HUYGHENS' PRINCIPLE AND THE DIFFRACTION OF LIGHT

Part I. Theoretical Considerations

By Sir C. V. Raman

(*Memoir No.* 118 *from the Raman Research Institute, Bangalore*-6)

Received July 23, 1959

1. Introduction

The first three chapters of Huyghens' Treatise on Light published in the year 1690 carry the titles "On Rays Propagated in Straight Lines"; "On Reflection"; "On Refraction". They seek to explain the familiar facts of geometrical optics on the basis of the idea that light consists of waves propagated outwards from the original source in the luminiferous medium. Huyghens based himself on certain hypotheses regarding the nature of the medium and on the nature of the light waves themselves. The adequacy of his explanations of the rectilinear propagation of light and of the geometric laws of reflection and refraction of light has necessarily to be viewed against the background of those hypotheses. The treatise of Huyghens exerted a great influence on the development of the wave-theory of light in later years. Surprisingly enough, however, his ideas were not correctly understood by those who followed him. They ascribed to Huyghens various statements not to be found in his treatise and then proceeded to criticise his ideas in an unjustifiable manner. In a recent memoir by the present author published in these *Proceedings*, a precis was given of the first three chapters of Huyghens' treatise and it was shown that, if we accept the hypotheses of Huyghens regarding the nature of the luminiferous medium and the nature of light, the rest of his argument forms a coherent and complete explanation of the facts which he sought to elucidate.

Treatises on physical optics devote a good deal of space to an exposition of what they refer to as "Huyghens' Principle" and the more advanced treatises proceed to sketch a mathematical development which is described as "The Rigorous Formulation of Huyghens' Principle" and associate it with the name of Kirchhoff. The readers of those treatises are invited to believe that Huyghens' Principle as thus formulated is an adequate basis for the explanation of the phenomena of the diffraction of light. However, in the author's memoir mentioned above, it was indicated that the explanations referred to are vitiated by the misunderstanding of the ideas of Huyghens on which they claim to be based. It is the object of the present memoir to go more fully into these matters and to show that the concepts of Huyghens can indeed be used as a framework for an elementary theory of diffraction, but that they lead to results which are wholly different from those indicated by the well-known formulæ of Kirchhoff. Extensive experimental studies of the phenomena of the diffraction of light have also been carried out by the author and the results of the same are reported in the second part of this memoir. They vindicate the present theoretical approach and contradict the consequences of Kirchhoff's theory, showing thereby that the latter is unsustainable.

2. THE WAVE-OPTICS OF HUYGHENS

When we examine the first three chapters of Huyghens' treatise, it becomes clear that a certain basic idea regarding the propagation of light underlies them all. What Huyghens sought to show in his first chapter is that in a homogeneous medium, each little piece of the primary wave emerging from a source of light is capable of travelling in a direction normal to itself more or less independently and that the primary wave-front is the locus or surface at which all the little pieces of which it is made up arrive at the same instant. The explanation of the laws of reflection and of refraction given in the second and third chapters proceeds on the same basis. Each piece of the original wave-front from the light-source on reaching the boundary between two media of which the optical properties are different finds itself unable to continue on its original course by reason of that fact. Such a situation arises, for example, if both media admit of the propagation of light through them but with different velocities. In these circumstances, the individual pieces of the wave take fresh paths which are different in the two media; the directions of travel in each case are such that the pieces of the original wave-front which are diverted from their path on reaching the boundary join up together again and form new wave-fronts in each medium. A simple geometric construction enabled Huyghens to ascertain the directions of travel which enable these conditions to be satisfied in each medium, from which again the geometric laws of reflection and refraction follow immediately. Huyghens' derivation of these laws is both simple and convincing. Regarded as a physical theory, it is also highly successful since it demonstrates that the refractive indices of the two media are in the inverse ratio of the velocities of light in them.

The concept of " partial waves " introduced by Huyghens plays a most important role in his arguments and indeed forms the hard core of his theory. It is not introduced *ad hoc* but is put forward with a definite physical justification wherever it is brought in and made use of. In the first chapter of the treatise in which the rectilinear propagation of light is sought to be explained, the partial waves considered are those which arise by reason of the structure which the luminiferous medium is assumed to possess. Each individual particle of the medium when disturbed by the passage of the primary wave becomes a source from which partial waves spread out in all directions, but these partial waves by reason of their excessive feebleness can produce a sensible effect only when a great number of them arrive simultaneously at a given point of observation: this again is only possible when the sources of the partial waves all lie on the straight line joining the original source and the point of observation. Accordingly, it is justifiable to regard the primary wave as having travelled out along that line and identify it with the summation of the partial waves of which the effects are superposed at the point of observation.

The partial waves considered in the second and third chapters of the treatise have a different origin. Here, Huyghens found himself compelled to introduce the idea that the elementary *areas* of the surface of separation between two media having different optical properties become sources of partial waves: the waves which go back into the first medium build up the reflected wave-front, while those which go forward into the second medium build the refracted wave-front.

