Analytic expression for Fresnel diffraction

Peng Wang and Yuguang Xu

Shanghai Institute of Optics and Fine Mechanics, Academia Sinica, P.O. Box 800-211, 201800 Shanghai, China

Wei Wang

North China Optical Instrument Factory, P.O. Box 2151, 100050 Beijing, China

Zhijian Wang

Changchun Institute of Optics and Fine Mechanics, P.O. Box 1003, 130022 Changchun, China

Received June 17, 1997; revised manuscript received October 2, 1997; accepted October 6, 1997

An analytic formula for the light field amplitude (and therefore also intensity) behind a circular aperture illuminated by spherical or plane waves in the Fresnel limit is given, and the essential difference between Fresnel diffraction and Fraunhofer diffraction is discussed. © 1998 Optical Society of America [S0740-3232(98)00403-7]

OCIS codes: 050.1940, 260.3160.

1. INTRODUCTION

This paper concerns diffraction, which has considerable significance in optics. Since the publication of Durnin's papers on nondiffracting beams,[1,2] diffraction has been the subject of many papers expressing different views.[3–9] The different views have resulted from different understandings of the physical nature of diffraction. Up to now, analytic expressions for Fraunhofer diffraction have been given. However, as far as we know, no similar solution to Fresnel diffraction has been presented. Numerical methods and the fractional Fourier transform have been used by some authors[10–17] to calculate the intensity distribution of Fresnel diffraction. In this paper we suggest an analytical expression that is convenient for calculating the amplitude and intensity of Fresnel diffraction. We also point out the essential difference between Fresnel diffraction and Fraunhofer diffraction. The latter occurs only under a special condition, i.e., $M = 0$ or $M/2b = 0$. Without this condition, Fresnel diffraction occurs.

2. GENERAL CASE: FRESNEL DIFFRACTION

By the general case we mean that the receiver screen fails to pass through the spherical-wave center; the diffraction that occurs is called Fresnel diffraction, which is divergent-wave diffraction, as shown in Fig. 1. The spherical-wave center and the receiver screen lie on opposite sides of the diffraction screen.

From the Fresnel–Kirchhoff diffraction formula, we know that

$$\tilde{E}_{(\mathrm{p})} = \frac{-i}{\lambda} \int \int_{\Sigma} f(\theta,\ \theta_0) \tilde{E}\ \frac{\exp(ikr)}{r}\ \mathrm{d}\sigma, \tag{1}$$

where $f(\theta,\ \theta_0)$ is an inclination factor:

$$f(\theta,\ \theta_0) = \tfrac{1}{2}(\cos\ \theta + \cos\ \theta_0), \tag{2}$$

$$\tilde{E} = \frac{A\ \exp(ikR)}{R}. \tag{3}$$

When $\theta_0 = \theta \approx 1$,

$$\tilde{E}_{(\mathrm{p})} = \frac{-i}{\lambda}\ \tilde{E} \int \int_{\Sigma} \frac{\exp(ikr)}{r}\ \mathrm{d}\sigma. \tag{4}$$

From Fig. 1 we obtain

$$r \approx b + \frac{R + b}{2Rb}\, q^2 + \frac{\rho^2}{2b} - \frac{q\rho}{b} \cos\ \alpha. \tag{5}$$

Let

$$M = \frac{R + b}{R}. \tag{6}$$

Then

$$r = b + \frac{M}{2b}\, q^2 + \frac{\rho^2}{2b} - \frac{q\rho}{b} \cos\ \alpha. \tag{7}$$

We also know that

$$\mathrm{d}\sigma = R^2 \sin\ u \mathrm{d}u \mathrm{d}\alpha. \tag{8}$$

When u is very small, then

$$\mathrm{d}\sigma \approx q \mathrm{d}q \mathrm{d}u. \tag{9}$$

Placing Eqs. (7) and (9) into Eq. (4), we obtain

$$\tilde{E}_{(\mathrm{p})} = \frac{-i}{\lambda}\ \tilde{E}\ \exp\left[ik\left(b + \frac{\rho^2}{2b}\right)\right] \int_0^a \times \exp\left(ik\ \frac{M}{2b}\, q^2\right) q \mathrm{d}q \int_0^{2\pi} \exp\left(-ik\ \frac{qp}{b} \cos\ \alpha\right) \mathrm{d}\alpha. \tag{10}$$

