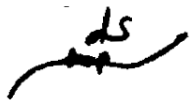
Lecture 22

Since the goal of geometrical optics is to describe propagation solely in terms of rays (and not even mentioning wave surfaces), we should try to obtain a differential equation which governs the ray path directly.

Consider the rate of change of $n\hat{s}$ along a ray:

$$\frac{d}{ds}(n\hat{s}) = \frac{d}{ds}(\nabla S) = \hat{s}\cdot\nabla(\nabla S) = \frac{\nabla S}{n}\cdot\nabla(\nabla S) = \frac{1}{2n}\nabla(\nabla S)^2 = \frac{1}{2n}\nabla n^2 = \nabla n$$



(ds is measured along the ray path)

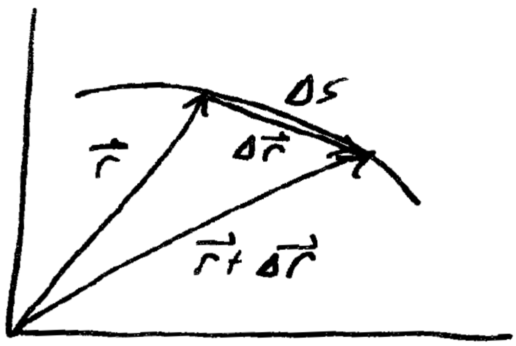
$$\nabla(\nabla S)^2 = \nabla[\nabla S\cdot\nabla S] = \nabla S\cdot\nabla(\nabla S) + \nabla S\cdot\nabla(\nabla S) = 2\nabla S\cdot\nabla(\nabla S)$$

So $\boxed{\frac{d}{ds}(n\hat{s}) = \nabla n}$ "ray equation "

Given $n(\vec{r})$ and an initial ray direction $\hat{s}_0$ of the light, this differential equation governs the path taken.

In words, the change in the optical path along the path is given by the gradient in the index of refraction.

- A reminder on differential line elements



$$\hat{S} = \frac{d\hat{r}}{ds} = \hat{x}\frac{dx}{ds} + \hat{y}\frac{dy}{ds} + \hat{z}\frac{dz}{ds}$$

ds= length element along curve

let's consider some very simple examples.

(i) n=constant => $\nabla n = 0$

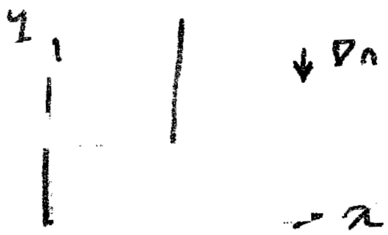
⇨ $\frac{d}{ds}(n\hat{s}) = 0$ => $\underline{\hat{s} = \text{constant}}$

⇨ Ray propagates in a straight line (as was the case for all the examples i-iii on P.163-4).

（ii） n varies along only one axis, e.g. the y axis, and the ray travels parallel to that axis .

$$n = n(y) \qquad \nabla n = \frac{dn}{dy}\hat{y}$$

$$\frac{d}{ds} = \frac{d}{dy}, \hat{S}_0 = \hat{y}$$



$$\rightarrow \frac{d}{dy}(n\hat{s}) = \frac{dn}{dy}\hat{y} \text{ ,so } \hat{s} = \hat{y}$$

As expected, the ray does not change direction.

However, as the index changes, the optical path length traversed in a unit distance changes directly with n.

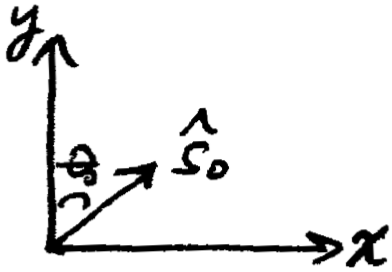
(ii) n varies only in one direction, e.g. n=n(y), and the ray starts out orthogonally, e.g. $\hat{S}_0 = \hat{x}$

$\overset{\hat{S}_0}{\longrightarrow} + \nabla n \uparrow = \hat{S}\nearrow$ or $\overset{\hat{S}_0}{\longrightarrow} + \nabla n \downarrow = \hat{S}\searrow$

