Teaching Geometrical Optics using Matrix Method

Sanat Karmakar,

Department of Physics, Jadavpur University, 188, Raja S. C. Mullick Road, Jadavpur, Kolkata 700032.

Abstract

Students in under graduate level often find difficult to understand the geometrical optics. Despite its enormous applications in designing optical equipments, starting from simple lens, compound lens to microscope and other optical equipments, this subject is neglected. Therefore, this article has primarily been targeted to under graduate students. The matrix scheme in geometrical optics is very useful and replaces complex optical system by a simple 2 $\times$ 2 matrix

Keywords: Geometrical optics, Matrix method.

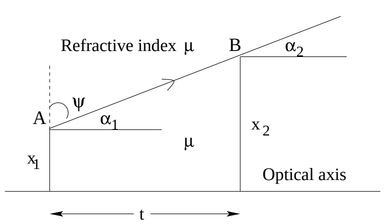
1 Introduction

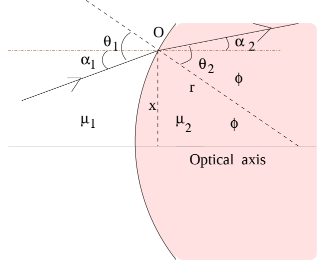
We enter into the realm of geometrical optics, if the wavelength of light is vanishingly small as compared to the dimension of the equipment available for their study. The concept of rays in geometrical optics is important which does not need knowledge of electromagnetic behavior of the light. Light rays in geometrical optics suffices to describe the image formation by a optical system. Drawing a ray diagram to specify the position of the image often difficult for a complicated optical system, such as combination of lenses. Matrix method with paraxial approximation provides a simple way of describing a image formation of complicated optical system. Stuents in under graduate level often find difficult to understand the aim or the purpose of the matrix method. In this article, I shall describe a step by step development of matrix formalism of optical system.

Let us consider the position of the point object (x_0, y_0) in object space and the position of the image (x_i, y_i) in the image space. These two conjugate points can be linked with the transformation matrix A. As the transformation occurs from one plane (object space) to another plane (image space), the matrix A must be 2×2.

$$\begin{bmatrix} x_i \\ y_i \end{bmatrix} = A \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \tag{1}$$

Where the transformation matrix A operates on the column matrix, specifying the position of the object gives the position of the image. The principle of reversibility demands that determinant of A must be unity. $A_{11}A_{22} - A_{12}A_{21} = 1$. Our job is to determine the matrix A. Any image formation process can be thought as the combination of alternate translation and refraction. Therefore, the matrix A is indeed related to translation and refraction operation. The matrices, representing translation and refraction which constitute the transformation matrix A are now necessary to find out. We represent an optical ray in a plane, by a two component vector $\begin{bmatrix} \alpha \\ x \end{bmatrix}$. It is convenient to replace the paraxial angle α by the optical direction cosine $\lambda = \mu \cos\psi = \mu \sin\alpha \simeq \mu\alpha$. Therefore, optical ray can be represented by a two component column vector as $\mathbf{X} = \begin{bmatrix} \lambda \\ x \end{bmatrix}$





Box -1

Ray transfer due to translation from A to B

$\mu\alpha_1 = \mu\alpha_2$; $\lambda_1 = \lambda_2$

$x_2 = x_1 + t\tan\alpha_1 = x_1 + t\alpha_1 = x_1 + \frac{t}{\mu}\lambda_1$;

$$\begin{bmatrix} \lambda_2 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{t}{\mu} & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ x_1 \end{bmatrix}$$

Ray tranfer due to refraction at O

$\theta_1 = \alpha_1 + \phi$; $\theta_2 = \alpha_2 + \phi$

Snell's law $\Rightarrow \mu_1 \sin\theta_1 = \mu_2 \sin\theta_2 \approx \mu_1\theta_1 = \mu_2\theta_2$

$\mu_1(\alpha_1 + \phi) = \mu_2(\alpha_2 + \phi); \phi \approx \frac{x}{r}$

$\mu_2\alpha_2 = \mu_1\alpha_1 - (\mu_2 - \mu_1)\phi = \mu_1\alpha_1 - \frac{(\mu_2 - \mu_1)}{r}x$

$$\lambda_2 = \lambda_1 - Px_1 \text{ ; } x_2 = x_1 \Rightarrow \begin{bmatrix} \lambda_2 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & -P \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ x_1 \end{bmatrix}$$

