# RAYS: THE EIKONAL TREATMENT OF GEOMETRIC OPTICS

Since ancient times, the notion of **ray or beam propagation** has been one of the most enduring and fundamental concepts in physics. As a zeroth order approximation we might consider a plane wave to be a model of a **beam** and its propagation vector to be a model of a **ray**. This is a reasonable start, but it is a much too restricted view and we can do much better. What we need is a solution to Maxwell's equations which is like a plane wave, but limited in spatial extent. One approach, the simplest, is called variously **ray, Gaussian or geometric optics**. To fully understand this approach in the context of Maxwell's equations, we start by writing the electric and magnetic fields in terms of what we call ***pseudo-simple* waves** -- *viz.*[1]

$$\vec{\mathbf{E}}(\vec{\mathbf{r}},\omega) = \vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)\,\exp\left[-j\,k_0\,\mathsf{S}(\vec{\mathbf{r}},\omega)\right] \qquad \text{[I-1a]}$$

$$\vec{\mathbf{H}}(\vec{\mathbf{r}},\omega) = \vec{\mathbf{h}}(\vec{\mathbf{r}},\omega)\,\exp\left[-j\,k_0\,\mathsf{S}(\vec{\mathbf{r}},\omega)\right] \qquad \text{[I-1b]}$$

where $k_0 = \omega\sqrt{\mu_0\,\varepsilon_0} = \omega/c$

It is assumed that $\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)$ and $\vec{\mathbf{h}}(\vec{\mathbf{r}},\omega)$ are weak functions of position. The scalar phase function $\mathsf{S}(\vec{\mathbf{r}},\omega)$ is the spatially varying phase of the *pseudo-simple* wave. For the cases of pseudo-plane waves and pseudo-spherical waves the phase function is given, respectively, by

$$k_0\,\mathsf{S}(\vec{\mathbf{r}},\omega) = x k_x + y k_y + z k_z \qquad \text{[I-2a]}$$

and

$$k_0\,\mathsf{S}(\vec{\mathbf{r}},\omega) = k_0\sqrt{x^2 + y^2 + z^2} \qquad \text{[I-2b]}$$

We now substitute these pseudo-simple wave expressions (*i.e.* Equations [I-1]) into Maxwell's equations to obtain

$$\exp\left[-j\,k_0\,\mathsf{S}(\vec{\mathbf{r}},\omega)\right]\left\{\vec{\ }\times\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega) - j\,k_0\,\vec{\ }\,\mathsf{S}(\vec{\mathbf{r}},\omega)\times\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)\right\} = -j\,\mu_0\,c k_0\,\vec{\mathbf{h}}(\vec{\mathbf{r}},\omega)\,\exp\left[-j\,k_0\,\mathsf{S}(\vec{\mathbf{r}},\omega)\right] \qquad \text{[I-3a]}$$

---

[1] See, for example, Max Born and Emil Wolf, *Principle of Optics*, Pergamon Press (1986), Chapter 3.

$$\exp\left[-j\, k_0 S(\vec{\mathbf{r}},\omega)\right]\left\{\vec{\nabla}\times\vec{\mathbf{h}}(\vec{\mathbf{r}},\omega) - j\, k_0 \vec{\nabla} S(\vec{\mathbf{r}},\omega)\times\vec{\mathbf{h}}(\vec{\mathbf{r}},\omega)\right\} = j\,\varepsilon(\vec{\mathbf{r}},\omega)\, c k_0\, \vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)\exp\left[-j\, k_0 S(\vec{\mathbf{r}},\omega)\right] \qquad \text{[I-3b]}$$

$$\exp\left[-j\, k_0 S(\vec{\mathbf{r}},\omega)\right]\left\{\vec{\nabla}\cdot\left[\varepsilon(\vec{\mathbf{r}},\omega)\,\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)\right] - j\, k_0\, \varepsilon(\vec{\mathbf{r}},\omega)\,\vec{\nabla} S(\vec{\mathbf{r}},\omega)\cdot\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)\right\} = 0 \qquad \text{[I-3c]}$$

$$\exp\left[-j\, k_0 S(\vec{\mathbf{r}},\omega)\right]\left\{\vec{\nabla}\cdot\vec{\mathbf{h}}(\vec{\mathbf{r}},\omega) - j\, k_0\, \vec{\nabla} S(\vec{\mathbf{r}},\omega)\cdot\vec{\mathbf{h}}(\vec{\mathbf{r}},\omega)\right\} = 0 \qquad \text{[I-3d]}$$

