

Interaction of Light With Subwavelength Structures

Masud Mansuripur, Armis R. Zakharian and Jerome V. Moloney

When a light field interacts with structures that have complex geometric features comparable in size to the wavelength of the light, it is not permissible to invoke the assumptions of the classical diffraction theory, which simplify the problem and allow for approximate solutions. For such cases, direct numerical solutions of the governing equations are sought through approximating the continuous time and space derivatives by the appropriate difference operators. The Finite Difference Time Domain (FDTD) method discretizes Maxwell's equations by using a central difference operator in both the time and space variables.¹ The \mathbf{E} and \mathbf{B} fields are then represented by their discrete values on the spatial grid, and are advanced in time in steps of Δt . The numerical solution thus obtained to Maxwell's equations (in conjunction with the relevant constitutive relations) provides a highly reliable representation of the electromagnetic field distribution in the space-time region under consideration.

This article presents examples of application of the FDTD method to problems involving the interaction between a focused beam of light and certain subwavelength structures of practical interest. A few general remarks concerning the nature of the FDTD method appear in the next section. This is followed by a description of the simulated system and two examples in which comparison is possible between the FDTD method and an alternative method of calculation. We then present simulation results for the case of a focused beam interacting with small pits and apertures in a thin film supported by a transparent substrate.

The FDTD method

The spatial unit cell used in three-dimensional (3D) FDTD simulations is shown in Fig. 1. The components of the vector fields \mathbf{E} and \mathbf{B} are located at different positions with respect to the cell center, so that every component of the electric field is surrounded by four circulating components of the magnetic field, and vice versa. Such a staggered mesh is motivated by the integral form of Maxwell's curl equations. The contour integrals of \mathbf{E} (\mathbf{B}) along the edges of the cell in Faraday's law (Ampere's law) circulate around the corresponding magnetic (electric) field component at the center of the cell face.

In 3D simulations at least six field components must be stored and updated at each grid point, which leads to considerable memory and CPU requirements for FDTD simulations. Fortunately, the time update of any field component involves only nearby fields located one or two cells away on the grid. This kind of locality in the physical space translates into computer memory access locality and allows for efficient implementation of the FDTD algorithm on many types of shared and distributed memory parallel platforms. Low-reflection absorbing boundary conditions that terminate the computational domain by a

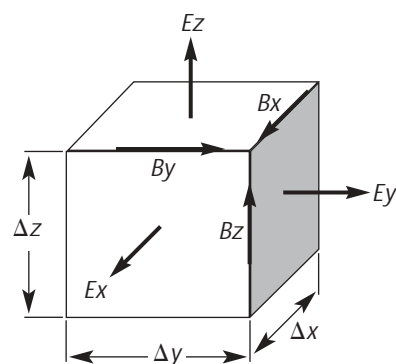


Figure 1. The unit cell of the FDTD mesh has dimensions $\Delta x \times \Delta y \times \Delta z$. The various components of the \mathbf{E} and \mathbf{B} fields are assigned to different locations on the unit cell. The staggered field components are shifted by a half-pixel in various directions.

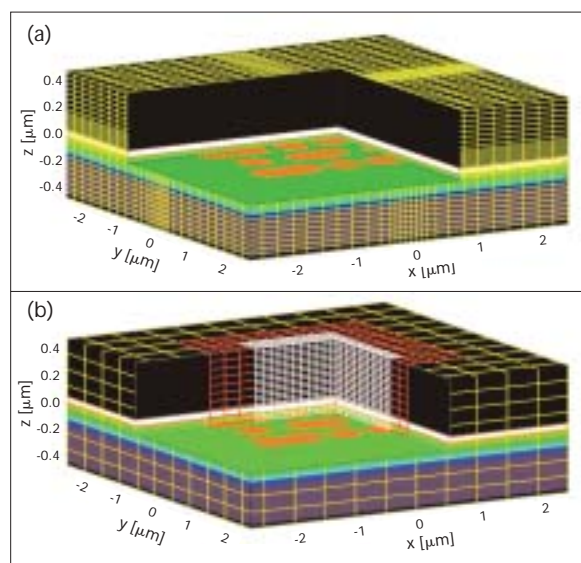


Figure 2. 3D computational domain for simulating the interaction between a focused beam of light and various marks (i.e., bumps or pits) on the surface of a multilayer data storage medium. (a) Non-uniform conformal grid; the grid-line density is higher near the center, where the focused beam and the multilayer stack are located. (b) Nested rectangular cells forming non-conformal hierarchical grid.

