

Three-dimensional intensity distribution near the focus in systems of different Fresnel numbers

Yajun Li

Institute of Electro-Optical Engineering, National Chiao Tung University, 1001 Ta Hsueh Road, Hsinchu, Taiwan 300, China

Emil Wolf*

Department of Physics and Astronomy, University of Rochester, Rochester, New York 14627

It was recently shown that, when a converging spherical wave is focused in a diffraction-limited system of sufficiently low Fresnel numbers, the point of maximum intensity does not coincide with the geometrical focus but is located closer to the exit pupil. In the present paper both qualitative and quantitative arguments are presented that elucidate the modifications that the whole three-dimensional structure of the diffracted field undergoes as the Fresnel number is gradually decreased. Contours of equal intensity in the focal region are presented for systems of selected Fresnel numbers, which focus uniform waves.

1. INTRODUCTION

In a number of publications that appeared in recent years¹⁻⁵ it was demonstrated that the classic theory regarding the structure of the focal region does not predict correct results under all circumstances. More specifically, it was found that, when a uniform, converging, monochromatic spherical wave is diffracted at a circular aperture, the classic theory is adequate for calculating the intensity distribution along the axis near the geometrical focus only if the Fresnel number of the focusing geometry [defined by Eq. (2.2) below] is large compared to unity.⁶ Since before the invention of the laser the Fresnel number of a focusing system was generally much larger than unity, no discrepancies with predictions of the classic theory were then found. However, in recent years systems have been developed, especially for use with laser beams, whose Fresnel numbers are of the order of unity or even smaller. Hence it is desirable to extend the classic theory of focusing to systems of this kind.

In the present paper we present isophotes (contours of equal intensity) in the focal region in systems of different Fresnel numbers. We begin with some general qualitative observations, which indicate why one must expect changes in the structure of the focal region as the Fresnel number N of the system changes gradually from large values ($N \gg 1$) to very small values ($N \ll 1$). The qualitative arguments also indicate that the classic theory of focusing reaches its limit of validity for systems whose Fresnel numbers are of the order of unity and that it breaks down completely when the Fresnel number is much smaller than unity. The qualitative considerations, which also indicate the general trend of the changes in the structure of the focal region as the Fresnel number decreases, are confirmed by results of detailed computations, presented in Section 4.

Only focusing of uniform waves is considered in this paper. However, the results may be expected to be similar for focused laser beams in systems in which the size of the diffracting

aperture is comparable with or smaller than the effective cross section of the laser beam that is incident upon it.

2. SOME QUALITATIVE CONSIDERATIONS

Let us consider a uniform, monochromatic, spherical wave converging to a focus F and diffracted at a circular aperture, of radius a , in an opaque screen. We denote by O the center of the aperture and by f the distance OF (Figs. 1 and 2). We assume that

$$a \gg \lambda, \quad \left(\frac{a}{f}\right)^2 \ll 1, \quad (2.1)$$

where λ is the wavelength. Let

$$N = a^2/\lambda f \quad (2.2)$$

be the Fresnel number of the focusing geometry.

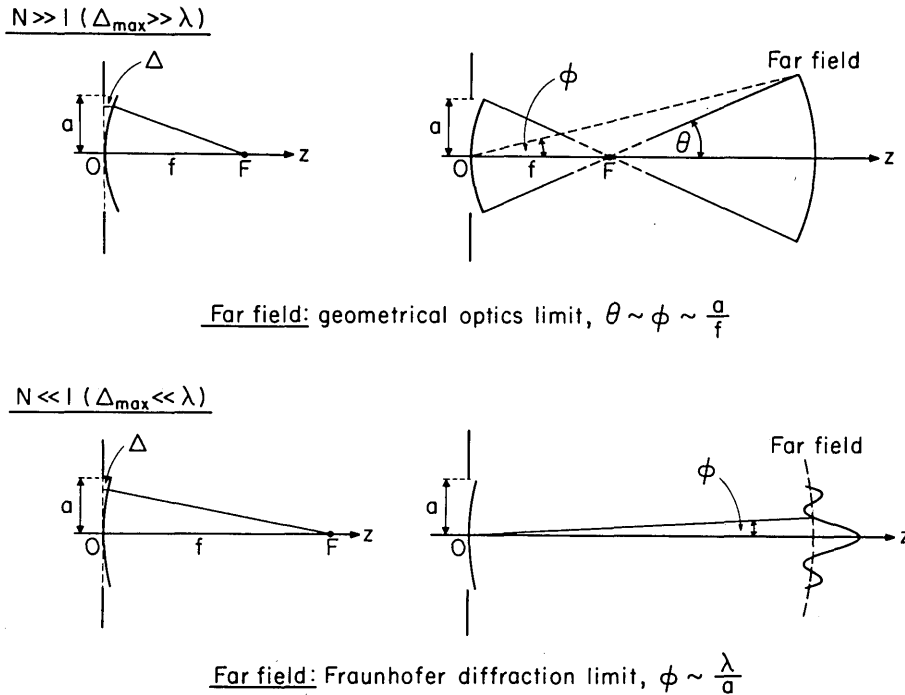
It follows from elementary calculations that the maximum, attained at the edge of the aperture, of the distance Δ between the spherical wave front through O and the plane of the aperture is, to a good approximation, given by