Using the partial integral method, we obtain

$$\tilde{E}_{(p)} = \tilde{E}_0 \exp\left[ik\,\frac{1}{2b}\,(\rho^2 + Ma^2)\right] \times \sum_{n=1}^{\infty}\left(-i\,\frac{Ma}{\rho}\right)^n J_n\left(k\,\frac{qa}{b}\right), \tag{11}$$

where

$$\tilde{E}_0 = A\,\frac{\exp[ik(R + b)]}{R + b} \tag{12}$$

and $\tilde{E}_0$ is the amplitude of point p_0 on the receiver screen when the light travels in straight lines.

Let

$$N = \frac{a^2}{\lambda b}, \tag{13}$$

where N is the Fresnel coefficient. For plane-wave diffraction, N is the half-period zone number. For spherical-wave diffraction, MN is the half-period zone number. Placing $k = 2\pi/\lambda$ and N into Eq. (11), we get

$$\tilde{E}_{(p)} = \tilde{E}_0 \exp\left[iMN\pi\left(1 + \frac{\rho^2}{Ma^2}\right)\right] \times \sum_{n=1}^{\infty}\left(-iM\,\frac{a}{\rho}\right)^n J_n\left(2N\pi\,\frac{\rho}{a}\right). \tag{14}$$

With the Bessel function, we obtain

$$u = \exp\left[\frac{x}{2}\left(t - \frac{1}{t}\right)\right] = \sum_{n=-\infty}^{\infty} J_n(x)t^n. \tag{15}$$

Let

$$x = 2N\pi\,\frac{\rho}{a}, \quad t = -iM\,\frac{a}{\rho}.$$

From $J_{-n}(x) = (-1)^n J_n(x)$ we obtain

$$u = J_0\left(2N\pi\,\frac{\rho}{a}\right) + \sum_{n=1}^{\infty}\left(-iM\,\frac{a}{\rho}\right)^n J_n\left(2N\pi\,\frac{\rho}{a}\right) + \sum_{n=1}^{\infty}\left(-i\,\frac{\rho}{Ma}\right)^n J_n\left(2N\pi\,\frac{\rho}{a}\right). \tag{16}$$

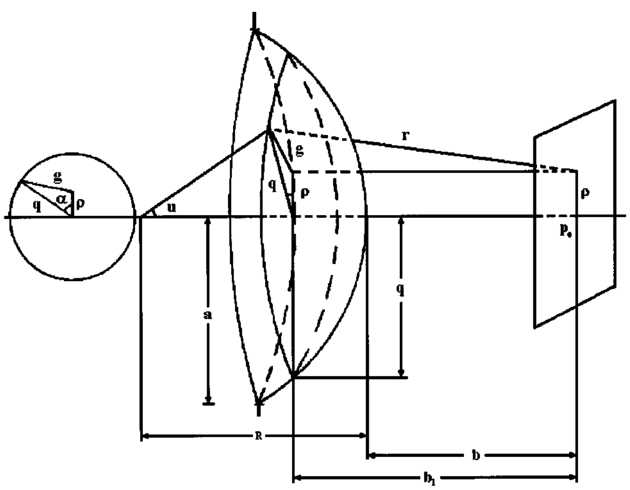

Fig. 1. Divergent-wave Fresnel diffraction.

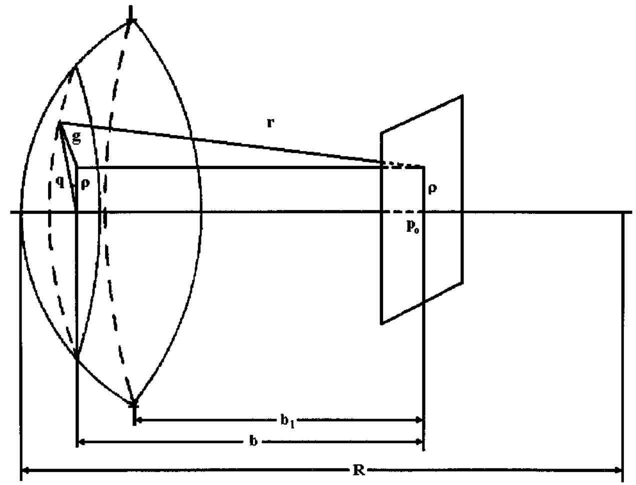

Fig. 2. Convergent-wave Fresnel diffraction.

