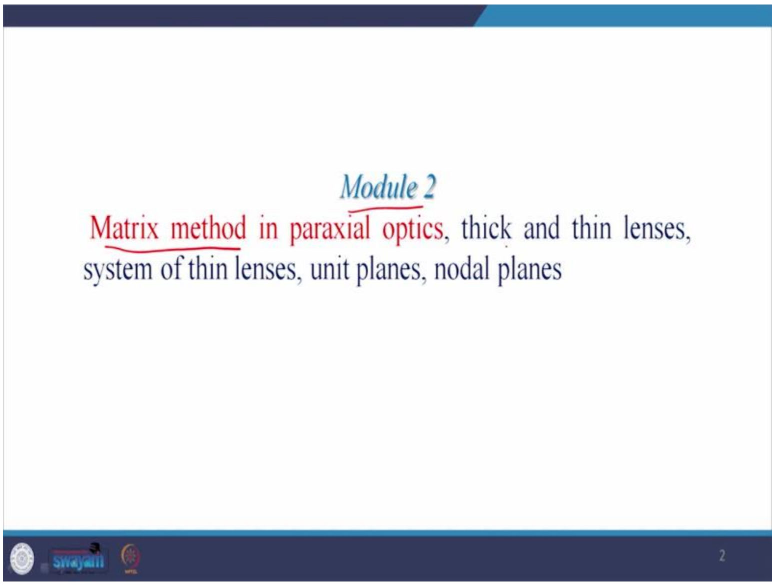
Applied Optics
Professor Akhilesh Kumar Mishra
Department of Physics
Indian Institute of Technology, Roorkee
Lecture 06
Matrix Method in Paraxial Optics - 1

Hello everyone, welcome again to my class. Today we will learn about a very important concept which is matrix method.

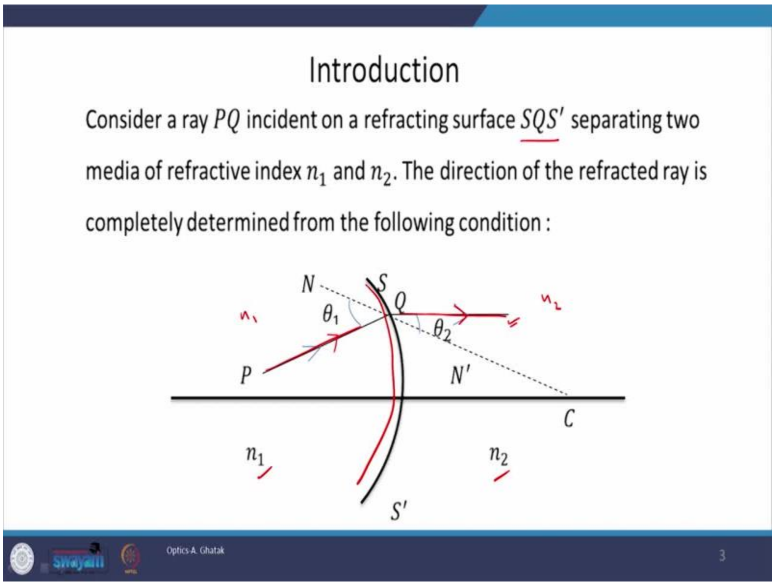
(Refer Slide Time: 0:35)



Now, we will start with matrix method. We were in module 1 and we talked about the reflection and refraction. And while talking about this refraction and reflection, we consider 2 types of interfaces, which is straight interface and spherical interfaces. We not only considered the single spherical interfaces, but we also talked about double interfaces wherein we covered refraction through thin lens. But whenever we do refraction through thin lens or suppose we have multiple lenses which are stacked together with certain spacing then every time we will have to use this lens formula for each interface, we will have to implement this lens formula which is quite cumbersome process.

