Alain J. Brizard
Saint Michael's College

Chapter One: Calculus of Variations

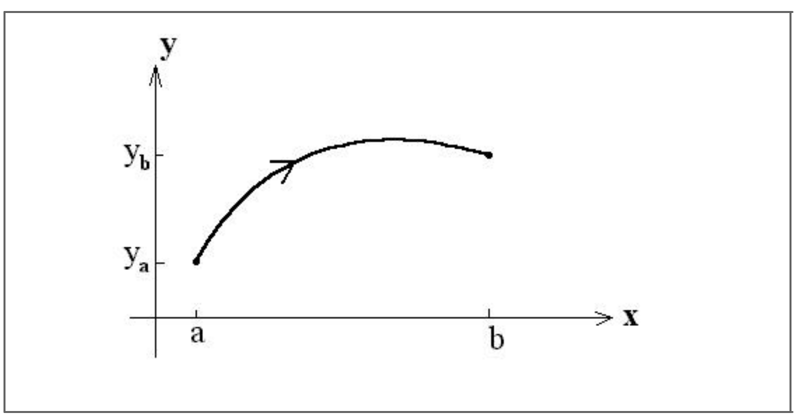
1 Principle of Least Time

Light travels in a straight line when it propagates in a uniform medium. Using the index of refraction $n_0 > 1$ of the uniform medium, the speed of light in the medium is expressed as $v_0 = c/n_0 < c$, where c is the speed of light in vacuum. When light propagates in a nonuniform medium, however, it travels along a path that *minimizes* the travel time between an initial point A (where a light ray is launched) and a final point B (where the light ray is received). Hence, the time taken by a light ray following a path γ from point A to point B is

$$T_\gamma = \int_\gamma c^{-1} n(s)\, ds,$$

where the refractive index $n(s)$ may depend on the position s along the path γ.

For ray propagation in two dimensions (x, y) in a medium with refractive index $n(y)$, an arbitrary point $(x,\ y = y(x))$ along the light path γ is parametrized by the x-coordinate, which starts at $A = (a, y_a)$ and ends at $B = (b, y_b)$. Note that the path γ is now represented by the mapping $y : x \mapsto y(x)$.



Along the path γ, the infinitesimal length element is $ds = \sqrt{1 + (y')^2}\, dx$ along the path $y(x)$ and the travel time

$$T[y] = \frac{1}{c} \int_a^b n(y)\, \sqrt{1 + (y')^2}\, dx \tag{1}$$

is now a *functional* of the mapping y (i.e., changing y changes the value of the integral $T[y]$). For the sake of convenience, we introduce the function

$$F(y, y'; x) \;=\; n(y)\,\sqrt{1 + (y')^2} \tag{2}$$

to denote the integrand of Eq. (1); here, we indicate an explicit dependence on x of $F(y, y'; x)$ for generality.

1.1 Euler's First Equation

To determine the path of least time, we introduce the *functional* derivative $\delta T[y]$ defined as

$$\delta T[y] \;=\; \frac{d}{d\epsilon} T[y + \epsilon\,\delta y]\bigg|_{\epsilon=0},$$

where $\delta y(x)$ is an arbitrary variation of the path $y(x)$ subject to the boundary conditions $\delta y(a) = 0 = \delta y(b)$ (i.e., the end points of the path are not affected by the variation). Using the expression for $T[y]$ in terms of the function $F(y, y'; x)$, we find

$$\delta T[y] \;=\; \frac{1}{c} \int_a^b \left[\, \delta y(x)\, \frac{\partial F}{\partial y(x)} \;+\; \delta y'(x)\, \frac{\partial F}{\partial y'(x)} \,\right] dx,$$

where $\delta y' = (\delta y)'$, which when integrated by parts becomes

$$\delta T[y] \;=\; \frac{1}{c} \int_a^b \delta y \left[\, \frac{\partial F}{\partial y} \;-\; \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) \right] \;+\; \frac{1}{c} \left[\, \delta y_b \left(\frac{\partial F}{\partial y'}\right)_b \;-\; \delta y_a \left(\frac{\partial F}{\partial y'}\right)_a \,\right].$$

Here, since the variation $\delta y(x)$ vanishes at the integration boundaries ($\delta y_b = 0 = \delta y_a$), we obtain

$$\delta T[y] \;=\; \frac{1}{c} \int_a^b \delta y \left[\, \frac{\partial F}{\partial y} \;-\; \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) \right]. \tag{3}$$

The condition that the path γ takes the *least time* corresponds to the variational principle $\delta T[y] = 0$, which yields the (first) Euler equation

$$\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) \;=\; \frac{\partial F}{\partial y}. \tag{4}$$

This ordinary differential equation for $y(x)$ yields a solution that gives the desired path of least time.