Perfectly Matched Layer (PML) allow the simulation of physical problems with open boundaries.²

Since the FDTD algorithm solves Maxwell's equations in the time domain, calculation for a broad range of frequencies is possible in a single simulation using a time-pulsed excitation. Other advantages include the possibility of modeling dispersive and nonlinear materials. An important property of the FDTD method is that it introduces no additional dissipation into the physical problem due to numerical discretization, and hence energy is conserved. However, the finite difference method contributes to a dispersion error. In the commonly used second order accurate implementation of FDTD, this error diminishes

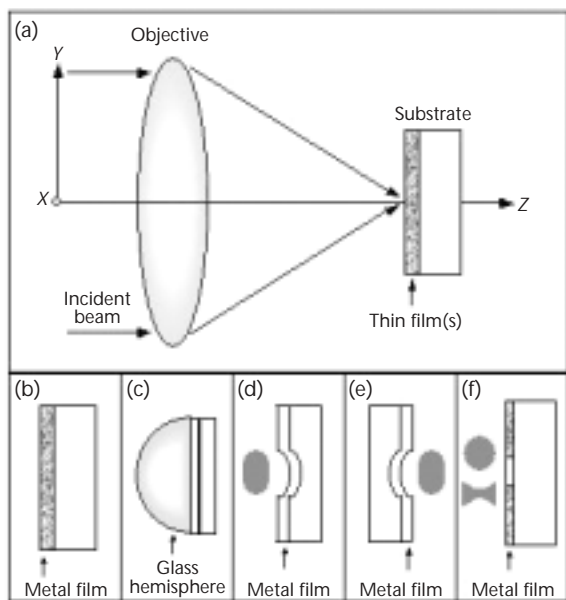


Figure 3. (a) Diagram of the simulated system, in which an aberration-free objective lens brings a coherent, monochromatic beam of light to focus. The bottom row shows the various samples used in the simulations. In (b), (d) and (e) a 50-nm-thick metal film is coated over a transparent substrate. In (c) the bilayer at the bottom of the glass hemispherical lens consists of quarter-wave-thick dielectric films. The spherocylindrical pits in (d) and (e) are 500 nm long, 300 nm wide, and 100 nm deep. In (f) the aperture in the 20-nm-thick metal film is circular in one simulation and bowtie-shaped in another. The circular hole's diameter is 400 nm, while the bowtie aperture is 400 nm long, 300 nm wide on each side, and 60 nm wide at the neck.

with cell size h as $O(h^2)$. In practice, therefore, to keep the numerical dispersion errors under control, a grid with about 30 points per wavelength is desired. The rather large number of points and iterations thus required for accurate results may render solution impractical for a problem with large spatial and/or temporal domain.

In many cases it is desirable to retain the efficiency of the FDTD scheme on the rectangular grids, but achieve higher resolution only in those regions of the computational domain where it is needed. The nonuniform grids allow one to vary a cell size in each coordinate direction, keeping the grid structured and conformal as in Fig. 2(a). A more efficient approach is to employ a collection of nested rectangular cells that form a nonconformal hierarchical grid, as in Fig. 2(b). Each successive nested level has a higher resolution, e.g., by a factor of two, than the previous level, allowing smaller cell sizes to be "focused" in the regions of interest (e.g., subwavelength features, photonic crystal microcavity, etc.). Inside each rectangular region the standard FDTD algorithm is applied, while at the boundaries between the grids an update scheme and interpolation must be employed to keep the method stable and accurate. In FDTD the time step Δt is proportional to the cell size, and hence the smallest time step is required on the grids with the highest resolution. Each grid can be updated with its own time step, the grids with cell size $2h$ doing half as many iterations as grids with cell size h .

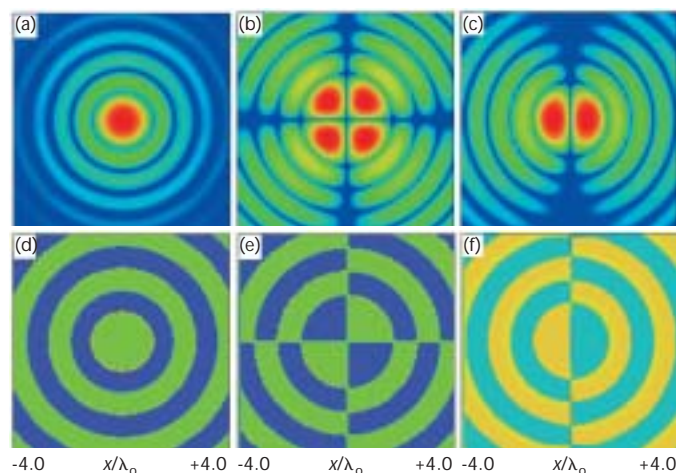


Figure 4. Plots of $\log_intensity_4$ (top) and phase (bottom) at the focal plane of the lens of Fig. 3(a). Left to right: E -field components along the X -, Y -, and Z -axes. At the entrance pupil the incident beam (wavelength = λ_0) is Gaussian with $1/e$ (amplitude) radius of $4000\lambda_0$, truncated at the lens aperture (radius = $3000\lambda_0$). The incident beam is linearly polarized along the X -axis, and its total optical power captured by the lens is unity, that is, $P_x = 1.0$, $P_y = 0.0$. (The power content of the E_z component at the focal plane is 8.3% of total power.)

