Curr. Sci. **28** 267–270 (1959)

The principle of Huyghens and the diffraction of light

SIR C V RAMAN

1. Introduction

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 especially their optical properties and their configuration in space, would determine these effects. Surprisingly enough, theories of diffraction have found general acceptance in which these factors receive very inadequate consideration. This situation is connected with the historical development of the subject and has arisen out of a misunderstanding of the ideas originally put forward by Huyghens in his celebrated Treatise on Light. A precis of the first three chapters of that treatise was given in a recent article in *Curr. Sci.*, and it was shown that the so-called principle of Huyghens as enunciated by later authors and made use by them as a basis for the theory of diffraction finds no warrant or support in the treatise. Huyghens did indeed introduce the concept of particular or partial waves and made effective use of it. But these partial waves of Huyghens had definite physical origins and the role which they played could therefore be readily understood. In these respects they differed radically from the ideas ascribed to him by later authors.

Theories clothed in the language of mathematical analysis have not infrequently found supporters and gained acceptance even though the physical ideas on which they are based are unsustainable. Kirchhoff's so-called rigorous formulation of the principle of Huyghens is a case of this kind. A statement often made and generally believed is that the Kirchhoff theory describes the experimental facts of the diffraction of light in a satisfactory manner. This belief has undoubtedly contributed to an uncritical acceptance of the ideas on which that theory is based. It is one of the objects of the present communication to show that it is indeed possible to make Huyghens' concept of partial waves the basis for a treatment of diffraction problems. This leads to results which are in agreement with the facts of experiment but are quite different from those indicated by the Kirchhoff theory. It follows that the latter theory is unsustainable and must accordingly be laid aside.

2. The wave-optics of Huyghens

Huyghens sought in his treatise to explain the three most familiar facts of geometrical optics on the basis of wave principles, viz., that the rays of light are propagated in straight lines; that the angles of incidence and reflection are equal; and that in refraction the ray is bent according to the law of sines. His explanations rest on the assumptions which he made regarding the structure of the luminiferous medium and the nature of light waves. His arguments led him to infer 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 together at the same instant. The same idea underlies Huyghens' explanation of the laws of reflection and refraction. Each piece of the original wave-front on reaching the boundary between two media is unable to continue on its original course by reason of the velocity of light being different in them. Accordingly it takes fresh paths, one in each of the two media, the direction of travel being such that the pieces of the original wave-front which are diverted from their path can all join up together again to form new wave-fronts in each medium. The latter requirement leads immediately to the equality of the angles of incidence and reflection in the first medium and to the law of sines for refraction into the second medium. This explanation was put into geometric form by Huyghens and is both simple and convincing. Regarded as a physical theory, it is 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.

Examining the ideas of Huyghens in detail, it becomes apparent that his explanation of the rectilinear propagation of light cannot possibly serve as a starting point for a theory of diffraction. On the other hand, his theory of reflection and refraction does offer itself as a basis. For, 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 impulses, we take account of their periodicity and also of the possibility of interferences between the secondary or partial waves having their origin at the different elements of area on the boundary, the restriction of the observable effect to precisely defined directions ceases to exist. In other words, the diffraction of light becomes a possibility.

That the diffraction of light stands in the closest relation to the phenomena of reflection and refraction is also otherwise obvious. As remarked earlier an obstacle of some kind in the path of a light-wave is a *sine-qua-non* for the manifestation of diffraction effects. A discontinuity in optical properties in the region traversed by the light represents such an obstacle, and if it exists over a sufficiently extended area, it would necessarily give rise to reflection and refraction.

3. The law of the secondary wave

A theory of diffraction which bases itself on the original ideas of Huyghens has accordingly to consider the 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. It is evident also that 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. These same circumstances would also determine the amplitude of the disturbance in the secondary waves sent out respectively into the two media.

A question of importance needing an answer is the manner of dependence on the angle of diffraction of the amplitude of the disturbance in the secondary waves. Considerations of an elementary nature enable us to deduce this. The projection of an element of area dS of the boundary on the surface of the enclosing hemisphere would be $dS \cos \phi$, ϕ being the angle of diffraction measured from the direction of the normal to the reflecting or refracting surface. This projected area would be a measure of the contribution which the element dS would make to the luminous effect observed in the direction ϕ. This would accordingly be a maximum in the direction of the normal ($\phi = 0$) and zero along the plane of the boundary ($\phi = \pi/2$). Hence in the expression for the amplitude of the effect due to each individual element, $\cos \phi$ would appear as a multiplying factor. At a sufficiently great distance from the diffracting surface, the angle of diffraction ϕ may be assumed to be the same for all its elements of area. It follows that when the expression for the intensity in the diffraction pattern is evaluated by a consideration of the interferences between the effects of the elementary areas, $\cos^2 \phi$ would appear in it as a multiplying factor.

The foregoing results are 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 refracted 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 opaque screens.