Where P $= \frac{(\mu_2 - \mu_1)}{r}$, r is the radius of curvature at O

2 Ray Transfer Matrix

The propagation of optical ray AB in the medium of refractive index μ is equivalent to the translation from A to B, separated by a distance t. Therefore, the ray transfer matrix describing the propagation over a distance t is

$$T = \begin{bmatrix} 1 & 0 \\ \frac{t}{\mu} & 1 \end{bmatrix} \tag{2}$$

However, propagation though a interface between two media of different refractive index will be followed by a refraction at the interface. The ray transfer matrix describing the refraction in a curved interface can easily be derived (See Box-1) and is given by

$$R = \begin{bmatrix} 1 & \frac{\mu_2-\mu_1}{r} \\ 0 & 1 \end{bmatrix} \tag{3}$$

where r is the radius of curvature at the point of refraction. Refraction at a planar interface ($r \rightarrow \infty$), the above matrix reduces to a unit matrix $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

As the reflection is a special case of refraction. Reflection matrix can easily be obtained by modifying Snell's law (μ_1 sinθ_i = μ_1 sinθ_r) as $\theta_i = \theta_r$ and μ_1 = - μ_2 = μ, where μ is the refractive index of the medium in which the mirror is immersed. Therefore, the reflection matrix is given by $\begin{bmatrix} 1 & \frac{2\mu}{r} \\ 0 & 1 \end{bmatrix}$

Note that determinant of matrices T and R is unity, as required.

3 System Matrix

The ray transfer matrix through a simple optical system, such as thin lens or thick lens can easily be obtained from the elementary T and R matrices. If the light propagates through a complex optical structure, such as system of lenses. then the entire ray transfer matrix, A, is the product of ray transfer matrices of individual optical components. By property of determinant (det (ABC)=det(A) det(B) det(C)), the matrix A has determinant unity as individual matrices have determinant one. This approach essentially reduces a very complicated optical system to a very simple 2 $\times$ 2 matrix.

Now, it is desirable to derive a system matrix for a complex optical system, such as combination of lenses. System matrix essentially describes the transfer of rays within the optical system which excludes both translation matrices in either side of the system (Fig. 1). Let the general form of system matrix be

$$\begin{bmatrix} b & -a \\ -d & c \end{bmatrix} \tag{4}$$

Therefore, if we multiply two translation matrix in either side of system matrix, a compete transformation matrix, A. This transformation will provide us image position given the position of the object without tracing the optical rays. Interestingly, the transformation matrix can be described in terms of total magnification and the effective focal length of the optical system. Sometime students would like to draw the ray diagram which will trace the image position, given the position of the object. In that it will be convenient to define certain planes, known as cardinal points. Positions of cardinal points can easily be determined from the transformation matrix. As an example, the system matrix for a thick lens (refractive index μ) of radii of

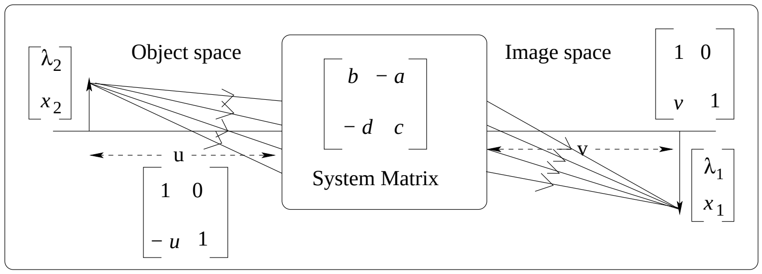


Figure 1: Schematic representation of image formation by optical system

curvature r_1 and r_2 and thickness t can be written as

$$\begin{bmatrix} b & -a \\ -d & c \end{bmatrix} = \begin{bmatrix} 1 & -P_1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -0 \\ \frac{t}{\mu} & 0 \end{bmatrix} \begin{bmatrix} 1 & -P_2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 - \frac{t}{\mu}P_2 & -P_1 - P_2(1 - \frac{t}{\mu}P_1 \\ \frac{t}{\mu} & 1 - \frac{t}{\mu}P_1 \end{bmatrix} \tag{5}$$

For a thin lens ($t \approx 0$), the above system matrix reduces to $\begin{bmatrix} 1 & -P_1 - P_2 \\ 0 & 1 \end{bmatrix}$. The element a can easily be identified as inverse of focal length of the lens.

