Analysis and numerical computation of diffraction of an optical field by a subwavelength-size aperture in a thick metallic screen by use of a volume integral equation

Kazuo Tanaka and Masahiro Tanaka

Diffraction of an optical field by an aperture in a thick metallic screen is analyzed numerically by use of a three-dimensional volume integral equation together with a generalized conjugate residual method and fast Fourier transformation. Numerical results were validated by reciprocity and the independence of the results of the truncated discretized volume size used in numerical calculations. Near and far fields of square, circular, and triangular apertures in a thick screen are obtained numerically. Some of the numerical results obtained in the present study agree with previously reported experimental results. The surface plasmon polaritons excited on the sidewalls of the aperture can explain the basic characteristics of near-field distribution of apertures. The Bethe–Bouwkamp theory was found to be insufficient to explain the basic characteristics of the near field around the subwavelength aperture in a practical metallic screen.

OCIS codes: 050.1220, 050.1960, 180.5810, 230.7370, 350.5400.

1. Introduction

The interaction between an object and the optical near field of an aperture in a metallic screen is one of the fundamental physical processes in near-field optics (NFO) technology,[1–4] and investigation of electromagnetic near fields around a subwavelength-size aperture in a metallic screen is important. The diffraction of electromagnetic waves by an aperture in an infinite metallic screen is one of the fundamental problems in electromagnetic theory and has been treated in many papers, including the papers of Bethe[5] and Bouwkamp[6] (BB). A number of authors have examined the case in which a screen, either infinitely thin or of finite thickness, is a perfect conductor.[7–13]

Several studies have analyzed by means of numerical and experimental methods the optical near fields and far fields of an aperture at the probe tip in a practical NFO system.[14–21] In these studies, the metal coating on the probe was treated as a dielectric of complex-valued permittivity of finite thickness. However, the physical characteristics of the optical field scattered by the aperture have not been investigated in detail. Recently, simulation of near-field scanning optical microscopy (NSOM) images of small particles in the illumination mode has revealed several interesting characteristics.[21] This previous study reported that the NSOM image characteristics cannot be accounted for with the BB model. Unfortunately, the small-aperture model of the previous investigation does not correspond to the practical structure of NSOM.

Analysis of the optical field scattered by the subwavelength-size aperture in a metallic screen is one of the important fundamental problems in NFO technology. In NFO, we must consider that the metallic screen has a finite thickness and the metal must be regarded as a dielectric media having complex-valued permittivity. In this paper we perform three-dimensional numerical simulations of the optical fields scattered by an aperture in a complex-valued dielectric screen of finite thickness using the volume integral equation (VIE) (dyadic Green's function) together with current numerical techniques.[22–30]

We first derive the VIE of the problem, solve the VIE numerically using current numerical techniques for large-scale matrix equations, and check the validity of the obtained numerical results. We then nu-

K. Tanaka (tanaka@tnk.info.gifu-u.ac.jp) and M. Tanaka (masahiro@tnk.info.gifu-u.ac.jp) are with the Department of Electronics and Computer Engineering, Gifu University, 1-1 Yanagido, Gifu 501-1193, Japan.

Received 28 March 2003; revised manuscript received 30 October 2003; accepted 3 December 2003.

0003-6935/04/081734-13$15.00/0

merically simulate the near-field distribution for square, circular, and triangular apertures. The surface plasmon polaritons (SPPs) excited on the sidewalls of the aperture can be used to explain the basic characteristics of near-field distribution of apertures. In addition, we investigate the dependence of the scattering cross section of the square aperture on the screen thickness and on the aperture area. The simulation results obtained in this study provide important information with regard to basic characteristics of NFO technology.

2. Volume Integral Equations

Here we consider the scattering problem of optical waves by an aperture in a thick metallic screen placed in a vacuum, as shown in Fig. 1. A small square aperture (SA) whose area is given by $a_x \times a_y$ was formed in a metallic screen (slab) of thickness w. The area of the metallic screen is infinite, and its relative complex-valued permittivity is given by ε_1. We adopt rectangular coordinates (x, y, z) and spherical coordinates (r, θ, ϕ), the origins of which are located at the geometric center of the SA on the upper screen surface, as shown in Fig. 1. A plane wave is assumed to be incident with the incident angle (θ_i, ϕ_i) from region (I) below the metallic screen, as shown in Fig. 1.

Here we solve the scattering problem using a VIE (Lippman–Schwinger equation in quantum physics). The VIE for the problem shown in Fig. 1 can be written as[31,32]

$$\mathbf{E}(\mathbf{x}) = k_0^2 \iiint_V [\varepsilon_r(\mathbf{x}') - 1]\bar{G}(\mathbf{x}|\mathbf{x}') \cdot \mathbf{E}(\mathbf{x}')dv' + \mathbf{E}^i(\mathbf{x}), \quad (1)$$

where $\mathbf{E}(\mathbf{x})$ is the total electric field, $\mathbf{E}^i(\mathbf{x})$ is the incident electric field, and $\bar{G}(\mathbf{x}|\mathbf{x}')$ is an electric-type free-space dyadic Green's function (tensor). The volume integral region V represents the entire space and $\varepsilon_r(\mathbf{x})$ represents the distribution of the relative permittivity, where $\varepsilon_r(\mathbf{x}) = \varepsilon_1$ inside the metallic screen and $\varepsilon_r(\mathbf{x}) = 1$ in the aperture and surrounding vacuum, as shown in Fig. 1. Since $\varepsilon_r(\mathbf{x}) - 1 = 0$ in exterior region (I) and region (II) in Fig. 1, region V in Eq. (1) can be regarded as the volume of the screen without an aperture. So we consider that V represents the infinite space defined by $-\infty < x < \infty$, $-\infty < y < \infty$ and $-w < z < 0$ hereafter in this paper. The VIE [Eq. (1)] has been applied to a number of NFO problems, revealing numerous interesting findings.[22–30] For the problem shown in Fig. 1, volume integral region V has an infinite volume. So, application of Eq. (1) to the problem directly is impossible. To overcome this difficulty, we consider the total electric field $\mathbf{E}(\mathbf{x})$ inside the metallic screen to be approximated by the field inside a dielectric slab without aperture at points far from the aperture, because the screen is a dissipative dielectric. The field inside the dielectric slab without aperture can be expressed analytically and rigorously by solving well-known problems of reflection and transmission with a

