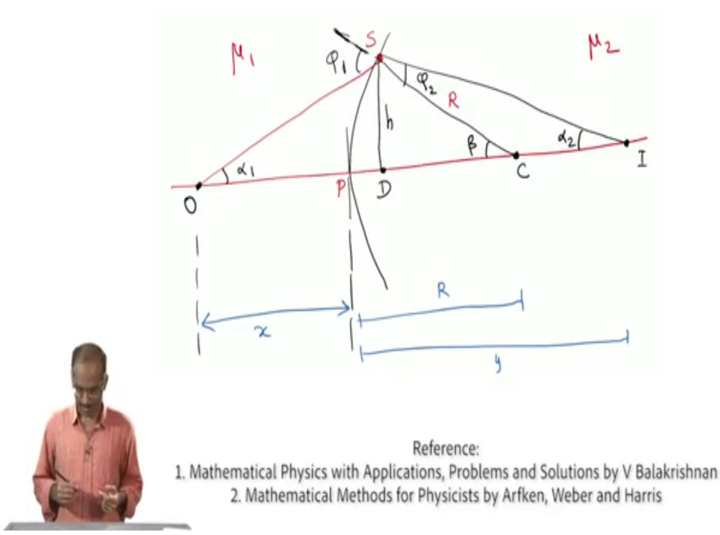
Waves and Oscillations
Prof. M S Santhanam
Department of Physics
Indian Institute of Science Education and Research, Pune

Lecture - 47
Waves in Optical Systems: Matrix Method

Let us continue our exploration of Waves in the Optical System. We are in the tenth week, and this is the fourth lecture in this topic. In todays lecture as well, we are going to stick to the limit of being able to describe waves through rays and also we will stick to paraxial approximation.

(Refer Slide Time: 00:37)

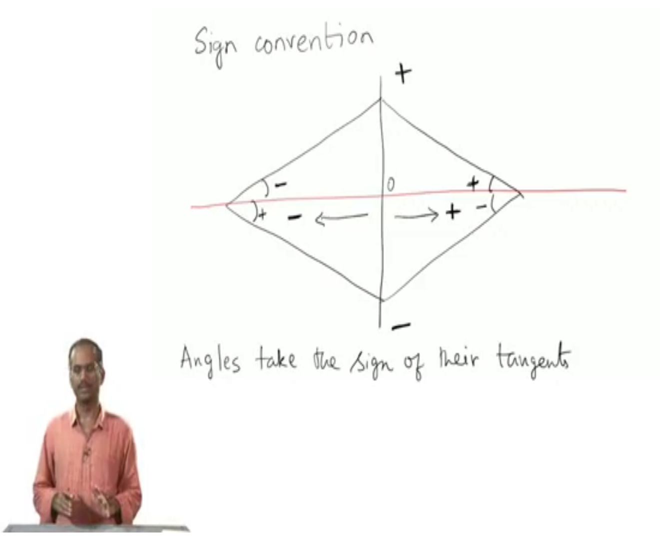


Unless, there is a reason for the ray to get refracted within a given medium, where the refractive index is a constant, the ray basically traces a straight line. So, which means that I need to worry about ray deviating from the straight line only at points where refraction takes place.

So, I can map this and rewrite this entire problem; elegantly using matrix methods, which is what we are going to do in today's lecture. So, I am going to do the following. One is very briefly introduce, how would a matrix method work; in the sense that what matrices are and so on. Again matrices is an very interesting mathematical object, in case you want to get a more deeper feel for what it is and have no idea what it is better to consult standard linear algebra textbook or chapters on matrices and standard mathematical physics books.

So, we are going to use Matrices in today's lecture.