We now apply the variational principle $\delta T[y] = 0$ for the case of Eq. (2), for which we find

$$\frac{\partial F}{\partial y'} \;=\; \frac{n(y)\,y'}{\sqrt{1 + (y')^2}} \quad \text{and} \quad \frac{\partial F}{\partial y} \;=\; n'(y)\,\sqrt{1 + (y')^2},$$

so that Euler's First Equation (4) yields

$$n(y)\, y'' \;-\; n'(y)\; \left(1 + (y')^2\right) \;=\; 0. \tag{5}$$

This equation can be expressed as

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

so that its solution is of the form

$$n(y) \;=\; \alpha \sqrt{1 + (y')^2}, \tag{6}$$

where α is a constant determined from the initial conditions of the light ray.

1.2 Euler's Second Equation

We derive Euler's Second equation as follows. First, we write the exact derivative dF/dx as

$$\frac{dF}{dx} \;=\; \frac{\partial F}{\partial x} \;+\; y' \;\frac{\partial F}{\partial y} \;+\; y'' \;\frac{\partial F}{\partial y'},$$

and substitute Eq. (4) to combine the last two terms so that we obtain Euler's Second equation

$$\frac{d}{dx} \left(F \;-\; y' \;\frac{\partial F}{\partial y'} \right) \;=\; \frac{\partial F}{\partial x}. \tag{7}$$

In the present case, the function $F(y, y'; x)$ given by Eq. (2) yields $\partial F/\partial x = 0$, and we find that

$$F \;-\; y' \;\frac{\partial F}{\partial y'} \;=\; \frac{n(y)}{\sqrt{1 + (y')^2}} \;=\; \text{constant},$$

and thus Eq. (7) implies Eq. (6).

1.3 Snell's Law

Eq. (6) states that as the light ray enters into a region of increased (decreased) refractive index its slope also increases (decreases). In particular, by substituting Eq. (5) into Eq. (6), we find

$$\alpha^2\, y'' \;=\; \frac{1}{2}\, \frac{dn^2(y)}{dy},$$

and hence the path of a light ray is concave upward (downward) where $n'(y)$ is positive (negative).

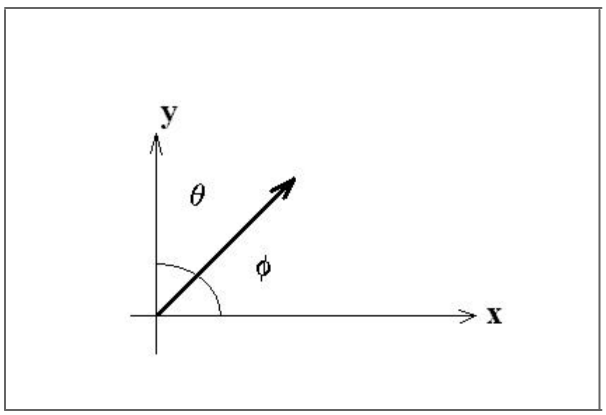
Let us now consider a light ray travelling from $(x, y) = (0, 0)$ at an angle ϕ_0 (measured from the x-axis) so that $y'(0) = \tan\phi_0$ is the slope at $x = 0$, assuming that $y(0) = 0$. Then the constant α is simply

$$\alpha \;=\; n_0\;\cos\phi_0,$$

where $n_0 = n(0)$ is the refractive index at $y(0) = 0$. Let $y'(x) = \tan\phi(x)$ be the slope of the light ray at $(x,\, y(x))$, then $\sqrt{1 + (y')^2} = \sec\phi$ and Eq. (6) becomes

$$n(y)\;\cos\phi \;=\; n_0\;\cos\phi_0,$$

which, when we substitute the complementary angle $\theta = \pi/2 - \phi$,



finally yields the standard from of Snell's Law: $n[y(x)]\;\sin\theta(x) = n_0\;\sin\theta_0$.

