# Fundamental Understanding of the Propagation of Light Using Geometry

**Ahmed K. Alsamarrai**

*Department of Physics, Faculty of Education, University of Baghdad, Baghdad, Iraq*

***The treatment of light as wave motion allows for a region of approximation in which the wavelength is considered to be negligible compared with the dimensions of the relevant components of the optical system. This region of approximation is called geometrical optics. When the wave character of the light may not be ignored, the field is known as physical optics. Since the wavelength of light is very small compared to ordinary objects, early unrefined observations of the behavior of a light beam passing through apertures or around obstacles in its path could be handled by geometrical optics.***

***Keywords***: *Optics, Light propagation, Geometrical optics, Optical path length*
**Received**: 5 October 2009, **Revised**: 31 October 2009, **Accepted**: 7 November 2009

Within the approximation represented by geometrical optics, light is understood to travel out from its source along straight lines or **rays**. The ray is simply the path along which energy is transmitted from one point to another in an optical system. The ray is a useful, although abstract, construct; perhaps the best approximation to a ray of light is a pencil-like laser beam. When a light ray traverses an optical system consisting of several homogeneous media in sequence, the optical path is a sequence of straight-line segments. The laws of geometrical optics that describe the subsequent direction of the rays are succinctly stated as:

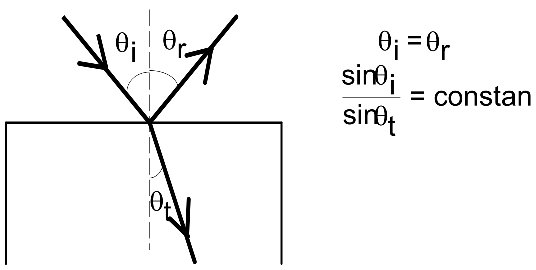
**Law of Reflection**: When a ray of light is reflected at an interface dividing two uniform media, the reflected ray remains within the **plane of incidence**, and the angle of reflection equals the angle of incidence. The plane of incidence includes the incident ray and the normal to the point of incidence.

**Law of Refraction (Snell's Law)**: When a ray of light is refracted at an interface dividing two uniform media, the transmitted ray remains within the plane of incidence and the sine of the angle of refraction is directly proportional to the sine of the angle of incidence.

These laws can be visually seen in the following figure

***Huygens' Principle***

The Dutch physicist Christian Huygens imagined each point of a propagating disturbance as capable of originating new pulses that contributed to the disturbance an instant later. To show how his model of light propagation implied the laws of geometrical optics, he formulated a principle which says that **each point on the leading surface of a wave disturbance may be regarded as a secondary source of spherical waves, which themselves progress with the speed of light in the medium and whose envelope at later times constitutes the new wavefront**.