(Refer Slide Time: 01:58)



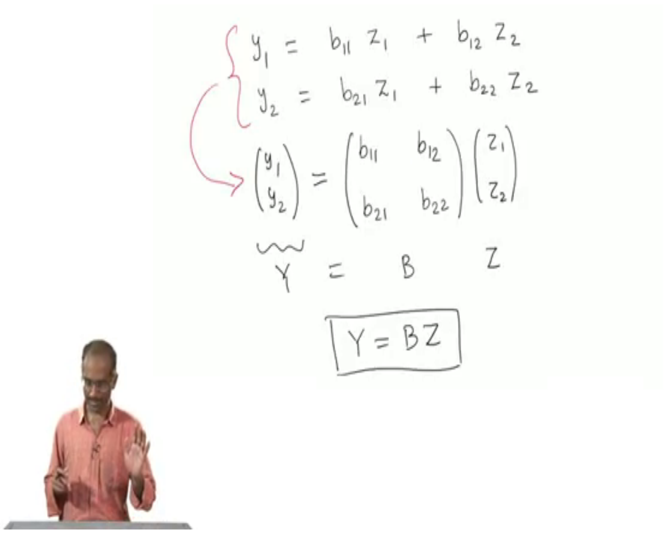
And as I go ahead, let me also remind you the sign convention once again, positive side from O is going to be plus for distances, similarly the negative sides is going to be minus for distances. And I have also indicated the convention for angles; here the angles take the sign of their tangents. With that let me step back from optics, physics, everything and just introduce the idea of matrix. So, I am going to do a very limited job here, in the sense of trying to tell you how matrix is a little more elegant way of representing a transformation.

Again matrix is very useful in many other context as well. This is not the only place where it is useful, it is really a beautiful way of looking at certain kind of things, but still, for our purposes we will look up to it as a transformation, ok. Here I have written for you two equations. So, if you look at it, it will look like two equations with two unknowns. So, if you treat y_1 and y_2 as unknowns, and a_{11}, a_{11}, a_{21} and a_{22} as some numbers, then it is precisely nothing more.

Here, I have written the same equation in a matrix form and if I give names like this, let me call it by capital $\mathbf{X}$, this vector $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and let me call this matrix capital $\mathbf{A}$, and this vector capital $\mathbf{Y}$. Then the corresponding equation written in very compact matrix notation would be; so this assumes that, we know how to multiply a matrix with a vector or a matrix with a matrix.

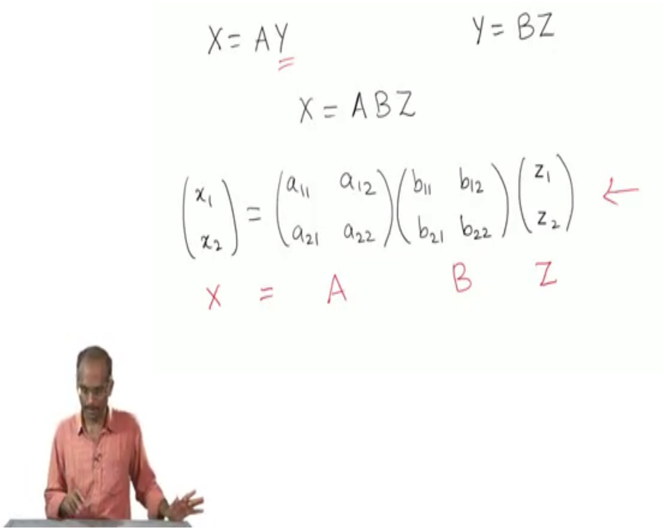
So, here what is involved on the right hand side is multiplication of a matrix with the vector. So, $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ is a column vector and I have this matrix $\mathbf{A}$ which is $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. And this product of this matrix $\mathbf{A}$ and this vector $\mathbf{Y}$ is nothing but, what is here. Let me also write y_1 and y_2 in terms of some other variable, let say z_1 and z_2.

(Refer Slide Time: 04:42)