The simulated system

The FDTD algorithm is quite powerful and can be applied to a wide variety of problems in electromagnetics. For demonstration purposes in this article, however, we confine our attention to a simple system involving the interaction between a focused beam of light and small (subwavelength) structures located in the focal region. Figure 3 shows a coherent, monochromatic beam of light (free-space wavelength = λ_0), brought to focus by an aberration-free objective lens (numerical aperture $NA = 0.6$, focal length $f = 5000\lambda_0$). The incident beam at the entrance pupil is linearly polarized along the X -axis, and the total optical power (i.e., integrated intensity) at the entrance pupil of the lens is set to unity. The sample typically consists of a thin film (or thin-film stack) coated on a transparent substrate; the various samples used in our simulations are depicted in Figs. 3(b-f). Detailed descriptions of these samples will be given in the context of the relevant simulations in the following sections. The focused spot may illuminate the sample directly, as in Figs. 3(b, d, f), or through a glass hemisphere (i.e., solid immersion lens) placed in contact with the sample, as in Fig. 3(c). When the hemispherical lens is present, the thin-film(s) may be coated directly on its flat facet, in which case the hemisphere acts as the sample substrate as well.

Figure 4 shows computed plots of intensity (top) and phase (bottom) at the focal plane of the lens depicted in Fig. 3(a).³ From left to right, these distributions represent the E -field components along the X -, Y -, and Z -axes. At the focal plane the peak intensities are in the ratio of $|E_x|^2 : |E_y|^2 : |E_z|^2 = 1000 : 0.4 : 45$. The various rings of the focused spot are phase-shifted by 180° relative to their adjacent neighbors, and the Z -component of the field is 90° out-of-phase relative to E_x and E_y . (In the remainder of this article we will omit the plots of E_z distribu-

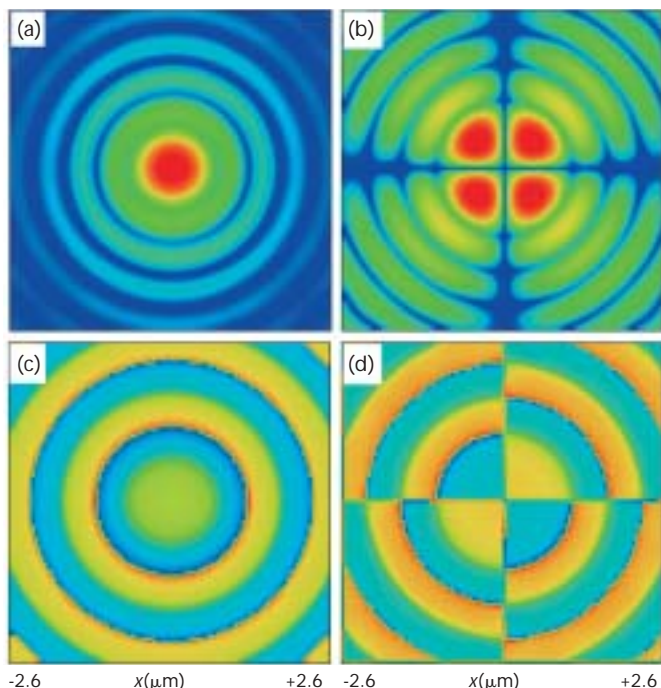


Figure 5. Plots of reflected \log_{10} intensity (top) and phase (bottom) from the metallic mirror depicted in Fig. 3(b) at $\lambda_0 = 650$ nm. The panels on the left-hand side correspond to E_x , while those on the right-hand side represent the E_y component of the reflected field. The front facet of the mirror is located a distance $z = 180$ nm beyond the focal plane.

tion, since E_z can always be computed from a knowledge of E_x and E_y distributions.)

The intensity and phase distributions in this article are plotted in an interval $x_{\min} \leq x \leq x_{\max}$ and $y_{\min} \leq y \leq y_{\max}$ of the XY -plane by assigning the color red to the maximum value of the function, blue to the minimum value, and the continuum of the white light spectrum to the values in between. The phase plots cover the range from -180° (blue) to $+180^\circ$ (red). The intensity distributions are first normalized by the peak value of the corresponding function, say, $I_{x\text{-peak}} = \text{Max}(|E_x|^2)$ within the displayed interval. The base 10 logarithm of the normalized function is then evaluated, and all pixel values below a certain level, say, $-\alpha$, are set equal to $-\alpha$. Displayed plots of \log_{10} intensity_α thus cover the range from $10^{-\alpha} I_{\text{peak}}$ (blue) to I_{peak} (red).

When the beam is focused through a hemispherical lens of refractive index n , as in Fig. 3(c), the same distribution as in Fig. 4 is found at the focal plane, but the spatial coordinates must shrink by a factor of n to account for the reduced wavelength ($\lambda = \lambda_0/n$) within the medium of the hemispherical lens. At the bottom of the hemisphere, therefore, the focused spot diameter is reduced by a factor of n compared to that shown in Fig. 4.