3. THE NATURE OF DIFFRACTION PHENOMENA

When we speak of the diffraction of light, we have in mind certain effects which are observed when the free propagation of light is modified or influenced by the presence of obstacles in its path. It is clear that the nature of the obstacles, including their optical properties and their configuration in space, would determine these effects and it follows that the factors referred to would play the leading role in any theory of diffraction. If, bearing this in mind, we seek to discover in the ideas of Huyghens a possible approach towards an understanding of the phenomena of diffraction, it becomes apparent that no such approach can be found in his explanation of the rectilinear propagation of light. On the other hand, his theory of reflection and refraction does offer itself as a basis. For, as already stated, it makes use of the idea that each element of area of the boundary between two media on which light is incident is a source of partial or secondary waves in the two media. Conceptually, these waves can diverge from each element in various directions, but the requirement imposed by the theory of Huyghens that the disturbances originating at the different elements of area should arrive simultaneously at a common wave-front fixes the actual direction of their movement. If, instead of considering light waves as simple pulses, we take account of their periodic nature and of the possibility of interferences between the secondary or partial waves having their origin at the different elements of area on the boundary between the two media, the restriction of the observable effect to precisely defined directions ceases to exist. In other words, the diffraction of light becomes a possibility.

Thus, we arrive at the important conclusion that a theory of diffraction which makes use of Huyghens' concept of partial waves has to base itself on the waves of that nature which arise in association with the reflection and refraction of light at the boundary between two media with different optical properties.

4. THE CHARACTER OF THE SECONDARY WAVES

Accordingly, we proceed to consider the partial or secondary waves having their origin at the elements of area of a boundary between two media of different refractive indices on which light is incident. There would clearly be two sets of such secondary waves travelling out respectively into the two media. The velocity of travel and the amplitude of the disturbance in the two sets being different, they must be considered as completely distinct from each other. If both media are isotropic, the configuration of the secondary waves in each medium would be hemispheres. The particular circumstances of the case, *viz.*, the refractive indices of the two media, the angle of incidence of the primary waves on the boundary and the state of polarisation of the incident light would determine the manner in which the energy of the incident radiation would be divided up between the reflected and refracted wave trains. The same circumstances would also determine the amplitude of the disturbance in the secondary waves sent out respectively into the two media.

The theory of diffraction proceeds by considering the secondary radiations emitted in different directions by the elements of area of the reflecting or refracting boundary and summing up their effects at the point of observation, having regard to their amplitudes and phases at that point. The first step in the theory is to write down a formal expression for the partial or secondary waves. Their amplitude would clearly be proportional to the amplitude of the primary disturbance incident on the boundary and reflected or refracted by it, as the case may be, and hence we have to find a dimensionless magnitude which expresses the proportionality. The quantities which might appear in it include, firstly, the element of area dS on the boundary,

r the radius vector joining the element with the point of observation, ϕ the angle between r and the normal to the plane of the boundary, and λ the wave-length of the light. It is reasonable to assume that the solid angle subtended by the element of area dS at the point of observation would determine its effect at that point. This solid angle is $dS.\cos\phi/r^2$. If we multiply this by the quantity r/λ, we obtain the dimensionless number $dS\cos\phi/\lambda r$, which exhibits an inverse proportionality both to r and to λ; the inverse proportionality with respect to r is to be expected for waves diverging outwards hemispherically, while the inverse proportionality with respect to λ is an indication that we are dealing with a wave-optical effect.

The important result emerging from the above is the appearance of $\cos\phi$ in the expression for the amplitude of the secondary waves. We shall refer to it as the *obliquity factor*. It has the value unity in the direction of the normal to the reflecting or refracting boundary, while it vanishes in any direction which lies in the plane of that boundary. It is evident that if we are concerned with a reflecting or refracting boundary of finite area and if the point of observation is at a sufficiently great distance from it, the angle ϕ may, without sensible error, be assumed to be the same for all the elements of area. It is evident also that in such a case, when we proceed to investigate the nature of the diffraction pattern at such distant point by summing up the effects of the elements of area with due regard to their phases and squaring the resultant amplitude to obtain the observed intensity, the square of the cosine of the obliquity, *viz.*, $\cos^2\phi$ would appear in it as a multiplying factor.

The result stated above is obviously of very general validity in respect of the diffraction patterns of the Fraunhofer class observed in various circumstances. All that is required is that the diffraction arises by reason of the limitation of the area of a plane surface at which light is reflected or through which it is transmitted; in the case of reflection, the material may be either a dielectric or a metal. It is not necessary that the surface should be continuous or that it should have uniform reflecting or transmitting power over the entire area. It might, for example, consist of several parallel strips, thus forming a plane diffraction grating. Further, since refraction at the boundary between two media which differ only infinitesimally in refractive index is equivalent to a simple transmission, it follows that the result would also be applicable to diffraction patterns of the Fraunhofer class arising from the passage of light through apertures in thin opaque screens.