$$\Delta_{\max} = a^2/2f. \quad (2.3)$$

It is seen from Eqs. (2.2) and (2.3) that

$$N = \Delta_{\max}/(\lambda/2), \quad (2.4)$$

i.e., the Fresnel number is equal to the value of Δ_{\max} expressed in units of half wavelength. Expressed differently, N represents the number of Fresnel zones (Ref. 7, Sec. 8.2) that fill the aperture when the aperture is viewed from the geometrical focus F . To obtain some indication about the order of magnitude of the Fresnel number in some typical cases, let us consider two simple examples. For a conventional focusing lens we might have $a = 1$ cm, $f = 5$ cm, $\lambda = 0.5$ μ m, and, according to Eq. (2.2), we then have $N = 4000$. On the other hand, for a long-focal-length infrared laser system we may



Far field: geometrical optics limit, $\theta \sim \phi \sim \frac{a}{f}$

Far field: Fraunhofer diffraction limit, $\phi \sim \frac{\lambda}{a}$

Fig. 1. Illustrating some qualitative differences in the structure of the far field generated by the diffraction of a uniform, converging monochromatic spherical wave at an aperture in an opaque screen in systems of large Fresnel numbers (upper figures) and low Fresnel numbers (lower figures).

have $a = 1 \text{ cm}$, $f = 3 \text{ m}$, $\lambda = 10 \mu\text{m}$, and, according to Eq. (2.2), $N = 3.3$ in this case.

It is well known that for conventional focusing systems (systems with $N \gg 1$) the maximum intensity in the diffracted field is at the geometrical focus F. Moreover, except close to the axis and in the vicinity of the edge of the geometrical shadow, the far field has the appearance of a cutoff portion of a uniform spherical wave, diverging from the geometrical focus^{8,9} F. This "spherical cap" subtends, in each cross section through the axis OF, the angle $2\theta = 2a/f$ at F; and, because it is located in the far zone, it subtends (within the accuracy of the asymptotic approximation of large distances from F) the angle $2\phi \sim 2\theta$ at the center O of the aperture:

$$\phi \sim \theta = a/f \tag{2.5}$$

(see the top right-hand diagram in Fig. 1). This structure of the far field is, of course, in agreement with predictions of geometrical optics.

Let us now consider the other extreme case, $N \ll 1$. For simplicity we will assume that the radius a of the aperture and the wavelength λ have the same values as before and that the small value of N is achieved by making the distance f between the plane of the aperture and the geometrical focus be sufficiently large. According to Eq. (2.4) we now have $\Delta_{\max} \ll \lambda/2$, and hence we are essentially dealing with the diffraction of a plane wave at the circular aperture. In such a case the far field has quite a different structure from that in systems that we just discussed (for which $N \gg 1$). It can no longer be described by geometrical optics but rather must be described by the theory of Fraunhofer diffraction. According to a well-known formula pertaining to Fraunhofer diffraction at a circular aperture, namely, the Airy diffraction formula (Ref. 7, Sec. 8.5.2), the far field will now have a sharp maximum in the forward (axial) direction and will rapidly decrease to zero in an oscillatory manner on each side of it. The effective angular spread of the radiation is given by the order-of-magnitude relation

$$\phi \approx \lambda/a. \tag{2.6}$$

Since, when $N \ll 1$, we are effectively dealing with the diffraction of a plane wave at a circular aperture, it seems intuitively clear that the intensity maximum will now be located in the vicinity of the center of the aperture¹⁰ and that the diffracted field, at distances sufficiently far away from the aperture, will behave as if it had been generated by a radiating source located in the region of the aperture.

Comparison of the two extreme cases that we have just briefly discussed indicates that, as we proceed from focusing systems of large Fresnel numbers N to focusing systems of small Fresnel numbers, the intensity maximum moves from its location at the geometrical focus F to some point in the vicinity of the center O of the aperture. It is clear that this

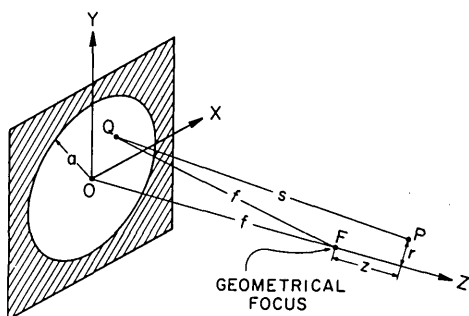


Fig. 2. Illustrating the notation used in determining the structure of the focal region. The point Q is located on a spherical wave front W of radius f , centered on F and passing through the center O of the aperture.

transition is accomplished in a continuous manner. If we assume, as seems reasonable, that the shift of the intensity maximum from its location at the geometrical focus is a monotonic function¹¹ of N (with a and λ being kept fixed), we may conclude that with decreasing values of the Fresnel number the point of maximum intensity will move toward the center of the aperture.¹² This is indeed what was found from explicit calculations³ and was readily confirmed by experiment.¹³