In some cases the beam must be focused onto the object of interest through a parallel plate cover glass or through the sample's substrate, as is the case for instance in Fig. 3(e). Under such circumstances, to obtain a focused spot free from spherical aberration, the objective lens must be designed for the specific thickness and refractive index of the substrate. Unlike focusing through a glass hemisphere, however, the focused spot inside a

cover plate (or flat substrate, as the case may be) has exactly the same dimensions as that obtained by focusing in air through an objective of the same NA. The reason is that, in passing from the air to the substrate through a flat interface, the effect of the reduced wavelength on the focused beam is exactly cancelled out by the reduced angle of the focused cone (Snell's law). The spot that illuminates the concave pit of Fig. 3(e) through the sample's substrate, therefore, has exactly the same size as that which directly illuminates the convex pit of Fig. 3(d).

Reflection from a metallic mirror

The mirror depicted in Fig. 3(b) is a 50-nm-thick metal film of complex refractive index $n + ik = 2.0 + i7.0$, coated over a transparent substrate of index $n = 1.5$. The large absorption coefficient k of the metal film ensures that the light does not reach the substrate; most of the incident light is therefore reflected, while a small fraction is absorbed in the metal. For the incident beam depicted in Fig. 4 at $\lambda_0 = 650$ nm, Fig. 5 shows computed plots of reflected intensity (top) and phase (bottom) obtained with the FDTD method. (The FDTD mesh size was $L_x = L_y = 12\lambda_0$, and the mirror's front facet was a distance $z = 180$ nm beyond the focal plane of the lens.)³ The integrated intensity of the reflected light over the XY -plane for the X - and Y -components of polarization may be defined as follows:

$$P_x = \iint |E_x|^2 dx dy, \quad P_y = \iint |E_y|^2 dx dy.$$

Using the FDTD method, we found $P_x = 0.85$, $P_y = 0.0016$ for the mirror of Fig. 3(b) illuminated with the focused spot of Fig. 4. To verify the accuracy of the FDTD method, we simulated the same system using an alternative method based on the superposition of plane-wave solutions to Maxwell's equations with matching boundary conditions at the various interfaces.³ The intensity and phase distributions thus obtained were visually indistinguishable from those shown in Fig. 5, and the corresponding integrated intensities were found to be $P_x = 0.86$, $P_y = 0.0018$. The slight differences between the two methods of computation reflect the cumulative effect of numerical errors inherent to the FDTD algorithm.

Similar simulations were performed for the sample of Fig. 3(b) illuminated through a glass hemisphere of index $n = 1.5$. (The FDTD mesh size in this case was $L_x = L_y = 8\lambda_0$, but the mirror's front facet remained at $z = 180$ nm beyond the focal plane of the lens.) The computed values of integrated intensity were $P_x = 0.78$, $P_y = 0.0019$. The corresponding quantities obtained with the alternative (and more accurate) method of plane-wave superposition were $P_x = 0.80$, $P_y = 0.0022$. Once again, comparison against a benchmark has shown the effect of small but cumulative numerical errors on the results of FDTD calculations.

Although the alternative method employed in the above examples is faster and more accurate than FDTD, it has the disadvantage of being restricted to geometries such as those in Figs. 3(b) and 3(c), where the sample consists of one or more homogeneous layers with flat surfaces/interfaces. As soon as inhomogeneities or nonuniformities are introduced, the computation method based on plane-wave superposition fails, and the

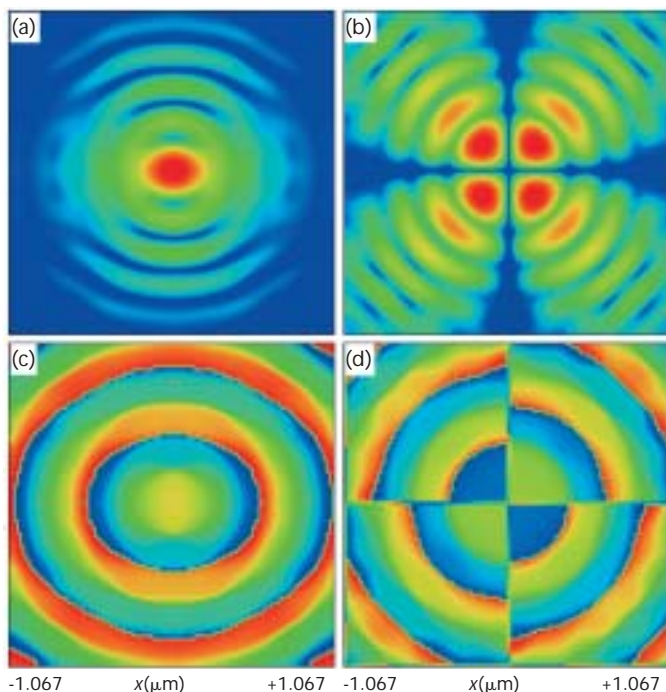


Figure 6. Plots of reflected log_intensity_3 (top) and phase (bottom) from the dielectric bilayer depicted in Fig. 3(c) at $\lambda_0 = 400$ nm. The panels on the left-hand side correspond to E_x , while those on the right-hand side represent the E_y component of the reflected field. The front facet of the stack is at $z_1 = 230$ nm beyond the focal plane.