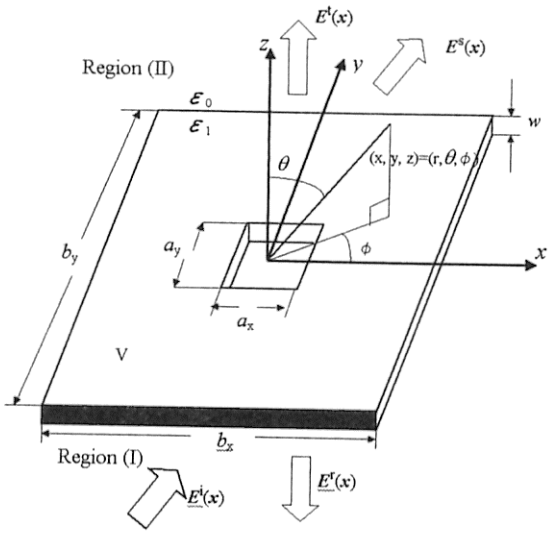

Fig. 1. Geometry of the problem. The optical plane wave is scattered by an aperture in the infinite metallic screen placed in the coordinate systems (x, y, z) and (r, ϕ, θ). Screens are of metal 1 ($\varepsilon_1 = -1.68 - j4.46$) and metal 2 ($\varepsilon_1 = -7.38 - j7.18$).

two-dimensional dielectric slab. We denote this field as $\mathbf{E}^{\text{slab}}(\mathbf{x})$ and hereafter refer to this field as the slab field. Namely, we make the following assumption for the electric field inside the metallic screen with an aperture:

$$\mathbf{E}(\mathbf{x}) = \mathbf{E}^c(\mathbf{x}) + \mathbf{E}^{\text{slab}}(\mathbf{x}), \quad (\mathbf{x} \in V). \quad (2)$$

The field denoted by $\mathbf{E}^c(\mathbf{x})$ in Eq. (2) is expected to be confined within the vicinity of the aperture. Substituting Eq. (2) into Eq. (1), we obtain

$$\mathbf{E}^c(\mathbf{x}) + \mathbf{E}^{\text{slab}}(\mathbf{x}) = k_0^2 \iiint_V [\varepsilon_r(\mathbf{x}') - 1]\bar{G}(\mathbf{x}|\mathbf{x}') \cdot \mathbf{E}^c(\mathbf{x}')dv' + k_0^2 \iiint_V [\varepsilon_r(\mathbf{x}') - 1]\bar{G}(\mathbf{x}|\mathbf{x}') \cdot \mathbf{E}^{\text{slab}}(\mathbf{x}')dv' + \mathbf{E}^i(\mathbf{x}), \ (\mathbf{x} \in V). \quad (3)$$

Note that slab field $\mathbf{E}^{\text{slab}}(\mathbf{x})$ must satisfy the following VIE:

$$\left.\begin{array}{l} \mathbf{E}^i(\mathbf{x}) + \mathbf{E}^r(\mathbf{x}):[\mathbf{x} \in (I)] \\ \mathbf{E}^{\text{slab}}(\mathbf{x}):[\mathbf{x} \in V] \\ \mathbf{E}^t(\mathbf{x}):[\mathbf{x} \in (II)] \end{array}\right\} = k_0^2 \iiint_V [\varepsilon_{\text{slab}}(\mathbf{x}') - 1]\bar{G}(\mathbf{x}|\mathbf{x}') \cdot \mathbf{E}^{\text{slab}}(\mathbf{x}')dv' + \mathbf{E}^i(\mathbf{x}), \quad (4)$$

where $\varepsilon_{\text{slab}}(\mathbf{x})$ represents the distribution of the relative permittivity of the infinite metallic screen (slab) without an aperture. Fields $\mathbf{E}^r(\mathbf{x})$ and $\mathbf{E}^t(\mathbf{x})$ represent the reflected plane wave in region (I) below the

metallic screen and the transmitted plane wave in region (II) above the screen, respectively, as shown in Fig. 1. Fields $\mathbf{E}^r(\mathbf{x})$, $\mathbf{E}^t(\mathbf{x})$, and $\mathbf{E}^{\text{slab}}(\mathbf{x})$ can be expressed analytically. Substituting Eq. (4) into Eq. (3) we can derive the VIE for field $\mathbf{E}^c(\mathbf{x})$ as

$$\mathbf{E}^c(\mathbf{x}) = k_0^2 \iiint_V [\varepsilon_r(\mathbf{x}') - 1]\bar{G}(\mathbf{x}|\mathbf{x}') \cdot \mathbf{E}^c(\mathbf{x}')\mathrm{d}v' + k_0^2 \iiint_V [\varepsilon_r(\mathbf{x}') - \varepsilon_{\text{slab}}(\mathbf{x}')]\bar{G}(\mathbf{x}|\mathbf{x}') \cdot \mathbf{E}^{\text{slab}}(\mathbf{x}')\mathrm{d}v'. \tag{5}$$

Comparing Eq. (5) with the original VIE Eq. (1), we can see that the unknown function $\mathbf{E}(\mathbf{x})$ in Eq. (1) is replaced by the field denoted by $\mathbf{E}^c(\mathbf{x})$ in Eq. (5) and that the incident wave $\mathbf{E}^i(\mathbf{x})$ in Eq. (1) is replaced by volume integration of the slab field over the aperture region, i.e., volume region V, where $[\varepsilon_r(\mathbf{x}) - \varepsilon_{\text{slab}}(\mathbf{x})] \neq 0$ is satisfied in Eq. (5). The basic structure of Eq. (5) is the same as that of Eq. (1), and Eq. (5) can be solved by the conventional method for solving a VIE. When no aperture is present in the metallic screen, the last term on the right-hand side (RHS) of Eq. (5) disappears and the solution for Eq. (5) is given by $\mathbf{E}^c(\mathbf{x}) = 0$. So, slab field $\mathbf{E}^{\text{slab}}(\mathbf{x})$ becomes the rigorous solution of the problem of Eq. (1). Since we can consider that the field denoted by $\mathbf{E}^c(\mathbf{x})$ in the screen becomes negligibly small at points far from the aperture, we can regard the infinite volume integral region V as a finite region and solve VIE (5) numerically.