In Fig. 1 the two extreme situations of systems with large and small Fresnel numbers N are illustrated by the upper and lower diagrams, respectively. They may be regarded as representing, in a sense, two modes of behavior of the diffracted field. One of the modes dominates when $N \gg 1$, the other when $N \ll 1$. We may expect, roughly speaking, that in general there will be a mixture of these two modes of behavior and that both will significantly influence the structure of the field when the angles ϕ , given by Eqs. (2.5) and (2.6), are of the same order of magnitude, i.e., when

$$a/f \approx \lambda/a. \tag{2.7}$$

If we recall the definition (2.2) of the Fresnel number N we see that Eq. (2.7) is equivalent to the condition

$$N \approx 1. \tag{2.8}$$

Hence our rough qualitative arguments indicate that, as the Fresnel number N of a focusing system is gradually decreased from large values ($N \gg 1$) to small values ($N \ll 1$), we may expect to find discrepancies with predictions of the classic theory of focusing when N attains values of the order of unity and that the classic theory will not be applicable when N has values much smaller than unity.

We will see shortly that these rough qualitative predictions are in agreement with results of detailed calculations.

3. EXPRESSIONS FOR THE FIELD IN THE FOCAL REGION

We again consider diffraction of a uniform, monochromatic, converging spherical wave at a circular aperture of radius a in an opaque screen (Fig. 2). Let

$$V^{(i)}(Q, t) = A \frac{e^{-ikf}}{f} e^{-i\omega t} \tag{3.1}$$

be the field distribution of the incident wave at a typical point Q on the wave front through the center O of the aperture. In Eq. (3.1) A is a constant and t denotes the time. With the assumptions expressed by Eqs. (2.1), the diffracted field at a point P is, according to the Huygens-Fresnel principle, given by [with the time-periodic factor $\exp(-i\omega t)$ omitted]

$$U_N(P) = -\frac{i}{\lambda} \frac{Ae^{-ikf}}{f} \iint_W \frac{e^{iks}}{s} dS, \tag{3.2}$$

where s denotes the distance QP and the integration extends over the wave front W filling the aperture. We attached the suffix N to the symbol $U(P)$ for the diffracted field to stress its dependence on the Fresnel number N , defined by Eq. (2.2).

Let (z, r, ψ) be the cylindrical coordinates of the field point P , referred to axes with origin at the geometrical focus F and with the z direction along OF (see Fig. 2). It is convenient to introduce the dimensionless variables¹⁴ u_N and v_N , which, together with ψ , specify the position of the field point P :

$$u_N = 2\pi N \frac{z/f}{1 + z/f}, \tag{3.3a}$$

$$v_N = 2\pi N \frac{r/a}{1 + z/f}. \tag{3.3b}$$

Because the field is rotationally symmetric with respect to the z axis, the diffracted field will not depend on the azimuthal angle ψ .

It was recently shown^{1,2,15} that, throughout a neighborhood of the geometrical focus, the diffracted field $U_N(P)$ may be expressed in the form

$$U_N(P) = B_N(u_N) \exp[i\Phi_N(u_N, v_N)] \times \int_0^1 J_0(v_N \rho) \exp(-iU_N \rho^2/2) \rho d\rho, \tag{3.4}$$

where J_0 is the Bessel function of the first kind and zero order,

$$B_N(u_N) = -\frac{2\pi i}{\lambda} \left(\frac{a}{f}\right)^2 \left(1 - \frac{u_N}{2\pi N}\right) A, \tag{3.5a}$$

and

$$\Phi_N(u_N, v_N) = \frac{1}{1 - u_N/2\pi N} \left[\left(\frac{f}{a}\right)^2 u_N + \frac{1}{4\pi N} v_N^2 \right]. \tag{3.5b}$$

The range of validity of expression (3.4) is discussed in Ref. 15.

Before proceeding further we note that the expression on the right-hand side of Eq. (3.4) is of the same mathematical form as the expression obtained by Lommel in his classic paper on the structure of the focal region,¹⁶ viz.,

$$U(P) = B \exp[i\Phi(u)] \int_0^1 J_0(v\rho) \exp(-i u \rho^2/2) \rho d\rho, \tag{3.6}$$

where

$$u = \frac{2\pi}{\lambda} \left(\frac{a}{f}\right)^2 z, \tag{3.7a}$$

$$v = \frac{2\pi}{\lambda} \left(\frac{a}{f}\right) r, \tag{3.7b}$$

$$B = -\frac{2\pi i}{\lambda} \left(\frac{a}{f}\right)^2 A, \tag{3.8a}$$

$$\Phi(u) = \left(\frac{f}{a}\right)^2 u. \tag{3.8b}$$

Now the following relations between the quantities u_N, v_N, B_N , and Φ_N of the present theory and the quantities u, v, B , and Φ of Lommel's theory can readily be deduced:

$$u_N = \frac{u}{1 + u/2\pi N}, \tag{3.9a}$$

$$v_N = \frac{v}{1 + u/2\pi N}, \tag{3.9b}$$

$$B_N(u_N) = \left(1 - \frac{u_N}{2\pi N}\right) B, \tag{3.10a}$$

$$\Phi_N(u_N, v_N) = \frac{1}{1 - u_N/2\pi N} \left[\Phi(u_N) + \frac{1}{4\pi N} v_N^2 \right]. \tag{3.10b}$$