(if n increases with height) (if n decreases with height)

⇨ The ray changes direction with propagation

(iv) Extending this example to the most general case, with n=n(y) and $\hat{S}_0 = \hat{x}\sin\theta_0 + \hat{y}\cos\theta_0$



The vector ray equation can be split into three component equations

x : $\frac{d}{ds}(n\sin\theta) = \frac{dn}{dx} = 0$

y : $\frac{d}{ds}(n\cos\theta) = \frac{dn}{dy} \neq 0$

z : $\frac{d}{ds}(n\gamma) = 0$ $\qquad (\hat{S} = \hat{x}\sin\theta_0 + \hat{y}\cos\theta_0 + \hat{z}\gamma)$

where γ = direction consine along z axis

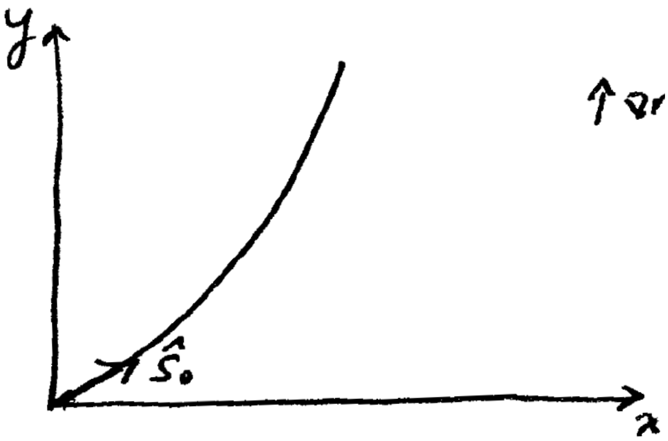
The z equation says $n\gamma$ = constant, but $\gamma_0 = 0$ since the ray is in the xy plane => γ =0

⇨ The ray stays in the xy plane.

The x equation gives $n\sin\theta = \text{constant} = n_0\sin\theta_0$

Thus Snell's law is locally obeyed

Note if $\frac{dn}{dy} > 0$, then the ray Gradually turns towards the y axis (just as your Snell's law Intuition says it should!)



Note that we can say more generally, that as a ray propagates, it bends towards the region of higher refraction index.

Note also that since $\hat{s}$ is a unit vector, the y-equation in our example (iv) does not give any information beyond that given by the x-equation.

We will use the ray equation to solve some representative problems in propagation in inhomogeneous media, but before more familiar principle of geometrical optics, namely Fermat's Principle, which we accomplish via the eikonal equation.

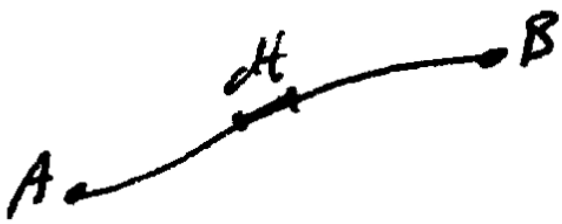
Fermat's Principle

There are various statements of Fermat's Principle, but the simplest (and loosest) is:

"Of all the paths light might take between two points, the actual path taken is the one that requires the least time."

The time required to go from A to B is

$$\int_A^B dt$$



Where the integral is along the path.

Now the differential time element is related to the optical path length element by $cdt = nds$

So $c\int_A^B dt = \int_A^B nd$ (so least time $\leftrightarrow$shortest optical path)

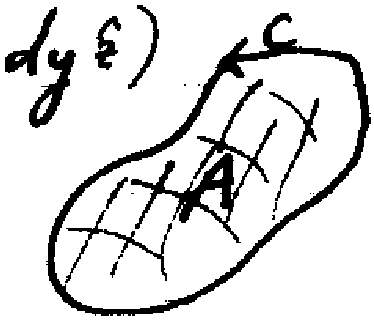
Therefore we need to consider the integral over the optical path length. In order to do this, we begin with the eikonal equation

$$\nabla S = n\hat{s}$$

From vector calculus, we know that the curl of any gradient is equal to zero, so we have

$$\nabla \times (\nabla S) = \nabla \times (n\hat{s}) = 0$$

We can integrate this over any open surface to obtain



$$\iint_A \nabla \times (n\hat{s}) d\vec{a} = 0 \quad \text{(e.g. } d\vec{a} = dxdy\hat{z} \text{)}$$

