the same result of these laws through the Principle of Least Action, which will make some dynamics problems easier to solve. This is a formalism which grew out of optics, so let's first consider it in this context.

> Fermat's Principle of Least Time: Light travels between two points so as to minimise its travel time.

We all know this law, and we see that in a uniform medium the time of travel is $t = \frac{d}{c}$ $d \to$ distance, $c \to$ speed of light

Now let's see if we can derive Snell's Law from this. Consider a beam of light propagating in two adjacent mediums:



The crossing point is $(x,0)$

The time of travel from (x_1, y_1) to (x_2, y_2) is given by

$$T[x] = \frac{d_1}{v_1} + \frac{d_2}{v_2}$$

Notice the square brackets. While this isn't strictly necessary here, it will be useful in later formulations, explained later.

$$\therefore \quad T[x] = \frac{\sqrt{(x-x_1)^2 + y_1^2}}{v_1} + \frac{\sqrt{(x-x_2)^2 + y_2^2}}{v_2}$$

So if we want to know the path of the light, we need $T[x]$ to be minimised.

$$T[x] = \frac{\{(x-x_1)^2+y_1^2\}^{1/2}}{v_1} + \frac{\{(x-x_2)^2+y_2^2\}^{1/2}}{v_2}$$

x is the only variable here, all others are fixed quantities so we will differentiate with respect to x.

$$\frac{dT}{dx} = \frac{\frac{1}{2}\{\dots\}^{-1/2}}{v_1}\cdot 2(x-x_1) + \frac{\frac{1}{2}\{\dots\}^{-1/2}}{v_2}\cdot 2(x-x_2)$$

$$= \frac{x-x_1}{v_1\sqrt{(x-x_1)^2+y_1^2}} + \frac{x-x_2}{v_2\sqrt{(x-x_2)^2+y_2^2}}$$

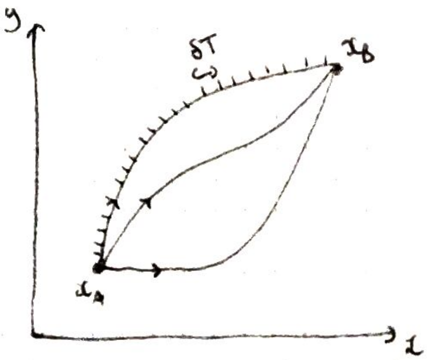
$$= \frac{1}{v_1}\underbrace{\left\{\frac{x-x_1}{\sqrt{(x-x_1)^2+y_1^2}}\right\}}_{\text{note that this is } \sin\theta_1} - \frac{1}{v_2}\underbrace{\left\{\frac{x_2-x}{\sqrt{(x_2-x)^2+y_2^2}}\right\}}_{\text{this is } \sin\theta_2}$$

$$\therefore \frac{dT}{dx} = \frac{1}{v_1}\sin\theta_1 - \frac{1}{v_2}\sin\theta_2$$ which for minimisation we require $=0$

$$\therefore \frac{1}{v_2}\sin\theta_2 = \frac{1}{v_1}\sin\theta_1$$ but $n = \frac{c}{v}$ so multiply both sides by c

$$\boxed{n_1\sin\theta_1 = n_2\sin\theta_2}$$ we have found Snell's Law.

This was a very simple example where there were two mediums of fixed refractive index. But what about a plane with varying refractive index? Here, the speed of light is dependent on the position on the plane, hence it is denoted $v(x,y)$



If we consider just one path for now, how do we know how long the light will take to travel along that path?

We divide the path into many infinitesimals, where light takes a time δT to travel along each infinitesimal.

Doing the same as before:

$$\delta T = \frac{\sqrt{\delta x^2 + \delta y^2}}{v(x,y)}$$

To find the total time along the path, we sum all the infinitesimals by taking $\delta T \to 0$ and integrating:

$$\int_0^T dT = \int_{x_A}^{x_B} \frac{\sqrt{dx^2+dy^2}}{v(x,y)} = \int \frac{\sqrt{1+(dy/dx)^2}}{v(x,y)}\, dx$$

Here, we pulled the $\sqrt{dx^2}$ outside

We can generalise this for any path by saying that

$\int_0^T dT = T[y(x)]$ where $y(x)$ is the path taken by light.