The basic VIE (5) is not appropriate for numerical evaluation. We must remove the singular point by performing volume integration over an infinitesimal sphere of which the center is located at observation point $\mathbf{x}$. This procedure is well known and can be found in the literature.[32–34] The integral equation can be rewritten as

$$-k_0^2 \iiint_V [\varepsilon_r(\mathbf{x}') - 1]\bar{G}(\mathbf{x}|\mathbf{x}') \cdot \mathbf{E}^c(\mathbf{x}')\mathrm{d}v' + [\varepsilon_r(\mathbf{x}) + 2]/3\mathbf{E}^c(\mathbf{x}) = k_0^2 \iiint_V [\varepsilon_r(\mathbf{x}') - \varepsilon_{\text{slab}}(\mathbf{x}')]\bar{G}(\mathbf{x}|\mathbf{x}') \cdot \mathbf{E}^{\text{slab}}(\mathbf{x}')\mathrm{d}v', \quad (\mathbf{x} \in V). \tag{6}$$

The integral in Eq. (6) denotes a principal-value integration over volume V. Multiplying both sides of Eq. (6) by $\varepsilon_r(\mathbf{x}) - 1$, we can obtain the following equation:

$$-k_0^2[\varepsilon_r(\mathbf{x}) - 1] \iiint_V \bar{G}(\mathbf{x}|\mathbf{x}') \cdot \chi^c(\mathbf{x}')\mathrm{d}v' + [\varepsilon_r(\mathbf{x}) + 2]/3\chi^c(\mathbf{x}) = k_0^2[\varepsilon_r(\mathbf{x}) - 1] \iiint_V [\varepsilon_r(\mathbf{x}') - \varepsilon_{\text{slab}}(\mathbf{x}')]\bar{G}(\mathbf{x}|\mathbf{x}') \cdot \mathbf{E}^{\text{slab}}(\mathbf{x}')\mathrm{d}v', \quad (\mathbf{x} \in V). \tag{7}$$

The unknown function of VEI (7) changes to the function

$$\chi^c(\mathbf{x}) = [\varepsilon_r(\mathbf{x}) - 1]\mathbf{E}^c(\mathbf{x}). \tag{8}$$

Since the relation between the unknown function $\chi^c(\mathbf{x})$ and the dyadic Green's function is given by the convolution in Eq. (7), we can use fast Fourier transformation (FFT) to solve Eq. (7) by the iteration technique.[35–37]

The unknown function $\chi^c(\mathbf{x})$ can be obtained numerically in Eq. (7), and the electric field $\mathbf{E}^c(\mathbf{x})$ can be obtained from Eq. (8). Once the electric field, denoted by $\mathbf{E}^c(\mathbf{x})$, in the metallic screen has been obtained numerically, the total electric field in region (II) above the metallic screen can be calculated from Eqs. (1) and (2) as

$$\begin{aligned}\mathbf{E}(\mathbf{x}) &= k_0^2 \iiint_V [\varepsilon_r(\mathbf{x}') - 1]\bar{G}(\mathbf{x}|\mathbf{x}') \cdot [\mathbf{E}^c(\mathbf{x}) + \mathbf{E}^{\text{slab}}(\mathbf{x})]\mathrm{d}v' + \mathbf{E}^i(\mathbf{x}) \\ &= k_0^2 \iiint_V [\varepsilon_r(\mathbf{x}') - 1]\bar{G}(\mathbf{x}|\mathbf{x}') \cdot \mathbf{E}^c(\mathbf{x}')\mathrm{d}v' \\ &\quad + k_0^2 \iiint_V [\varepsilon_r(\mathbf{x}') - \varepsilon_{\text{slab}}(\mathbf{x}')]\bar{G}(\mathbf{x}|\mathbf{x}') \cdot \mathbf{E}^{\text{slab}}(\mathbf{x}')\mathrm{d}v' + \mathbf{E}^t(\mathbf{x}), \quad \mathbf{x} \in (\text{II}).\end{aligned} \tag{9}$$

For the derivation of Eq. (9) we used Eq. (4). The scattered far field $\mathbf{E}^s(r, \theta, \phi)$ produced by the aperture in the metallic screen can be expressed as (see Appendix A)

$$\mathbf{E}^s(r, \theta, \phi) \cong [\exp(-jk_0r)/(k_0r)]\mathbf{F}(\theta, \phi), \quad k_0r \gg 1, \tag{10}$$

where $\mathbf{F}(\theta, \phi)$ is the scattering coefficient and can be written as

$$\mathbf{F}(\theta, \phi) = -\mathbf{i}_r \times \mathbf{i}_r \times \left(\frac{1}{4\pi}\right)\left[k_0^3 \iiint_V [\varepsilon_r(\mathbf{x}') - 1]\mathbf{E}^c(\mathbf{x}')\exp(jk_0\mathbf{x}' \cdot \mathbf{i}_r)\mathrm{d}v' + k_0^3 \iiint_V [\varepsilon_r(\mathbf{x}') - \varepsilon_{\text{slab}}(\mathbf{x}')]\mathbf{E}^{\text{slab}}(\mathbf{x}')\exp(jk_0\mathbf{x}' \cdot \mathbf{i}_r)\mathrm{d}v'\right]. \tag{11}$$

In Eqs. (10) and (11) polar coordinates (r, θ, ϕ) are as shown in Fig. 1; $\mathbf{i}_r$ is a unit vector in the radial direction in these polar coordinates. The scattering cross section W of the aperture can be written as

$$W = \int_0^{\pi/2} \int_0^{2\pi} |F(\theta, \phi)|^2 \sin\theta \mathrm{d}\theta\mathrm{d}\phi. \tag{12}$$