From the character of orthogonal polynomial, we obtain

$$p = \frac{\sum_{n=1}^{\infty}\left(-i\,\frac{\rho}{Ma}\right)^n J_n\left(2N\pi\,\frac{\rho}{a}\right)}{\sum_{n=1}^{\infty}\left(-iM\,\frac{a}{\rho}\right)^n J_n\left(2N\pi\,\frac{\rho}{a}\right)} = \frac{\exp\left(\frac{\rho^2}{M^2a^2}\right) - 1}{e - 1}. \tag{17}$$

Thus

$$u = J_0\left(2N\pi\,\frac{\rho}{a}\right) + (1 + p)\sum_{n=1}^{\infty}\left(-iM\,\frac{a}{\rho}\right)^n J_n\left(2N\pi\,\frac{\rho}{a}\right). \tag{18}$$

We also know that

$$u = \exp\left\{N\pi\,\frac{\rho}{a}\left[\left(-iM\,\frac{a}{\rho}\right) - \left(-iM\,\frac{a}{\rho}\right)^{-1}\right]\right\} = \exp\left[-iMN\pi\left(1 + \frac{\rho^2}{M^2a^2}\right)\right]. \tag{19}$$

From Eqs. (14), (18) and (19), we obtain

$$\tilde{E}_{(p)} = D\tilde{E}_0\left\{\exp\left[iN\pi\,\frac{\rho^2}{a^2}\left(1 - \frac{1}{M}\right)\right] - \exp\left[iMN\pi\left(1 + \frac{\rho^2}{Ma^2}\right)\right]J_0\left(2N\pi\,\frac{\rho}{a}\right)\right\}, \tag{20}$$

where

$$D = \frac{1}{1 + p} = \frac{e - 1}{e + \exp\left(\frac{\rho^2}{M^2a^2}\right) - 2}, \tag{21}$$

which is called amplitude decay coefficient.