2 Application of the Principle of Least Time

As an application of the Principle of Least Time, we consider the propagation of a light ray in a medium with refractive index $n(y) = n_0\,(1 - \beta\, y)$ exhibiting a constant gradient $n'(y) = -\,n_0\,\beta$. Once again with $\alpha = n_0\cos\phi_0$ and $y'(0) = \tan\phi_0$, Eq. (6) becomes

$$1 \;-\; \beta\, y \;=\; \cos\phi_0\,\sqrt{1 + (y')^2}.$$

By separating dy and dx we obtain

$$dx \;=\; \frac{\cos\phi_0\; dy}{\sqrt{(1 - \beta\, y)^2 - \cos^2\phi_0}}. \tag{8}$$

We now use the trigonometric substitution

$$1 - \beta\, y = \cos\phi_0\;\sec\theta, \tag{9}$$

with $\theta = \phi_0$ at $y = 0$, to find

$$dy \;=\; -\,\frac{\cos\phi_0}{\beta}\;\sec\theta\,\tan\theta\;d\theta$$

and

$$\sqrt{(1-\beta\,y)^2 - \cos^2\phi_0} \;=\; \cos\phi_0\;\tan\theta,$$

so that Eq. (8) becomes

$$dx \;=\; -\,\frac{\cos\phi_0}{\beta}\;\sec\theta\;d\theta. \tag{10}$$

The solution to this equation, with $x = 0$ when $\theta = \phi_0$, is

$$x \;=\; -\,\frac{\cos\phi_0}{\beta}\;\ln\left(\frac{\sec\theta + \tan\theta}{\sec\phi_0 + \tan\phi_0}\right). \tag{11}$$

If we now define

$$\psi \;=\; \frac{\beta\,x}{\cos\phi_0} \;-\; \ln(\sec\phi_0 + \tan\phi_0),$$

Eq. (11) becomes

$$\sec\theta \;+\; \sqrt{\sec^2\theta \;-\; 1} \;=\; e^{-\,\psi},$$

which can be solved for $\sec\theta$ as

$$\sec\theta \;=\; \cosh\psi \;=\; \cosh\left(\frac{\beta\,x}{\cos\phi_0} \;-\; \ln(\sec\phi_0 + \tan\phi_0)\right),$$

so that the light ray follows a path $y(x)$ obtained from Eq. (9) as

$$y(x;\beta) \;=\; \frac{1}{\beta} \;-\; \frac{\cos\phi_0}{\beta}\;\cosh\left(\frac{\beta\,x}{\cos\phi_0} \;-\; \ln(\sec\phi_0 + \tan\phi_0)\right). \tag{12}$$

Note that, using the identities

$$\cosh\left[\ln(\sec\phi_0 + \tan\phi_0)\right] \;=\; \sec\phi_0,$$

$$\sinh\left[\ln(\sec\phi_0 + \tan\phi_0)\right] \;=\; \tan\phi_0,$$

we can check that in the uniform case ($\beta = 0$) we recover the expected result

$$\lim_{\beta\to 0}\; y(x;\beta) \;=\; (\tan\phi_0)\,x.$$

Next, we observe that $y(x;\beta)$ exhibits a single maximum located at $x = \overline{x}(\beta)$. Solving for $\overline{x}(\beta)$ from $y'(\overline{x};\beta) = 0$, we obtain

$$\tanh\left(\frac{\beta\,\overline{x}}{\cos\phi_0}\right) \;=\; \sin\phi_0,$$

or

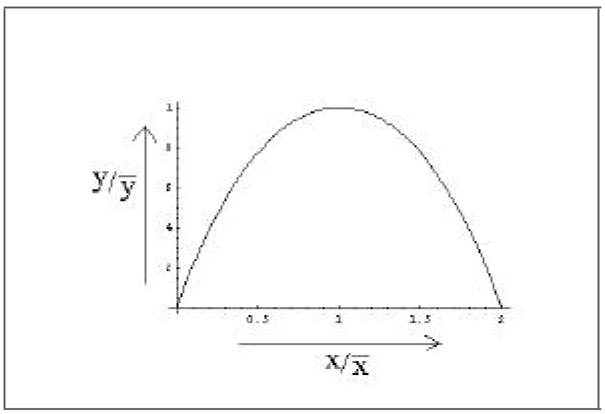
$$\overline{x}(\beta) \;=\; \frac{\cos\phi_0}{\beta}\;\ln(\sec\phi_0+\tan\phi_0), \tag{13}$$

and hence

$$y(x;\beta) \;=\; \frac{1}{\beta} \;-\; \frac{\cos\phi_0}{\beta}\;\cosh\left(\frac{\beta}{\cos\phi_0}\,(x-\overline{x})\right),$$

and

$$\overline{y}(\beta) \;=\; y(\overline{x};\beta) \;=\; \frac{1-\cos\phi_0}{\beta}. \tag{14}$$



Plot of $y/\overline{y}$ versus $x/\overline{x}$ for $\phi_0 = \pi/3$.