It follows at once from Eqs. (3.9) and (3.10) that, as $N \rightarrow \infty$, with u and v being kept fixed,¹⁷

$$u_N \sim u, \quad v_N \sim v, \quad (3.11)$$

$$B_N(u_N) \sim B, \quad \Phi_N(u_N, v_N) \sim \Phi(u), \quad (3.12)$$

and the expression on the right-hand side of Eq. (3.4) then reduces to Lommel's expression (3.6), i.e., as $N \rightarrow \infty$,

$$U_N(P) \sim U(P), \quad (3.13)$$

where the symbol \sim indicates limit in the asymptotic sense. Thus we conclude that Lommel's solution represents the asymptotic limit of our solution for focusing systems of large Fresnel numbers.

Since the mathematical structure of the integral in expression (3.4) and in Lommel's expression (3.6) is the same, we may develop the integral on the right-hand side of Eq. (3.4) in the same type of series as that first employed by Lommel in analyzing consequences of his solution.¹⁶ We then obtain for the field $U_N(P)$ in the focal region the following expression:

$$U_N(P) = (1/2)B_N(u_N)\{\exp[i\Phi_N(u_N, v_N)]\} \\ \times [C(u_N, v_N) - iS(u_N, v_N)], \quad (3.14)$$

where

$$C_N(u_N, v_N) = \frac{\cos(u_N/2)}{(u_N/2)} U_1(u_N, v_N) \\ + \frac{\sin(u_N/2)}{(u_N/2)} U_2(u_N, v_N), \quad (3.15a)$$

$$S_N(u_N, v_N) = \frac{\sin(u_N/2)}{(u_N/2)} U_1(u_N, v_N) \\ - \frac{\cos(u_N/2)}{(u_N/2)} U_2(u_N, v_N), \quad (3.15b)$$

and U_1 and U_2 are two of the Lommel functions:

$$U_n(u_N, v_N) = \sum_{s=0}^{\infty} (-1)^s \left(\frac{u_N}{v_N}\right)^{n+2s} J_{n+2s}(v_N), \quad (3.16)$$

$J_m(v_N)$ being a Bessel function of the first kind and of order m .

Although the formulas are valid in the neighborhood of the geometrical focus F, they are convenient for computations only when $|u_N/v_N| < 1$, i.e., when $|z|/r < f/a$; this is the region of the geometrical shadow. When $|u_N/v_N| > 1$, i.e., when $|z|/r > f/a$, the field point P is in the geometrically illuminated region, and it is then more convenient to use the following alternative expressions for the functions $C(u_N, v_N)$ and $S(u_N, v_N)$ that appear in Eq. (3.14):

$$C(u_N, v_N) = \frac{2}{u_N} \sin \frac{v_N^2}{2u_N} + \frac{\sin(u_N/2)}{(u_N/2)} V_0(u_N, v_N) \\ - \frac{\cos(u_N/2)}{(u_N/2)} V_1(u_N, v_N), \quad (3.17a)$$

$$S(u_N, v_N) = \frac{2}{u_N} \cos \frac{v_N^2}{2u_N} - \frac{\cos(u_N/2)}{(u_N/2)} V_0(u_N, v_N) \\ - \frac{\sin(u_N/2)}{(u_N/2)} V_1(u_N, v_N), \quad (3.17b)$$

where V_0 and V_1 are two of the Lommel functions¹⁸:

$$V_n(u_N, v_N) = \sum_{s=0}^{\infty} (-1)^s \left(\frac{v_N}{u_N}\right)^{n+2s} J_{n+2s}(v_N). \quad (3.18)$$

4. THE STRUCTURE OF THE FOCAL REGION

The close formal similarity of expression (3.4) of the present theory and expression (3.6) of Lommel's theory makes it possible to draw at once some interesting conclusions.

We note that the variables u and v of Lommel's theory depend linearly on the two cylindrical coordinates (z, r) of the field point [Eqs. (3.7)]. On the other hand, the variables u_N and v_N of the present theory are nonlinear functions of z and r [Eqs. (3.3)]. In the linear regime, the effect of changing any of the parameters λ , a , or f is equivalent to *scaling*; i.e., if one of the parameters is changed, the field distribution $U(P)$ is modified only in the simple manner of becoming spatially magnified (with magnification ≥ 1) in the longitudinal (z) direction or the transverse (r) direction or both. This is, however, not so with the field distribution $U_N(P)$. The nonlinear relationship (3.3) between the cylindrical coordinates (z, r) and the dimensionless variables (u_N, v_N) implies a more drastic change. This is illustrated in Fig. 3. In the top left-hand diagram of Fig. 3 a rectangular mesh formed by the z - and r -coordinate lines has been superimposed upon the well-known isophote diagram (diagram formed by contours of equal intensity $I = |U|^2$) in the region of focus, calculated from Lommel's theory. Because of the linearity, the mesh also represents lines of constant values of u (vertical lines) and of constant values of v (horizontal lines). The other three diagrams in Fig. 3 show the corresponding lines $u_N = \text{constant}$, $v_N = \text{constant}$ in focusing systems of Fresnel numbers $N = 100$, $N = 10$, and $N = 1$, respectively, plotted in the z, r plane.

