Vector Diffraction Theory of Focusing in Systems of High Numerical Aperture

The diffraction image formed by a perfect lens is well known as the "Airy Disc". Many engineers may not realize that this familiar concept requires modification when large numerical apertures and polarization are involved. The following technical note indicates some of the effects for an extreme F/0.52 imaging situation, as might be found in an optical data storage system. This article also discusses some of the techniques needed for computing diffraction effects in such extreme, but more frequently occurring, conditions.

The classical theory of diffraction, which maintains that the distribution of light at the focal plane of a lens is the Fourier transform of the distribution at its entrance pupil, is applicable to lenses of moderate numerical aperture (NA). The incident beam, of course, must be monochromatic and coherent, but its polarization state is irrelevant since the classical theory is a scalar theory. If the incident beam happens to be a plane wave and the lens free from aberrations, then the focused spot will have the well-known Airy pattern. When the incident beam is Gaussian, the focused spot will also be Gaussian, since this particular profile is preserved under Fourier transformation. In general, arbitrary distributions of the incident beam (with or without aberrations and defocus) can be transformed numerically to yield the distribution in the vicinity of focus using the fast Fourier transform (FFT) algorithm.

There are two basic reasons for the applicability of the classical scalar theory to systems of moderate NA. The first is that bending of the rays by the focusing element(s) is fairly small, causing the electromagnetic field vectors (**E** and **B**) before and after the lens to have more or less the same orientations. A scalar amplitude assigned to each point on the emergent wavefront from a low/moderate-NA system is therefore sufficient to describe its electromagnetic state, whereas in the high-NA regime one can no longer ignore the vectorial nature of light. The second reason for the success of the classical theory within its proper limits is that a certain integral (that which represents the decomposition of a converging wavefront into its plane-wave constituents) submits to evaluation by the method of stationary phase approximation. The remaining integral (that which represents superposition of plane-waves arriving at the focal plane) is then calculated with the aid of Fourier transformation. When the stationary phase technique fails, so does the classical theory, as is evidenced, for instance, in systems of very low numerical aperture: The well-known focal shift phenomenon[1] is but one manifestation of the failure of the stationary phase approximation in low-NA systems.

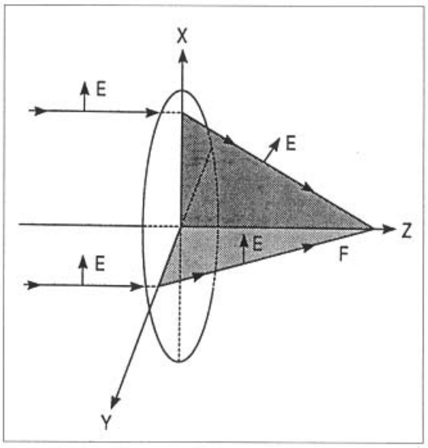

FIGURE 1. FOCUSING OF LINEARLY POLARIZED LIGHT BY A HIGH-NA LENS CAUSES BENDING OF POLARIZATION VECTORS. THE AMOUNT AND DIRECTION OF BENDING DEPEND ON THE COORDINATES OF THE RAY.

In the stationary phase approximation, the plane-wave spectrum of the convergent beam at the exit pupil coincides with the light amplitude distribution at that pupil, thus enabling each geometric-optical ray to represent one plane-wave of the spectrum, namely, that which propagates in the direction of the ray.[2] This correspondence between rays and plane-waves, which is an important feature of many diffraction problems, is therefore understood to be a direct consequence of the stationary phase approximation. Now, let θ be the angle between a converging ray in the image space and the optical axis at the focal point. Since the projection of the wave vector **k** on the exit pupil has length $k \sin\theta$, whereas the intersection of the ray with the pupil occurs at a radius of $r = f \tan\theta$, in order to convert from light amplitude distribution to the corresponding plane-wave spectrum, one must compress the distribution function at the exit pupil. Aside from a trivial scaling of the aperture's radius by the focal length f, the radial compression must assign to $r = \sin\theta$ the value of the function at $r = \tan\theta$, followed by proper normalization to preserve the integrated intensity. The compressed distribution is therefore confined to a disk of radius NA = $\sin\theta_{max}$, where θ_{max} is the angle subtended by the rim of the exit pupil at the focal point. It must be emphasized that this scaling, compression, and normalization procedure is not merely justified on heuristic grounds, but is a rigorous consequence of the stationary phase approximation itself.[3]