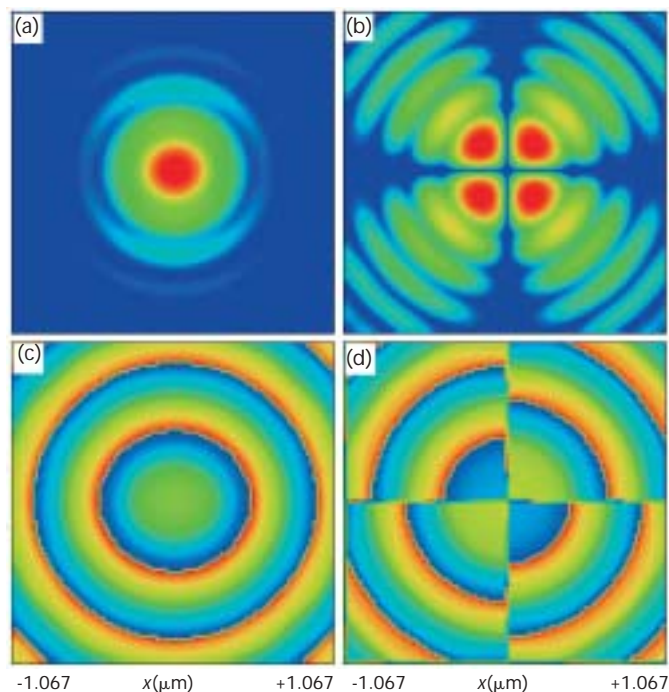


Figure 7. Same as Fig. 6 but for the transmitted beam. The distance from the rear of the stack to the plane where the transmitted beam is observed is $z_2 = 30$ nm.

FDTD method becomes an attractive (though costly) candidate for numerical solution of Maxwell's equations.

Reflection and transmission at a dielectric bilayer

The sample depicted in Fig. 3(c) consists of two quarter-wave-thick dielectric layers coated at the bottom of a glass hemisphere of index $n = 1.5$. The layer directly in contact with the hemisphere has $n = 2.0$, $d = 50$ nm, while the other layer has $n = 1.5$, $d = 67$ nm. Since the layers are homogeneous and the interfaces are flat, the method of computation based on plane-wave superposition may be used once again to check the accuracy of the FDTD simulations.

In our FDTD calculations of the bilayer stack of Fig. 3(c) the incident focused beam had $\lambda_0 = 400$ nm, the mesh size was $L_x = L_y = 8.08\lambda_0$, and the distance from the focal plane to the top of the stack was $z_1 = 230$ nm, while that from the bottom of the stack to the plane in which the transmitted beam is observed was $z_2 = 30$ nm. Figures 6 and 7 show computed plots of intensity and phase for the reflected and transmitted fields, respectively. The corresponding distributions obtained with the alternative method of plane-wave superposition were visually indistinguishable from those in Figs. 6 and 7. The integrated values of reflected intensity are $P_x = 0.022$ (0.019 with the alternative method) and $P_y = 0.0026$ (both methods). The corresponding quantities for the transmitted beam are $P_x = 0.97$ (1.01 with the alternative method) and $P_y = 0.01$ (both methods). Once again the FDTD method is seen to be adequate for these types of calculation, provided that a few percentage point

deviation from the exact solution (caused by discretization and numerical errors) is deemed acceptable.

Reflection from convex and concave pits

The substrate shown in cross-section in Fig. 3(d) is embossed with a sphero-cylindrical pit having a length of 500 nm along X , width of 300 nm along Y , and depth of 100 nm along Z (the profile of the pit in the XY -plane can also be seen in Fig. 2). The substrate's index is $n = 1.5$, and the metal film's thickness and complex index are $d = 50$ nm, $n + ik = 2.0 + 7.0i$. In our FDTD simulations the incident wavelength was $\lambda_0 = 650$ nm, the mesh size was $L_x = L_y = 12\lambda_0$, and the front facet of the metal film was at $z = 280$ nm beyond the focal plane.³ Figure 8 shows computed plots of reflected intensity and phase from a pit whose center has been displaced by $\Delta x = -250$ nm from the center of the focused spot. The integrated values of reflected intensity are $P_x = 0.82$, $P_y = 0.0025$.