5. VERIFICATION OF THE OBLIQUITY LAW

Any elementary treatment of diffraction theory can only be expected to be valid when the linear dimensions of the diffracting aperture are large compared with the wave-length of the light. As the angular spread of the diffraction pattern would in these circumstances be small, an experimental test of the law of the secondary wave might seem impracticable. Fortunately, however, this is not the case. For, the angle of diffraction ϕ is measured from the direction of the normal to the aperture and hence when the incidence of the light on the aperture is oblique, ϕ may be large enough for the factor $\cos^2 \phi$ to vary rapidly over the area of the diffraction pattern. Further, at such settings the diffraction patterns are spread out over a fairly wide angular range even when the dimensions of the aperture are many times larger than the wave-length. In these circumstances, the effect of the $\cos^2 \phi$ factor on the distribution of the intensity in the pattern becomes conspicuous and can indeed easily be observed and measured.

We may illustrate these remarks by considering a simple case, *viz.*, a diffracting aperture which is a plane strip bounded by parallel straight edges. As is well known, when the effects due to the infinitesimal elements of such an aperture are summed up, the expression obtained for the intensity in its Fraunhofer pattern includes a factor of the form $\sin^2 \zeta/\zeta^2$. This factor has a maximum value when $\zeta = 0$, and vanishes when $\zeta = \pm \pi, \pm 2\pi, \pm 3\pi$, etc. Since the value of $\sin^2 \zeta/\zeta^2$ is unaltered by a reversal of the sign of ζ, the graph of the function when set out with ζ as the abscissa is a symmetric curve in which the maxima on either side intermediate between the zero values are of equal intensity. The factor $\cos^2 \phi$ by which the expression for the intensity is multiplied would, however, modify this situation to an extent determined by the circumstances of the case.

In the particular case of normal incidence of the light on the aperture, $\zeta = \pi a \sin \phi/\lambda$, a being the width of the aperture, λ the wave-length and ϕ the angle of diffraction as already defined. More generally, when the light is incident on the aperture at an angle θ in a plane normal to its edges, $\zeta = \pi a (\sin \phi - \sin \theta)/\lambda$. Differentiating this, we obtain $d\zeta = \pi a/\lambda . \cos \phi \, d\phi$. Hence, as the incidence is made more oblique and $\cos \phi$ diminishes in value, the angular spread of the pattern determined by the increments of $d\phi$ becomes greater. The bands for which ϕ is greater than θ would also appear more widely spaced than those for which ϕ is less than θ. In these circumstances, the multiplying factor $\cos^2 \phi$ would have a very conspicuous influence on the distribution of intensity in the pattern. The bands for which ϕ is greater than θ would be much less intense than those for which ϕ is less than θ; indeed as ϕ approaches the limiting value $\pi/2$, the intensity in the former cases would become vanishingly small.

6. THE RESULTS OF EXPERIMENTAL STUDY

The present theory of diffraction and that of Kirchhoff thus differ fundamentally in the observable results which they indicate. This is scarcely a matter for surprise since they approach the diffraction problem from completely different points of view. Whereas the diffracting body plays the leading role in the present theory, it is not considered at all in the Kirchhoff formulation; the latter is based on the idea that the primary radiation from a source situated in free space can be expressed as a summation of secondary radiations from a set of sources distributed over a closed surface in free space enclosing the point of observation. The present theory leads to the result that the amplitude of the secondary waves vanishes along the plane of the surface at which they originate and increases progressively as we move away from that plane towards the direction of its normal. On the other hand, the Kirchhoff formulation indicates that the secondary waves have a maximum amplitude in the forward direction of the light rays from the original source and zero amplitude in the backward direction. The difference between the consequences of the two theories is of such a striking character that it is a simple matter by means of experimental study to decide between them.

In view of the importance of the issue here raised for a correct understanding of the theory of the diffraction of light, numerous experimental studies have been carried out by the writer. Diffracting apertures of various sizes ranging from several centimetres down to fractions of a millimetre have been employed. The angles of incidence of the light on the apertures have been varied from normal incidence up to grazing incidence. The circumstances in which the diffraction manifests itself have also been varied to include various cases, *e.g.*, the reflection of light at a plane surface of a dielectric or metal, the emergence of light after refraction through a transparent medium at various angles, the internal reflection of light within a transparent medium at various incidences and the transmission of light through apertures in thin opaque screens. The cases investigated include both simple and multiple apertures and plane diffraction gratings prepared by various techniques and operating by reflection as also by transmission. In all the cases investigated, the consequences of the present theoretical approach are completely vindicated by the facts of observation.

7. SUMMARY

The conventional treatment of diffraction problems based on the so-called Principle of Huyghens as analytically formulated by Kirchhoff is based on a misunderstanding of the original ideas of Huyghens regarding the propagation of light. It seeks to express the luminous effect due to the primary source as a summation of the effects of secondary sources situated on the elementary areas of a surface in free space enclosing the point of observation. Since the diffraction of light is a consequence of the presence of obstacles in the path of the light waves, the optical character of the obstacles and their configuration in space are of the very essence of the problem. It follows that the approach adopted in the Kirchhoff theory is misconceived and erroneous. It is however possible to base a theory of diffraction on Huyghens' concept of partial waves and his explanation of the reflection and refraction of light at the boundary between media with different optical properties. But this leads to results which differ fundamentally from those indicated by Kirchhoff's formula. It is shown that the issue which thus presents itself is readily capable of experimental test.