Rearranging, we obtain

$$\vec{\nabla} S(\vec{\mathbf{r}},\omega)\times\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega) - \mu_0\, c\, \vec{\mathbf{h}}(\vec{\mathbf{r}},\omega) = \left[j\, k_0\right]^{-1}\vec{\nabla}\times\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega) \qquad \text{[I-4a]}$$

$$\vec{\nabla} S(\vec{\mathbf{r}},\omega)\times\vec{\mathbf{h}}(\vec{\mathbf{r}},\omega) + \varepsilon(\vec{\mathbf{r}},\omega)\, c\, \vec{\mathbf{e}}(\vec{\mathbf{r}},\omega) = \left[j\, k_0\right]^{-1}\vec{\nabla}\times\vec{\mathbf{h}}(\vec{\mathbf{r}},\omega) \qquad \text{[I-4b]}$$

$$\vec{\nabla} S(\vec{\mathbf{r}},\omega)\cdot\left[\varepsilon(\vec{\mathbf{r}},\omega)\,\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)\right] = \left[j\, k_0\right]^{-1}\vec{\nabla}\cdot\left[\varepsilon(\vec{\mathbf{r}},\omega)\,\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)\right] \qquad \text{[I-4c]}$$

$$\vec{\nabla} S(\vec{\mathbf{r}},\omega)\cdot\vec{\mathbf{h}}(\vec{\mathbf{r}},\omega) = \left[j\, k_0\right]^{-1}\vec{\nabla}\cdot\vec{\mathbf{h}}(\vec{\mathbf{r}},\omega) \qquad \text{[I-4d]}$$

In the **ray**, **Gaussian** or **geometric approximation** we assume that we may neglect the RHS's of these equations. To get something useful we operate on the first equation -- *i.e.* Equation [I-4a] -- with the operator "$\left[\mu_0\, c\right]^{-1}\vec{\nabla} S(\vec{\mathbf{r}},\omega)\times$" as follows:

$$\left[\mu_0\, c\right]^{-1}\vec{\nabla} S(\vec{\mathbf{r}},\omega)\times\{\text{ Equation [I-4a]}\} \qquad \text{[I-5a]}$$

$$\left[\mu_0\, c\right]^{-1}\vec{\nabla} S(\vec{\mathbf{r}},\omega)\times\left\{\vec{\nabla} S(\vec{\mathbf{r}},\omega)\times\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega) - \mu_0\, c\, \vec{\mathbf{h}}(\vec{\mathbf{r}},\omega)\right\} \approx 0 \qquad \text{[I-5b]}$$

Applying the "abc = bac - cab" rule[2] we obtain

---

[2] That is, using $\vec{\mathbf{a}}\times\left(\vec{\mathbf{b}}\times\vec{\mathbf{c}}\right) = \vec{\mathbf{b}}\left(\vec{\mathbf{a}}\cdot\vec{\mathbf{c}}\right) - \vec{\mathbf{c}}\left(\vec{\mathbf{a}}\cdot\vec{\mathbf{b}}\right)$.

$$\begin{aligned}&\left[\mu_0 c\right]^{-1}\Big\{\vec{\nabla}\,\mathrm{S}(\vec{\mathbf{r}},\omega)\left[\vec{\nabla}\,\mathrm{S}(\vec{\mathbf{r}},\omega)\;\;\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)\right]\\ &\qquad -\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)\left|\vec{\nabla}\,\mathrm{S}(\vec{\mathbf{r}},\omega)\right|^2-\mu_0\,\mathrm{c}\;\vec{\nabla}\,\mathrm{S}(\vec{\mathbf{r}},\omega)\times\vec{\mathbf{h}}(\vec{\mathbf{r}},\omega)\Big\}\approx 0\end{aligned} \qquad \text{[I-5c]}$$

which becomes upon substitution from the second Equation [I-4b]

$$\begin{aligned}&\left[\mu_0 c\right]^{-1}\Big\{\vec{\nabla}\,\mathrm{S}(\vec{\mathbf{r}},\omega)\left[\vec{\nabla}\,\mathrm{S}(\vec{\mathbf{r}},\omega)\;\;\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)\right]\\ &\qquad -\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)\left|\vec{\nabla}\,\mathrm{S}(\vec{\mathbf{r}},\omega)\right|^2\Big\}+\varepsilon(\vec{\mathbf{r}},\omega)\,\mathrm{c}\,\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)\approx 0\end{aligned} \qquad \text{[I-5d]}$$

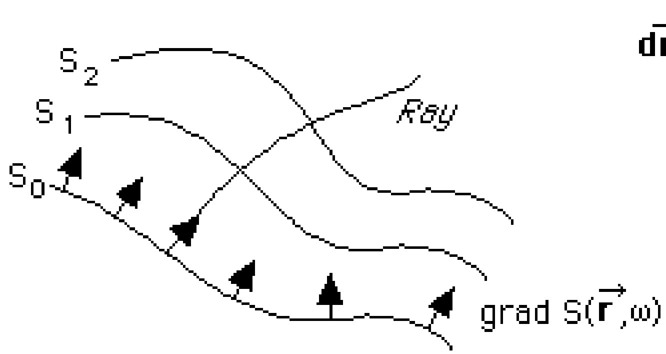
From Equation [I-4c], we see that the first term vanishes in the geometric approximation -- *i.e.* if we neglect the term $\left[j\,\mathrm{k}_0\right]^{-1}\vec{\nabla}\;\left[\varepsilon(\vec{\mathbf{r}},\omega)\,\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)\right]$. Therefore, for non-vanishing $\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)$ we obtain the following **reduction** of Maxwell's equations:

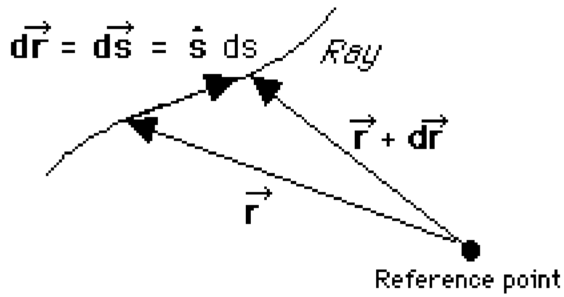
$$\left|\vec{\nabla}\,\mathrm{S}(\vec{\mathbf{r}},\omega)\right|^2\approx\varepsilon(\vec{\mathbf{r}},\omega)\,\mu_0\,c^2=n^2(\vec{\mathbf{r}},\omega) \qquad \text{[I-6]}$$

where $n(\vec{\mathbf{r}},\omega)$ is the index of refraction. More explicitly, we may write an equation for a "ray vector" -- *i.e.* the tangent to a space curve orthogonal to the surfaces of constant $\mathrm{S}(\vec{\mathbf{r}},\omega)$

$$\boxed{\vec{\nabla}\,\mathrm{S}(\vec{\mathbf{r}},\omega)=n(\vec{\mathbf{r}},\omega)\,\hat{\mathbf{s}}=n(\vec{\mathbf{r}},\omega)\,\frac{d\,\vec{\mathbf{r}}}{ds}} \qquad \text{[I-7]}$$

We illustrate the geometric relationships below:




We may now derive the all important **eikonal equation.** To that end, we first take a derivative along the ray direction -- *viz*.

$$\frac{d}{ds}\left[\vec{\nabla} S(\vec{\mathbf{r}},\omega)\right] = \frac{d}{ds}\left[n(\vec{\mathbf{r}},\omega)\,\frac{d\vec{\mathbf{r}}}{ds}\right] \qquad \text{[I-8a]}$$

However, from the definition of the **grad** operator we know that

$$d\left[\vec{\nabla} S(\vec{\mathbf{r}},\omega)\right] = d\vec{\mathbf{r}} \cdot \vec{\nabla}\left[\vec{\nabla} S(\vec{\mathbf{r}},\omega)\right]$$

so that

$$\frac{d\vec{\mathbf{r}}}{ds} \cdot \vec{\nabla}\left[\vec{\nabla} S(\vec{\mathbf{r}},\omega)\right] = \left\{\frac{1}{n(\vec{\mathbf{r}},\omega)}\,\vec{\nabla} S(\vec{\mathbf{r}},\omega)\right\} \cdot \vec{\nabla}\left[\vec{\nabla} S(\vec{\mathbf{r}},\omega)\right] = \frac{d}{ds}\left[n(\vec{\mathbf{r}},\omega)\,\frac{d\vec{\mathbf{r}}}{ds}\right] \qquad \text{[I-8b]}$$

or

$$\frac{1}{2\,n(\vec{\mathbf{r}},\omega)}\,\vec{\nabla}\left\{\vec{\nabla} S(\vec{\mathbf{r}},\omega) \cdot \vec{\nabla} S(\vec{\mathbf{r}},\omega)\right\} = \frac{1}{2\,n(\vec{\mathbf{r}},\omega)}\,\vec{\nabla}\left\{n^2(\vec{\mathbf{r}},\omega)\right\} = \vec{\nabla} n(\vec{\mathbf{r}},\omega) = \frac{d}{ds}\left[n(\vec{\mathbf{r}},\omega)\,\frac{d\vec{\mathbf{r}}}{ds}\right] \qquad \text{[I-8c]}$$

Thus we have obtain the ***eikonal*** [3] equation for the ray vector -- *viz.*

$$\frac{d}{ds}\left[n(\vec{\mathbf{r}},\omega)\,\frac{d\vec{\mathbf{r}}}{ds}\right]=\vec{\nabla}\,n(\vec{\mathbf{r}},\omega) \qquad \text{[I-9]}$$

## APPLICATIONS OF THE EIKONAL EQUATION: THE "ABCD" RAY MATRICES

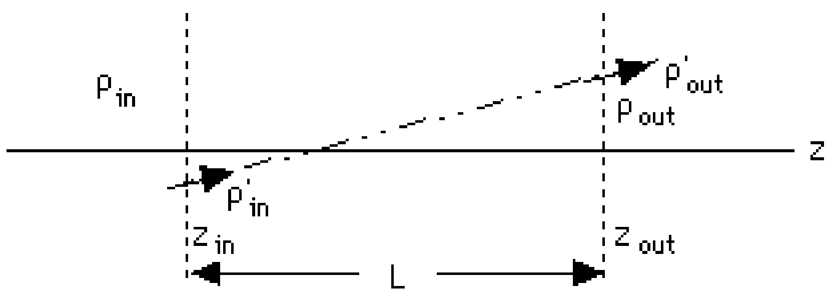
1. Uniform dielectric medium -- *i.e.* $n(\vec{\mathbf{r}},\omega)$ is a constant so that $\frac{d}{ds}\left[\frac{d\vec{\mathbf{r}}}{ds}\right]=0$. so that the ray **must be a straight line** which may be written $\vec{\mathbf{r}}=s\,\vec{\mathbf{a}}+\vec{\mathbf{b}}$ .