Stokes's theorem gives

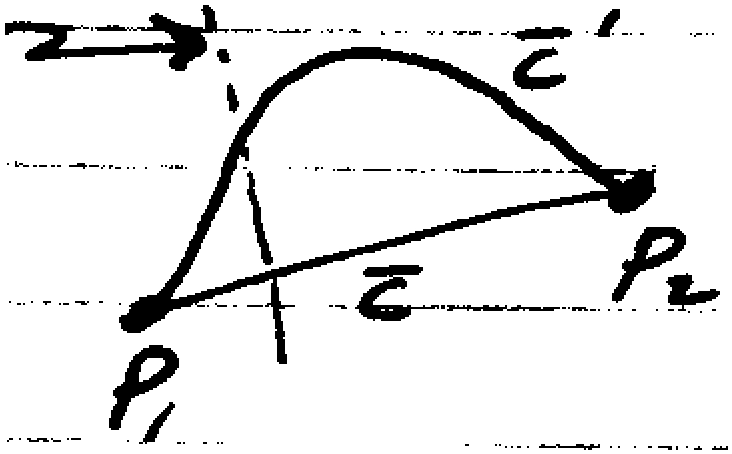
$$\iint_A \nabla \times (n\hat{s}) d\vec{a} = \oint_C n\hat{s} d\vec{r} = 0$$

Where C is the closed curve bounding A.

This result is known as the Lagrange integral invariant .

Lagrange's integral invariant gives us an easy proof of Fermat's Principle:

- Consider a ray which passes through points P_1 and P_2 along a curve $\bar{C}$
- Consider another curve $\bar{C}'$ also passing through P_1 and P_2 wavefront



- applying lagrange's invariant to the loop:

$$\int_{\bar{C}} n\hat{s}\cdot d\vec{r} - \int_{\bar{C}'} n\hat{s}\cdot d\vec{r} = 0$$

minus sign since $\int_{\bar{C}'} = \int_{P_1}^{P_2}$

• along $\bar{C}, \hat{s} \| d\vec{r}$ since $\bar{C}$ is a ray , and $\hat{s}$ lies along the ray ($\perp$ to wavefronts)

$$\Rightarrow \int_{\bar{C}} nds = \int_{\bar{C}'} n\hat{s}\cdot d\vec{r}$$

• by the triangle inequality $\hat{s}d\vec{r} \le ds$

(Equality if and only if the curve lies along a ray, i.e. $\perp$ to wavefronts)

$$\Rightarrow \int_{\bar{C}} nds \le \int_{\bar{C}'} nds$$

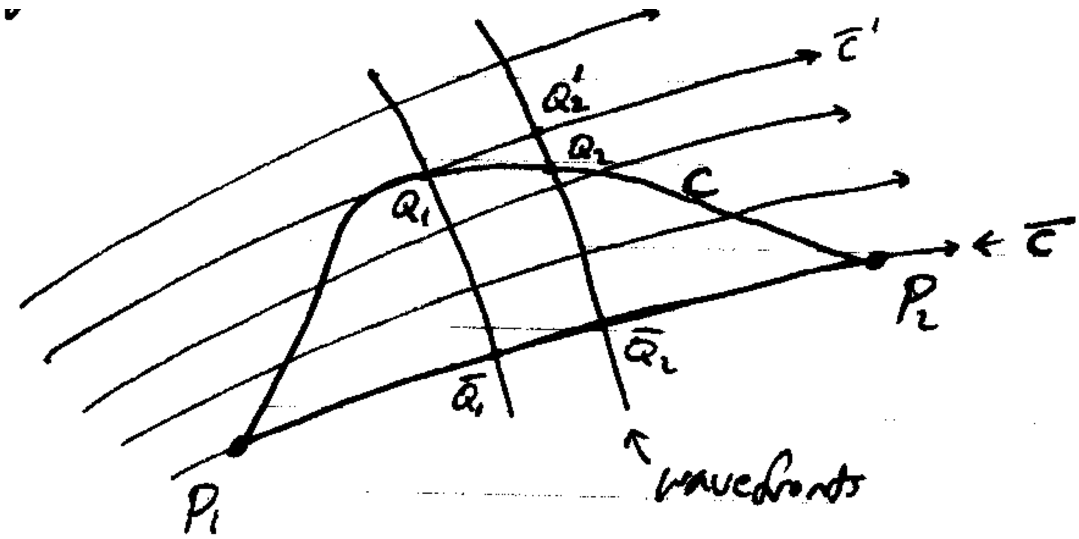
With equality if $\bar{C} = \bar{C}'$ at every point