For lenses of low/moderate numerical aperture (say, NA $\leq$0.3) the difference between $\sin\theta$ and $\tan\theta$ is negligible, and the effects of compression can be ignored. At the exit pupil,

continued on page 74

Masud Mansuripur, *Optical Sciences Center, University of Arizona, Tucson, Ariz.*

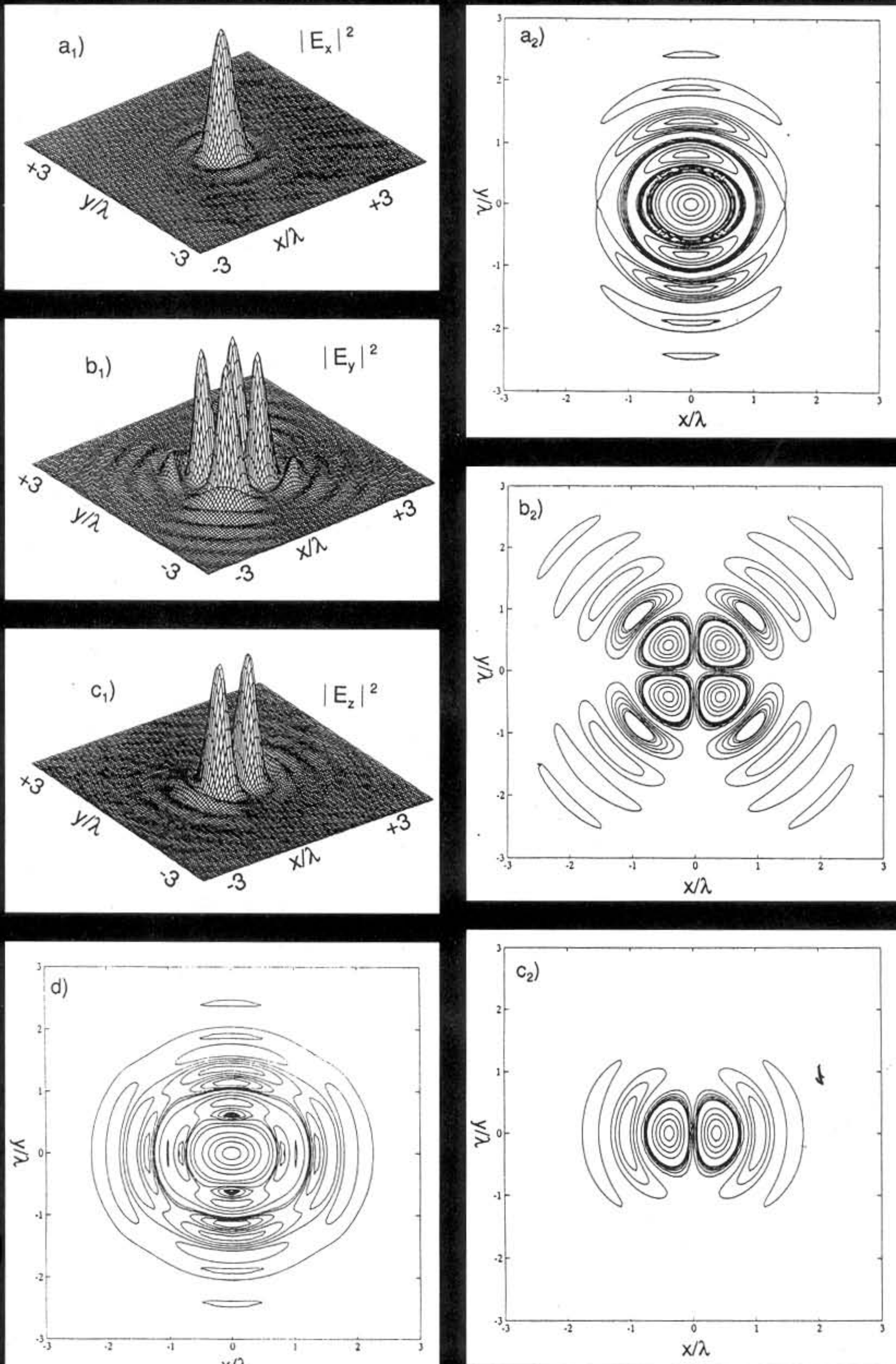

FIGURE 2. INTENSITY PROFILES OF THE VARIOUS COMPONENTS OF POLARIZATION AT THE FOCAL PLANE OF A LENS (NA = 0.966, f = 3000λ). FOR BEST VIEWING OF THE 3D PLOTS, THE VERTICAL SCALE IS CHOSEN DIFFERENTLY IN THE THREE CASES: THE PEAKS IN A_1, B_1 AND C_1 (WHICH CORRESPOND, RESPECTIVELY, TO THE X, Y, AND Z COMPONENTS OF POLARIZATION) ARE IN THE RATIO OF 2.12 : 0.013 : 0.35. THE CORRESPONDING CONTOUR PLOTS ARE SHOWN IN A_2, B_2, AND C_2. THE CONTOUR PLOT IN D REPRESENTS THE SUM OF THE THREE PROFILES, I.E., THE TOTAL E-FIELD ENERGY DENSITY DISTRIBUTION IN THE FOCAL PLANE.

Engineering
continued from page 72

the plane-wave spectrum of these lenses is usually the same as the incident distribution at the entrance pupil, modified only by the presence of aberrations. For lenses of high numerical aperture, on the other hand, it is necessary to obtain first the exit pupil distribution[4] from the knowledge of lens characteristics and the entrance pupil distribution, before proceeding to the compression operation. Noteworthy in this respect is the aplanatic lens that, by virtue of satisfying Abbe's sine condition, guarantees that the compressed exit pupil function is identical with the distribution at the entrance pupil.

To account for the polarization effects at high numerical aperture, one usually ignores transmission losses at the various surfaces of a lens, assuming that a ray goes through the system unattenuated but with its polarization vector bent in accordance with the known laws of refraction.[2-5] (The assumption of losslessness is not necessary here, but it simplifies the problem by enabling the polarization state of individual rays at the exit pupil to be determined solely based on their coordinates, without requiring any knowledge of the lens structure.)