   In the two-dimensional paraxial approximation, we assume that $s\approx z$ and write

$$\rho_{\text{out}}=\rho_{\text{in}}+L\left.\frac{d\rho}{dz}\right|_{\text{in}} \qquad \text{[I-10a]}$$

$$\left.\frac{d\rho}{dz}\right|_{\text{out}}=\left.\frac{d\rho}{dz}\right|_{\text{in}} \qquad \text{[I-10b]}$$

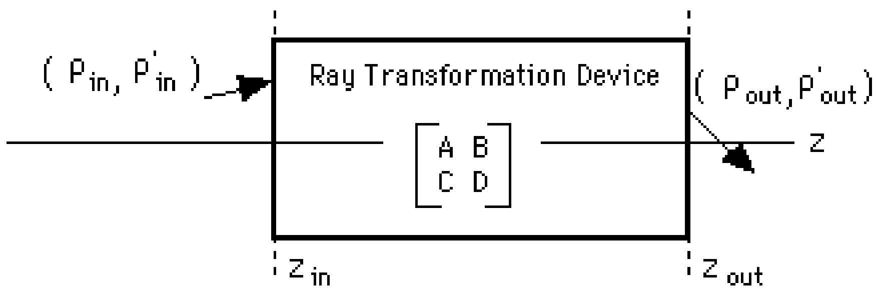
where $\vec{\rho}=\rho\,\hat{\rho}=x\,\hat{\mathbf{x}}+y\,\hat{\mathbf{y}}$ .




[3] The *eikonal* (from the Greek: εικϖν means *image*) was introduced in 1895 by H. Bruns.

We may write results of this sort in the form of the famous and highly useful **ray transform** or **ABCD matrix** -- *viz*.

$$\begin{bmatrix} \rho_{out} \\ \rho'_{out} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \rho_{in} \\ \rho'_{in} \end{bmatrix} \qquad \text{[I-11]}$$




In the case of a uniform dielectric

$$\begin{bmatrix} \rho_{out} \\ \rho'_{out} \end{bmatrix} = \begin{bmatrix} 1 & L \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \rho_{in} \\ \rho'_{in} \end{bmatrix} \qquad \text{[I-12]}$$

so that A = 1, B = L, C = 0, and D = 1

2. A dielectric discontinuity: Starting with Equation [I-7] and noting, once again, that *curl grad* $\{ \text{anything} \} = \vec\nabla \times \vec\nabla \{ \text{anything} \} = 0$ we see that

$$\vec\nabla \times \vec\nabla S(\vec{\mathbf{r}},\omega) = \vec\nabla \times \{ n(\vec{\mathbf{r}},\omega)\, \hat{\mathbf{s}} \} = 0 \qquad \text{[I-13]}$$

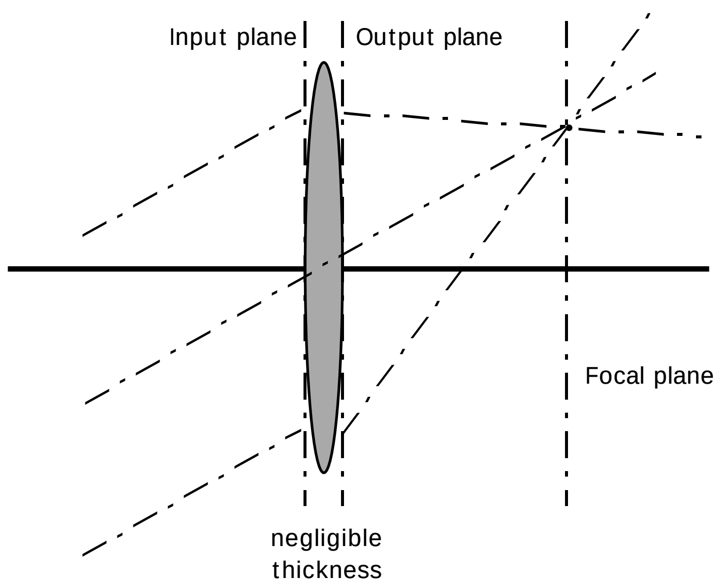
which is identical to the continuity (*saltus*) condition on the electric and magnetic fields at a dielectric interface! Hence $\{ n(\vec{\mathbf{r}},\omega)\, \hat{\mathbf{s}} \}_{tangent}$ is continuous across the dielectric boundary so that $n_1 \sin\theta_1 = n_1 \sin\theta_1$ -- *i.e.* Snell's law! This result in the paraxial approximation may be written in ray matrix form as

$$\begin{bmatrix} \rho_{\text{out}} \\ \rho'_{\text{out}} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \rho_{\text{in}} \\ \rho'_{\text{in}} \end{bmatrix} \approx \begin{bmatrix} 1 & 0 \\ 0 & n_{\text{in}}/n_{\text{out}} \end{bmatrix} \begin{bmatrix} \rho_{\text{in}} \\ \rho'_{\text{in}} \end{bmatrix} \qquad \text{[I-14]}$$

where we have used the approximation that $\sin\theta \approx \tan\theta \approx \theta$.