In Fig. 4 the isophotes $I_N \equiv |U_N|^2 = \text{constant}$, normalized to unity at the geometrical focus, are shown for focusing systems of selected Fresnel numbers. We see that with decreasing N , the point of maximum intensity moves from its coincidence with the geometrical focus toward the center of the aperture and that the isophote diagrams approach more and more those of a complicated radiation field generated by a source in the vicinity of the center of the aperture. These quantitative results are in agreement with the conclusions that we reached by rough qualitative arguments in Section 2.

Finally we note that, when the point $P(z, r, \psi)$ is located in the geometrical focal plane $z = 0$, one has, according to Eqs. (3.3), (3.7), and (2.2),

$$u_N = u = 0, \quad v_N = v, \quad (4.1)$$

$$B_N(0) = B, \quad \Phi_N(0, v_N) = \Phi(0) + \frac{1}{4\pi N} v^2. \quad (4.2)$$

On substituting from Eqs. (4.1) and (4.2) into Eq. (3.4) and on comparing the resulting expression with Eq. (3.6) we see that

$$U_N(P)|_{z=0} = U(P)|_{z=0} \exp\left(\frac{iv^2}{4\pi N}\right). \quad (4.3)$$

This formula implies that, irrespective of the value of the Fresnel number N , the amplitude of the field at any point in the geometrical focal plane has the same value as given by Lommel's theory and that the corresponding phase factor differs from the phase factor of Lommel's theory by the amount $v^2/4\pi N$. It is an immediate consequence of the first of these two results that, *irrespective of the Fresnel number N of the system, the intensity distribution in the focal plane*