In words:

"The path of wavefront normals is also the path of least time (or equivalently shortest optical path length)."

In order to obtain Fermat's Principle from this, we follow a clever argument given in Born + Wolf 3.3.2.

1. Consider a family of rays, along one of which light propagates from P_1 to P_2.



$\bar{C}$ = actual path from P_1 to P_2.

C = arbitrary curve (a possible path) joining P_1 to P_2.

Two adjacent wavefronts intersect $\bar{C}$ at $\bar{Q}_1$ and $\bar{Q}_2$, and C at Q_1 and Q_2.

Q_2' is on ray $\bar{C}'$ which passes through Q_1 as shown .

2. Apply Lagrange's integral invariant to the triangle $Q_1Q_2Q_2'$

$$\left(\mathrm{n}\hat{s}\cdot d\vec{r}\right)_{Q_1Q_2}+\left(\mathrm{n}\hat{s}\cdot d\vec{r}\right)_{Q_2Q_2'}-\left(nds\right)_{Q_1Q_2'}=0$$

3. Last time, we saw that $\hat{s}\perp$wavefronts, so

$$\left(\mathrm{n}\hat{s}\cdot d\vec{r}\right)_{Q_2Q_2'}=0$$

$$\left(\mathrm{n}\hat{s}\cdot d\vec{r}\right)_{Q_1Q_2}=\left(nds\right)_{Q_1Q_2'}$$

4. We also saw last time that the optical path length between any two pairs of points on two wavefronts is the same, i.e.

$$\left(nds\right)_{Q_1Q_2'}=(nds)_{\bar{Q}_1\bar{Q}_2}$$

Thus $\left(\mathrm{n}\hat{s}\cdot d\vec{r}\right)_{Q_1Q_2}=(nds)_{\bar{Q}_1\bar{Q}_2}$

5. Triangle inequality (or definition of dot product of two vectors) implies

$$\left(\mathrm{n}\hat{s}\cdot d\vec{r}\right)_{Q_1Q_2} \le (nds)_{Q_1Q_2}$$

Equality is obtained only when $d\vec{r}$ is parallel to $\hat{s}$, and that is true only along a ray.

6. Combining the results in 4 and 5 yields

$$(nds)_{\bar{Q}_1\bar{Q}_2} \le (nds)_{Q_1Q_2}$$

7. This is true for all segment (adjacent wavefronts) between P_1 and P_2, so we can integrate, yielding

$$\int_{\bar{C}} nds \le \int_C nds$$

Equality would be obtained if (and only if) $\left(\mathrm{n}\hat{s}\cdot d\vec{r}\right)_{Q_1Q_2} = (nds)_{Q_1Q_2}$, i.e. $\hat{s}$ and $d\vec{r}$ are parallel at every point on the ray .

This is true only if Q_1Q_2 is an actual ray

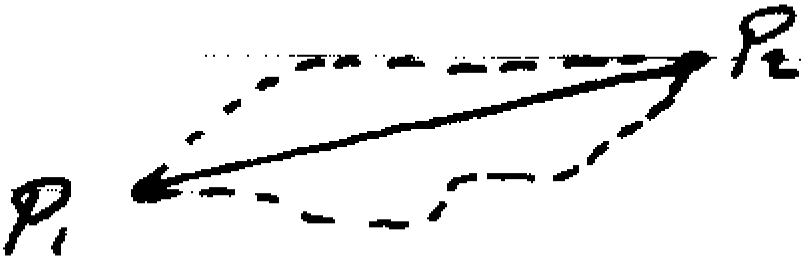
(which is the case on our picture only if $Q_2 = Q_2'$ in every segment of the path, and hence if $C = \bar{C}$!)