4 Transformation matrix A

With reference to the Fig. 1, the postion of the image can be written as

$$\begin{bmatrix} \lambda_2 \\ x_2 \end{bmatrix} = A \begin{bmatrix} \lambda_1 \\ x_1 \end{bmatrix} \tag{6}$$

Where,

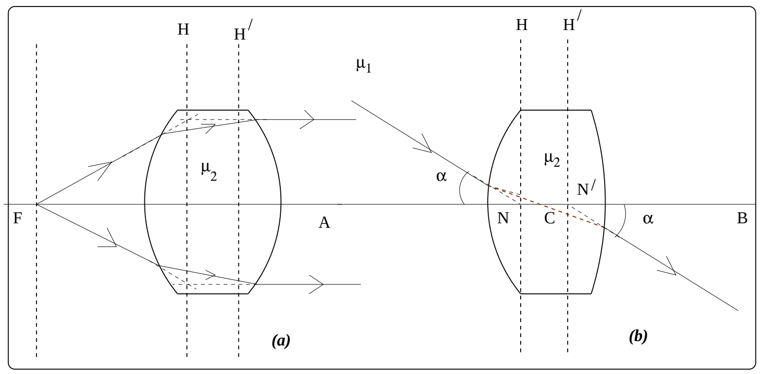
$$A = \begin{bmatrix} 1 & 0 \\ v & 1 \end{bmatrix} \begin{bmatrix} b & -a \\ -d & c \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -u & 1 \end{bmatrix} = \begin{bmatrix} b + au & -a \\ 0 & c - av \end{bmatrix} \tag{7}$$

For an axial object, $x_1 = 0$. therefore, in the image space x_2 must be zero. This condition implies that the matrix element A_{21} must be zero. The magnification of the optical system, described by M $= \frac{x_2}{x_1} = c - av = A_{22}$. As the determinant of A is unity, the element $A_{11} = b + au = \frac{1}{M}$. It can be easily shown that the matrix element A_{12} can be described as the effective focal length of the optical system. Therefore,

$$\begin{bmatrix} M & -a \\ 0 & 1/M \end{bmatrix} \tag{8}$$

5 Principal Points and Planes

The concept of principal planes allows us to characterise complex optical system, such as microscope objectives, eyepieces, camera lenses, projector lenses etc. These optical equipments are basically composed of compound lens. Therefore, imaging properties of the compound lens can be characterised, once the focal length and principal planes are known. As already discussed, elements of a transformation matrix can be described in terms of magnification of the object and the focal length of the optical system. I shall now define and then represent the principal planes in terms of elements of a transformation matrix. Principal planes are two conjugate



(a)

Figure 2: (a) Primary and (b) location of optical centre (C), principal points (N and N′) and principal planes (H and H′). In case of thick lens, the conjugate points N and N′) are also known as nodal points or unit points if the refractive indices in both side of the lens are same. These planes are normal to the optical axis (AB) and passing through the nodal points or principal points

planes having positive unit lateral magnifications. In Fig. 2, H and H′ are two principal plane and the point of intersection of these planes with the optical axis are termed as principal points or unit points. Therefore, any ray approaching to this points emerges parallel to the incident ray provided the refractive indices in both sides are same. In this case, nodal points which represents two conjugate points having unit positive angular magnification will be the same points or planes as that of unit points or planes.

It is convenient to measure the object (u) and image distance (v) from the principle point so as to obtain the familiar relationship between object and image distances, i.e, $\frac{1}{v} - \frac{1}{u} = a = \frac{1}{f}$. We can readily determine the position of the first unit plane d_{u1} and second unit plane (d_{u2}) from the elements of a system matrix. By definition of principal planes, M = 1, Therefore $b + ad_{u1} = 1$, $d_{u1} = \frac{1-b}{a}$ and $d_{u2} = \frac{c-1}{a}$.

6 Concluding remarks

The matrix method for geometrical optics describes complex optical by simple 2×2 matrix. It is very useful in understanding operation of optical system and also design of complex optical system from individual optical components. In matrix scheme one can easily determine the positions of the image and principal planes, focal planes etc without drawing the complicated ray diagram. However, it is limited the paraxial approximation in particular. Non- paraxial rays lead to aberration of the optical system. Therefore, this matrix scheme does not provide the information about image quality.

Suggested Reading

[1] A. Gerrard and J. M. Burch, *Introduction to matrix method in Optics*, John Wiley and Sons, New York (USA), 1975.

[2] F. A. Jenkins and H. E. White, *Fundamentals of Optics*, McGrow-Hill International Editions, pp. 44-99.

[3] A. Ghatak, *Optics*,Tata McGraw Hill, New Delhi, pp. 87-110, 1992.

[4] M. Born and W. Wolf, *Principles of Optics*, Pergamon Press, Oxford, 1975.