3. A "<u>Thin" lenses</u>: In passing we note that the ray matrix of a thin lens is given by or, perhaps more accurately, a thin lens is essentially defined by the matrix equation

$$\begin{bmatrix} \rho_{\text{out}} \\ \rho'_{\text{out}} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \rho_{\text{in}} \\ \rho'_{\text{in}} \end{bmatrix} \approx \begin{bmatrix} 1 & 0 \\ -f^{-1} & 1 \end{bmatrix} \begin{bmatrix} \rho_{\text{in}} \\ \rho'_{\text{in}} \end{bmatrix} \qquad \text{[I-15]}$$



4. Axially symmetric <u>GRIN media</u>: Consider the use of GRaded INdex technology to obtain an axially symmetric variation in the index of refraction of the form [4]

$$n(\rho) = n_M \left[ 1 - \Delta \left( \frac{\rho}{a} \right)^m \right] \qquad \text{[I-16]}$$

Within such a GRIN rod, write $\vec{\mathbf{r}} = \rho\,\hat{\rho} + z\,\hat{\mathbf{z}}$ for the ray coordinates and $\nabla n(\vec{\mathbf{r}}) = \hat{\rho}\,\frac{d}{d\rho}\,n(\rho)$ for the index variation. Using the eikonal equation -- *i.e.* Equation [I-9] -- in the paraxial approximation, we find

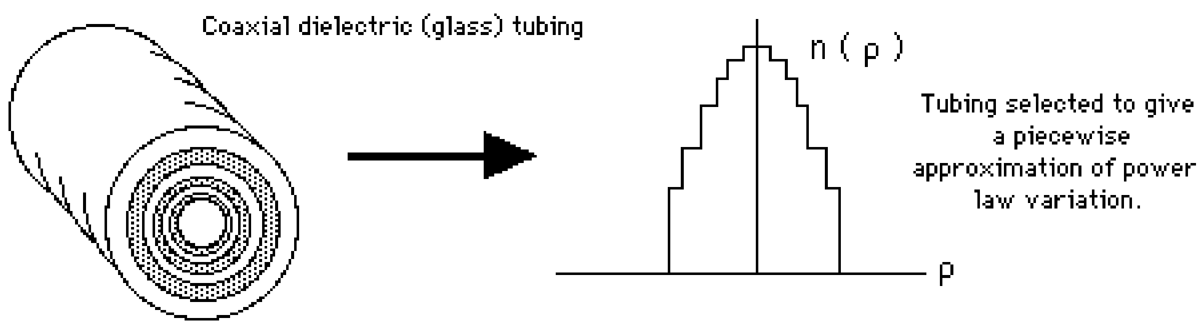
$$\nabla n(\rho) = \frac{d}{ds}\left[ n(\rho)\,\frac{d\vec{\mathbf{r}}}{ds} \right] \approx \frac{d}{dz}\left[ n(\rho)\,\frac{d\vec{\mathbf{r}}}{dz} \right] = \frac{d}{dz}\left\{ n(\rho) \left[ \frac{d\rho}{dz}\,\hat{\rho} + \hat{\mathbf{z}} \right] \right\} \qquad \text{[I-17a]}$$

or

$$\frac{d}{d\rho}\,n(\rho) \approx \frac{d}{dz}\left[ n(\rho)\,\frac{d\rho}{dz} \right] \qquad \text{[I-17b]}$$

---

4 **<u>A note on GRIN technology</u>**: In GRIN technology one may build up a glass rod with a specific radial index of refraction distribution by fusing a sequence of coaxially arranged glass tubes with appropriate index and diameter as illustrated in the following:



Therefore

$$\frac{d^2}{dz^2}\rho \approx \frac{1}{n(\rho)}\frac{d}{d\rho}n(\rho) = \frac{d}{d\rho}\ln[n(\rho)] \qquad \text{[I-17c]}$$

or

$$\begin{aligned} \frac{d^2}{dz^2}\rho &\approx \frac{1}{n(\rho)}\frac{d}{d\rho}\left\{ n_M\left[1-\Delta\left(\frac{\rho}{a}\right)^m\right]\right\} \\ &\approx \frac{1}{n(\rho)}\left[-n_M\frac{m\,\Delta}{a}\left(\frac{\rho}{a}\right)^{m-1}\right] \\ &\approx -\frac{m\,\Delta}{a}\left(\frac{\rho}{a}\right)^{m-1} \end{aligned} \qquad \text{[I-17d]}$$