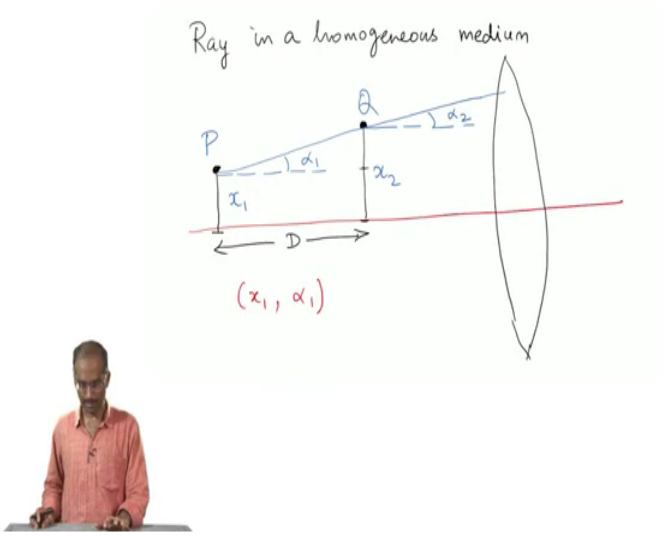
Now, I have these two sets of equation y_1 and y_2, written in terms of z_1 and z_2. But now, let me also write it as a matrix equation. So, the same set of two equations, now I have written the matrix form. Again I write it in a compact notation, $\mathbf{Y} = \mathbf{BZ}$.

(Refer Slide Time: 05:03)



Now, I can substitute for **Y** from here. So, **X** can be directly related to **Z** in the following way; **X** will be, so here when I write A into B. I should know how to multiply these two matrices.

(Refer Slide Time: 05:22)

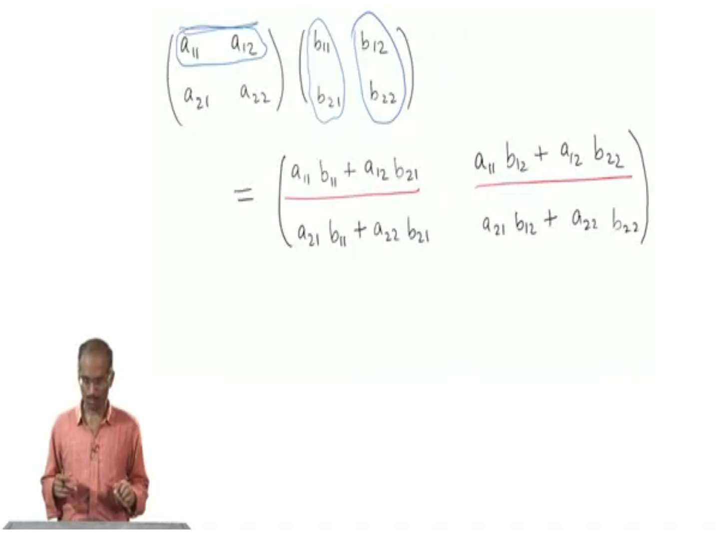


So, here I have written the product of two matrices capital A and capital B. The basis of treating rays using matrix methods is actually shown in this diagram that I have just drawn. The ray between P and Q or even a little further can be described by the quantity

(x_1, α_1). So, x_1 is the height of that point from the axis, axis of your optical system and α_1 is of course, the angle that it makes with the horizontal. At Q; you could characterize them with pair (x_2, α_2) for the same reason. It is equivalent to saying that straight line can be characterized by just two numbers; one the intercept and other is the slope.

So, every time there is a refraction then, you need new set of such pair of numbers to describe the ray. We are relating one set of (x_1, α_1) to another set of some other pair of numbers, whenever there is a refraction that is taking place.

(Refer Slide Time: 06:46)