Thus we have the conclusion that the optical path length taken by a ray is shorter than any other path from P_1 and P_2.

This could be considered an alternative, but completely equivalent, statement of Fermat's Principle.

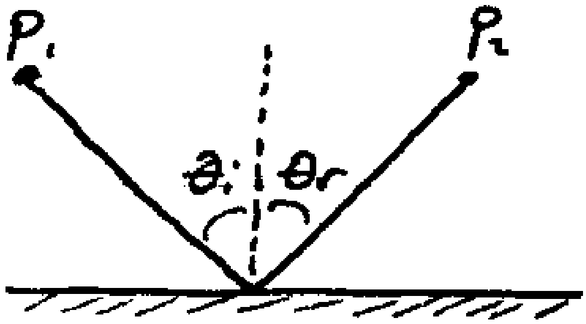
Trivial example: n = constant (homogeneous medium)

Fermat => light travels in a straight line



Law of Reflection

It is a simple exercise in freshman calculus to verify that the shortest path in which light can go from P_1 to P_2 by bouncing off a surface satisfies $\theta_i = \theta_r$.

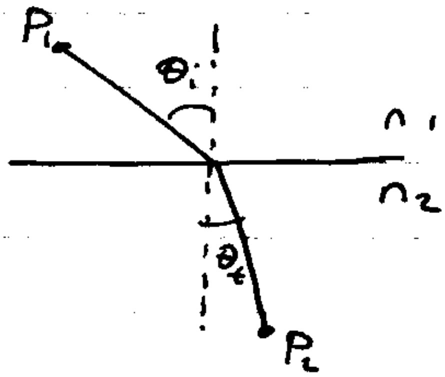


(For proof, see Guenther P.136).

Snell's Law

It is an equally simple exercise to show that the shortest optical path from P_1 to P_2 satisfies

$$n_1 \sin\theta_1 = n_2 \sin\theta_2$$

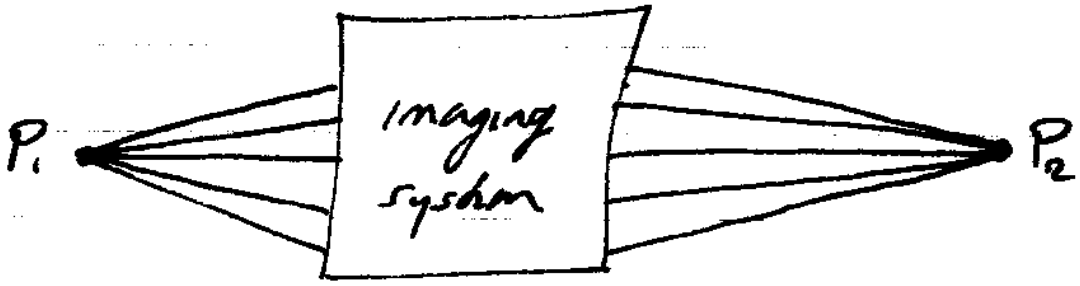


(For proof, see Guenther P.137- you should go through it if you've never seen it before.)

Intuition: recall the "drowning swimmer" analogy! What path will minimize the time to get to the drowning swimmer? You run more of the distance than you swim, since $V_1 > V_2$, so you run on the path satisfying $\frac{1}{V_1}\sin\theta_1 = \frac{1}{V_2}\sin\theta_2$.

Fermat and Optical Imaging

If all the rays in a certain region of point P_1 (the "object") converge to a point P_2, then P_2 is said to be an image of P_1. An optical system that collects the rays from P_1 and redirects them to P_2 is called an imaging system.



Thus, according to Fermat's Principle, all the rays from P_1 to P_2 must travel exactly the same path length. (And they are all, of course, the minimum path length.)