Doubtless, the simplest and most valuable GRIN materials are designed so that m = 2 -- *i.e.* what is usually called **parabolic** or **quadratic** material -- wherein

$$\frac{d^2}{dz^2}\rho \approx -\frac{2\,\Delta}{a}\left(\frac{\rho}{a}\right)^{2-1} = -\frac{2\,\Delta}{a^2}\rho = -\kappa^2\rho \qquad \text{[I-18]}$$

so that

$$\rho(z) = \rho_{\text{in}}\cos(\kappa z) + \frac{\rho'_{\text{in}}}{\kappa}\sin(\kappa z) \qquad \text{[I-19a]}$$

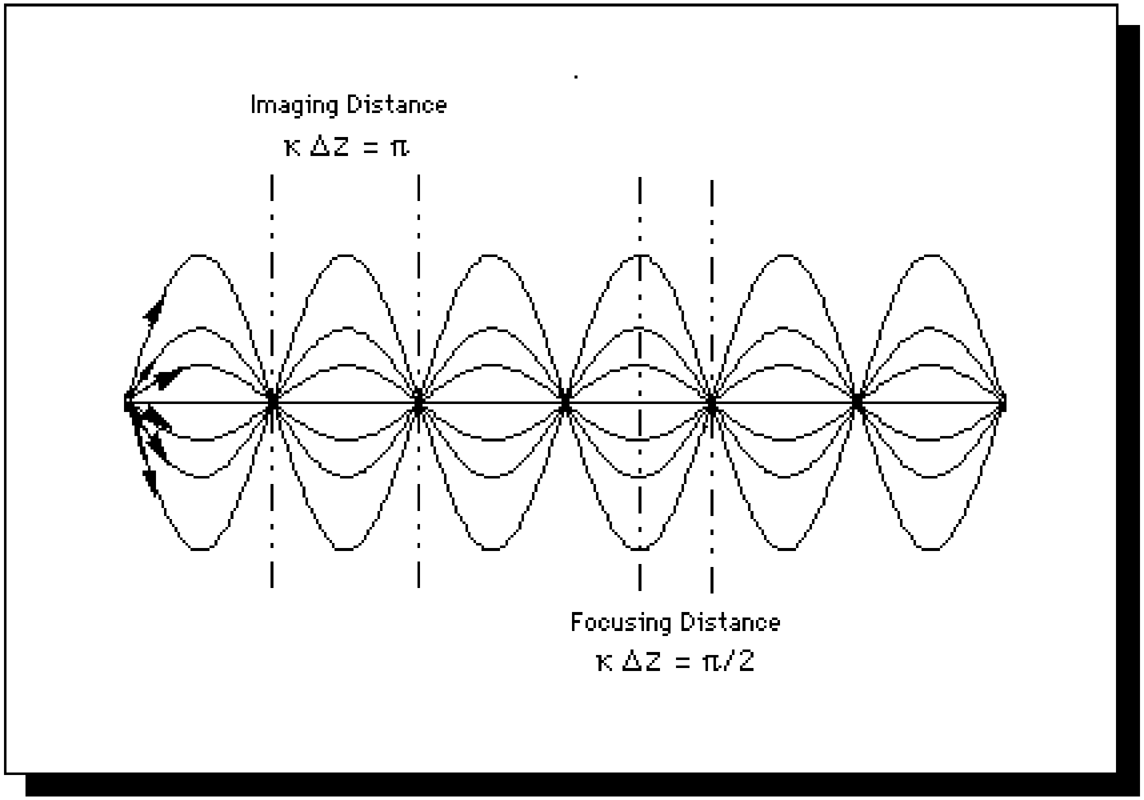
$$\rho'(z) = -\rho_{\text{in}}\,\kappa\,\sin(\kappa z) + \rho'_{\text{in}}\cos(\kappa z) \qquad \text{[I-19b]}$$

So that in terms of a ray transform matrix the GRIN rod is represented by

$$\begin{bmatrix}\rho_{\text{out}}\\ \rho'_{\text{out}}\end{bmatrix} = \begin{bmatrix}A & B\\ C & D\end{bmatrix}\begin{bmatrix}\rho_{\text{in}}\\ \rho'_{\text{in}}\end{bmatrix} \approx \begin{bmatrix}\cos(\kappa z) & \kappa^{-1}\sin(\kappa z)\\ -\kappa\sin(\kappa z) & \cos(\kappa z)\end{bmatrix}\begin{bmatrix}\rho_{\text{in}}\\ \rho'_{\text{in}}\end{bmatrix} \qquad \text{[I-20]}$$

where $\kappa = \sqrt{2\,\Delta/a^2}$ .

**Ray trajectories confined in a GRIN rod.**




1. Mirages: Air adjacent to a hot surface rises in temperature and becomes less dense. Thus over a flat hot surface, such as a desert expanse or a sun drenched roadway, air density **locally** increases with height and the average **refractive index** may be approximated by a simple linear variation of the form

$$n(x) = n_g \{1 + \kappa \ x\} \qquad \text{[I-21]}$$

where $x$ is the vertical height above the planar surface, $n_g$ is the refractive index at ground level, and $\kappa$ is a positive constant.