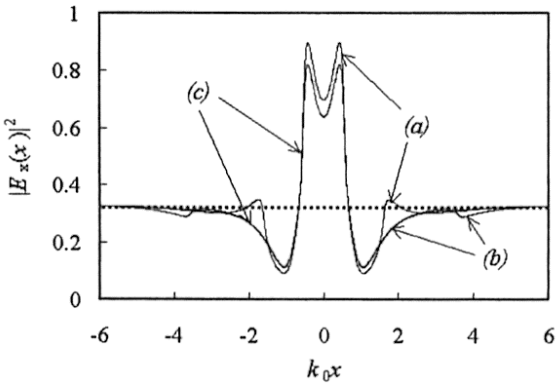

Fig. 2. Dependence of near-field distributions $|E_x(k_0x, 0.0, 0.1)|^2$ above the aperture on the discretized volume size of the screen used for the numerical calculation: (a) $k_0b_x \times k_0b_y \times k_0w = 3.2 \times 3.2 \times 0.3$, (b) $k_0b_x \times k_0b_y \times k_0w = 7.2 \times 7.2 \times 0.3$, (c) $k_0b_x \times k_0b_y \times k_0w = 11.2 \times 11.2 \times 0.3$. The screen is metal 1 ($\varepsilon_1 = -1.68 - j4.46$). The surrounding region is a vacuum ($\varepsilon_0 = 1$), and the size of the SA is $k_0a_x = k_0a_y = 1.2$.

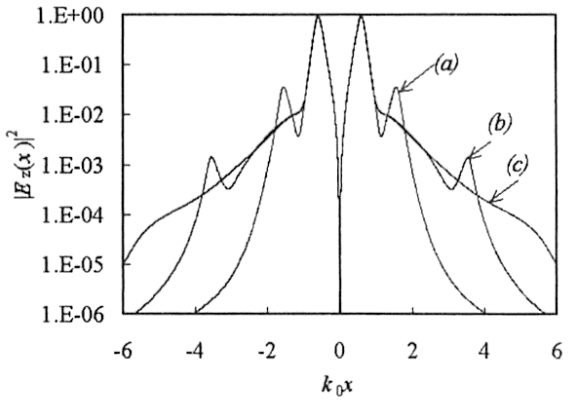

Fig. 3. Dependence of near-field distributions $|E_z(k_0x, 0.0, 0.1)|^2$ above the aperture on the discretized volume size of the screen used for the numerical calculation: (a) $k_0b_x \times k_0b_y \times k_0w = 3.2 \times 3.2 \times 0.3$, (b) $k_0b_x \times k_0b_y \times k_0w = 7.2 \times 7.2 \times 0.3$, (c) $k_0b_x \times k_0b_y \times k_0w = 11.2 \times 11.2 \times 0.3$. The screen is metal 1 ($\varepsilon_1 = -1.68 - j4.46$). The surrounding region is a vacuum ($\varepsilon_0 = 1$), and the size of the SA is $k_0a_x = k_0a_y = 1.2$.

We can regard W as the power transmitted through the aperture in the metallic screen into region (II) for the case of an incident plane wave of unit amplitude.

3. Confirmation of Numerical Results

In this paper we fix the following parameters: wavelength λ, incident angle $\theta_i = 0$, the incident electric vector $\mathbf{E}^i(\mathbf{x})$ is parallel to the x axis in Fig. 1, the SA size is $k_0a_x = k_0a_y = 1.2$ (approximately $0.19\lambda \times 0.19\lambda$) and the complex permittivities of the metallic screen are given by $\varepsilon_1 = -1.68 - j4.46$ (metal 1) and $\varepsilon_1 = -7.38 - j7.18$ (metal 2). The techniques for numerically solving a VIE such as Eq. (7) are well established in the field of computational electromagnetic theory.[31–37] Here we employ the method of moment using the pulse function as a basis function and the delta function as a testing function in the discretization of VIE (7).[32] The width of the basis pulse function (size of cubes used as a discretized element) is $k_0\delta_x = k_0\delta_y = k_0\delta_z = 0.1$, and the number of unknown resultant linear systems exceeds one million for the problems in this paper. We solved the resultant large-scale system of linear equations using an iteration method referred to as the generalized conjugate residual method combined with FFT.[38,39] The finite-size volume integral region of $b_x \times b_y \times w$ was used in the practical numerical calculation of Eq. (7) as shown in Fig. 1. We confirmed the numerical results by using the following procedure.

A. Invariance of the Near Field with Respect to the Truncated Volume Size

Since the solution of Eq. (7) must be that for an infinitely large metallic screen on the x–y plane, the numerical results must be independent of the truncated volume size of the screen on the x–y plane given by $b_x \times b_y$ in the numerical calculation. In Figs. 2 and 3 the resultant near-field intensities $|E_x(k_0x, 0.0, 0.1)|^2$ and $|E_z(k_0x, 0.0, 0.1)|^2$ obtained on the line parallel to the x axis above the aperture are shown as parameter $k_0b_x = k_0b_y$. The field intensity $|E_y|^2$ was smaller than one tenth of $|E_x|^2$ and $|E_z|^2$ and so is omitted in this calculation. The screen is metal 1 and the thickness of the screen is given by $k_0w = 0.3$ (approximately 0.05λ). From Figs. 2 and 3, when the truncated size of the screen used in the numerical calculation is greater than that given by $k_0b_x \times k_0b_y = 11.2 \times 11.2$ (approximately $1.78\lambda \times 1.78\lambda$), we can see that calculated near-field intensities are independent of the size of the truncated screen size. Note that the ordinate of Fig. 3 is a logarithmic scale. Results of the screen of metal 2 are similar to those of metal 1.

The validity of the assumption in Eq. (2) is based on the physically reasonable idea that can be stated as the effects of the aperture on the fields, i.e., $\mathbf{E}^c(\mathbf{x})$, will vanish at points far from the aperture inside the dissipative dielectric screen. However, when the decay of $\mathbf{E}^c(\mathbf{x})$ is small, we must use a large truncated discretized volume size and the size of the system of linear equations could exceed the ability of the system. SPPs excited on the screen surfaces parallel to the x–y plane have a somewhat small attenuation constant. However, when the aperture size is small and the incident angle is nearly vertical to the screen surface, we can expect the amplitude of an excited SPP on the screen surfaces parallel to the x–y plane to be small. In Fig. 2, at points far from the aperture, the x component of the total near-field intensity is close to $|T|^2 = 0.32$, that is, the power transmission coefficient of the slab without an aperture shown by a broken line in Fig. 2. The z component of $\mathbf{E}^c(\mathbf{x})$, which does not exist above the screen without an aperture for x-directed incident polarization, decays fast at points far from the aperture as shown in Fig. 3. This result also shows that $\mathbf{E}^c(\mathbf{x})$ decays effectively at points far from the aperture and SPPs excited on the screen surface by an aperture are small. These results show the validity of