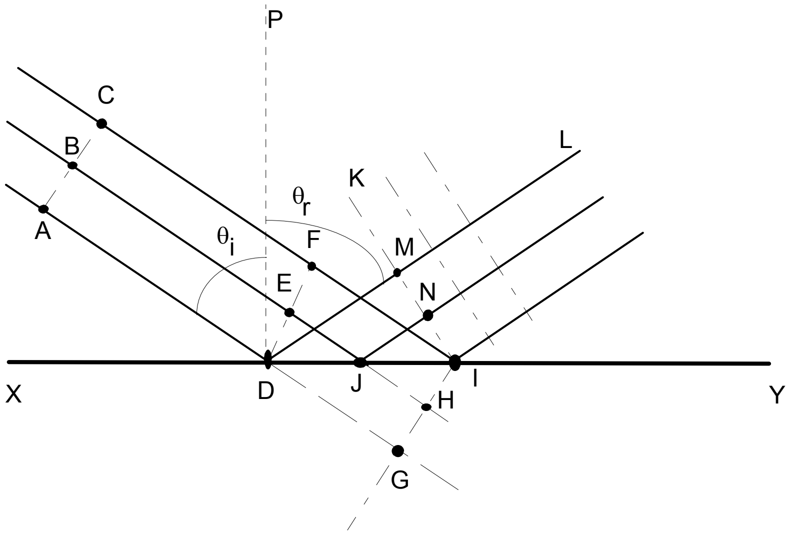



Notice that the new wavefront is tangent to each wavelet at a single point. According to Huygens, the remainder of each wavelet is to be disregarded in the application of the principle. In so disregarding the effectiveness of the overlapping wavelets, Huygens avoided the possibility of diffraction of the light into the region of geometric shadow. Huygens also ignored the wavefront formed by the back half of the wavelets, since these wavefronts implied a light disturbance traveling in the opposite direction. Despite weaknesses in this model, remedied later by Fresnel and others, Huygens was able to apply his principle to prove the laws of both refection and refraction.

ISSN 1999-656X  

***Law of Reflection from Huygen's Principle***

The figure illustrates Huygen's construction for a narrow, parallel beam of light to prove the law of reflection. Huygen's principle must be modified to accommodate the case in which a wavefront, such as *AC*, encounters a plane interface, such as *XY*, at an angle. Here the angle of incidence of the rays *AD*, *BE*, and *CF* relative to the perpendicular *PD* is $\theta_i$. Since points along the plane wavefront do not arrive at the interface simultaneously, allowance is made for these differences in constructing the wavelets that determine the reflected wavefront.



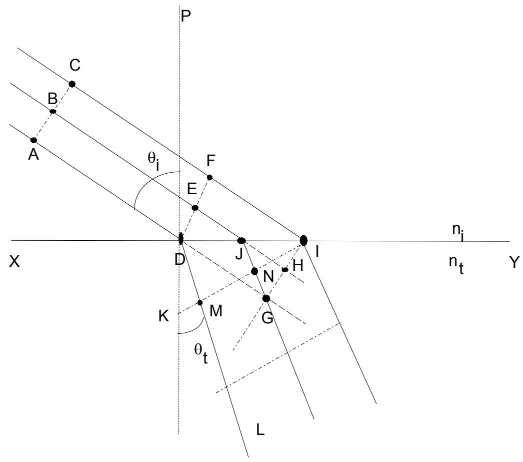

If the interface *XY* were not present, the Huygens construction would produce the wavefront *GI* at the instance ray *CF* reached the interface at *I*. The intrusion of the reflecting surface, however, means that during the same time interval required for ray *CF* to progress from *F* to *I*, ray *BE* has progressed from *E* to *J* and then a distance equivalent to *JH* after reflection. Thus a wavelet of radius *JH* centered at *J* is drawn above the reflecting surface. Similarly, a wavelet of radius *DG* is drawn centered at *D* to represent the propagation after reflection of the lower part of the beam. The new wavefront, which must now be tangent to these wavelets at points *M* and *N*, and include the point *I*, is shown as *KI* in the figure. A representative reflected ray is *DL*, shown perpendicular to the reflected wavefront. The normal *PD* drawn for this ray is used to define angles of incidence and reflection for the beam. The construction clearly shows the equivalence between the angles of incidence and reflection.

***Law of Refraction using Huygen's Principle***

Similarly, we can use a Huygens construction to illustrate the law of refraction.

Here we must take into account a different speed of light in the upper and lower media. If the speed of light in vacuum is *c*, we express the speed in the upper medium by the ratio $c/n_i$, where $n_i$ is the refractive index. Similarly, the speed of light in the lower medium is $c/n_t$. The points *D*, *E* and *F* on the incident wavefront arrive at points *D*, *J* and *I* of the plane interface *XY* at different times. In the absence of the refracting surface, the wavefront *GI* is formed at the instant ray *DF* reaches *I*. During the progress of ray *CF* from *F* to *I* in time *t*, however, the ray *AD* has entered the lower medium, where the speed is different. Thus if the distance *DG* is $v_i t$, a wavelet of radius $v_t t$ is constructed with center at *D*. The radius *DM* can also be expressed as