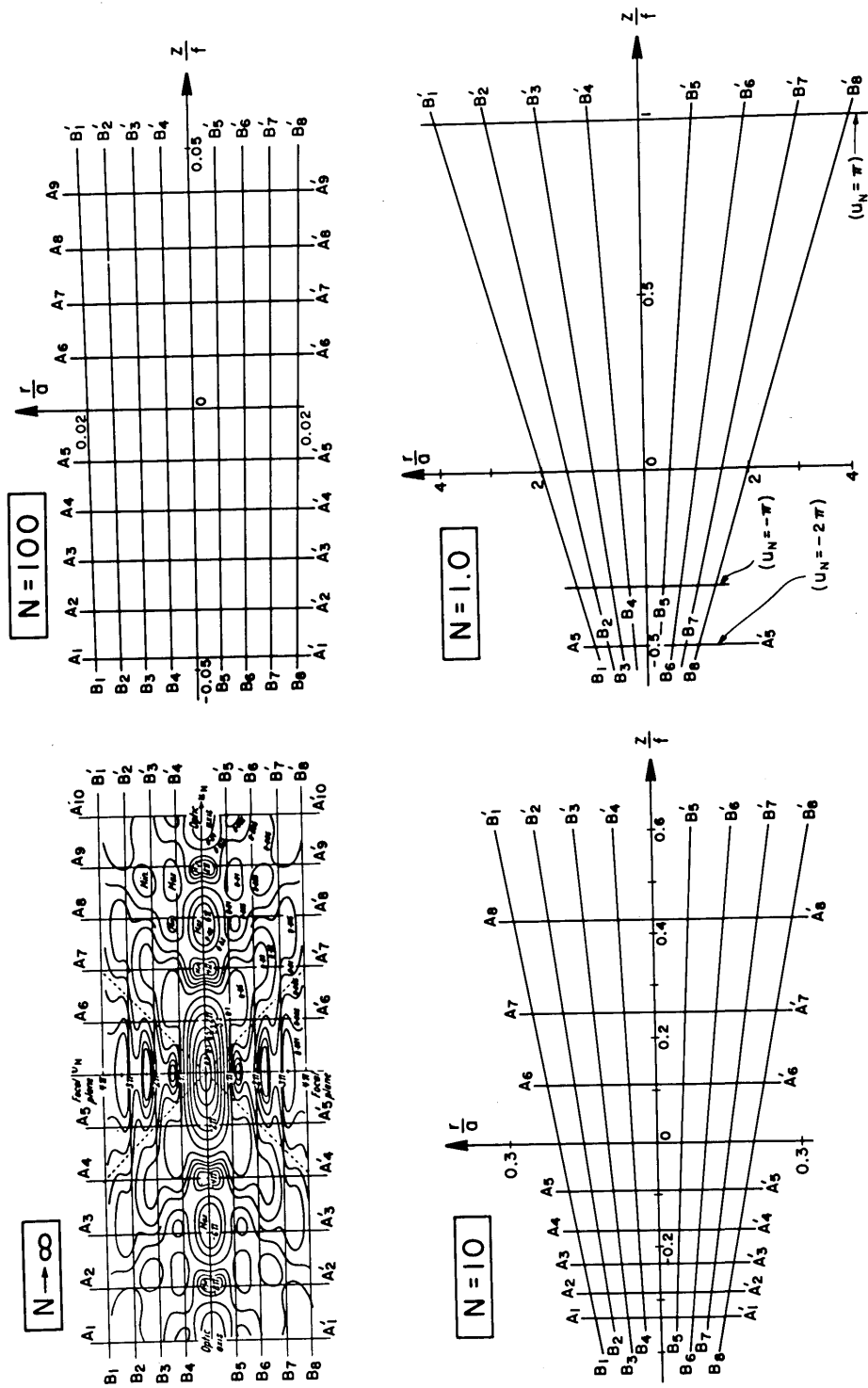
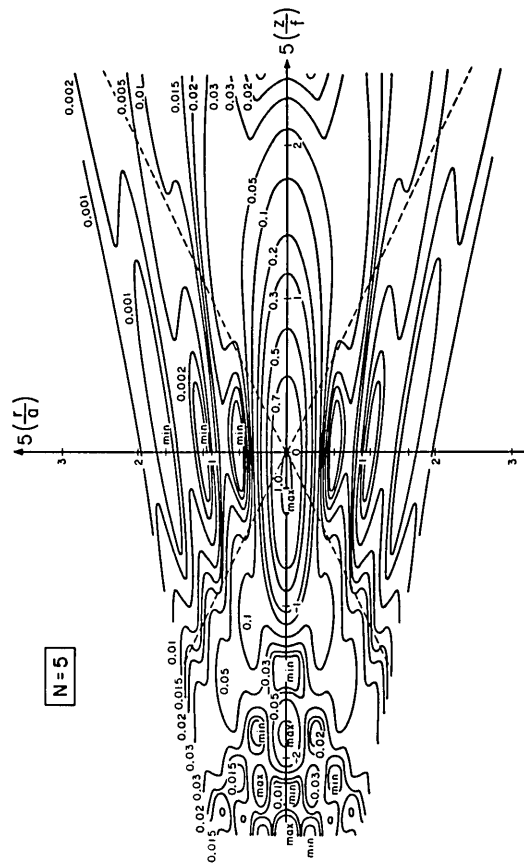
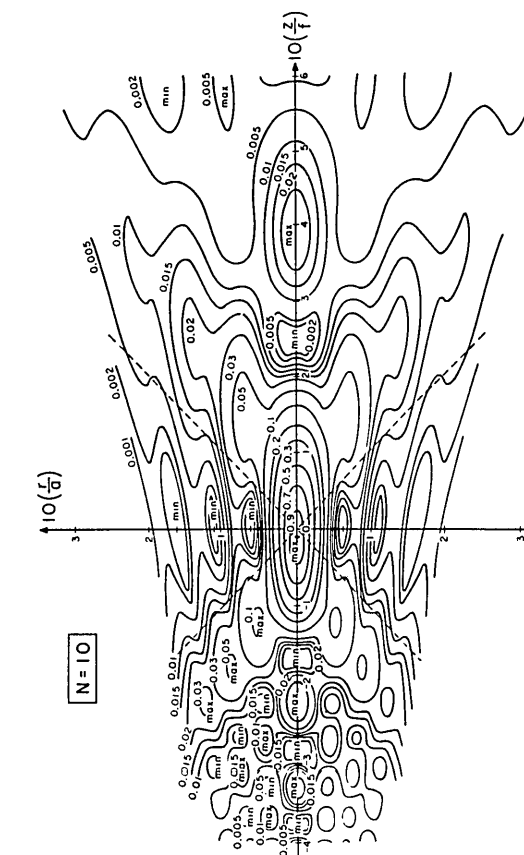
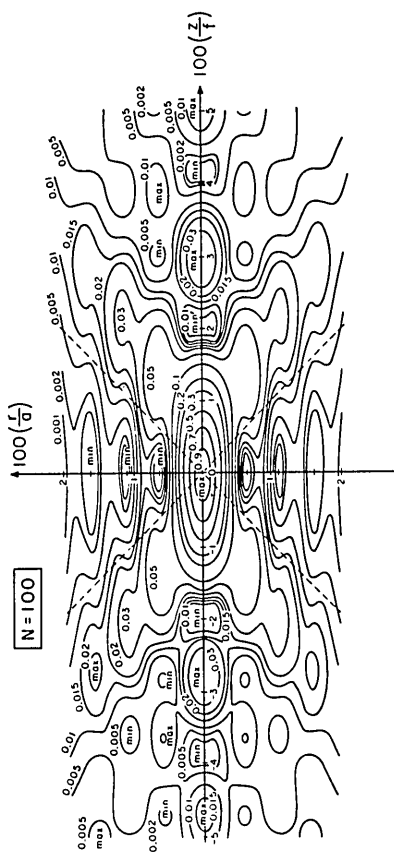


Fig. 3. Illustrating consequences of the nonlinear transformation [Eqs. (3.3)]. Some of the coordinate lines $u_N = \text{constant}$ (vertical) and $v_N = \text{constant}$ (horizontal or inclined to the z axis), plotted in the z, r plane, for systems of different Fresnel numbers. Effects of the transformation on the isophote diagrams that are included in the top left-hand figure ($N \rightarrow \infty$) are shown in Fig. 4.



