Optik 124 (2013) 6122–6125

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

A fresh outlook for Fresnel diffraction

Narahari Joshi*

University of Los Andes, Merida, Venezuela

ARTICLE INFO

Article history:
Received 2 December 2012
Accepted 28 April 2013

Keywords:
Diffraction
Circular aperture
Huygens wave front
Field quantization
Wave propagation

ABSTRACT

Fresnel theory, with certain level of assumptions, explains satisfactorily observed diffraction patterns. It is a scalar theory and basically it deals special way of superposition of fields originated from several points of Huygens wave front. In the present paper a total new approach is proposed according to which the interacting electric fields are not originated from the secondary emitters of the same wave front, instead, they are originated from the successive Huygens wave fronts. The proposed approach is explained with the help of the circular aperture. This new view supports Fresnels theory without a set of assumptions.

1. Introduction

The diffraction patterns recorded at short distances are explained on the basis of Fresnel theory with the consideration that waves are propagated in a spherical form. To explain the observed data certain assumptions are made, some of them are understandable, meanwhile, others need reconsideration. Understanding the details of the diffraction phenomena is an important issue as it deals with fundamental aspects of radiation such as its nature and propagation in space and time. Here we are addressing, a simple case, namely the diffraction with a circular aperture with a totally different approach.

Fresnel diffraction is obtained when either the source of radiation or the observing screen (or both) are at a short distance and the concept of near field propagation is valid [1,2]. Near field diffraction is defined with Fresnel number F

$$F = \frac{a^2}{\lambda D} \tag{1}$$

where a is the radius of the aperture, D is the distance between the observing screen and the aperture and λ is the wavelength with which the diffraction pattern is recorded. When $F \geq 1$, the diffraction is considered as a near field and is explained with the help of a spherical wave propagation.

It is well accepted that radiation propagates by means of Huygens wave front which acts as a continuous emitter of spherical secondary wavelets in all direction with uniform intensity. However, experimental data confirm that radiation does not flow in backward directions, therefore, Kirchhoff correction factor [1],

$$K(\theta) = \frac{1}{2}[1 + \cos(\theta)] \tag{2}$$

is introduced in an ad hoc manner to explain why radiation does not flow in the backward direction. To take into account directional dependence, Eq. (2) is launched and obviously for $\theta = \pi$, the term vanishes and the intensity in the backward direction is 0. The intensity at any point is estimated by constructing imaginary half period zones on a spherical wave front.

If the intensity is to be estimated at point O, then half zones are obtained by constructing circles of radius OP + 3λ/2, OP + λ, OP + 3/2λ, etc. where OP is the distance between O and the center of the aperture P. The optical field strength received at point O from each successive zone is out of phase by π because of the optical path differences and on this basis the intensity of the diffraction patterns is estimated. The details of the calculations were provided earlier [1–3]. It is based on the assumption that within each zone, there is no phase difference and the integration of all optical disturbances can be taken into account by a single wave of certain intensity which differs from the next (or previous) by exactly 180°.

It is clear that the contribution in the intensity from the first zone is positive, and then from the second zone is negative and so on. By using the Kirchhoff's correction factor [3] and with some assumptions, the obtained results are explained satisfactorily with the theoretically expected values and the theory is considered to be in excellent agreement with the reported experimental results.

Basically, in interference and diffraction phenomena, the optical disturbances or electric fields are added from rays originated from the same wave front, i.e. emitters are in the same phase but they have different path lengths. This means that they are not reaching

* Tel.: +58 2742661880.
E-mail address: joshi@ula.ve

0030-4026/$ – see front matter © 2013 Elsevier GmbH. All rights reserved.
http://dx.doi.org/10.1016/j.ijleo.2013.04.084

at the same time to the point O (or the observer) and hence it is not appropriate to add optical disturbances or intensity of electric and magnetic fields vectorially, particularly when there are periodic terms (like sine or cosine) involved. This is not accurate when one wants to determine the positions of maxima and minima from waves coming at different times, even though they are emitted in the same phase at the same time. How the interference or the interaction can be examined when one wave is ahead with respect to other by a finite distance. The advanced wave will interact with the detecting system (photographic plate or detector) before the delayed wave will meet the first one. This is particularly true when waves are not continuous. In short, the scalar diffraction theory for electromagnetic waves is totally an inadequate approach. In spite of this fundamental problem, the theory works miraculously and Fresnel's diffraction is applied in several fields with a remarkable success.

The one and only reason to understand this discrepancy could be that the addition of waves are not from the same wave front but from different wave fronts which arrive at the same time, it means they are located at the same distance from the observer. The points on Huygens wave front are continuous emitters of wavelets and naturally at the point of the observer, the waves coming at the same time should be added independent of where they are originated from the same wave front or different wave fronts. In a continuous wave emission process, a series of waves are emitted from the source. The interference or diffraction takes place from the waves reaching at the same time (not from the same wave front). This is achieved by the rays originated from successive fronts which are separated by intervals of $\lambda/2$. It is a totally different and unconventional approach.