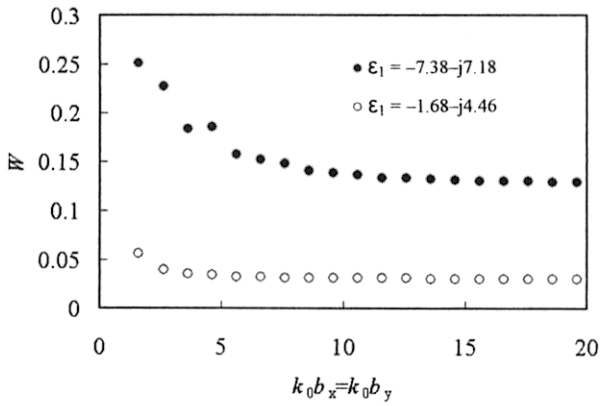

Fig. 4. Dependence of scattering cross section W of the SA on the discretized volume size used for the numerical calculation. The screen thickness is fixed at $k_0w = 0.3$. The size of the SA is $k_0a_x = k_0a_y = 1.2$.

the idea in the mathematical formulation and the code introduced in this paper.

B. Invariance of Scattering Cross Section W with Respect to the Truncated Volume Size

We have also confirmed the invariance of scattering cross section W with respect to the truncated size of screen $k_0b_x \times k_0b_y$ used for the numerical calculation. In Fig. 4 the dependences of scattering coefficient W on the truncated size used in volume integration $k_0b_x = k_0b_y$ are shown. The thickness of the slab is $k_0w = 0.3$. From Fig. 4 we can see that calculated scattering coefficient W is independent of the truncated volume size of the screen, when the truncated volume size is sufficiently large. These results show the numerical convergence of the numerical results solved by the mathematical formulation shown in this paper and the code used in this paper.

C. Reciprocity

We use the reciprocity relation, which is valid for the scattering problem of a dissipative medium, to confirm the numerical results. Since the concrete expression of the reciprocity relation for the scattering problem shown in Fig. 1 has not yet been reported, we derive the reciprocity relation for the problem shown in Fig. 1 in Appendix B as

$$\mathbf{R}_1 \cdot \mathbf{F}_2(\pi - \theta_1, \phi_1) + \mathbf{T}_1 \cdot \mathbf{F}_2(\theta_1, \phi_1) = \mathbf{R}_2 \cdot \mathbf{F}_1(\pi - \theta_2, \phi_2) + \mathbf{T}_2 \cdot \mathbf{F}_1(\theta_2, \phi_2), \quad (13)$$

where $\mathbf{R}_1$, $\mathbf{T}_1$ and $\mathbf{R}_2$, $\mathbf{T}_2$ represent the vector reflection and transmission coefficients of plane waves with a screen (slab) without an aperture for incident angles of (θ_1, ϕ_1) and (θ_2, ϕ_2), respectively (see Appendix B). Similarly, $\mathbf{F}_1(\theta, \phi)$ and $\mathbf{F}_2(\theta, \phi)$ represent the scattering coefficients given by Eq. (11) for incident angles of (θ_1, ϕ_1) and (θ_2, ϕ_2), respectively (see Appendix A). We solved the problem by using a volume size of $k_0b_x \times k_0b_y \times k_0w = 20.0 \times 20.0 \times 0.3$ and calculated the RHS and left-hand side (LHS) of Eq. (13) for the case of $\theta_2 = 5.0$ (deg) and $\theta_2 = 10.0$ (deg) under fixed values of $\theta_1 = \phi_1 = \phi_2 = 0$ for metal 1 and metal 2 screens. The results are shown in Table 1. Based on the numerical results in Table 1, when incident angle θ_i is less than 5 deg, the numerical results satisfy reciprocity relation (13) to an acceptable degree of accuracy. These results demonstrate the validity of the numerical evaluation for small values of incident angles θ_i. The reciprocity discrepancy for large values of incident angle θ_i could be due to the excitation of SPPs that propagate along the screen surfaces to infinity. Under these conditions, large truncated volume size is required for the numerical evaluation, which was difficult in our system. All the numerical results hereafter are those for the case of normal incidence of $\theta_i = 0$ and $\phi_i = 0$.

Table 1. Verification of Reciprocity of the Problem ($\theta_1 = \phi_1 = \phi_2 = 0$)

θ_2 (deg)	RHS of Eq. (13)	LHS of Eq. (13)
5.0 (metal 1)	$-0.029862 + j0.046526$	$-0.030988 + j0.046524$
10.0 (metal 1)	$-0.028227 + j0.045531$	$-0.030676 + j0.046151$
5.0 (metal 2)	$-0.145388 + j0.027705$	$-0.155554 + j0.024608$
10.0 (metal 2)	$-0.140179 + j0.029362$	$-0.154507 + j0.024806$

4. Near-Field Distributions of an Aperture

A. Square Aperture

We first consider the case of a SA of $k_0a_x = k_0a_y = 1.2$ of thickness $k_0w = 1.9$ (approximately 0.30λ) in Figs. 5 and 6. Distribution of total near-field intensities $|\mathbf{E}(k_0x, k_0y, 0.1)|^2$ and their components $|E_x(k_0x, k_0y, 0.1)|^2$, $|E_y(k_0x, k_0y, 0.1)|^2$, and $|E_z(k_0x, k_0y, 0.1)|^2$ above the SA are shown in Fig. 5 for the metal 1 screen and in Fig. 6 for the metal 2 screen, respectively, for the x-directed incident polarization in Fig. 1. All the distributions in this paper are those on the plane parallel to the x–y plane placed at $k_0z = 0.1$ (approximately 0.016λ) above the screen. We note that the transmitted plane wave is negligibly small for a screen thickness of $k_0w = 1.9$ for both the metal 1 and the metal 2 screens. Note the difference in scale range between Figs. 5 and 6. From these figures, we conclude the following:

(a) Two main spots of high optical intensity are created along the side boundaries, which are perpendicular to the incident polarization of the aperture. The optical intensities just above the aperture center are smaller than those at these two spots.