The intensity distribution is

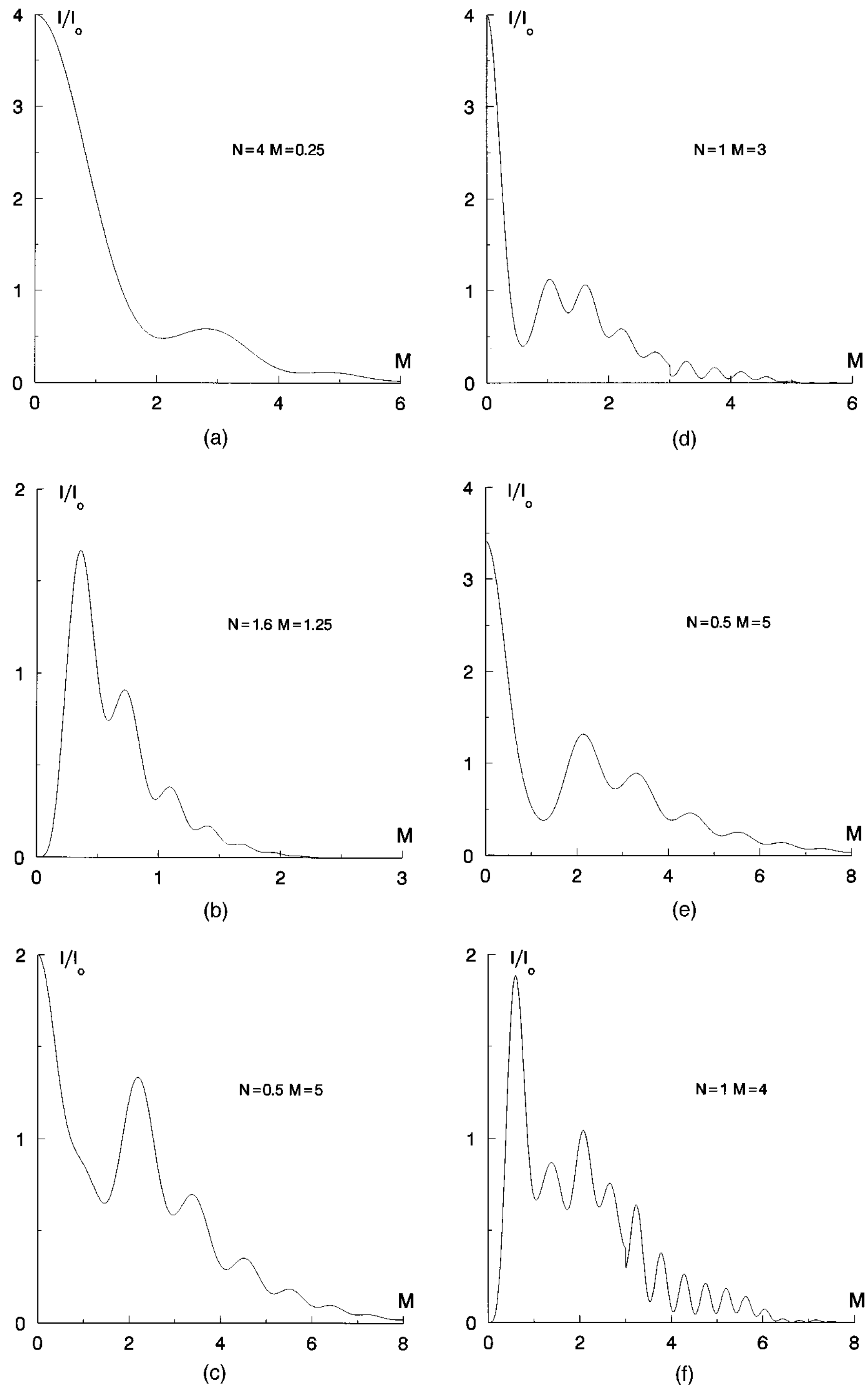

Fig. 3. Intensity distribution of Fresnel diffraction.

$$I_{(\mathrm{p})} = D^2 I_0 \left\{ 1 - 2J_0\left(2N\pi \frac{\rho}{a}\right) \cos\left[\left(1 + \frac{\rho^2}{M^2a^2}\right) MN\pi\right] + {J_0}^2\left(2N\pi \frac{\rho}{a}\right)\right\}. \tag{22}$$

Equation (20) is the amplitude formula for Fresnel diffraction, and Eq. (22) is the intensity distribution formula for Fresnel diffraction.

In calculating the intensity of point p_0, under the Fresnel diffraction condition, since $\rho = 0$ and $D = 1$ Eq. (22) is transformed to

$$I_{(p_0)} = 4I_0 \sin^2\left(\frac{MN}{2}\pi\right) = 4I_0 \sin^2\left(\frac{R+b}{2Rb} a^2 \frac{\pi}{\lambda}\right), \tag{23}$$

where

$$a = \sqrt{\frac{mRb}{R + b}\lambda} \quad \text{(bright half-period zone)},$$

$$a = \sqrt{\frac{2mRb}{R + b}\lambda} \quad \text{(dark whole-period zone)},$$

and m is a positive integer.

The conclusions above are identical with the results obtained by means of half-period zone construction and the graphical vector method.

Figure 2 shows convergent-wave diffraction; thus

$$\tilde{E} = A\,\frac{\exp(-ikR)}{R};\ \tilde{E}_0 = A\,\frac{\exp[-ik(R - b)]}{R - b};$$

$$M = \frac{R - b}{R}.$$

When there is plane-wave diffraction, $M = 1$, and Eq. (23) is transformed to

$$I_{(\mathrm{p})} = D^2 I_0\left\{1 - 2J_0\left(2N\pi\,\frac{\rho}{a}\right)\cos\left[\left(1 + \frac{\rho^2}{a^2}\right)N\pi\right] + {J_0}^2\left(2N\pi\,\frac{\rho}{a}\right)\right\}. \tag{24}$$

The intensity distributions of Fresnel diffraction are shown in Fig. 3 for different M and N.