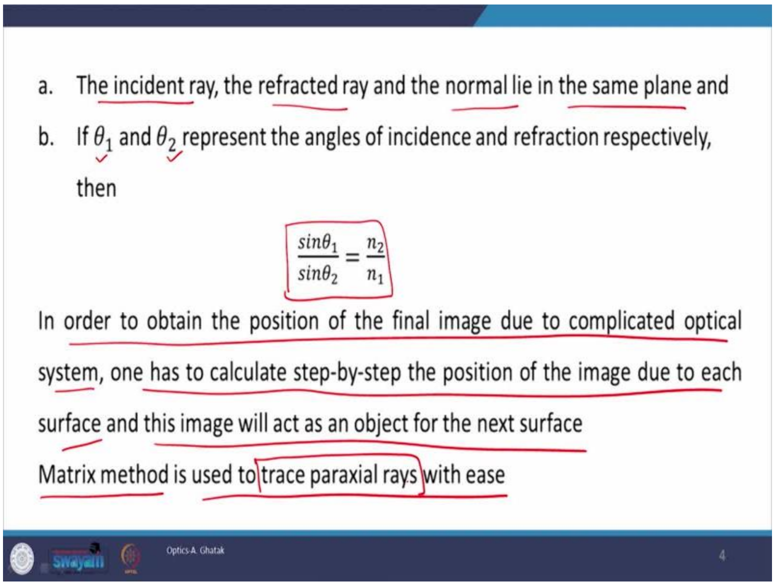
Therefore, whenever we have a series of lenses, we implement matrix method which ease out the calculations. Now, in this module, we have matrix method in paraxial optics, then, we will talk about thick and thin lenses, then we will talk about system of thin lenses, unit planes and nodal planes. But to start with we will talk about matrix method in paraxial optics.

(Refer Slide Time: 2:10)



But before that, let us consider a spherical refracting surface which is represented by SQS' this is our spherical refracting surface and the ray PQ is met to incident. It falls on this refracting surface and then it goes in a certain direction after getting refracted and this refracting surface separating two media, one is of refractive index n_1 and other is n_2, these are needless to say, we know these things. These are shown here, the direction of this refracted ray, it is completely determined by two conditions which we have imbibed very thoroughly and what are these two conditions.

(Refer Slide Time: 2:56)



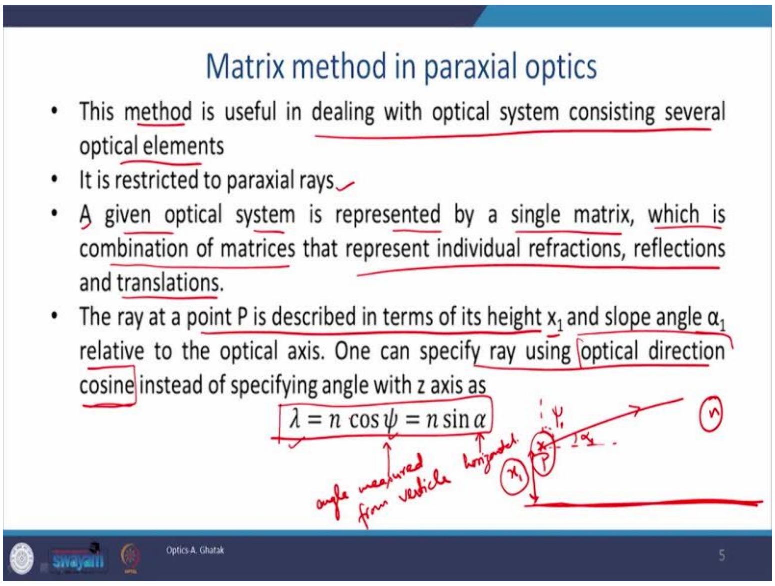
These two conditions are written here, the first is the incident ray, the refracted ray and the normal to the interface, they all three lie in the same plane and what is the second condition, the second condition is that angle of incidence and angle of refraction which areθ_1 and θ_2 respectively. They should be such that $sin\theta_1/sin\theta_2 = n_2/n_1$, this is your Snell's law. If we implement these 2 conditions, then we can easily decide the direction of the refracted ray.