We may use the **eikonal equation** to find an equation for the approximate ray trajectory -- *i.e.* an equation for ray height $x$ as a function of ground distance $z$ -- of a light ray launched from a height $x_o$ and at an angle $\theta_o$ with respect to the surface of the earth. Therefore,

$$\frac{d}{ds}\left[n(\vec{\mathbf{r}},\omega)\,\frac{d\vec{\mathbf{r}}}{ds}\right] = \vec{\nabla}\, n(\vec{\mathbf{r}},\omega) \qquad \frac{d^2 x}{dz^2} = \frac{1}{n(x,\omega)}\,\frac{d}{dx}\, n(x,\omega) \qquad \text{[I-22a]}$$

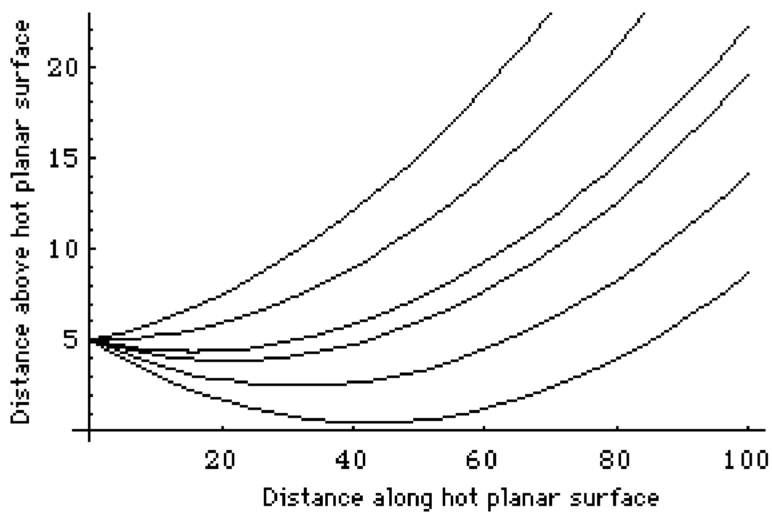
or from Equation [I-21}

$$\frac{d^2 x}{dz^2} \approx \kappa \;. \qquad \text{[I-22b]}$$

Thus, the ray trajectory is given by

$$\vec{\mathbf{r}}(z) = \left[\frac{\kappa}{2}\, z^2 + \tan\theta_0\, z + x_0\right]\hat{\mathbf{x}} + z\,\hat{\mathbf{z}} \qquad \text{[I-23]}$$

**Ray trajectories diverted by a hot surface**



## AN ALTERNATIVE DERIVATION OF EIKONAL EQUATION:

### FERMAT'S PRINCIPLE [5]

Like most laws of physics, the equations of geometric optics can be derived from a **variation principle**. In this context the variation principle is called the Fermat principle which states that a ray always chooses a trajectory that minimizes[6] the optical path length -- *viz.*

$$\int_{P_1}^{P_2} n(x, y, z)\, ds = \text{minimum} \qquad \text{[I-24]}$$

where the line element, $ds$, is measured along a ray and the two end-points $P_1$ and $P_2$ are fixed in space.[7] Analysis of the variation problem is facilitated by choosing the projected coordinate z as the new variable of integration. Accordingly,

$$ds = \sqrt{d\,x^2 + d\,y^2 + dz^2} = \sqrt{1 + x'^2 + y'^2}\;\; dz \;\;, \qquad \text{[I-25]}$$

where $x' = \dfrac{dx}{d\,z}$ and $y' = \dfrac{d\,y}{dz}$, Fermat's variation principle is transformed into the more familiar **Lagrangian form** -- *viz.*

$$\int_{P_1}^{P_2} L(x, y, x', y')\, dz = \text{minimum} \qquad \text{[I-26]}$$

where

$$L(x, y, x', y') = n(x, y, z)\,\sqrt{1 + x'^2 + y'^2}\;. \qquad \text{[I-27]}$$

---

5 See, for example, Dietrich Marcuse, *Light Transmission Optics*, Van Nostrand Reinhold (1972).

6 More precisely, the path must be a local *extremum* and in rare cases may, in fact, be a maximum. See R. Y. Luneberg, *Mathematical Theory of Optics*, University of California Press, Berkeley and Los Angeles (1964).

7 From Equation [ II-7 ] we see that

$$S(P_2) \quad S(P_1) = \int_{P_1}^{P_2} n(x, y, z)\, ds\;\;.$$

The minimization procedure is then well-known in the ***variational calculus*** and leads to the famous **Euler-Lagrangian equations** -- *i.e.*

$$\frac{d}{dz}\frac{\partial L}{\partial x'} - \frac{\partial L}{\partial x} = 0 \qquad \text{[I-28a]}$$

$$\frac{d}{dz}\frac{\partial L}{\partial y'} - \frac{\partial L}{\partial y} = 0 \qquad \text{[I-28b]}$$

When applied to the **Fermat Lagrangian**, as defined in Equation [I-24], these equations yield

$$\frac{d}{dz}\frac{n x'}{\sqrt{1 + x'^2 + y'^2}} = \sqrt{1 + x'^2 + y'^2}\,\frac{\partial n}{\partial x} \qquad \text{[I-29a]}$$

$$\frac{d}{dz}\frac{n y'}{\sqrt{1 + x'^2 + y'^2}} = \sqrt{1 + x'^2 + y'^2}\,\frac{\partial n}{\partial y} \; . \qquad \text{[I-29b]}$$