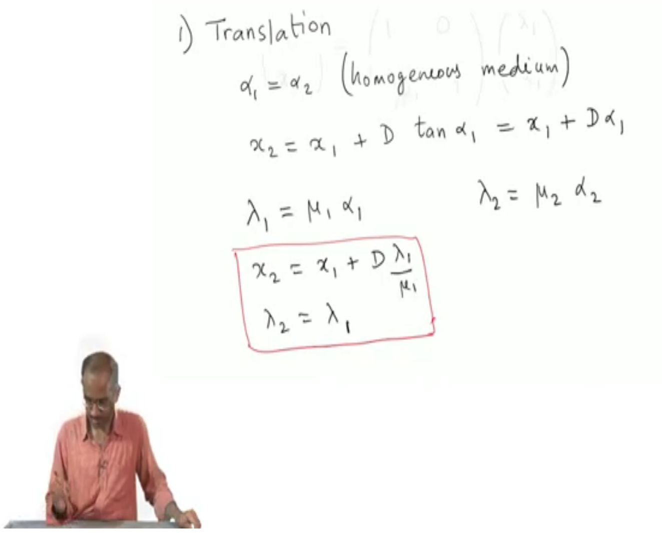


So, just as we did a little while ago for the case of matrices. So, I had one set of $\mathbf{X}$ being related to $\mathbf{Y}$; that is x_1, x_2 related to y_1, y_2. And then I had y_1, y_2 related to z_1, z_2. So, in this case, I can directly relate $\mathbf{X}$ to $\mathbf{Z}$. You could in general see such transformations in rays as only two things; one is there is a translation the ray goes from point A to point B. So, that is translation, no refraction involved because, we have assume that it is in a homogeneous medium.

But in principle, you can have refractions as well. We need to know, how to describe translation, we need to know how to describe refraction. And if all this can be done using matrix language, well and good then we could elegantly handle all these within the framework of matrices. To obtain a matrix form for translation, let me refer back to this figure and also indicate this distance by capital D.

So, let us say my ray goes from point P to point Q and the distance, or the horizontal distance that it travels is D and there is also a vertical distance that it has traversed; which is $x_2 - x_1$. And in general, we have taken these two angles to be different α_1 and α_2, but if the medium is homogeneous; we know that α_1 has to be equal to α_2.

(Refer Slide Time: 08:17)

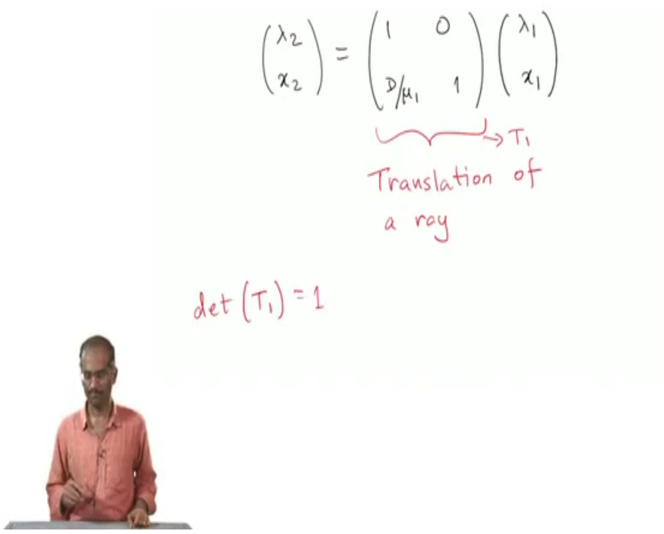


Because, he said in a homogeneous medium; $\alpha_1 = \alpha_2$. Now from the geometry of figure, I can write an expression for x_2 in terms of x_1. And if you stick to paraxial approximation, this would simply become; I introduce a variable λ_1, which will be $\mu_1 \times \lambda_1$ and λ_2 is $\mu_2 \times \lambda_2$.

Let me introduce a variable λ_1, which will be equal to $\mu_1 \times \alpha_1$ and also another variable λ_2 which will be equal to $\mu_2 \times \alpha_2$. We have already said that, we are working with homogeneous medium in which case, $\mu_1 = \mu_2$, $\alpha_1 = \alpha_2$ and hence λ_1 will be equal to λ_2. And let me also write the expression for x_2 which will be $x_1 + D\alpha_1$.