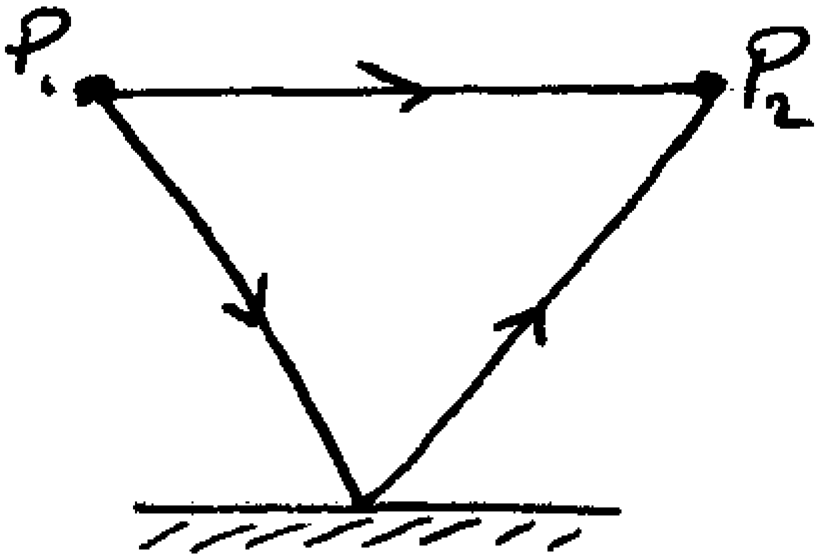
In fact, we shall see that one way to calculate or design an imaging system is to demand that all the rays have the same path length.

Generalized Fermat's Principle

We begin by noting that of our imaging system, a given ray has a ray immediately next to it which also goes from P_1 to P_2 with the same path length.

Thus the path length of our given ray is not a local minimum.

Note also that for our reflection problem, the ray path satisfying $\theta_i = \theta_r$ is not even a global minimum, although it is a local minimum. The global minimum is the straight line from P_1 to P_2!



(Of course, the straight line from P_1 to P_2 is also a local minimum.)

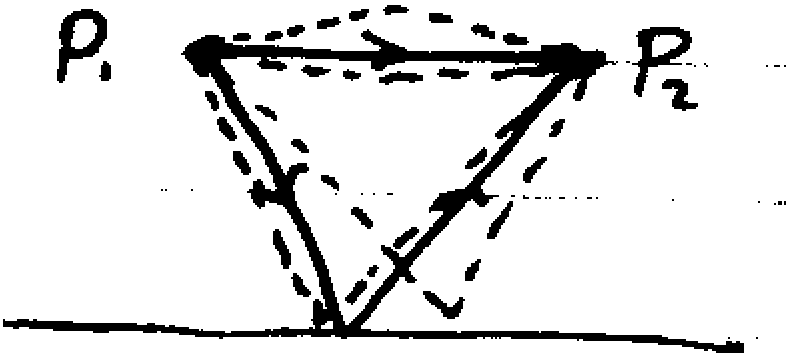
Thus it is more accurate to give a more general statement of Fermat's Principle:

The optical path traversed from P_1 to P_2 is that for which the integral

$$\int_{P_1}^{P_2} n ds$$

is stationary .

By "stationary", we mean that an arbitrarily small displacement of the path about the true path will show that the true path is a local minimum, or result in no change in the optical path.



Mathematically, this is often expressed as

$$\boxed{\delta\int_{P_1}^{P_2} n ds = 0}$$

This is the proper mathematical statement of Fermat's Principle.

(note the conceptual similarity to the first derivative =0 giving the stationary points of a function.)

Fermat's Principle thus stated is very useful conceptually, and provides a framework for understanding many optical phenomena. As a calculational tool, however, it is not so useful. The general problem of determining a path for which the integral is stationary forms the field of mathematics called the calculus of variations.

We need to use only the central result of the calculus of variations. This is the following.