2. Field quantization and Huygen's principle

This above mentioned point of view is exactly expected from the field quantization process [4–6]. Recent investigation [5] shows that the dual nature of photons is not only for the interaction [7] but also in the propagation process [5,6,8]. According to quantum field theory, electromagnetic fields are converted in to particles (and vice versa) with the help of creation and annihilation operators defined as [4,5]

$$a = \frac{\omega q + ip}{\sqrt{2\omega}} \tag{3a}$$

$$a^{+} = \frac{\omega q - ip}{\sqrt{2\omega}} \tag{3b}$$

where q is a generalized coordinate and p is the momentum operator given by

$$p = i\frac{\partial}{\partial x} \quad (c = \hbar = 1, \quad \text{units used in quantum field theory})$$

where ω stands for the frequency of the oscillation. This means that to generate photons from the fields, the source should be periodic with frequency ω. In the present case the fields (electric and magnetic) are periodic, therefore, photons are created and annihilated with the help of Eqs. (3a) and (3b). The details were given earlier by Joshi [5].

Thus, in space and time the fields (electric and magnetic) are converted into photons and again converted back into fields. The sum of the energy associated with fields and particles is conserved. Obviously, when the field energy is zero, the photon density is maximum and the wave front is formed. This means that at the nodes, $\lambda/2$, λ, $3\lambda/2$, the wave fronts are created and therefore, the separation between two successive wave fronts is $\lambda/2$. The present outlook is based on quantum electrodynamics and it not only supports the formation of Huygens wave fronts but also suggests the separation between them. This is the first time that such a conclusion has been reported about the separation between two successive wave fronts [5]. Application of Huygens–Fresnel principle to estimate the intensity at the point of the observer is based on summation of the amplitude originated from each point located on the primary wave front. In order to achieve the agreement with the observed experimental data the amplitude is multiplied by $1/\lambda$ and Kirchhoff's inclination factor [1]. Multiplication by $1/\lambda$ indicates that the intensity is reduced by the factor $1/\lambda$ and this indicates that wave fronts are separated by λ. Therefore, the approximation made in the intensity calculation by Fresnel theory is also fully supported by the present approach.

In short, the observed diffraction pattern is explained with the help of summation of the field originated from different wave fronts but they reach at the observation point at exactly the same time with different phases. This means that optical disturbances are positive from the one, then it is negative from the previous. This is predicted by the theory and obviously leads to the same results like Fresnel's zone.

3. Diffraction by a circular aperture

Let us consider the well studied and simple case of diffraction by a circular aperture. Fig. 1 shows the planer view of Huygens wave fronts originated from the source S and separated by the distance $\lambda/2$. The figure is not to the scale. Now let us examine the intensity at the point O when the aperture is very small. The radiation reaching at the point O, at time t_1 is from the secondary emitters which are at a distance OP which leads to a bright spot. The intensity could be higher if there would not have been an aperture. This is understandable because the intensity is not only due to the exact point at the center but also from the central circular region of the secondary emitters. Let us consider that the radius of the aperture is increased in such a way that the wave front located at the back of the aperture at a distance $\lambda/2$ will also comes into play. In this case, the equidistance points from O are B and B_1 located on the first wave front meanwhile b and b_1 are located on the previous wave front. This means the optical disturbances reaching to the observer at time t_1 are out of phase and obviously the dark spot will appear at the center. It means that a distance O to B is equal to OP + $\lambda/2$. This is exactly the Fresnel's first zone for which the dark spot is observed. If the aperture is further increased, to include the contribution from the previous wave front separated from the aperture by distance λ, then again the bright spot at the center will appear. In this case, the contribution from the first wave front will be canceled with the second one and the optical disturbances will be only from the third wave front it means from the points c_2 and c_3 which are in phase. It will happen when the OC will be OP + λ. This again corresponds to

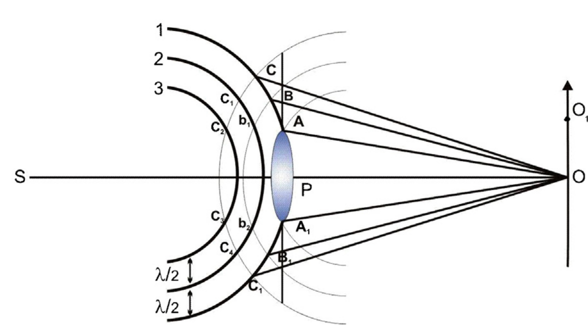

Fig. 1. A sectional view of the intersections of spherical Huygens fronts (shown by numbers 1, 2 and 3) and the imaginary spheres of radius OA, OB and OC where A, B and C are the opening of the apertures. When the radius of the aperture is increased, higher number of secondary emitters participate in the diffraction pattern originated from different wave fronts giving white or black circles at the center.