(b) These two spots exist in the region outside boundaries of the aperture.

(c) The main component of the electric fields at these two spots of high intensity is the z component (perpendicular to the screen).

(d) The basic characteristics of the intensity distributions for metals 1 and 2 are similar. Therefore, we hereafter show only the intensity distributions of the metal 2 screen.

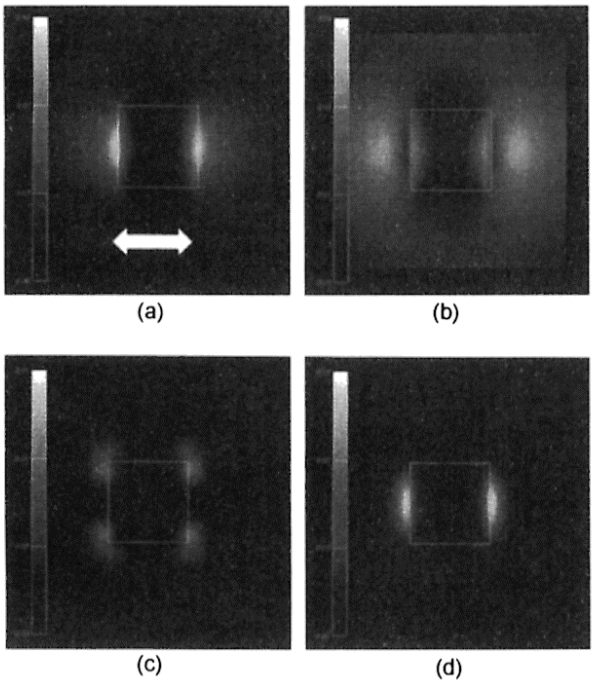

Fig. 5. Electric near-field distributions normalized by the intensity of the incident wave on the plane parallel to the x–y plane placed at $k_0z = 0.1$ for a SA in a metal 1 screen ($\varepsilon_1 = -1.68 - j4.46$): (a) $|\mathbf{E}(k_0x, k_0y, 0.1)|^2$, (b) $|E_x(k_0x, k_0y, 0.1)|^2$, (c) $|E_y(k_0x, k_0y, 0.1)|^2$, (d) $|E_z(k_0x, k_0y, 0.1)|^2$. The thickness of the screen is $k_0w = 1.9$. The square indicates the $k_0a_x = k_0a_y = 1.2$ aperture. The arrow in (a) indicates incident polarization. The intensity ranges are (a) (0–0.1), (b) (0–0.04), (c) (0–0.01), (d) (0–0.08).

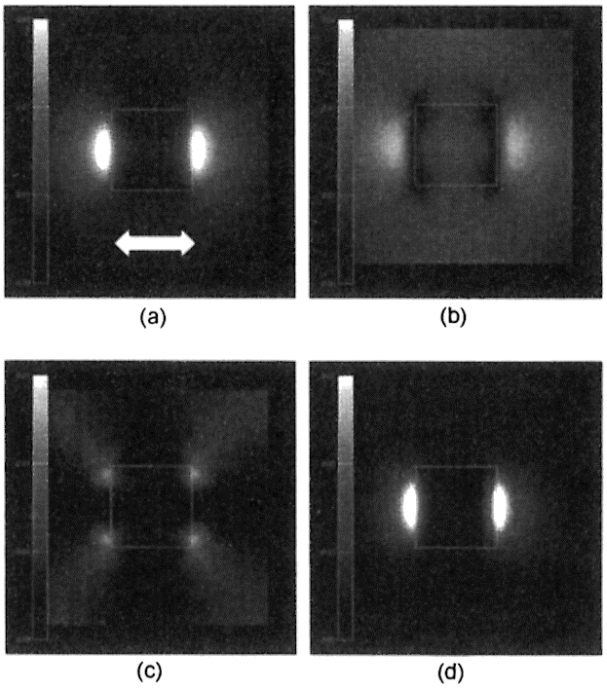

Fig. 6. Electric near-field distributions normalized by the intensity of the incident wave on the plane parallel to the x–y plane placed at $k_0z = 0.1$ for a SA in a metal 2 screen ($\varepsilon_1 = -7.38 - j7.18$): (a) $|\mathbf{E}(k_0x, k_0y, 0.1)|^2$, (b) $|E_x(k_0x, k_0y, 0.1)|^2$, (c) $|E_y(k_0x, k_0y, 0.1)|^2$, (d) $|E_z(k_0x, k_0y, 0.1)|^2$. The thickness of the screen is $k_0w = 1.9$. The square indicates the $k_0a_x = k_0a_y = 1.2$ aperture. The arrow in (a) indicates incident polarization. The intensity ranges are (a) (0–0.1), (b) (0–0.04), (c) (0–0.01), (d) (0–0.08).

B. Circular Aperture

Distributions of near-field intensity for the case of a circular aperture (CA) having an area equal to that of the SA of $k_0a_x \times k_0a_y = 1.2 \times 1.2$ in a metal 2 screen of thickness $k_0w = 1.9$. The total near-field intensities $|\mathbf{E}(k_0x, k_0y, 0.1)|^2$ on the plane at $k_0z = 0.1$ are shown in Fig. 7. The basic characteristics of the intensity distributions agree with those of the BB theory[5,6] for an infinitely thin perfect conductor. However, the details of the intensity distributions shown in Fig. 7 differ from those of the BB theory. In the BB theory two primary spots exist in the range inside the CA, and their main electric field component is parallel to the incident polarization. However, as shown in Fig. 7, for a sufficiently thick metallic screen, the two primary spots are located in the region outside the CA, and their main electric field component is perpendicular to the screen for the sufficiently thick metallic screen. These characteristics are identical to characteristics (a), (b), and (c) for the SAs shown in Figs. 5 and 6. The results of Fig. 7 agree qualitatively with the experimental results reported by Molenda *et al.*[40] and Höppener *et al.*[41] The average intensity for the CA in Fig. 7 is smaller than that for the SA in Fig. 6 under the condition of equal aperture area.