The pit in the above example is similar to those embossed on the plastic substrate of a compact disk (CD) or a digital versatile disk (DVD). However, the focused laser beam in a CD or DVD player does not shine directly onto the pit; rather, the beam arrives through the plastic disk substrate as in Fig. 3(e). We simulated this case at $\lambda_0 = 650$ nm with an FDTD mesh of dimensions $L_x = L_y = 8\lambda_0$; the front facet of the metal film was at $z = 280$ nm beyond the focal plane. Figure 9 shows computed plots of reflected intensity and phase from the pit of Fig. 3(e) when the pit center is displaced by $\Delta x = -250$ nm from the center of the focused spot. The computed values of integrated intensity in this case are $P_x = 0.77$, $P_y = 0.022$. A comparison of $|E_x|^2$ distributions in Figs. 8 and 9 reveals that, whereas the

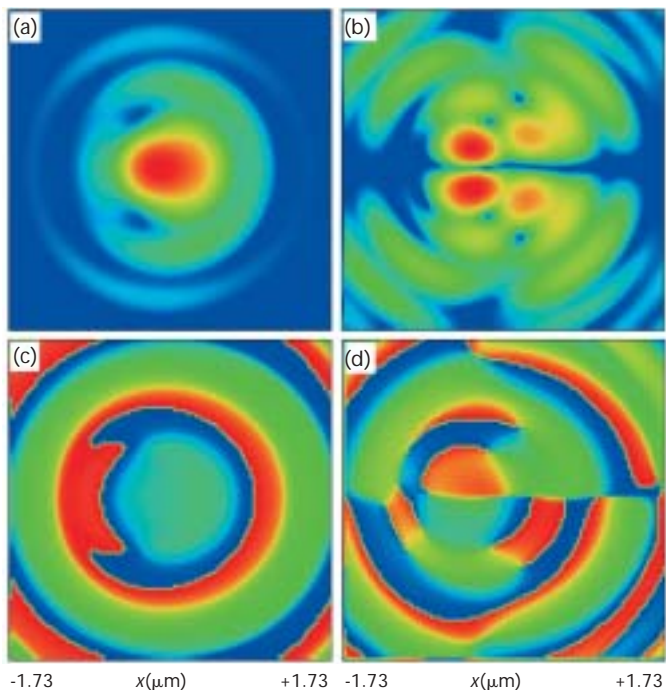


Figure 8. Plots of reflected log_intensity_3 (*top*) and phase (*bottom*) from the convex pit in the sample depicted in Fig. 3(d) at $\lambda_o = 650$ nm. The panels on the left-hand side correspond to E_x , while those on the right-hand side represent the E_y component of the reflected field. The pit center is 250 nm to the left of the focused spot center.

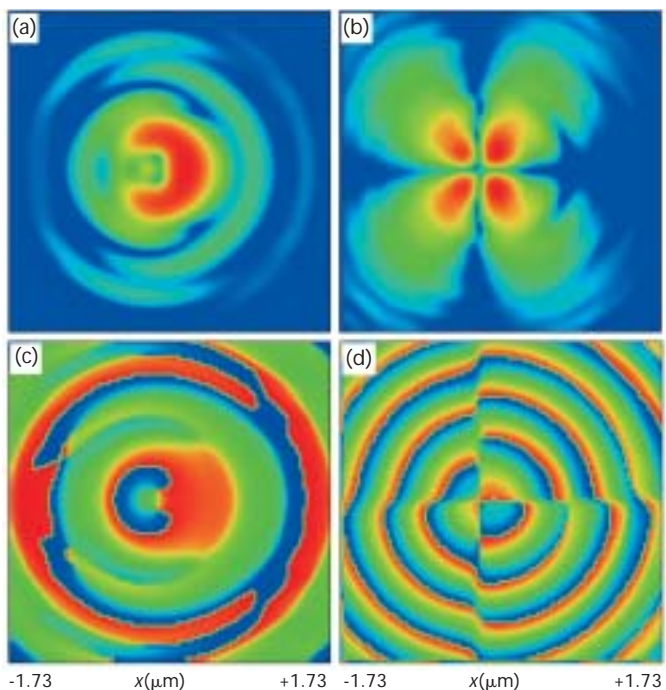


Figure 9. Same as Fig. 8 but for the sample of Fig. 3(e). The objective is now corrected for the thickness and refractive index of the substrate, so the beam focused on this concave pit continues to be the diffraction-limited spot shown in Fig. 4.