Now, as I said earlier, now, in order to obtain the position of the final image due to complicated optical systems, what do I mean by complicated optical systems? Many lenses if we stack together many lenses and then shine a ray on it, then it would be very difficult to calculate the direction of the refracted ray, which finally comes out of this complicated optical system or this stack of lenses.

Because at each interface you will have to exercise these two conditions, at each interface we have to calculate what is θ_1? What is θ_2? What is n_1? What is n_2? And they may keep changing and then it is a time consuming process and one has to calculate step by step the position of the image due to each surface and then this image will act as an object for the next surface.

This is a time consuming here. And therefore, to deal with these situation, the people devised a very quick method and this is called matrix method and it is used to trace paraxial rays with ease. In paraxial optics, it is used to trace the paraxial rays it can quickly tell us the direction of the refracted ray irrespective of the complication in optical system involved. With this we will move towards the matrix method in paraxial optics.

(Refer Slide Time: 5:28)



This is a method which is useful in dealing with the optical system consisting of several optical elements like many convex lenses, many concave lenses or combination of the two. It is restricted to the paraxial rays all these geometrical optics which we are studying about we are studying it in the domain of paraxial optics here. Paraxial optics means, we are only considering the rays which are close to the lens axis and which makes very small angle with the lens axis and whenever we deviate from the paraxial optics aberration appears in the image here and aberration is entirely different topic it is not in the purview of this course.

Therefore, to avoid aberration, we will only consider rays in paraxial approximation. Next point is, a given optical system is represented by a single matrix, in this matrix method any optical system, be it a concave lens, be it a convex lens or a single lens system would be represented by a single matrix and it would be a 2 by 2 matrix, not only this lenses, but also a translation or between the two lenses if there is some gap in which ray is travelling then this gap is also represented by a matrix. For each lens one matrix for each gap one matrix.

And what would be the single matrix? This matrix would be a combination of matrices that represents the individual refraction, reflection and translation. Here what is being said is that we have a combination of lens system each lens is represented by one matrix each translation is represented by another type of matrix and then the combined matrix would be the multiplication of these matrices or let me reframe it. We are given an optical system in this optical system each reflection would be represented by one matrix, each refraction would be represented by another matrix, the translation would be represented by another matrix.

And if we have a combination of lens system then we would be have a series of refraction and translation and the final resultant matrix would be the multiplication of all these matrices. Now, when a ray travel, suppose, this is the ray which is travelling and it is inclined at angle α with the horizontal then we may also say that it is inclined at angle ψ with the vertical. Suppose this is point P at which these discussions are done. Now a ray at a point P is described in terms of its height x_1 and slope angle α_1, what does this height and slope angle means now? This ray is travelling then it must be travelling towards or from some lens system.

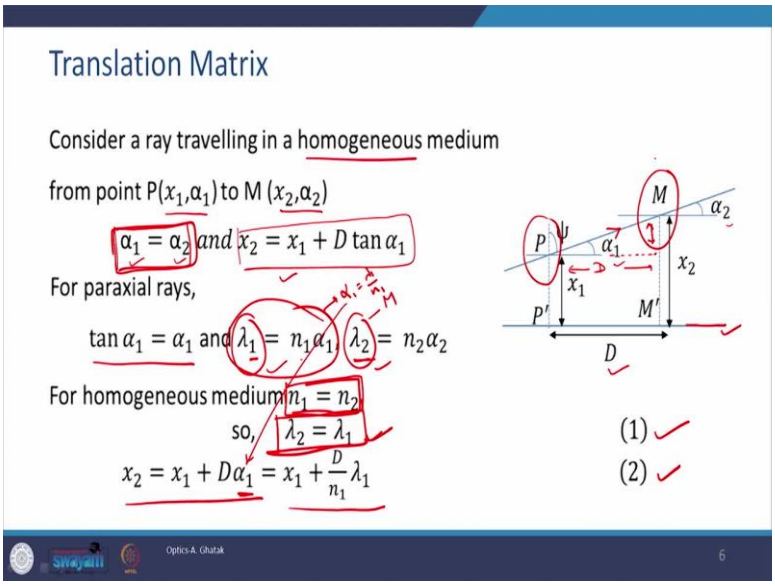
Suppose that it is going towards some lens system and this horizontal axis represents the axis of the lens and if this P point is at a height x_1 from the axis of the lens then the point P can be designated by two independent variables. I repeat there is a ray which is travelling in some medium of some refractive index say n and this horizontal line represents the axis of some lens system and we are measuring things with respect to this horizontal line. This ray is inclined at angle α with this horizontal line. Let us pick a point P on this ray.