C. Triangular Aperture

Distribution of near-field intensity for the case of a triangular aperture (TA) having an area equal to that of the SA of $k_0a_x \times k_0a_y = 1.2 \times 1.2$ in a metal 2 screen of $k_0w = 1.9$ thickness. The total near-field intensities $|\mathbf{E}(k_0x, k_0y, 0.1)|^2$ on the plane at $k_0z = 0.1$ are shown in Fig. 8(a) for x-directed incident polarization and in Fig. 8(b) for y-directed incident polarization. Note that the scale ranges in Figs. 8(a) and 8(b) are double those of Figs. 5–7. The average intensities for the TA are larger than those for the SA and CA under the condition of equal aperture area. Naber *et al.*[42] reported optical intensity distributions for a TA made

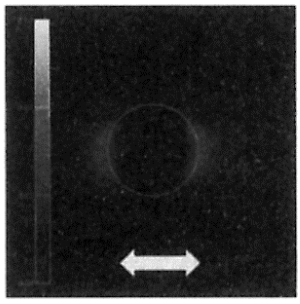

Fig. 7. Total electric near-field distributions normalized by the intensity of an incident wave on the plane parallel to the x–y plane placed at $k_0z = 0.1$ for a CA in a metal 2 screen ($\varepsilon_1 = -7.38 - j7.18$). The thickness of the screen is $k_0w = 1.9$. The circle indicates the aperture whose area is the same as that of a SA of $k_0a_x = k_0a_y = 1.2$. The arrow indicates incident polarization. The intensity ranges from 0 to 0.1.

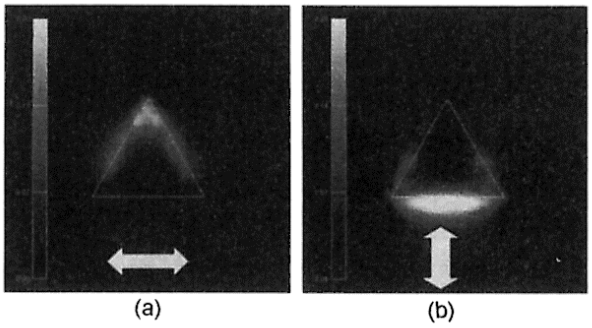

Fig. 8. Total electric near-field distributions $|\mathbf{E}(k_0x, k_0y, 0.1)|^2$ normalized by the intensity of the incident wave on the plane parallel to the x–y plane placed at $k_0z = 0.1$ for TAs in a metal 2 screen ($\varepsilon_1 = -7.38 - j7.18$). The thickness of the screen is $k_0w = 1.9$. The triangle indicates the aperture whose area is the same as that of a SA of $k_0a_x = k_0a_y = 1.2$. The arrows indicate incident polarization. The intensity ranges from 0 to 0.2.

on the probe tip as measured experimentally by use of an excellent technique that involves a fluorescent nanosphere. Although the parameters used in their experiments differ from those we used, their experimentally obtained characteristics of the intensity distributions for a TA agree well with the numerical results shown in Figs. 8(a) and 8(b). They reported that an enhanced single spot near the base of the TA appears for a specific incident polarization. This single spot can be observed for y-directed incident polarization in Fig. 8(b). These results indicate the validity of the numerical evaluation in this paper.

5. Explanation of Near-Field Distributions by Surface Plasmon Polaritons

Since the screen thickness is much larger than the skin depth of metal 2 in the given wavelength in Figs. 5–8, the two primary spots for a SA and a CA and a single spot for a TA are not due to the waves that propagate directly inside the screen material. Considering that the main electric field component of the spot is perpendicular to the screen and exists along the boundary perpendicular to the incident polarization, it is possible to explain these spots by the SPPs that progress along the sidewalls inside the aperture as shown in Fig. 9. We consider only

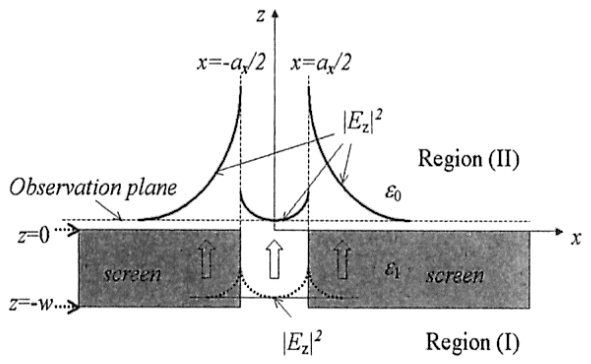

Fig. 9. Explanation of the numerical results obtained with SPPs on the sidewalls of the aperture.

the case of a SA for simplicity. We assume that the field near $y = 0$ inside the aperture is due to SPPs excited on the sidewalls at $x = \pm a_x/2$. The electric field of SPPs on the walls have components E_x and E_z and numerical results have this property approximately. When the aperture is not in a cutoff condition, i.e., $k_0a_y \gg \pi$, we can consider that E_z is smaller than E_x inside the aperture from a two-dimensional calculation of SPPs.[43] However, the aperture is very small and is under a cutoff condition, i.e., $k_0a_y \ll \pi$, shown in this paper, E_x inside the aperture is suppressed because Poynting vector $E_x \times H_y$ must be small under a cutoff condition. So, it is reasonable that E_z is not much smaller than E_x inside the aperture. The field $|E_z|^2$ must be continuous across the walls at $x = \pm a_x/2$ inside the aperture from boundary conditions and decrease exponentially from both walls inside the aperture and the metallic screen in the range of $-w < z < 0$ shown by the dotted curves in Fig. 9. So, a maximum optical intensity exists at both walls at $x = \pm a_x/2$ in the $-w < z < 0$ range. Numerical results agree with these characteristics and show the validity of the assumption.