1. Suppose we parameterize a path by σ
2. The path is some unknown function $g(\sigma)$, which goes from P_1 to P_2
3. We want to find the stationary points of $\int_{P_1}^{P_2} F[\sigma, g(\sigma), \frac{dg}{d\sigma}]d\sigma$ where F

is a known “functional ” of σ and the unknown function g

4. Theorem : a necessary and sufficient condition that the integral be stationary is that

$$\frac{d}{d\sigma}\left(\frac{\partial F}{\partial g'}\right)-\frac{\partial F}{\partial g}=0$$

This is known as the “Euler differential equation” associated with the stationary values of the integral.

Thus we do not solve for the path via the Fermat integral directly, but by solving the associated Euler differential equation.

We want the stationary points of

$$I=\int_{P_1}^{P_2} nds$$

We have n = n (x, y, z), and

$$ds=\sqrt{(dx)^2+(dy)^2+(dz)^2}$$ = length element along path

We begin by parameterizing the path

$$x=x(\sigma), y=y(\sigma), z=z(\sigma)$$

Where σ = normalized distance along path =>

$$ds=\sqrt{\left(\frac{dx}{d\sigma}\right)^2+\left(\frac{dy}{d\sigma}\right)^2+\left(\frac{dz}{d\sigma}\right)^2}d\sigma$$

Where $\sigma = 0$ at P_1

$\sigma = 0$ at P_2

⇨ $$I=\int_0^1 F(x,y,z,x',y',z')d\sigma$$

Where $F(x,y,z,x',y',z')=n(x,y,z)\sqrt{(x')^2+(y')^2+(z')^2}$

With $x'=dx/d\sigma, y'=dy/d\sigma, z'=dz/d\sigma$

Let's consider the Euler eqn. associated with x

$$\frac{d}{d\sigma}\left(\frac{\partial F}{\partial x'}\right)-\frac{\partial F}{\partial x}=0$$

$$\frac{\partial F}{\partial x'}=n\frac{1}{2\sqrt{(x')^2+(y')^2+(z')^2}}2x'=\frac{nx'}{\sqrt{(x')^2+(y')^2+(z')^2}},$$

$$\frac{\partial F}{\partial x}=\sqrt{(x')^2+(y')^2+(z')^2}\frac{\partial n}{\partial x}$$

$$\Rightarrow\frac{d}{d\sigma}\left[\frac{nx'}{\sqrt{(x')^2+(y')^2+(z')^2}}\right]-\sqrt{(x')^2+(y')^2+(z')^2}\frac{\partial n}{\partial x}=0$$

$$\frac{1}{\sqrt{(x')^2+(y')^2+(z')^2}}\frac{d}{d\sigma}\left[\frac{nx'}{\sqrt{(x')^2+(y')^2+(z')^2}}\right]=\frac{\partial n}{\partial x}$$

Now a path length elements ds is related to $d\sigma$ by

$$ds=\sqrt{\left(\frac{dx}{d\sigma}\right)^2+\left(\frac{dy}{d\sigma}\right)^2+\left(\frac{dz}{d\sigma}\right)^2}d\sigma=\sqrt{(x')^2+(y')^2+(z')^2}d\sigma$$

Also note that

$$\frac{x'}{\sqrt{(x')^2+(y')^2+(z')^2}}=\alpha$$ = direction cosine that ds makes with x axis

$$\Rightarrow\frac{d}{ds}(n\alpha)=\frac{\partial n}{\partial x}$$

Doing the same thing for y and z gives

$$\frac{d}{ds}(n\beta)=\frac{\partial n}{\partial y}$$

$$\frac{d}{ds}(n\gamma)=\frac{\partial n}{\partial z}$$

β,γ =y, z direction cosines

These three equations can be expressed in vector form as

$$\boxed{\frac{d}{ds}(n\hat{s})=\nabla n}$$

This is just the ray equation that we obtained earlier from the eikonal equation!

In other words, the equation for the ray path that satisfies Fermat's Principle is equivalent to the Euler differential equation for the stationary points of the integral $\int nds$.

Thus we have proved that the ray equation is completely equivalent to Fermat's Principle.

To summarize:

Maxwell's Wave equation

↓ (harmonic w)

Helmholtz eqn.

↓ $(\lambda \to 0)$

Eikonal eqn. → ray equation

↓

Fermat's Principle → ray equation