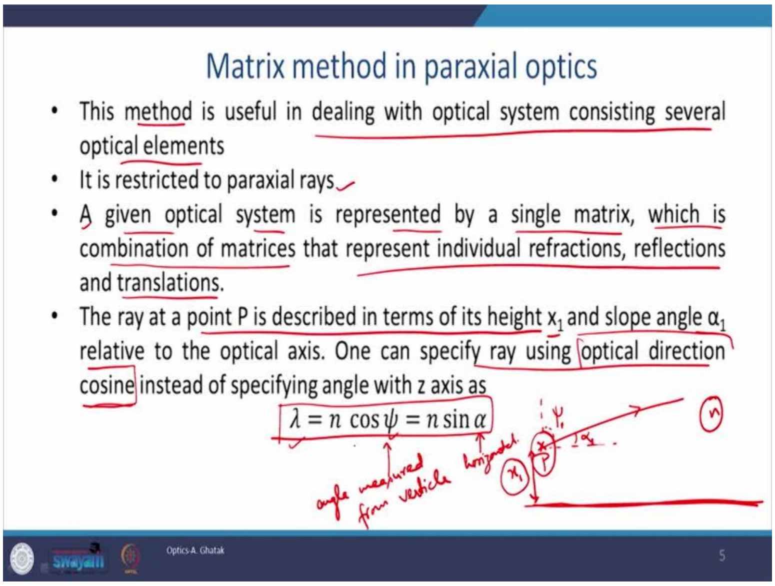
Now to designate this point P on the ray, we require two parameter. One is the inclination of the ray with the horizontal and the second is the distance of point P from the horizontal line, this distance here is x_1 and the angle suppose it is α_1, this α_1 is from the horizontal and the from vertical angle ψ_1 say. Now instead of using two parameters x_1 and α_1, one can specify a ray using optical direction cosines, the word is optical direction cosines, instead of specifying angle, we can use optical direction cosines what is optical direction cosines and how it is defined? It is defined through this relation, $\lambda = ncos\psi = nsin\alpha$.

ψ is the angle from the vertical and α is the angle from the horizontal this is angle measured from vertical and this is an angle measured from horizontal. I repeat, suppose we have a ray which is travelling in certain direction then to specify a point on this ray, suppose this point is P, we just need to know the inclination of this point P from the horizontal and the distance of point P from the axis of lens system or optical system. These are the two parameters which are enough to specify point P.

Now, instead of using angle one can also define optical direction cosine, optical direction cosine is defined by λ and it is given as $\lambda = ncos\psi.$ n is refractive index of the medium and ψ is the angle from the vertical and equivalently it can also be written as $\lambda = nsin\alpha.$ n is again refractive index of the medium and α is now the angle from the horizontal, this is λ is called direction cosine.

