Snippets of Physics

12. Paraxial Optics and Lenses

T Padmanabhan

T Padmanabhan works at IUCAA, Pune and is interested in all areas of theoretical physics, especially those which have something to do with gravity.

Discovering unexpected connections between completely different phenomena is always a delight in physics. In this and the next installment, we will look at one such connection between two unlikely phenomena: propagation of light and path integral approach to quantum mechanics!

The propagation of light – which is just an electromagnetic wave – is governed by a wave equation. The electric field and the magnetic field obey a wave equation, the solution to which describes the propagation of light in any specific context. In this installment we look at the wave nature of light from a particular point of view which we will connect up with a seemingly different phenomenon in the next installment.

For our purpose the vector nature of the electromagnetic field is not relevant (since we will not be interested, e.g., in the polarization of the light.) Hence we will just deal with one component of the relevant vector field – let us call it $A(t, \mathbf{x})$ – which satisfies the wave equation. The basic solution to the wave equation $\Box A = 0$ is described by the (real and imaginary parts of the) function $\exp i[\mathbf{k} \cdot \mathbf{x} - \omega t]$. Here $\mathbf{k}$ denotes the direction of propagation of the wave which also determines its frequency through the dispersion relation $\omega = |\mathbf{k}|c$. Since the wave equation is linear in A, we can superpose the solutions with different values of $\mathbf{k}$, each with an amplitude $F_1(\mathbf{k})$, say. This leads to a solution of the form:

$$A(t, \mathbf{x}) = \int F_1(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} e^{-i\omega t} \frac{\mathrm{d}^3 k}{(2\pi)^3}. \qquad (1)$$

Keywords
Optics, waves.

We now want to specialize to a situation which arises in the study of optical phenomena quite often where we are concerned with waves which are propagating, by and large, in some given direction, say along the positive z-axis. (For example, consider the study of diffraction by a circular hole in a screen which is located in the $z = 0$ plane. We will consider, in such a context, light incident on the screen from the left and getting diffracted.) Mathematically, this means that the function $F_1(\mathbf{k})$ is nonzero only for wave vectors with $k_z > 0$. Further, since the wave has a definite frequency ω, the magnitude of the wave vector is fixed at the value ω/c. It follows that one of the components of the wave vector, say k_z, can be expressed in terms of the other three. So, the function F_1 has the structure

$$F_1(k_z, \mathbf{k}_\perp) = 2\pi f(\mathbf{k}_\perp)\delta_D\left(k_z - \sqrt{\omega^2/c^2 - \mathbf{k}_\perp^2}\right), \quad (2)$$

where the subscript $\perp$ denotes the components of the vector in the transverse $x - y$ plane. Note that, in general, we could have had $k_z = \pm\sqrt{\omega^2/c^2 - \mathbf{k}_\perp^2}$ and we have consciously picked out one with $k_z > 0$.

Substituting this expression in (1) we find that $A(t; z, \mathbf{x}_\perp)$ can be written in the form $a(z, \mathbf{x}_\perp)e^{-i\omega t}$ (in which we have separated out the oscillations in time) where

$$a(z, \mathbf{x}_\perp)$$

$$= \int \frac{\mathrm{d}^2\mathbf{k}_\perp}{(2\pi)^2} f(\mathbf{k}_\perp) e^{i\mathbf{k}_\perp \cdot \mathbf{x}_\perp} \exp\left[\frac{iz}{c}\sqrt{\omega^2 - c^2 k_\perp^2}\right]. \quad (3)$$

Since the time variation of a monochromatic wave is always $\exp(-i\omega t)$, we shall ignore this factor and concentrate on the spatial dependence of the amplitude, $a(z, \mathbf{x}_\perp)$.

To proceed further we shall consider the case in which all the components building up the wave are travelling essentially along the positive z-axis with a small transverse spread. For such a wave travelling, by and large, along the z direction, the transverse components of $\mathbf{k}$ are small compared to its magnitude; that is, $c^2k_\perp^2 \ll \omega^2$. Using the Taylor series

$$\sqrt{\omega^2 - c^2k_\perp^2} \cong \omega\left(1 - \frac{1}{2}\frac{c^2k_\perp^2}{\omega^2}\right) = \omega - \frac{1}{2}\frac{c^2k_\perp^2}{\omega}, \qquad (4)$$

in (3), we find that

$$a(z, \mathbf{x}_\perp) \equiv e^{i\omega z/c}\int \frac{\mathrm{d}^2\mathbf{k}_\perp}{(2\pi)^2} f(\mathbf{k}_\perp)\exp\left[i\left(\mathbf{k}_\perp\cdot\mathbf{x}_\perp - (c/2\omega)k_\perp^2 z\right)\right]. \qquad (5)$$