NB:// We use square brackets here since T is a functional, which means its value depends on a function $y(x)$.

$$\therefore\ T[y(x)] = \int \frac{1}{v(x,y)} \cdot \sqrt{1+\left(\frac{dy}{dx}\right)^2}\, dx$$

if we let $\dot{y} = \frac{dy}{dx}$, we can express the integrand as a "Lagrangian":

$$\boxed{L(y,\dot{y},x) = \frac{1}{v(x,y)}\sqrt{1+\dot{y}^2}}$$

so $T[y(x)] = \int_{x_A}^{x_B} L(y,\dot{y},x)\, dx$

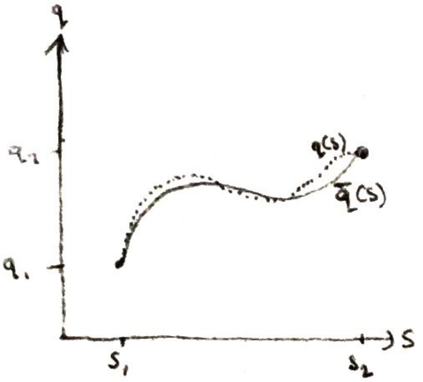
So how do we find the minimal $T[y(x)]$. This seems rather difficult since we would have to test out an infinite number of paths $y(x)$ to find the lowest T.

Actually, there is a mathematical miracle we can use. It turns out the problem of finding the "true path" of light, i.e. the one that minimises T, is equivalent to solving the Euler-Lagrange Equation:

$$\boxed{\frac{d}{dx}\left(\frac{\partial L}{\partial \dot{y}}\right) - \frac{\partial L}{\partial y} = 0}$$

A note on the Euler Lagrange Equation

So where does this magical Euler-Lagrange Equation come from?
Consider an arbitrary curve on the q-s plane:



Suppose that we want to minimise some quantity $S[q(s)]$ that is a functional that depends on the path $q(s)$. $S[q(s)]$ is called an "action"

As we did before, we can construct an integral that uses the Lagrangian.

The Lagrangian is a number that takes a value at each point on the curve. $S[q(s)] = \int_{s_1}^{s_2} L(q, \dot{q}, s)\, ds$

We will call the curve that minimises $S[q(s)]$ $\bar{q}(s)$
Consider another curve that lies very close to $\bar{q}(s)$. We will call this curve $q(s)$.

$$\therefore \quad q(s) = \bar{q}(s) + \delta q(s)$$

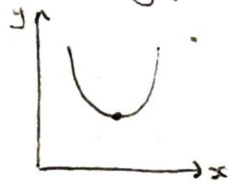
We have the same endpoints so $\delta q(s_1) = \delta q(s_2) = 0$

The change in action between the two curves is:

$$\delta S = S[\bar{q} + \delta q] - S[\bar{q}]$$

As explained in the book, we could say that $\delta S \approx 0$.

A quick note on an upcoming assumption. Consider a curve with a turning point:



Now consider a small deviation in x, δx about the minimum. We could say the change in y $\delta y \approx 0$
This only works for deviations about the turning point.

Let's calculate δS from $S[\bar{q} + \delta q]$

$$S[\bar{q} + \delta q] = \int_{s_1}^{s_2} L(\bar{q} + \delta q, \dot{\bar{q}} + \delta \dot{q}, s)\, ds$$

Let's do a 1st order taylor expansion

$$S[\bar{q} + \delta q] = \int_{s_1}^{s_2} \underbrace{L(\bar{q}, \dot{\bar{q}}, s)}_{\text{This is just } S[\bar{q}]} + \delta \dot{q} \frac{\partial L}{\partial \dot{q}} + \delta q \frac{\partial L}{\partial q}\, ds$$

$$\text{so:} \quad S[\bar{q} + \delta q] = S[\bar{q}] + \int_{s_1}^{s_2} \delta \dot{q} \frac{\partial L}{\partial \dot{q}} + \delta q \frac{\partial L}{\partial q}\, ds$$

$$S[\bar{q}+\delta q] - S[\bar{q}] = \underbrace{\int_{s_1}^{s_2} \delta\dot{q}\, \frac{\partial L}{\partial \dot{q}}\, ds}_{\text{Let's consider just this part first}} + \int_{s_1}^{s_2} \delta q\, \frac{\partial L}{\partial q}\, ds$$

$$\int_{s_1}^{s_2} \delta\dot{q}\, \frac{\partial L}{\partial \dot{q}}\, ds = \underbrace{\left[\delta q\, \frac{\partial L}{\partial \dot{q}}\right]_{s_1}^{s_2}}_{=0 \text{ since } \delta q \text{ is } 0 \text{ at } s_1 \text{ and } s_2} - \int_{s_1}^{s_2} \delta q\, \frac{d}{ds}\left(\frac{\partial L}{\partial \dot{q}}\right) ds$$

from integration by parts

So:

$$S[\bar{q}+\delta q] - S[\bar{q}] = -\int_{s_1}^{s_2} \delta q \left(\frac{d}{ds}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q}\right) ds$$

we need $S[\bar{q}+\delta q] - S[\bar{q}] = 0$ which can only be true if

$$\boxed{\frac{d}{ds}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} = 0}$$

Thus we have found the Euler - Lagrange equation

if we have more than one dimension, we have:

$$\boxed{\frac{d}{ds}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0}$$

where i denotes the component of the dimension.