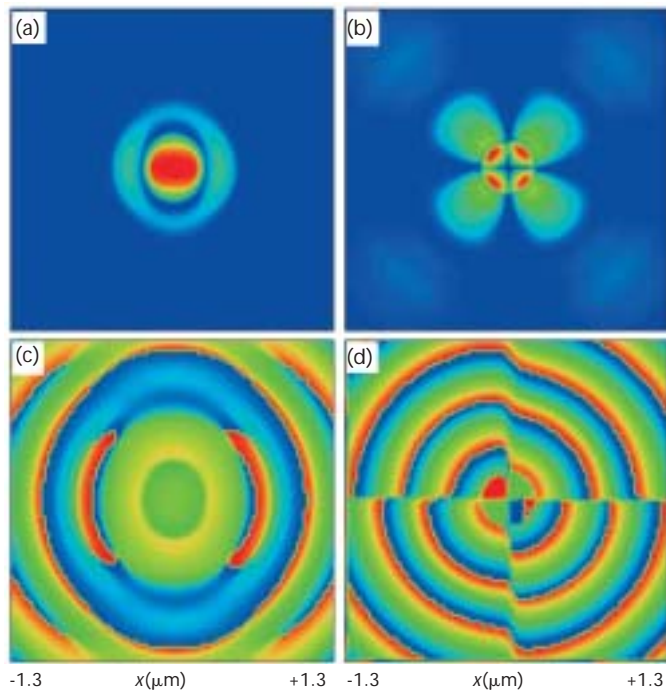


Figure 10. Plots of transmitted log_intensity_3 (*top*) and phase (*bottom*) at $\lambda_o = 650$ nm through the thin metal film depicted in Fig. 3(f). The film contains a 400-nm-diameter circular aperture at its center. The panels on the left-hand side correspond to E_x , while those on the right-hand side represent the E_y component of the field. The observation plane is 20 nm past the interface between the film and the substrate.

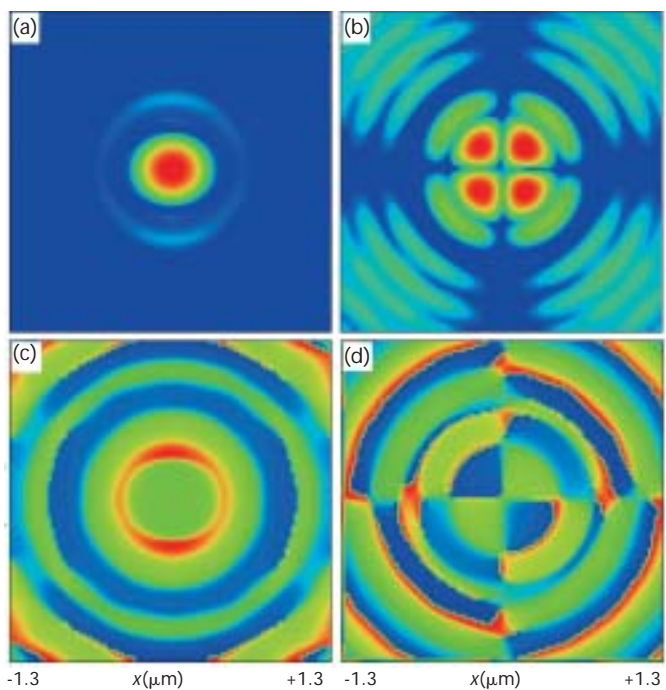


Figure 11. Same as Fig. 10 but with evanescent field components filtered out.

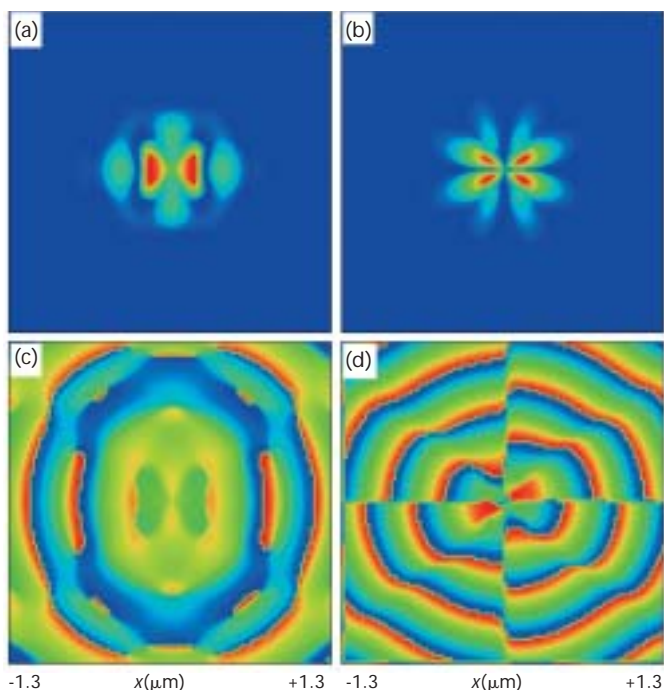


Figure 12. Plots of transmitted $\log_intensity_3$ (top) and phase (bottom) through the 20-nm-thick metal film depicted in Fig. 3(f). The bowtie aperture at the center of the film is 400 nm long along X and 300 nm-wide along Y ; the rectangular neck of the bowtie is 60 nm wide. The panels on the left-hand side correspond to E_x , while those on the right-hand side represent the E_y component of the transmitted field.