This equation describes the propagation of a wave along the positive z-axis with a small spread in the transverse direction. The function $f(\mathbf{k}_\perp)$ can be determined by a simple Fourier transform if the amplitude $a(z', \mathbf{x}'_\perp)$ is given at some location z'. Doing this, we can relate the amplitudes of the wave at two planes with coordinates z and z' by

$$a(z, \mathbf{x}_\perp) = e^{i\omega(z-z')/c}\int \mathrm{d}^2\mathbf{x}'_\perp a(z', \mathbf{x}'_\perp)\; G\left(z - z'; \mathbf{x}_\perp - \mathbf{x}'_\perp\right), \qquad (6)$$

where

$$\begin{aligned} G\left(z - z'; \mathbf{x}_\perp - \mathbf{x}'_\perp\right) &= \int \frac{\mathrm{d}^2\mathbf{k}_\perp}{(2\pi)^2} e^{i\mathbf{k}_\perp\cdot\left(\mathbf{x}_\perp - \mathbf{x}'_\perp\right)} e^{-(ic/2\omega)k_\perp^2(z-z')} \\ &= \left(\frac{\omega}{2\pi i c}\right)\frac{1}{|z - z'|}\exp\left[\frac{i\omega}{2c}\frac{\left(\mathbf{x}_\perp - \mathbf{x}'_\perp\right)^2}{(z - z')}\right]. \qquad (7) \end{aligned}$$

The function G may be thought of as a propagator which propagates the amplitude from the location $(z', \mathbf{x}'_\perp)$ to the location $(z, \mathbf{x}_\perp)$. The factor $e^{i\omega(z-z')/c}$ in (6) does not contribute to the intensity and we will drop it when not necessary.

A little thought shows that we have achieved something quite interesting. We know that the amplitude satisfies a second order differential equation (viz. the wave equation) and hence its evolution cannot be determined by just knowing the amplitude (ie., one single function, $a(z', \mathbf{x}'_\perp)$) at a given location $(z', \mathbf{x}'_\perp)$. This could be done in (6) only because of the assumption that the wave is travelling essentially forward in the z direction. The actual form of the propagator depends on the assumption that the transverse components of the wave vector are small compared to k_z. The study of wave propagation under these approximations is called *paraxial optics*. (We shall see in the next installment that all these expressions have interesting connections with the path integral propagator in quantum mechanics – which will emerge as the paraxial optics of relativistic field theory!)

The actual form of the propagator depends on the assumption that the transverse components of the wave vector are small compared to k_z. The study of wave propagation under these approximations is called *paraxial optics*.

Let us take a closer look at the structure of the propagator G. It introduces a factor $|z - z'|^{-1}$ to the amplitude and, more importantly, contributes an amount

$$\phi = \frac{\omega}{2c} \frac{(\mathbf{x}_\perp - \mathbf{x}'_\perp)^2}{(z - z')} \tag{8}$$

to the phase. The change in the amplitude merely reflects r^{-2} fall-off of the intensity (which is proportional to the square of the amplitude) of the wave. But what does the phase factor mean? To understand the origin of the change in phase, note that a path difference Δs between two points in space will introduce a phase difference of $k\Delta s$ in a propagating wave. In our case, it is clear that the phase difference is

$$k\Delta s = \frac{\omega}{c}\left[\sqrt{\left(\mathbf{x}_\perp - \mathbf{x}'_\perp\right)^2 + \left(z-z'\right)^2} - \left(z-z'\right)\right]$$
$$\cong \frac{\omega}{c}\left[\frac{1}{2}\frac{\left(\mathbf{x}_\perp - \mathbf{x}'_\perp\right)^2}{\left(z-z'\right)}\right], \qquad (9)$$

provided the transverse displacements are small compared to the longitudinal distance – an assumption which is central to paraxial optics. With hindsight we could have guessed the form of G without doing any algebra! In paraxial optics, it introduces a phase corresponding to the path difference and decreases the amplitude to take into account the normal spread of the wave.