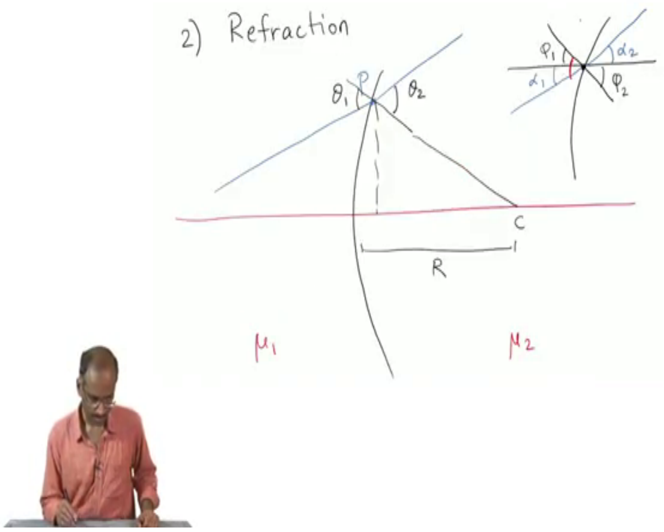
Now, if you see the left hand side of this equation; it is x_2 and λ_2, which relates to the second point Q. And on the right hand side is everything relating to the first point P here, which is x_1, α_1 and λ_1. So, essentially, I have written information about the second point Q in terms of parameters of the first point. Before I write this in matrix notation, let me also make one change here I can write α_1 as $\frac{\lambda_1}{\mu_1}$. So, let me replace this α_1 by $\frac{\lambda_1}{\mu_1}$.

(Refer Slide Time: 10:08)



So, this gives me the relation between $x_2\lambda_2$ and $x_1\lambda_1$ and it is now written in matrix form. So, you can think of this matrix that I have written down here, as a matrix that represents translation. So, it represents translation of a ray from a point that is described in terms of $x_1\lambda_1$ to another point which is described in terms of $x_2\lambda_2$. If I call this matrix T_1, then you will notice that the determinant of T_1 is equal to 1 itself.

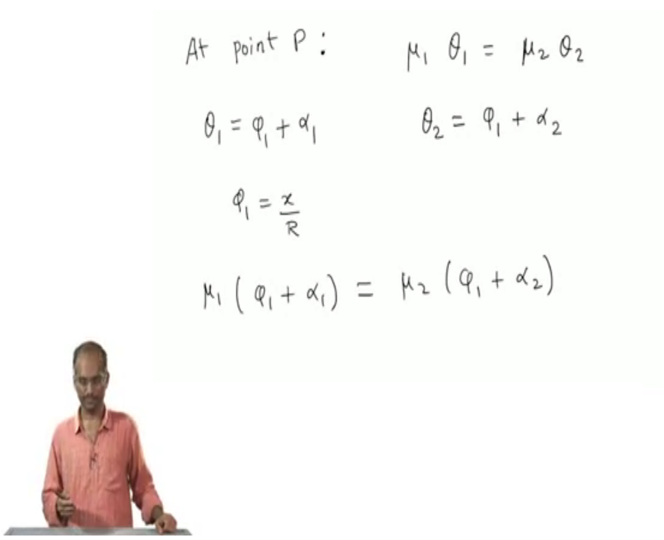
(Refer Slide Time: 10:54)



Here I have a refracting surface like I have shown in black line and red line represents the axis; let me call it by P, the point at which refraction of the ray takes place.

So, let me blow up this part and draw it here. And let say that, this is a line that is parallel to the x axis and let us say that this is the incident ray and this is the refracted ray. And now let me also draw the normal at this point. So, our aim is to relate $x_1\lambda_1$ to $x_2\lambda_2$, like we did for translation we are going to do that for refraction as well using all these angles.

(Refer Slide Time: 12:00)