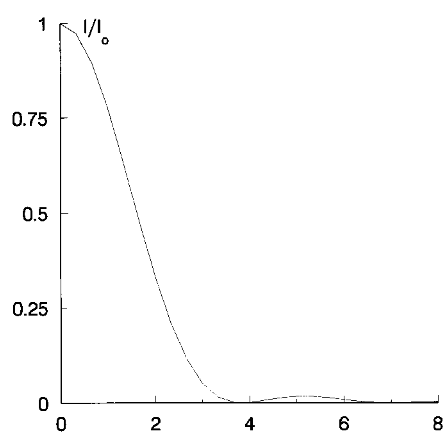

Fig. 4. Intensity distribution of Fraunhofer diffraction.

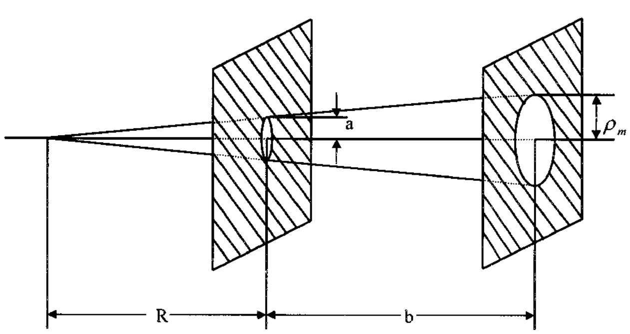

Fig. 5. Diagrammatic sketch of M value.

3. SPECIAL CASE: FRAUNHOFER DIFFRACTION

Only when the receiver screen passes through the spherical-wave center will there be convergent-wave diffraction. Now $R = b$, $M = 0$, and Eq. (10) is transformed to

$$\tilde{E}_{(\mathrm{p})} = \frac{-i}{\lambda}\,\tilde{E}\,\exp\left[ik\left(b + \frac{\rho^2}{2b}\right)\right]\int_0^a q\,\mathrm{d}q\int_0^{2\pi} \times \exp\left(-ik\,\frac{q\rho}{b}\cos\alpha\right)\mathrm{d}\alpha. \tag{25}$$

The integral is simple. We obtain

$$\tilde{E}_{(\mathrm{p})} = \pi a^2\tilde{E}_0\left[\frac{2J_1\left(2N\pi\,\dfrac{\rho}{a}\right)}{2N\pi\,\dfrac{\rho}{a}}\right] = \pi a^2 E_0\left[\frac{2J_1(x)}{x}\right]. \tag{26}$$

The intensity is

$$I_{(\mathrm{p})} = I_0\left[\frac{2J_1(x)}{x}\right]^2, \tag{27}$$

where

$$I_0 = \pi^2 a^4(\tilde{E}_0\tilde{E}_0{}^*). \tag{28}$$

The intensity distribution of Fraunhofer diffraction is shown in Fig. 4.

4. CONCLUSION

For the straight propagation of light (no diffraction taken into consideration), M means the ratio of the radius ρ_m of the disk on the receiver screen to the radius a of the circular aperture, as shown in Fig. 5. With a divergent spherical wave, $M > 1$; with a convergent spherical wave, $M < 1$; and for plane waves, $M = 1$. Now the equation $M = 0$ is obtained when there are convergent spherical waves and the receiver screen passes through the center of the spherical wave, and with the diffraction of plane waves, the equation $M/2b = 0$ is obtained when the receiver screen is located at infinite distance. In such a case ($M = 0$ or $M/2b = 0$) Fraunhofer diffraction occurs; otherwise, there is Fresnel diffraction. And so it may be said that Fraunhofer diffraction is a special occurrence of Fresnel diffraction. It is the intensity distribution of Fraunhofer diffraction that is observed in most optical systems in which the receiver elements are located on the image plane ($M = 0$), and therefore Fraunhofer diffraction is of great importance in engineering. The distinction between Fresnel diffraction and Fraunhofer diffraction can be seen in formula (10): When $M = 0$ or $M/2b = 0$, then $\exp[ik(M/2b)q^2] = 1$ and the integral becomes simple, which is why the analytic expression for Fraunhofer diffraction was found earlier. When $M \neq 0$ or $M/2b \neq 0$, the integral is complex; however, this paper presents the first (to our knowledge) analytic expression for Fresnel diffraction obtained by strictly mathematical derivation.