Using Equation [I-22]we see that the **Euler-Lagrangian** equations may be expressed in the vector form as

$$\left\{ \frac{d}{ds}\left[ n\frac{dx}{ds} \right], \frac{d}{ds}\left[ n\frac{dy}{ds} \right] \right\} = \left\{ \frac{\partial n}{\partial x}, \frac{\partial n}{\partial y} \right\} \qquad \text{[I-30]}$$

**which is precisely the content of Equation [I-9]-- QED**.

## HAMILTONIAN FORMULATION OF RAY OPTICS

The analogy between ray optics and particle mechanics is most striking when the equations of ray optics are expressed in Hamiltonian form.[8] To that end, we define the **generalized momentum** which is canonically conjugate to x and y by the vector equation

8 The formal theory of optical systems was developed by Sir W. R. Hamilton in 1828-37.

$$\{p_x, p_y\} = \left\{ \frac{\partial L}{\partial x}, \frac{\partial L}{\partial y} \right\} . \qquad \text{[I-31]}$$

The Hamiltonian is then define in terms of the generalized momentum by the relation

$$H(x, y, p_x, p_y) \equiv p_x x' + p_y y' - L(x, y, x', y') . \qquad \text{[I-32]}$$

With the assumed functional dependence of the Hamiltonian, we form the derivatives

$$\frac{\partial H}{\partial p_x} = x' + p_x \frac{\partial x'}{\partial p_x} + p_y \frac{\partial y'}{\partial p_x} - \frac{\partial L}{\partial x'} \frac{\partial x'}{\partial p_x} - \frac{\partial L}{\partial y'} \frac{\partial y'}{\partial p_x} \qquad \text{[I-33a]}$$

$$\frac{\partial H}{\partial p_y} = p_x \frac{\partial x'}{\partial p_y} + y' + p_y \frac{\partial y'}{\partial p_y} - \frac{\partial L}{\partial x'} \frac{\partial x'}{\partial p_y} - \frac{\partial L}{\partial y'} \frac{\partial y'}{\partial p_y} . \qquad \text{[I-33b]}$$

Given the definitional relationships embodied in Equation [I-28]we see that these expression reduce to one set of Hamilton's equation -- *viz*.

$$\left\{ \frac{dx}{dz}, \frac{dy}{dz} \right\} = \left\{ \frac{\partial H}{\partial p_x}, \frac{\partial H}{\partial p_y} \right\} . \qquad \text{[I-34]}$$

The other set of Hamilton's equation -- *viz*.

$$\left\{ \frac{dp_x}{dz}, \frac{dp_y}{dz} \right\} = \left\{ -\frac{\partial H}{\partial x}, -\frac{\partial H}{\partial y} \right\} . \qquad \text{[I-35]}$$

follow directly from the Euler-Lagrangian equations -- *i.e.* Equations [I-25a]and [I-25b]-- and the definitions embodied in Equation [I-28]. Using the Fermat Lagrangian we see that

$$\left\{p_x, p_y\right\} = \left\{\frac{\partial L}{\partial x}, \frac{\partial L}{\partial y}\right\} = \left\{\frac{n x'}{\sqrt{1+x'^2+y'^2}}, \frac{n y'}{\sqrt{1+x'^2+y'^2}}\right\} \qquad \text{[I-36]}$$

and consequently that we may solve for $\{x', y'\}$ in terms of $\{p_x, p_y\}$ as

$$\{x', y'\} = \left\{\frac{p_x}{\sqrt{n^2 - p_x^2 - p_y^2}}, \frac{p_y}{\sqrt{n^2 - p_x^2 - p_y^2}}\right\} \qquad \text{[I-37]}$$

Substituting into Equation [I-29], we find an expression for the **Fermat** or **ray optics Hamiltonian** -- *viz.*

$$H = -\sqrt{n^2 - p_x^2 - p_y^2} \ . \qquad \text{[I-38]}$$

which resembles the mechanical Hamilton of a relativistic particle -- *i.e.*

$$c\sqrt{m_0^2 c^2 + p_x^2 + p_y^2 + p_z^2} \ .$$

But the analogy is even stronger in the paraxial approximation where the Hamiltonian is approximated by an expression which is identical in form with the Hamiltonian of a non-relativistic particle -- *viz.*

$$H = -n\sqrt{1 - \frac{p_x^2 + p_y^2}{n^2}} \approx \frac{p_x^2 + p_y^2}{2\langle n \rangle} - n \qquad \text{[I-39]}$$

when $p_x, p_y << \langle n \rangle$ .[9]

---

[9] Applying the quantization rules of quantum mechanics to these Hamiltonians, we can go full circle and recover wave optics from ray optics. Equation [ II-35 ] leads directly to the equivalent of the relativistic Klein-Gordon equation while the equivalent of the nonrelativistic Schrödinger equation follows directly from Equation [ II-36 ].