(Refer Slide Time: 12:43)





Now, suppose we have this axis of symmetry which is horizontal and this is the direction of the ray. Now, let us pick two points on the ray P and M. Here P and M are two points on the ray and this point P is at a distance x_1 from the horizontal and point M is at a distance x_2 from the horizontal, these are the heights of the points P and M. I repeat x_1 and x_2 are heights of point P and point M respectively from this horizontal line and these points, the horizontal separation between these point P and M is D, capital D. And it is also assumed that point P and M, the ray which is passing through point P and M they are inclined at angles α_1 and α_2 with the horizontal respectively.

Now, under these conditions, we can define the coordinate of P and M. The coordinate of P would be x_1 and α_1, x_1 is the distance from the horizontal and α_1 is the angle from the horizontal and similarly its co-ordinate would be x_2 and α_2.

The ray is supposed to be travelling in a medium which is homogeneous. Now, since the medium is homogeneous therefore, α_1 of course would be equal to α_2. And what would be the relation between x_1 and x_2. We can easily calculate, $x_2 = x_1 + D\ tan\alpha_1$.. If you take tangent of α_1 than you will get the relation between this vertical distance and capital D and this gives us the relation between x_2 and D_1.

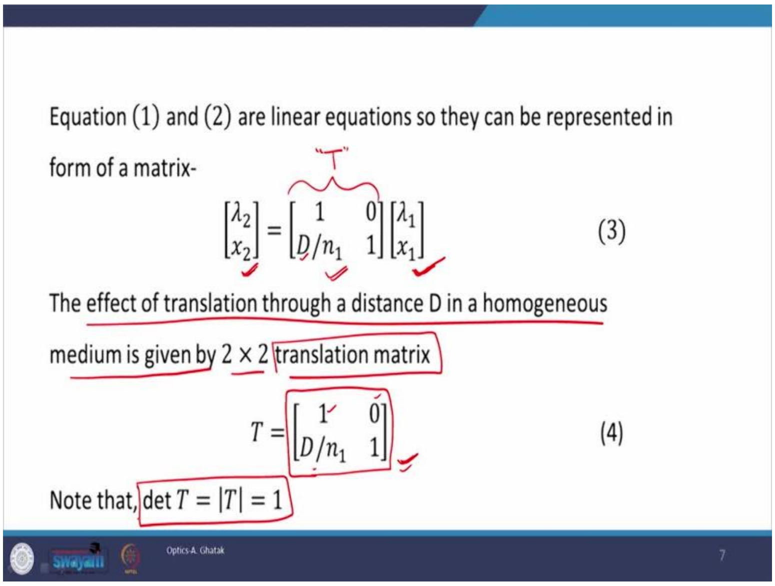
But we are in the paraxial domain. Therefore, $tan\alpha_1 = \alpha_1$ here therefore, we can safely write $tan\alpha_1 = \alpha_1$. Therefore λ_1 which is the direction cosine which we defined here in our previous slide, here we can say that λ=nα and instead of writing sinα, we will write $n\alpha$ because $sin\alpha$ would be equal to α in paraxial domain.

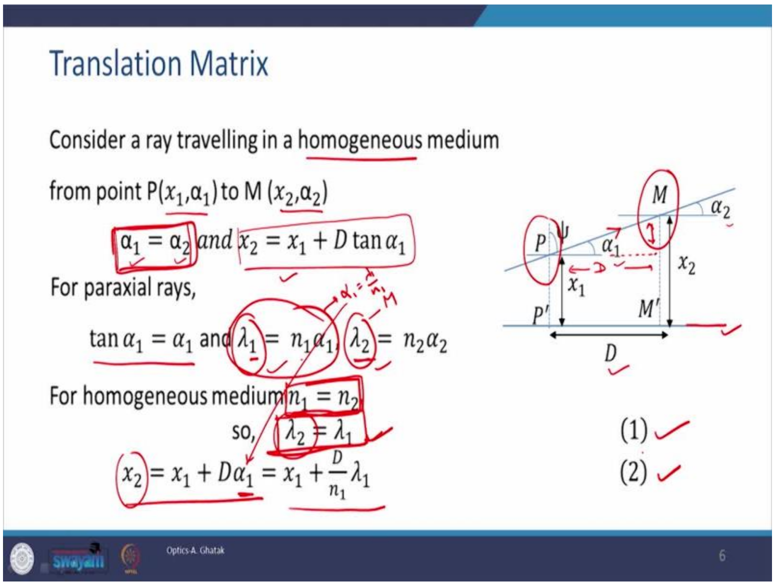
For point P, we always use a subscript 1 here, for point P subscript 1 is used therefore, for point P its direction cosine $\lambda_1 = n_1\alpha_1$. Similarly, for point M, its direction cosine will be given by λ_2 and for the expression for $\lambda_2 = n_2\alpha_2$. As I said before the medium is homogeneous therefore, the refractive index at point P and around point P would be same as the refractive index at point M. Therefore, $n_1 = n_2$and if $n_1 = n_2$ and $\alpha_1 = \alpha_2$. With these two relations, $\lambda_2 = \lambda_1$, we can easily say that $\lambda_2 = \lambda_1$.