Equation (6) allows one to compute the wave amplitude at any location on the plane $z = z_2$, if the amplitude on a plane $z = z_1 < z_2$ is given. To see it in action, let us apply it to a standard situation, which arises quite often in optics. A wave front propagates freely up to a plane $z = z_1$ where it passes through an optical system (say a lens, screen with a hole, atmosphere, etc.) which modifies the wave in a particular fashion. The optical system extends from $z = z_1$ to $z = z_2$ and the wave propagates freely for $z > z_2$. We will be interested in the amplitude at $z > z_2$, given the amplitude at $z < z_1$. It is clear that our equation (6) can be used to propagate the amplitude from some initial plane $z = z_\mathrm{O} < z_1$ to $z = z_1$ and from $z = z_2$ to some final plane $z = z_\mathrm{I} > z_2$. (The subscripts O and I stand for object and image, based on the idea of the optical system being a lens). The propagation of wave from z_1 to z_2 depends entirely on the optical system and, in fact, defines the particular optical system. An optical system is called *linear* if the output is linear in input. In such a case, the amplitudes at the exit point of the optical system is related to the amplitude at the entrance point by a relation of the kind

$$a\left(z_2, \mathbf{x}_2\right) = \int \mathrm{d}^2\mathbf{x}_1 P\left(z_2, z_1; \mathbf{x}_2, \mathbf{x}_1\right) a\left(z_1, \mathbf{x}_1\right), \qquad (10)$$

where the functional form of P defines the kind of optical system. (Here and in what follows, we shall omit the subscript $\perp$ with the understanding that the vector $\mathbf{x}$ is in the transverse plane and is two dimensional.) In this case, the amplitude at the image plane can be expressed in terms of the amplitude at the object plane by the relation

$$a\left(z_{\mathrm{I}}, \mathbf{x}_{\mathrm{I}}\right)=\int \mathrm{d}^{2} \mathbf{x}_{\mathrm{O}} \mathcal{G}\left(z_{\mathrm{I}}, z_{\mathrm{O}} ; \mathbf{x}_{\mathrm{I}}, \mathbf{x}_{\mathrm{O}}\right) a\left(z_{\mathrm{O}}, \mathbf{x}_{\mathrm{O}}\right), \quad (11)$$

where

$$\begin{aligned} &\mathcal{G}\left(z_{\mathrm{I}}, z_{\mathrm{O}} ; \mathbf{x}_{\mathrm{I}}, \mathbf{x}_{\mathrm{O}}\right) \\ &=\int \mathrm{d}^{2} \mathbf{x}_{2} \mathrm{d}^{2} \mathbf{x}_{1}\ G\left(z_{\mathrm{I}}-z_{2}, \mathbf{x}_{\mathrm{I}}-\mathbf{x}_{2}\right)\ P\left(z_{2}, z_{1} ; \mathbf{x}_{2}, \mathbf{x}_{1}\right) \\ &\qquad\qquad \times\ G\left(z_{1}-z_{\mathrm{O}}, \mathbf{x}_{1}-\mathbf{x}_{\mathrm{O}}\right). \end{aligned} \quad (12)$$

Given the properties of any linear optical system, one can compute the quantity P, and thus evaluate $\mathcal{G}$ and determine the properties of wave propagation.

As a simple example let us find out the form of the function P for a convex lens. If the lens is sufficiently thin, P will be nonzero only at the plane of the lens $z_2 = z_1 = z_{\mathrm{L}}$. Since the lens does not absorb radiation, it cannot change the amplitude $|a(z_{\mathrm{L}}, \mathbf{x}_{\mathrm{L}})|$ of the incident wave and can only modify the phase. Therefore, P must have the form $P = \exp[i\theta(\mathbf{x}_{\mathrm{L}})]$. Then the amplitude at the image plane is given by

$$\begin{aligned} &a\left(z_{\mathrm{I}}, \mathbf{x}_{\mathrm{I}}\right) \\ &=\int \mathrm{d}^{2} \mathbf{x}_{\mathrm{L}} a\left(z_{\mathrm{L}}, \mathbf{x}_{\mathrm{L}}\right) P\left(z_{\mathrm{L}}, \mathbf{x}_{\mathrm{L}}\right) G\left(z_{\mathrm{I}}-z_{\mathrm{L}}, \mathbf{x}_{\mathrm{I}}-\mathbf{x}_{\mathrm{L}}\right) \\ &=a \int \mathrm{d}^{2} \mathbf{x}_{\mathrm{L}} e^{i \theta\left(z_{\mathrm{L}}, \mathbf{x}_{\mathrm{L}}\right)} G\left(z_{\mathrm{I}}-z_{\mathrm{L}}, \mathbf{x}_{\mathrm{I}}-\mathbf{x}_{\mathrm{L}}\right), \end{aligned} \quad (13)$$