For a linearly-polarized incident beam, Figure 1 shows the bending of the **E**-vector at two azimuthal positions. The ray at the top of the lens contributes both an X and a Z component to the distribution in the image space, whereas the ray in the YZ plane contributes only an X component. By the same token, rays intermediate between those shown here will contribute to the polarization along all three axes. In the stationary phase approximation, each ray is associated with a single plane-wave, the three polarization components of which may be treated independently. Therefore, for each of the X,Y,Z components, a single superposition integral (*i.e.*, Fourier transform) yields the sought-after distribution in the focal plane.[3]

The technique described above is quite general and can be applied to arbitrary incident distributions with arbitrary polarization states, and can take into consideration various lens aberrations, including substantial amounts of defocus. Computed results for an aberration-free, aplanatic lens with $f = 3000\lambda$ and NA = $sin(75°)$ = 0.966 are shown in Figure 2. The assumed geometry in these calculations is that depicted in Figure 1, where the incident beam is a uniform plane-wave with linear polarization along the X axis. Frames (a_1), (b_1), and (c_1) in Figure 2 are intensity plots for the X, Y, and Z components of polarization in the focal plane; their peak intensities are in the ratio of 2.12 : 0.013 : 0.35. The four-fold symmetry of the Y component and two-fold symmetry of the Z component are consistent with one's expectations based on ray bending arguments. The corresponding contour plots are shown in frames (a_2), (b_2), and (c_2). The contour plot of the total electric field energy density $E_x^2 + E_y^2 + E_z^2$ in Figure 2*(d)* shows an elliptical profile with the major axis parallel to the direction of the incident polarization. (Richards and Wolf[5] obtained a similar plot using a somewhat different formulation of the diffraction problem.)

The computations reported here were performed on a VAX-station using a 512 × 512 square mesh and required no more than three minutes of CPU time.

References

1. V.N. Mahajan, "Axial irradiance and optimum focusing of laser beams," *Appl. Opt.* **22**, 1983, 3042-3053.
2. J.J. Stamnes, *Waves in Focal Regions*, Adam Hilger Publishing Co., Bristol, 1986.
3. M. Mansuripur, "Certain computational aspects of vector diffraction problems," *JOSA A*, **6**, 1989, 786-805.
4. H.H. Hopkins, "The Airy disk formula for systems of high relative aperture," *Proc. Phys. Soc.*, London, **55**, 1943, 116-128.
5. B. Richards and E. Wolf, "Elecromagnetic diffraction in optical systems: Structure of the image field in an aplanatic system," *Proc. Roy. Soc. Ser. A.* **253**, 1959, 358-379.