It was firmly believed in classical documents that if $k(M/2b)a^2 \ll \pi$, there will be Fraunhofer diffraction.[18–20]

Certainly, the intensity distribution of Fresnel diffraction that has been obtained when MN is very small (such as $MN \leqslant 1$) is similar to Fraunhofer diffraction, as shown in Fig. 3(a), but it still is Fresnel diffraction.

It should be noticed that the expression has been derived under the condition that the Kirchhoff inclination factor is thought to approximate one. MN means the Fresnel half-period zone. With a plane wave, $MN = N$, N is known as the Fresnel number in some of the literature. N is in fact the Fresnel half-period zone that appears when plane waves are diffracted. The diagram of the intensity distribution resulting from calculations in this paper is equivalent to those given in Refs. 11 and 14. The principles given here have also been applied to the calculation of the intensity distribution of lasers and to the design of a diffraction grating with high chromatic dispersion ($\Delta\lambda = 0.005$ nm) and wider slits (0.016 mm).

REFERENCES

1. J. Durnin, J. J. Miceli, Jr., and J. H. Eberly, "Diffraction-free beams," Phys. Rev. Lett. **58**, 1499–1501 (1987).
2. J. Durnin, "Exact solutions for nondiffracting beams. I. The scalar theory," J. Opt. Soc. Am. A **4**, 651–654 (1987).
3. P. Sprangle and B. Hafizi, "Comment on nondiffracting beams," Phys. Rev. Lett. **66**, 837–838 (1991).
4. S. Wang, "On principles of diffraction," Optik (Stuttgart) **100**, 143–144 (1995).
5. S. Ruschin, "Modified Bessel nondiffracting beams," J. Opt. Soc. Am. A **11**, 3224–3228 (1994).
6. P. L. Overfelt and C. S. Kenney, "Comparison of the propagation characteristics of Bessel, Bessel–Gauss, and Gaussian beams diffracted by a circular aperture," J. Opt. Soc. Am. A **8**, 732–745 (1991).
7. Z. Jiang, Q. Lu, and Z. Liu, "Several questions related to diffraction," Appl. Lasers (China) **6**, 266–270 (1994).
8. J. Gu, "Discussion on π jump at boundary in diffraction phenomena," Appl. Lasers (China) **6**, 270–272 (1994).
9. P. Chen and J. Qin, "Discussion about the non-diffraction and super-diffraction," Appl. Lasers (China) **6**, 274–276 (1994).
10. W. H. Southwell, "Validity of the Fresnel approximation in the near field," J. Opt. Soc. Am. **71**, 7–14 (1981).
11. A. M. Steane and H. N. Rutt, "Diffraction calculations in the near field and validity of the Fresnel approximation," J. Opt. Soc. Am. A **6**, 1809–1814 (1989).
12. C. J. R. Sheppard and M. Hrynevych, "Diffraction by a circular aperture: a generalization of Fresnel diffraction theory," J. Opt. Soc. Am. A **9**, 274–281 (1992).
13. Carl R. Schultheisz, "Numerical solution of the Huygens–Fresnel–Kirchhoff diffraction of spherical waves by a circular aperture," J. Opt. Soc. Am. A **11**, 774–778 (1989).
14. H. G. Kraus, "Huygens–Fresnel–Kirchoff wave-front diffraction formulation: spherical waves," J. Opt. Soc. Am. A **6**, 1196–1205 (1989).
15. D. E. Dauger, "Simulation and study of Fresnel diffraction for arbitrary two-dimensional apertures," Comput. Phys. **10**, 591–604 (1996).
16. C. de Jong, "Precise mathematical description of the diffraction pattern beyond a square aperture in an absorbing medium," Opt. Eng. **30**, 641–645 (1991).
17. H. Haman, "Efficient Fresnel-transform algorithm based on fractional Fresnel diffraction," J. Opt. Soc. Am. A **12**, 1920–1931 (1995).
18. M. Born and E. Wolf, *Principles of Optics*, 6th ed. (Pergamon, New York, 1980), Chap. 8.
19. Q. Liang, *Physical Optics* (Publishing House of the Mechanics Industry, Beijing, China, 1986), p. 167.
20. K. Zhao and X. Chan, *Optics* (Publishing House of Beijing University, Beijing, China, 1982), p. 192.