4. 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 wavelength 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 wavelength. 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 obliquity factor $\cos^2 \phi$ appearing in the expression for the intensity 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, $\zeta = \pi\alpha \sin \phi/\lambda$, α being the width of the aperture, λ the wavelength 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, $\zeta = \pi\alpha(\sin\phi - \sin\theta)/\lambda$. Differentiating this, we obtain $d\zeta = \pi\alpha/\lambda \cos\phi \, d\phi$. 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

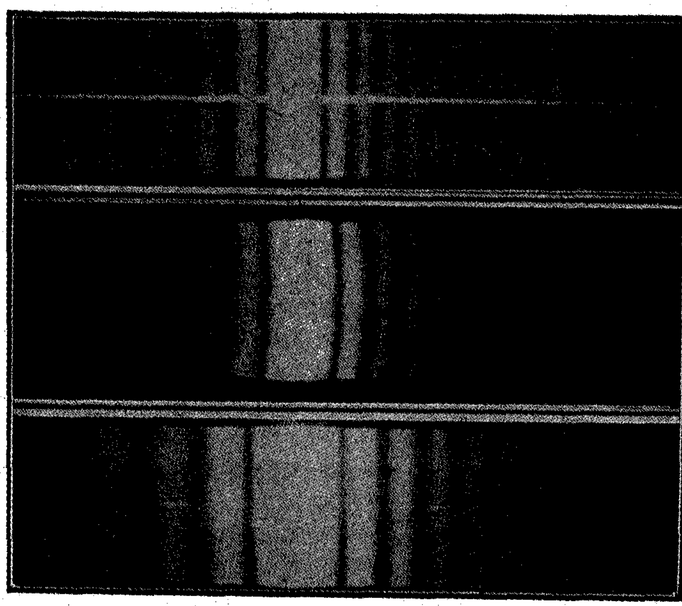

Figure 1. Diffraction of light by rectilinear apertures.

greater than θ would also appear more widely spaced than those for which ϕ is less than θ. In these circumstances, the obliquity factor $\cos^2 \phi$ would have a very conspicuous influence on the character of 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.

5. 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 or aperture 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 in free space can be represented as an integral in which the elements of area of a surface enclosing the primary source function as sources of secondary waves. The present theory leads to the result that the amplitude of the secondary waves emitted by the elements of the diffracting aperture vanishes in the plane of the aperture 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 incident light rays and zero amplitude in the backward direction. The difference is of such a striking character that it is a simple matter by means of experimental study to decide between the two theories.

In view of the importance of the issue here raised for a correct understanding of the theory of the diffraction of light, an extended series of 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 right 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 incidences beyond the critical angle, and the transmission of light through apertures in plane opaque screens. The cases investigated include both simple and multiple apertures and plane diffraction

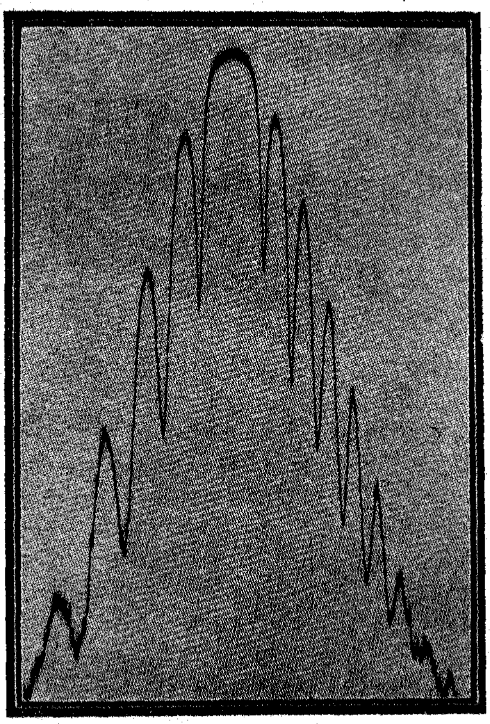

Figure 2. Microphotometer record of figure 1(*c*).

gratings prepared by various techniques and operating by reflection as also by transmission. It will suffice here to state that in all the cases investigated, the consequences of the present theoretical approach have been completely vindicated by the facts of observation.

Figure 1 (*a*, *b* and *c*) in the text are photographs of the diffraction of light at oblique incidences by a rectilinear aperture obtained by three different techniques. Figure 1(*a*) represents the diffraction pattern of the Fraunhofer class obtained with the monochromatic light of a sodium lamp *reflected* by a plane polished surface of glass one millimetre wide at oblique incidences. Figure 1(*b*) represents a diffraction pattern observed when light emerges obliquely after *refraction* through a prism of glass, the rear face of which was covered up by an opaque film of silver except for a narrow slit with parallel edges scratched out of it. Figure 1(*c*) represents the diffraction pattern *transmitted* obliquely through a rectilinear slit formed by the edges of two razor blades held parallel to each other. It will be seen that all the three photographs show the characteristic features indicated by theory and discussed in the third paragraph of section 4 above. It will be noticed that in each case the intensity of the diffraction bands falls off rapidly to zero on the side where they are broader and the number visible is quite small, while on the other side a great many fringes are seen, the intensity of which falls off very slowly. A microphotometer record of the pattern reproduced as figure 1(*c*) appears as figure 2 in the text. The record shows very conspicuously the great difference in the intensities of the corresponding bands on either side of the central maximum.