where we have used the fact that the amplitude $a(z_{\mathrm{L}}, \mathbf{x}_{\mathrm{L}})$ on the lens plane is constant for a plane wave incident from a large distance. To determine the form of $\theta(\mathbf{x}_{\mathrm{L}})$, we use the basic defining property of a lens of focal length f: If a plane wavefront of constant intensity is incident on the lens plane $z = z_{\rm L}$, the rays will be focused at a point $z_{\rm I} = z_{\rm L} + f$, when the wave nature of the light is ignored. In the limit of zero wavelength for the wave, most of the contribution to the integral in (13) comes from points at which the phase of the integrand is stationary. Since the phase of G is $(k/2)[(\Delta\mathbf{x})^2/\Delta z]$, the principle of stationary phase gives the equation

$$\frac{\partial\theta}{\partial\mathbf{x}_{\rm L}} = \frac{k}{f}\left(\mathbf{x}_{\rm I} - \mathbf{x}_{\rm L}\right), \qquad (14)$$

where $f = z_{\rm I} - z_{\rm L}$. For the image to be formed along the z-axis, this equation should be satisfied for $\mathbf{x}_{\rm I} = 0$. Setting $\mathbf{x}_{\rm I} = 0$ and integrating this equation we find that $\theta = (-kx_{\rm L}^2/2f)$ and

$$P\left(\mathbf{x}_{\rm L}\right) = \exp\left(-\frac{ik}{2f}x_{\rm L}^2\right). \qquad (15)$$

Thus the effect of a lens is to introduce a phase variation which is quadratic in the transverse coordinates. Such a lens will focus the light to a point on the z-axis, in the limit of zero wavelength.

A geometrical interpretation of this result is given in *Figure* 1. The constant phase surfaces are planes to the left of the lens and they are arcs of circles (centered on the focus F) on the right side of the lens. Changing the plane to a circle (of radius f) at $z = z_{\rm L}$ introduces a path difference of $\Delta l = [f - (f^2 - x_{\rm L}^2)^{1/2}] \simeq (x_{\rm L}^2/2f)$ at a transverse distance $x_{\rm L}$. This corresponds to a phase difference $k\Delta l = (kx_{\rm L}^2/2f) = \theta$ introduced by the lens.

The effect of a lens is to introduce a phase variation which is quadratic in the transverse coordinates. Such a lens will focus the light to a point on the *z*-axis, in the limit of zero wavelength.

Let us next consider the effect of this lens on a point source of radiation along the z-axis at $z = z_{\rm O}$. [That is, we take the initial amplitude to be $a(z_{\rm O}, \mathbf{x}_{\rm O}) \propto \delta_D(\mathbf{x}_{\rm O})$.] This can be obtained by first propagating the field from $z_{\rm O}$ to $z_{\rm L}$, modifying the phase due to the lens at $z =$

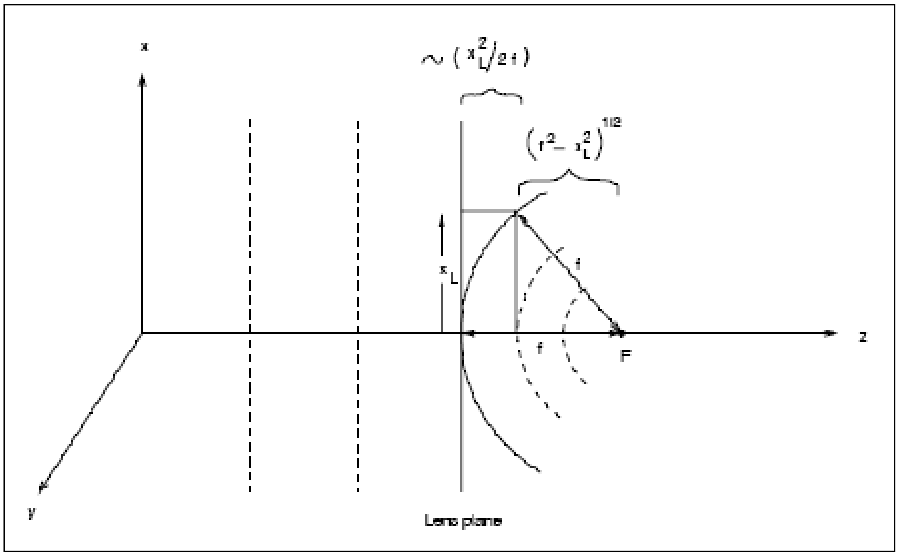

Figure 1. The focusing action of a convex lens in terms of the phase change of wave fronts.