Let's do an example with light:

Light in the Atmosphere

The speed of light in air depends on refractive index, which in turn depends on density of air (so height above ground, h):

$$v(h) = \frac{c}{n(h)}$$

We previously stated the Lagrangian as

$$L(y, \dot{y}, x) = \frac{1}{v(x,y)} \sqrt{1+\dot{y}^2}$$

we can make this specific for our case.

so $L = \frac{n(h)}{c}\sqrt{1+\dot{h}^2}$

Note here that L is independent of x.

so $$\frac{dL}{dx} = \underbrace{\frac{\partial L}{\partial x}}_{=0 \text{ since } L \text{ is independent of } x} + \frac{\partial L}{\partial h}\dot{h} + \frac{\partial L}{\partial \dot{h}}\ddot{h}$$

$$\Rightarrow \frac{dL}{dx} = \frac{\partial L}{\partial h}\dot{h} + \frac{\partial L}{\partial \dot{h}}\ddot{h}$$

Now we can make use of the Euler-Lagrange equation:

$$\frac{d}{ds}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} = 0$$

$$\Rightarrow \frac{d}{dx}\left(\frac{\partial L}{\partial \dot{h}}\right) = \frac{\partial L}{\partial h}$$ sub this into what we found earlier

$$\frac{dL}{dx} = \frac{d}{dx}\left(\frac{\partial L}{\partial \dot{h}}\right)\dot{h} + \frac{\partial L}{\partial \dot{h}}\ddot{h}$$

$$\frac{dL}{dx} - \frac{d}{dx}\left(\frac{\partial L}{\partial \dot{h}}\right)\dot{h} - \frac{\partial L}{\partial \dot{h}}\ddot{h} = 0 \quad \Rightarrow \frac{d}{dx}\left(L - \dot{h}\frac{\partial L}{\partial \dot{h}}\right) = 0$$

from product rule

$$\Rightarrow L - \dot{h}\frac{\partial L}{\partial \dot{h}} = \text{constant}$$ which we can call D

$$\Rightarrow \frac{n(h)}{c}\sqrt{1+\dot{h}^2} - \frac{\dot{h}^2 n(h)}{c\sqrt{1+\dot{h}^2}} = D$$

so $$D = \frac{n}{c\sqrt{1+\dot{h}^2}}$$

which gives us $$\frac{dh}{dx} = \sqrt{\frac{n^2}{D^2} - 1}$$

so $$x - x_0 = \int_{h_0}^{h} \frac{dh}{\sqrt{\frac{n^2}{D^2} - 1}}$$

This is our final answer. It tells us the horizontal distance the light travels when it moves between two heights.

What does this mean exactly?

Imagine a ray of light that begins moving horizontally (so that $\dot{h} = 0$) at $h = 0$ in an atmosphere where:

$$n(h) = n_0 - \lambda h \qquad \text{for some constant } \lambda$$

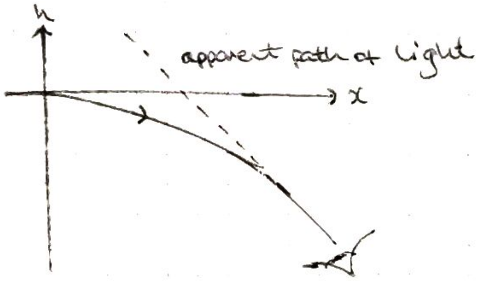
$$x = \int \frac{dh}{\sqrt{\frac{(n_0 - \lambda h)^2}{D^2} - 1}} \quad \Rightarrow \quad x = -\int \frac{D}{\lambda} \, d\phi = -\frac{D}{\lambda}\phi + c$$

where we have made the substitution $n_0 - \lambda h = D \cosh \phi$

so $$\underline{h = \frac{n_0}{\lambda}\left(1 - \cosh \frac{\lambda x}{n_0}\right)}$$

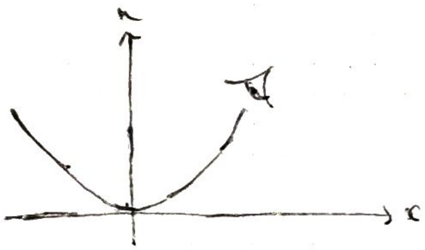
This tells us the equation of the path taken by light.

so, if λ is positive:



so an object would appear taller than it is

if λ is negative:



So a mirage of the sky is observed on the ground.

λ is dependent on temperature. Usually, high temperature implies negative λ.