Let's start with Fermat (~1600) and his observation that light tends to travel in straight lines, but "bends" at an interface. He observed that the bending could be explained if light follows the path of least time.

Given \( h_1, h_2, v_1, v_2 \), find \( x \) such that \( t \) is minimum.

\[
\begin{align*}
\ell_1 &= \frac{1}{v_1} \sqrt{h_1^2 + x^2} \\
\ell_2 &= \frac{1}{v_2} \sqrt{(d-x)^2 + h_2^2} \\
\ell &= \ell_1 + \ell_2
\end{align*}
\]

\[
\frac{d\ell}{dx} = \frac{x}{v_1 \sqrt{h_1^2 + x^2}} - \frac{(d-x)}{v_2 \sqrt{(d-x)^2 + h_2^2}} = 0
\]

\[
\tan \theta_1 = \frac{x}{\sqrt{h_1^2 + x^2}} \\
\tan \theta_2 = \frac{(d-x)}{\sqrt{(d-x)^2 + h_2^2}}
\]

Then: \( \frac{\tan \theta_1}{v_1} = \frac{\tan \theta_2}{v_2} = \frac{v_1}{v_2} \Rightarrow v_1 \sin \theta_1 = v_2 \sin \theta_2 \)
Now let's see if we can "relativize" Fermat's principle. First of all, which "time" should we extremize?

The only distinguished time in relativity is the proper time which we defined as:

$$\tau = \int_A^B \sqrt{-g_{\mu\nu} \, dx^\mu \, dx^\nu} = \int_A^B \sqrt{-g \, dt^2 + dx^2}.$$  

We want the path $x(t)$ which extremizes $\tau$, so we set $\delta \tau = 0$ under variations which vanish at $A, B$.

$$\delta \tau = \int_A^B \delta f(v) \, dt = \int_A^B \left( \frac{\delta f}{\delta v} \right) \delta v^i \, dt = \int_A^B \frac{\partial f}{\partial v^i} \delta v^i \, dt = \int_A^B \left( \frac{\partial f}{\partial v^i} \right) \delta v^i \, dt$$

Integrating by parts using:

$$\frac{\partial f}{\partial v^i} \left[ \frac{\partial f}{\partial v^i} \right] = \frac{\partial f}{\partial v^i} \frac{\partial f}{\partial v^i} = \frac{\partial f}{\partial v^i} \frac{\partial f}{\partial v^i}$$

then:

$$\delta \tau = \int_A^B \frac{\partial f}{\partial v^i} \delta v^i \, dt - \int_A^B \frac{\partial f}{\partial v^i} \delta v^i \, dt = 0$$

$$\frac{\partial f}{\partial v^i} \left|_A^B \right. = 0$$

$$\frac{\partial f}{\partial v^i} = 0$$

$$\Rightarrow \delta v^i = 0$$

Which is just an constant velocity path from $A$ to $B$. 

$$\nabla \cdot \mathbf{v} = 0$$
What if we are not in flat spacetime? Or use coordinates other than (ct, x, y, z)?

\[ Z = \int_A \sqrt{-g} \, d^4x \, d^4\nu \]  

Now since there are no preferred coordinates we really should be parametrizing the path \( \nu(t) \) and \( t \).

\[ Z = \int_A \sqrt{-g} \, \frac{d\nu}{dt} \, dt = \int_A \sqrt{-g} \, d\nu \]

We want:

\[ \int_A \sqrt{-g} \, d\nu \]

But recall:

\[ g_{\nu\nu} \frac{d\nu}{dt} \frac{d\omega}{dt} = U^\mu U_{\mu} = -1 \]

so then:

\[ \int_A \sqrt{-g} \, d\nu = 0 \]

which is much easier to handle than the form with \( \int_A \sqrt{-g} \, d\nu \).

Consider varying:

\[ x^\lambda \rightarrow x^\lambda + \delta x^\lambda \]

\[ g_{\nu\mu} \rightarrow g_{\nu\mu} + \frac{\partial g_{\nu\mu}}{\partial x^\lambda} \delta x^\lambda \]

Since the metric varies over spacetime, i.e., depends on \( x^\lambda \), different paths will experience different forms of the metric.

Then:

\[ \delta I = \frac{1}{2} \int_A \left( g_{\nu\mu} \frac{\partial x^\lambda}{\partial x^\rho} \frac{\partial x^\rho}{\partial x^\sigma} \delta x^\lambda \right) \, dt = \int_A \frac{\partial}{\partial t} \left[ \frac{\partial}{\partial \nu} \left( g_{\nu\mu} \frac{\partial x^\lambda}{\partial x^\rho} \frac{\partial x^\rho}{\partial x^\sigma} \delta x^\lambda \right) \right] \, dt \]

We can integrate the last two terms using:

\[ \frac{\partial}{\partial t} \left[ g_{\nu\mu} \frac{\partial x^\lambda}{\partial x^\rho} \frac{\partial x^\rho}{\partial x^\sigma} \delta x^\lambda \right] = \frac{\partial g_{\nu\mu}}{\partial x^\alpha} \delta x^\alpha + \frac{\partial^2 g_{\nu\mu}}{\partial x^\alpha \partial x^\beta} \delta x^\alpha \delta x^\beta \]

So:

\[ \delta I = \int_A \frac{\partial g_{\nu\mu}}{\partial x^\alpha} \delta x^\alpha \, dt \]

Altogether we have:

\[ \delta I = \frac{1}{2} \int_A \left[ \frac{\partial g_{\nu\mu}}{\partial x^\alpha} \frac{\partial x^\alpha}{\partial \nu} \delta x^\mu - g_{\nu\mu} \frac{\partial^2 g_{\nu\mu}}{\partial x^\alpha \partial x^\beta} \delta x^\alpha \delta x^\beta - g_{\mu\nu} \frac{\partial^2 g_{\nu\mu}}{\partial x^\alpha \partial x^\beta} \delta x^\alpha \delta x^\beta \right] \, dt \]

But for arbitrary \( \delta x^\lambda \) this requires:

\[ -2 g_{\nu\mu} \frac{\partial^2 g_{\nu\mu}}{\partial x^\alpha \partial x^\beta} \delta x^\alpha \delta x^\beta = 0 \]

Apply this to both sides:

\[ \frac{\partial^2 g_{\nu\mu}}{\partial x^\alpha \partial x^\beta} \delta x^\alpha \delta x^\beta = 0 \]

Recall \( g g_{\nu\mu} = g^\nu \) or using:

\[ \gamma_{\nu\mu} = \frac{1}{2} g^\rho \left( \partial_\rho g_{\nu\mu} + \partial_\nu g_{\rho \mu} - \partial_\mu g_{\rho \nu} \right) \]

we have:

\[ \frac{\partial^\nu g_{\mu \nu}}{\partial x^\rho} + \frac{\partial g_{\nu\mu}}{\partial x^\rho} \frac{\partial x^\rho}{\partial x^\sigma} \delta x^\sigma = 0 \]

The geodesic equation!
What about good old \( S = \int L \, dt \) with \( L = T - U \)?

Well recall that for SR we extremized \( E = \int_A \sqrt{1 - \frac{v^2}{c^2}} \, dt = \int_A \sqrt{1 - \frac{x^2}{c^2}} \, dt \)

For \( \frac{v}{c} \ll 1 \) this is \( E \approx \int_A \sqrt{1 + \frac{v^2}{c^2}} \, dt \approx \int_A \left( 1 + \frac{1}{2} \left( \frac{x}{v} \right)^2 + \cdots \right) dt \approx \int_A \frac{1}{2} v^2 \, dt \)

\( \int T \, dt \)