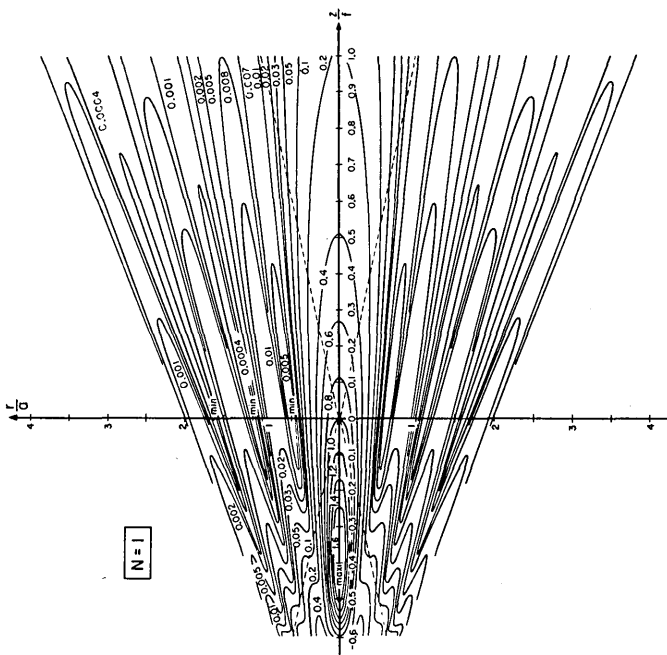
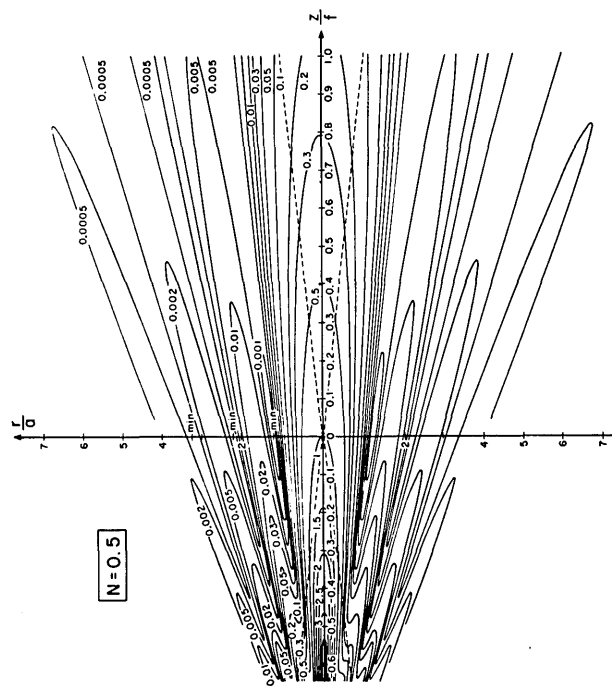
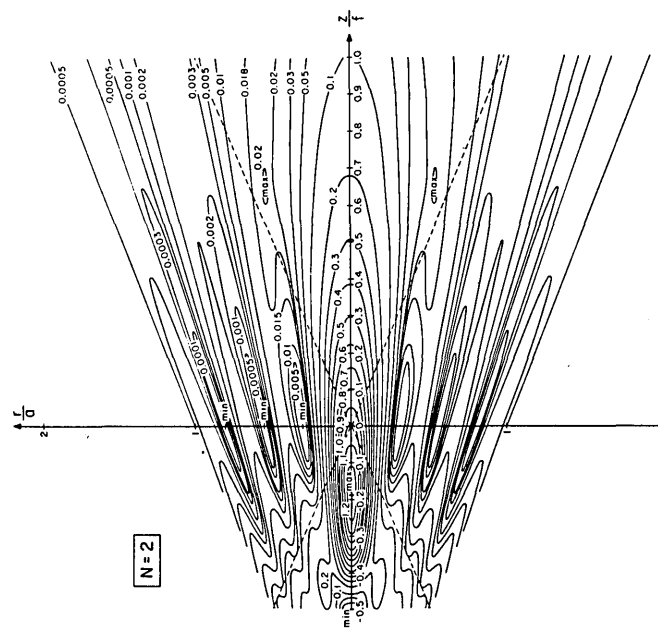