$$DM = v_t t = v_t\left(\frac{DG}{v_i}\right) = \left(\frac{n_i}{n_t}\right)DG$$



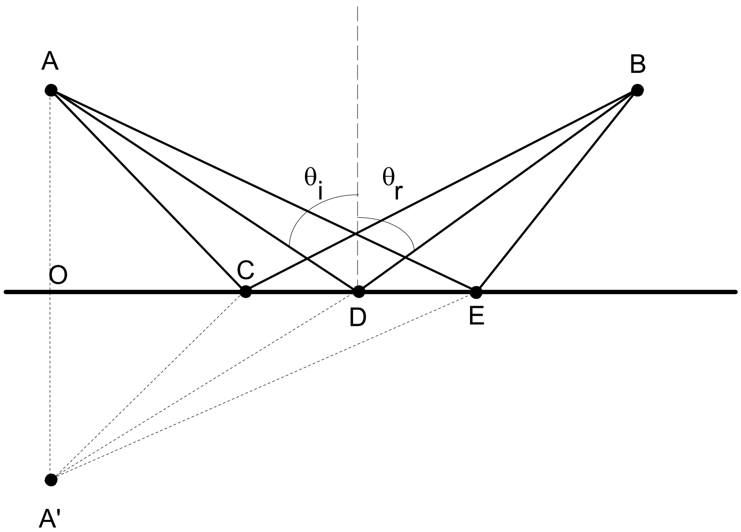
Similarly, a wavelet of radius $(n_i/n_t)JH$ is drawn centered at *J*. The new wavefront *KI* includes point *I* on the interface and is tangent to the two wavelets at points *M* and *N*. The geometric relationship between the angles $\theta_i$ and $\theta_t$, formed by the representative incident ray *AD* and refracted ray *DL*, is **Snell's law**, which may be expressed as

$$n_i \sin\theta_i = n_t \sin\theta_t \qquad (1)$$

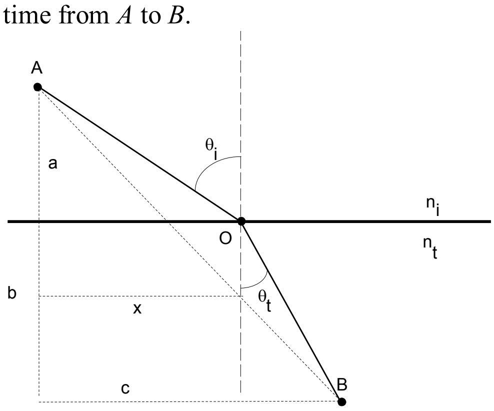
***Fermat's Principle***

The laws of geometrical optics can also be derived from a different fundamental hypothesis. Let us suppose that nature is economical, and thus requires that the time required for light to travel from point *A* to *B* is the minimum time required. To prove the law of reflection, we use the fact that, for propagation in the same medium, the velocity is a constant and this minimizing the time is the same as minimizing the distance traveled. Consider the following drawing.

Three possible paths from *A* to *B* are shown. Let's look at the arbitrary path *ACB*. If point *A'* is constructed on the perpendicular *AO* such that *AO=A'O*, the right triangles *AOC* and *A'OC* are equal. Thus *AC=A'C* and the distance traveled by the ray of light from *A* to *B* via *C* is the same distance from *A'* to *B* via *C*. The shortest distance from *A'* to *B* is obviously the straight line *A'DB*, so the path *ADB* is the correct choice taken by the actual light ray. Geometry shows that for this path, $\theta_i=\theta_r$. Also note that to maintain *A'DB* as a single straight line, the reflected ray must remain within the plane of incidence.




We can also prove the law of refraction. If the light travels more slowly in the second medium, light bends at the interface so as to take a path that favors a shorter time in the second medium, thereby minimizing the overall transit time from *A* to *B*.