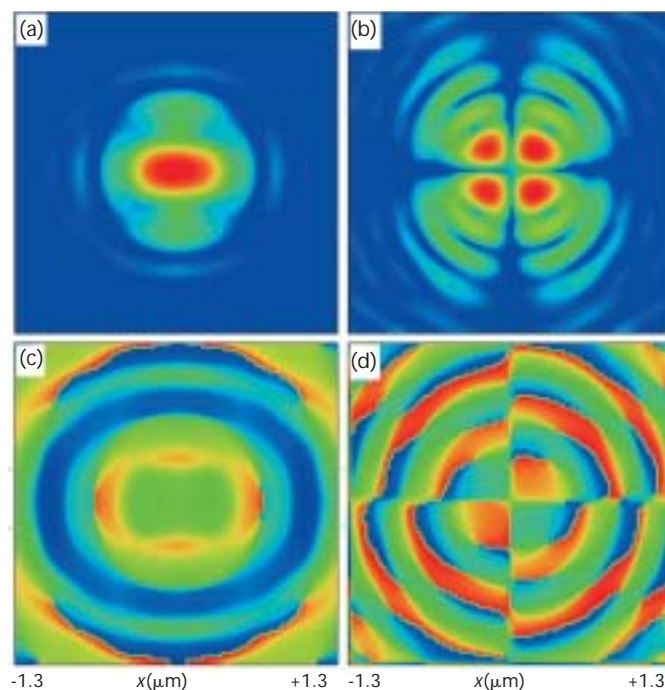


Figure 13. Same as Fig. 12 but with evanescent field components filtered out.

convex pit of Fig. 3(d) tends to concentrate the incoming rays toward the pit center, the concave pit of Fig. 3(e) disperses these rays away from the center.

Transmission through small apertures

Figure 3(f) shows a 20-nm-thick metal film ($n + ik = 2.0 + i7.0$) with an air-filled hole at the center, coated on a glass substrate of index $n = 1.5$. The hole is either a 400-nm-diameter circular aperture, or a bowtie-shaped aperture 400 nm long and 300 nm wide. In our FDTD simulations of these apertures the wavelength was $\lambda_0 = 650$ nm, the mesh size was $L_x = L_y = 12\lambda_0$, the front facet of the metal film was at $z_1 = 77$ nm beyond the focal plane, and the distance from the rear facet of the film to the plane in which the transmitted beam is observed was $z_2 = 20$ nm.

Figure 10 shows computed plots of the transmitted intensity and phase for the sample of Fig. 3(f) containing a circular aperture. Note that, despite its large absorption coefficient, the metal film is not thick enough to completely block the incident beam. Thus, in addition to the light that passes through the aperture, a weak ring of light is also transmitted through the film. The integrated values of transmitted intensity are $P_x = 0.35$, $P_y = 0.0086$. Since the focused cone of light consists of p - as well as s -polarized rays, the difference in sample reflectivity for these differently polarized rays at oblique incidence is partly responsible for the elongated shape of the transmitted intensity profile in Fig. 10(a). The proximity of the observation plane to the

aperture ensures that the transmitted field contains a mixture of propagating as well as evanescent plane waves. If these evanescent components are filtered out, then the remaining field will propagate undiminished to the far field. The filtered field in the same observation plane (i.e., at $z_2 = 20$ nm beyond the interface between the metal film and the substrate) is shown in Fig. 11. The integrated intensity of the X - and Y -components of polarization in these calculations is found to be $P_x = 0.31$, $P_y = 0.002$.

For the bowtie aperture in the thin-film sample of Fig. 3(f), computed plots of transmitted intensity and phase are shown in Fig. 12. The computed integrated intensities in this case are $P_x = 0.19$, $P_y = 0.04$. When the evanescent content of the transmitted field is filtered out, the distributions shown in Fig. 13 are obtained. (The integrated intensity values now drop to $P_x = 0.09$, $P_y = 0.005$.) Note that the bowtie shape of the aperture is no longer discernible in the filtered transmitted beam, ostensibly because the fine features of this aperture contribute primarily to the evanescent field.

References

1. K. S. Yee, *IEEE Trans. Antennas and Prop.* **14**, 302-7 (1966).
2. A. Taflov and S. C. Hagness, *Computational Electrodynamics*, Artech House (2000).
3. The computer simulations reported in this article were performed by Sim3D_Max™, a FDTD-based program developed by A. R. Zakharian, and by DIFFRACT™, a product of MM Research, Inc., Tucson, Arizona.

OPN contributing editor Masud Mansuripur <masud@u.arizona.edu> is a professor of optical sciences at the University of Arizona in Tucson. His collection of past OPN columns has been published in *Classical Optics and Its Applications* (Cambridge University Press, UK, 2002). Aramis Zakharian and Jerome Moloney are with the Department of Mathematics and the Optical Sciences Center at the University of Arizona.