Fig. 4. Isophotes (contours of the intensity) in a meridional plane in the neighborhood of the geometrical focus of a uniform, converging, monochromatic spherical wave diffracted at a circular aperture in an opaque screen, in systems of different Fresnel number N . The intensity is normalized to unity at the geometrical focus. The dotted lines represent the boundary of the geometrical shadow. The first isophote diagram (for $N = 100$) is essentially the same as that obtained from Lommel's classic theory (Fig. 2 of Ref. 19 and Fig. 8.41 of Ref. 7).



is given by the Airy-pattern formula [Ref. 7, p. 441, Eqs. (25) and (22)]

$$I_N(P)|_{z=0} \equiv |U_N(P)|_{z=0}^2 = \left[\frac{2J_1(v)}{v} \right]^2 I_0, \quad (4.4)$$

where

$$I_0 = \left(\frac{\pi a^2 |A|}{\lambda f^2} \right)^2 \quad (4.5)$$

is the intensity at the geometrical focus.

ACKNOWLEDGMENTS

The main part of this research was carried out while E. Wolf was on leave of absence at Schlumberger-Doll Research in Ridgefield, Connecticut. Preliminary results were presented on October 22, 1982, at the Annual Meeting of the Optical Society of America in Tucson, Arizona [J. Opt. Soc. Am. 72, 1818 (A) (1982)].

* E. Wolf is also with the Institute of Optics, University of Rochester.

REFERENCES

1. J. H. Erkkila and M. E. Rogers, "Diffracted fields in the focal volume of a converging wave," J. Opt. Soc. Am. 71, 904-905 (1981).
2. J. J. Stamnes and B. Spjelkavik, "Focusing at small angular apertures in the Debye and Kirchhoff approximations," Opt. Commun. 40, 81-85 (1981).
3. Y. Li and E. Wolf, "Focal shifts in diffracted converging spherical waves," Opt. Commun. 39, 211-215 (1981).
4. Y. Li, "Dependence of the focal shift on Fresnel number and f number," J. Opt. Soc. Am. 72, 770-774 (1982).
5. M. P. Givens, "Focal shifts in diffracted converging spherical waves," Opt. Commun. 41, 145-148 (1982).
6. Similar results were also found in connection with focused Gaussian beams. For a discussion of this subject and for pertinent references, see Y. Li and E. Wolf, "Focal shift in focused truncated Gaussian beams," Opt. Commun. 42, 151-156 (1982).
7. M. Born and E. Wolf, *Principles of Optics*, 6th ed. (Pergamon, Oxford, 1980).
8. E. Wolf and Y. Li, "Conditions for the validity of the Debye integral representation of focused fields," Opt. Commun. 39, 205-210 (1981).
9. G. C. Sherman and W. C. Chew, "Aperture and far-field distributions expressed by the Debye integral representation of focused fields," J. Opt. Soc. Am. 72, 1076-1083 (1982).
10. In this connection it seems worthwhile to mention that none of the currently available theories of diffraction at an aperture in an *opaque* screen appears to be adequate to predict the structure of the diffracted field in the near zone of the aperture; nor are any experimental results available to elucidate this question.
11. For the same reason as indicated in Ref. 10, this assumption may not be valid when the maximum reaches the immediate vicinity of the aperture.
12. A somewhat paradoxical aspect of this situation should be noted. As the Fresnel number N becomes smaller, with a and λ being kept fixed, the distance f increases, i.e., the geometrical focus F moves farther away from the aperture; however, the point of maximum intensity moves in the opposite direction, i.e., closer to the aperture.
13. Y. Li and H. Platzer, "An experimental investigation of diffraction patterns in low-Fresnel-number focusing systems," Opt. Acta 30, 1621-1643 (1983).
14. The choice of these variables is suggested by detailed calculations given, for example, in Ref. 15. As will be seen shortly, they reduce to the usual dimensionless variables of the classic theory of Lommel in the limit of large Fresnel numbers.
15. Y. Li, "Encircled energy for systems of different Fresnel numbers," Optik 64, 207-218 (1983).
16. E. Lommel, "Die Beugungerscheinungen einer kreisrunden Oeffnung und eines kreisrunden Schirmchens," Abh. Bayer. Akad. Math. Naturwiss. Kl. 15, 233-328, (1885). The main part of Lommel's analysis is presented (in English) in Sec. 8.8 of Ref. 7.
17. This limiting procedure is equivalent to keeping the field point $P(z, r, \psi)$ as well as the wavelength λ and the angular semiaperture a/f fixed and letting $f \rightarrow \infty$. That the resulting expression is to be interpreted as an asymptotic rather than an ordinary approximation is strongly suggested by recent discussions of the Debye representation of focused fields,^{8,9} which is known to be equivalent to Lommel's representation under the usual circumstances.
18. When the suffix n is a negative integer, Lommel's original definition of the function V_n differs from that given by Eq. (3.18) by a factor $(-1)^n$.
19. E. H. Linfoot and E. Wolf, "Phase distribution near focus in an aberration-free diffraction image," Proc. Phys. Soc. London Sect. B 69, 823-832 (1956).