W know now the relation between λ_2 and λ_1 and we know now the relation between x_2 and x_1. Here under paraxial approximation we replaced tanα_1 with α_1. α_1 from this relation this is equal to λ_1/n_1. You substitute this expression of α_1 from here to here then you get $x_2 = x_1 + D/n_1 \times \lambda_1$. These are the 2 relations which we got equation number 1 and equation number 2. Equation number 1 is $\lambda_2 = \lambda_1$ and equation number 2 is $x_2 = x_1 + D/n_1 \times \lambda_1$.

Now, you see that on the left hand side of equation number 1 and 2 it is λ_2 and x_2 while on the right hand side of equation number 1 and 2 it is λ_1 and x_1. Now, the ray is travelling from P to M now, once we know the coordinates of P then we can also know the coordinates of M. How? through equation 1 and 2 here. Now, we can write equation 1 and 2 in form of a matrix we can replace 1 and 2 with a single matrix.

(Refer Slide Time: 17:52)





How to do this? Here it is a linear equation though we can write on left hand side $\begin{bmatrix}\lambda_2\\x_2\end{bmatrix}$ is equal to some coefficient and then $\begin{bmatrix}\lambda_1\\x_1\end{bmatrix}$, λ_1 and x_1 is input. Once you know the input you can calculate output. To calculate the output you will have to multiply the input with some matrix. What is this matrix? We can fill this matrix once we know what our equation 1 and 2 and this is how we have filled it. Now, you know that coordinates at one point then after translating a certain distance D you can calculate the coordinates on the other point.

The effect of translation through a distance D in a homogeneous medium is therefore given by this 2 by 2 translation matrix, this is your translation matrix. Once we know the coordinates of a ray at some point in a homogeneous medium then we can calculate the coordinates of a point which is farther from this initial point and where the ray will go after certain distance. And how to calculate the coordinates of this farther point, the other point? By using this translation matrix, this translation matrix will give us, this translation matrix has the information of this translation.

And what are the elements of this transfer matrix? It is a 2 by 2 matrix, the first element is 1 second is 0, third is D/n_1, D is the distance between the two points on the ray and n_1 is the refractive index of the medium. And the very important and notable property of this translation matrix is that determinant of T is equal to 1, very important property. Then what we learn today is that, if a ray is travelling in homogeneous medium of certain refractive index, then this translation, this travel can be replaced by a 2 by 2 translation matrix. And once we know the coordinates, then just by implementing this translation matrix, we can know the effect of translation on the ray.

And using this translation matrix, we can trace the path of the ray. Today we learn about translation matrix which takes into account the translation in a homogeneous medium and therefore, if a ray is translating in a homogeneous medium, then using this translation matrix we can easily calculate or we can easily trace the ray path. The same thing can also be done for refraction through an interface which divides the two media of different refractive index. We will talk about the matrix for refraction in the next class. Thank you all.