z_{L} and propagating it further to some point z with the transverse coordinate set to zero. The net result is given by

$$a(z,0) = -\frac{k^2}{4\pi^2 uv}\int \mathrm{d}^2\mathbf{x}_{\mathrm{L}} \exp\left(-\frac{ik}{2f}x_{\mathrm{L}}^2\right)\cdot \exp\left[\frac{ikx_{\mathrm{L}}^2}{2u}+\frac{ikx_{\mathrm{L}}^2}{2v}\right], \tag{16}$$

where $u = z_{\mathrm{L}} - z_{\mathrm{O}}$ and $v = z - z_{\mathrm{L}}$. In the limit of zero wavelength (called *ray optics*), the maximum contribution to this integral can again be obtained by setting the variation of the phase to zero. This gives

$$-\frac{k}{f}\mathbf{x}_{\mathrm{L}} + \frac{k}{u}\mathbf{x}_{\mathrm{L}} + \frac{k}{v}\mathbf{x}_{\mathrm{L}} = 0, \tag{17}$$

or

$$\frac{1}{u}+\frac{1}{v}=\frac{1}{f}, \tag{18}$$

which is a familiar formula in the theory of lenses.

The above result was obtained in the limit of ray optics. To study the wave propagation through the lens we note that the action of a lens on the phase of an initial intensity distribution is governed by the integral

$$a(z,\mathbf{x}) \propto \int \mathrm{d}^2\mathbf{x}_\mathrm{L} a\left(z_\mathrm{L},\mathbf{x}_\mathrm{L}\right) \exp\left(-\frac{ik}{2f}x_\mathrm{L}^2\right) \times \exp\left[\frac{ik}{2(z-z_\mathrm{L})}\left(\mathbf{x}-\mathbf{x}_\mathrm{L}\right)^2\right], \quad (19)$$

where $a(z_\mathrm{L},\mathbf{x}_\mathrm{L})$ is the incident amplitude on the lens, the first exponential gives the distortion in phase produced by the lens and the second exponential gives the propagation amplitude z_L to z. On the focal plane, which is a plane located at a distance f from the lens, at $z = z_\mathrm{L}+f$, the second exponential characterizing the propagation becomes

$$\exp\frac{ik\left(\mathbf{x}-\mathbf{x}_\mathrm{L}\right)^2}{2(z-z_\mathrm{L})} = \exp\frac{ik}{2f}\left(x^2+x_\mathrm{L}^2-2\mathbf{x}\cdot\mathbf{x}_\mathrm{L}\right). \quad (20)$$

The quadratic term $(ikx_\mathrm{L}^2/2f)$ in the propagation amplitude is now precisely cancelled by the phase distortion introduced by the lens, so that the resultant amplitude can be written as

$$a(z_\mathrm{L}+f,\mathbf{x}) \propto \exp\left(\frac{ik}{2f}x^2\right)\int \mathrm{d}^2\mathbf{x}_\mathrm{L} a\left(z_\mathrm{L},\mathbf{x}_\mathrm{L}\right) \times \exp\left(\frac{ik}{f}\mathbf{x}\cdot\mathbf{x}_\mathrm{L}\right). \quad (21)$$

The intensity at the focal plane is given by $|a(z_\mathrm{L}+f,\mathbf{x})|^2$ in which the phase factor $\exp[ikx^2/2f]$ does not contribute. We then get the rather cute result that the lens essentially produces – on the focal plane – the two-dimensional Fourier transform of the incident amplitude!

Address for Correspondence
T Padmanabhan
IUCAA, Post Bag 4
Pune University Campus
Ganeshkhind
Pune 411 007, India.
Email:
paddy@iucaa.ernet.in
nabhan@iucaa.ernet.in

Suggested Reading

[1] **Several textbooks in optics and electromagnetic theory describe these aspects; see, for example, T Padmanabhan, *Theoretical Astrophysics – Vol.I* (*Astrophysical Processes*), Chapter 3, Cambridge University Press, 2000.**