Mathematically, we are required to minimize the total time

$$t=\frac{AO}{v_i}+\frac{OB}{v_t}=\frac{\sqrt{a^2+x^2}}{v_i}+\frac{\sqrt{b^2+(c-x)^2}}{v_t} \qquad (2)$$

Since other choices of path change the position of the point *O* and therefore the distance *x*, we can minimize the time by setting *dt/dx*=0:

$$0=\frac{dt}{dx}=\frac{x}{v_i\sqrt{a^2+x^2}}-\frac{c-x}{v_t\sqrt{b^2+(c-x)^2}} = \frac{\sin\theta_i}{v_i}-\frac{\sin\theta_t}{v_t} \qquad (3)$$

where the last step used the relationships shown in the figure. Introducing the refractive indices of the media, we arrive at Snell's law

$$n_i \sin\theta_i = n_t \sin\theta_t \qquad (4)$$

ISSN 1999-656X 

Fermat's principle, like that of Huygens, required refinement to achieve more general applicability. Situations exist where the actual path taken by a light ray may represent a maximum time or even one of many possible paths, all requiring equal time. As an example of the latter case, consider light propagating from one focus to the other inside an ellipsoidal mirror, along any of an infinite number of possible paths. Since the ellipse is the locus of all points whose combined distances from the two foci remain constant, all paths are indeed of equal time. A more precise statement of Fermat's principle, which requires merely an extremum relative to neighboring paths, may be given as follows: **The actual path taken by a light ray in its propagation between two given points in an optical system is such as to make its optical path equal, in the first approximation, to other paths closely adjacent to the actual one**.

With this formulation, Fermat's principle falls in the class of problems called variational calculus, a technique which determines the form of a function that minimizes a definite integral. In optics, the definite integral is the integral of the time required for the transit of a light ray from starting to finishing points.

***Optical Path Length***

Suppose that we have a stratified material composed of m layers, each having a different index of refraction. The transit time across the layers will then be

$$t = \frac{s_1}{v_1} + \frac{s_2}{v_2} + \cdots + \frac{s_m}{v_m} = \sum_{i=1}^{m} \frac{s_i}{v_i} = \frac{1}{c}\sum_{i=1}^{m} n_i s_i \qquad (5)$$

where the summation is called the **optical path length** traversed by the ray. Clearly for an inhomogeneous medium where $n$ is a function of position, the summation must be changed to an integral

$$(OPL) = \int n(s)ds$$

Since the optical path length is related to the time, we can restate Fermat's principle again as **a light ray in going from point *A* to point *B* must traverse an optical path length that is stationary with respect to variations of that path**.

***Optical Reversibility***

Consider applying Fermat's principle to an optical system. Since the time must be minimized, we see that the same path is predicted regardless of whether we start at $A$ and travel to $B$, or start at $B$ and travel to $A$. In general, any actual ray of light in an optical system, if reversed in direction, will retrace the same path backward. Before discussing the formation of images in a general way, let's look at the simplest - and experimentally, the most accessible - case of images formed by plane mirrors. In this context it is important to distinguish between **specular reflection** from a perfectly smooth surface and **diffuse reflection** from a granular or rough surface. Specular reflection occurs when all the rays of a parallel beam incident on the surface obey the law of reflection from a plane surface and therefore reflect as a parallel beam. In the case of diffuse reflection, although the law of reflection holds locally, the microscopically granular surface results in reflected rays in various directions and thus a diffuse scattering of the originally parallel rays of light. Every plane surface will produce some such scattering, since a perfectly smooth surface is not obtainable in reality. In many cases, however, the diffuse scattering is small and we can approximate the reflection as specular reflection.

Consider the specular reflection of a single light ray from the $x$-$y$ plane. By the law of reflection, the reflected ray remains within the plane of incidence, making equal angles with the normal at the point of contact. If the path is resolved into components, it is clear that the direction of the incident ray is altered only by reflection along the $z$ direction, and then in such a way that its $z$ component is simply reversed. If the direction of the incident ray is described by its unit vector $\hat{r}_1 = (x, y, z)$, then the reflection causes

$$\hat{r}_1 = (x, y, z) \rightarrow \hat{r}_2(x, y, -z) \qquad (6)$$

It follows that if a ray is incident from such a direction as to reflect sequentially from all three coordinate planes, then

$$\hat{r}_1 = (x, y, z) \rightarrow \hat{r}_2(-x, -y, -z) \qquad (7)$$

and the ray returns precisely parallel to the line of its original approach. A network of such corner reflectors ensures the exact return of a beam of light.