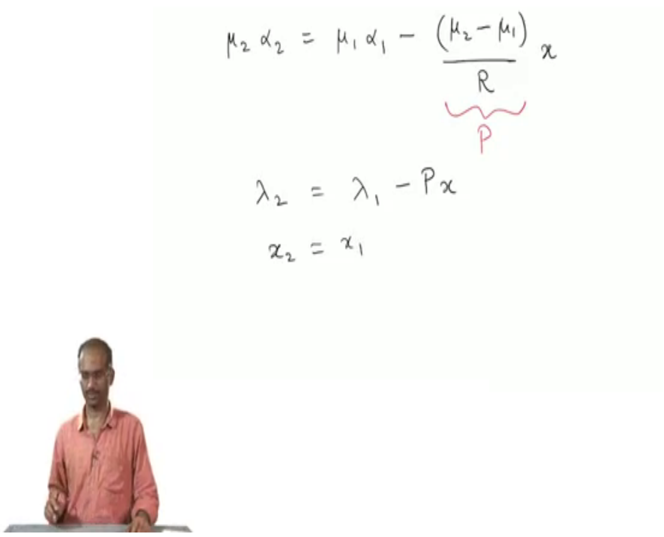


So, at point P; what we need to do is to use the law of refraction at point P, $\mu_1\theta_1 = \mu_2\theta_2$. So, you remember that, we are going from one medium to another. So, this will be medium μ_1 and on right hand side its medium μ_2. So, what I have done is to assume paraxial approximation and write the law of refraction; $\mu_1\theta_1 = \mu_2\theta_2$. θ_1 is simply this entire angle here, which I am now showing in red and that is equal to $\alpha_1 + \phi_1$. So, that is my first of the equations; θ_1 is $\phi_1 + \alpha_2$, it is easy to figure out that, θ_2 is simply $\phi_1 + \alpha_2$.

So, I have these two equations for θ_1 and θ_2. One another quantity that I will need is the radius of curvature. So, let me extend this back here, this will go to the center of a circle of which this curved surface is apart, hence this is going to represent for me the radius of curvature R.

And let me also draw perpendicular here to this point and call this as x, if you take the paraxial approximation. The distance between this point here and this base point here would be so small that you can take the distance from this first point to see as R itself. In that case ϕ_1 will be equal to $\frac{x}{R}$. So, we replace ϕ_1 in terms of $\frac{x}{R}$. If I do that I will get the following relation.

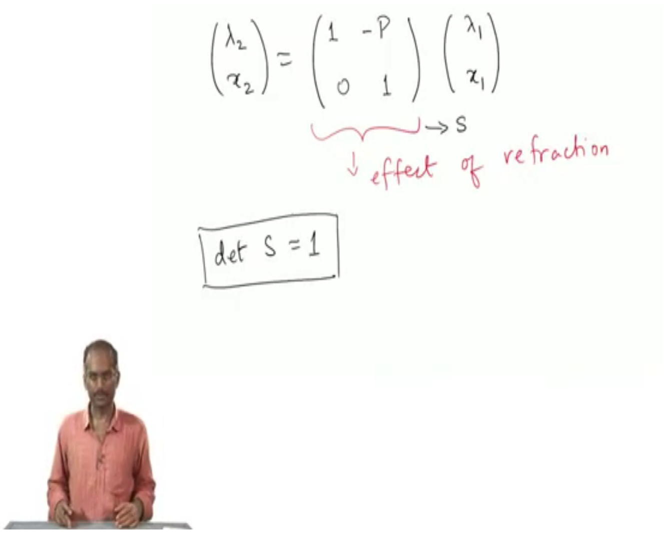
(Refer Slide Time: 14:04)



Now, you can recognize that, this quantity here is simply what we can call the power of the refracting surface in which case I could write this as. So, if you remember the relation that we got for translation, you will notice that we related λ_2, x_2 to λ_1, x_1. How do you get that distance from the axis in this case?

So, go back to this figure that we drew for refraction. You will notice that, there is only one of them, in the sense that, there are no two different values; there is no x_1 and x_2. Instead what we have is at the point of refraction there is only one height or one distance from the red line, the axis. So, which means that in this case x_2 and x_1 are exactly the same so, I am going to have a x_2 being equal to x_1 itself.

(Refer Slide Time: 14:56)



So, now I have this form for matrix form for refraction. So, it tells me like, if I start a ray which is described by x_1, λ_1 after refraction what happens to it. So, it is described by x_2, λ_2. And how are they related? They are related by this equation; specifically the important part here is this matrix, which gives you the effect of refraction. Now, you can trace rays across several lenses, refracting surfaces of various kind, but you know that whatever be the complexity of your optical system; it is either translation or a refraction.

So, now we have isolated the effect of translation; isolated the effect of refraction. So, all I need to do is to lookup in what sequence they occur Translation and refraction and just multiply the appropriate matrices and I can very easily get the final effect, if I call this matrix S, then determinant of S is 1 in this case as well. So, what is left; now is to do some examples and we shall proceed with examples in the next module.