Let us consider the distribution of $|E_z|^2$ on the observation plane just above and close to the screen in the free space ($z > 0$) shown by solid curves in Fig. 9. The field $|E_z|^2$ just above the aperture is similar to that inside the aperture because there is no change of relative permittivity across the plane at $z = 0$. However, $|E_z|^2$ just above the metallic screen is enhanced by the factor $|\varepsilon_1/\varepsilon_0|^2$ because electric fluxes perpendicular to the screen surface must be continuous across the screen surface at $z = 0$, i.e., $\varepsilon_1E_z = \varepsilon_0E_z$, shown in Fig. 9. Furthermore, E_z in the free space ($z > 0$) must distribute continuously because there are no aperture boundaries in this region. Therefore, positions of the maximum optical intensity (primary spots) $|E_z|^2$ above the metallic screen in free space are shifted to the outer direction from aperture boundaries at $x = \pm a_x/2$. The enhanced field $|E_z|^2$ outside the boundary can become larger than $|E_x|^2$ just above the aperture on the observation plane in Fig. 9. For a SA, numerical results show that the distributions of E_z just below the x–y plane at $z = 0$ along the y axis near the walls at $x = \pm a_x$ can be approximated by $\cos(\pi/a_yy)$, which is determined by the boundary condition on the walls at $y = \pm a_y/2$. So, from the above-mentioned consideration we can consider that two spots in the near-field distribution of a SA and a CA as shown in Figs. 5–7 are due to the SPPs excited on the sidewalls of the aperture. The SPPs can be excited most effectively on the sidewall that is straight and is perfectly perpendicular to the given incident polarization. Since there is only one straight sidewall that satisfies the above condition for a TA in Fig. 8(b), it is reasonable that only one spot can be observed as shown in Fig. 8(b). We can also understand that the near-field intensity of the CA shown in Fig. 7 is smaller than that of the SA shown in Fig. 6, because there is no straight sidewall that is perfectly perpendicular to the given incident polarization for the CA in Fig. 7. The propagation of SPPs

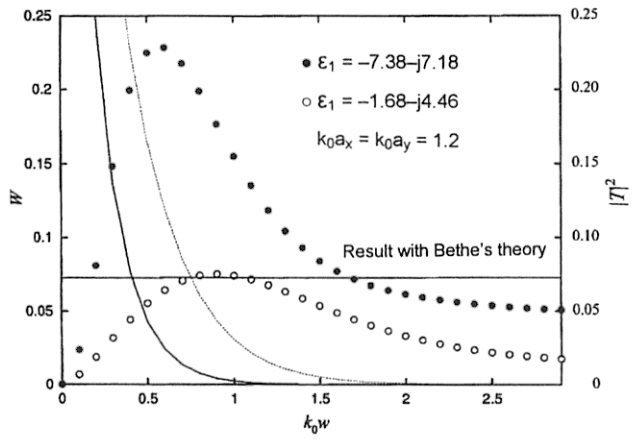

Fig. 10. Dependence of scattering cross section W of a SA on screen thickness k_0w. The size of the SA is $k_0a_x = k_0a_y = 1.2$. The solid and dotted curves represent the power transmission coefficients $|\mathbf{T}|^2$ of the metal 2 and metal 1 slabs without an aperture, respectively, of thickness k_0w. The solid line represents Bethe's results for the scattering cross section of a CA having an area equal to that of the SA.

along the walls inside the aperture under cutoff conditions requires further investigation. However, we can state that these results reveal that the BB theory is insufficient for analysis of the aperture problem in a thick metallic screen that is often used in NFO, because the BB theory completely neglects the effects of SPPs. Using the characteristic of SPPs excited on the sidewalls of an aperture, it is possible to make an aperture that gives high intensity and a small spot size of near-field intensity in a thick metallic screen.[44]

6. Dependence of Scattering Cross Section *W* on Screen Thickness and Aperture Area

The dependence of a transmitted optical far field through a SA, i.e., scattering cross section W of the SA, on screen thickness k_0w is important for application of the small aperture to NFO technology. The results for a metal 2 screen are represented by filled circles in Fig. 10, and those for a metal 1 screen are represented by open circles. In Fig. 10 the power transmission coefficients $|\mathbf{T}|^2$ of a metallic slab without aperture, where $\mathbf{T}$ is the vector transmission coefficient of the metallic slab without aperture defined by Eq. (A7) in Appendix A, are shown for the metal 2 slab by a solid curve and those for the metal 1 slab are shown by a dotted curve. Bethe's results for the scattering cross section and the results for a circular aperture having an area equal to that of the SA are shown. The thickness that yields maximum scattering cross section W is found to be $k_0w \cong 0.6$ for the metal 2 screen and $k_0w \cong 0.9$ for the metal 1 screen in Fig. 10. When the screen is thinner than the skin depth, the scattering cross section W can be small because most of the incident energy is transmitted through the screen as the transmitted plane wave. When the screen is sufficiently thick, the transmitted plane wave becomes small, as shown in Fig. 10, and the transmitted power through the aperture, i.e., W, must be small. Therefore, the existence of a screen thickness that produces a maximum scattering cross section W is reasonable. The thickness that produces a maximum scattering cross section W will be useful in the design of an aperture for the probe tip of NFO technology.

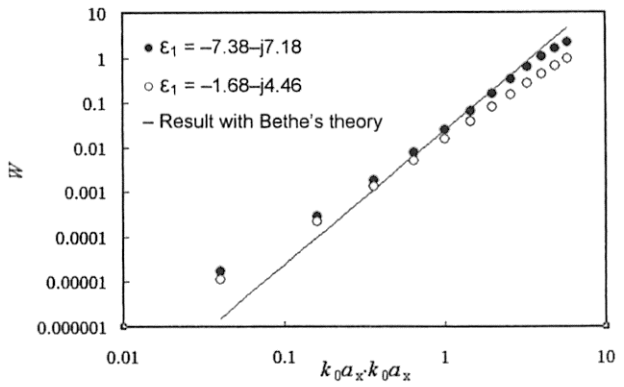

Fig. 11. Dependence of scattering cross section W of a SA in a thick metallic screen having an aperture area of $k_0a_x \times k_0a_y$. The screen thickness is $k_0w = 1.9$. The solid line represents the scattering cross section as calculated by use of Bethe's theory.

The dependences of scattering cross section W on the area of the SA are shown in Fig. 11 by open circles for the metal 1 screen and by filled circles for the metal 2 screen, for a screen of large thickness $k_0w = 1.9$. The solid line represents Bethe's results for a circular aperture of area equal to that of the SA. These characteristics do not differ greatly from Bethe's result.

7. Conclusions

The scattering of optical waves by a small aperture in a thick metallic screen has been analyzed by use of a volume integral equation together with the generalized conjugate residual method and FFT. The near-field distribution above the aperture in the thick metallic screen differs from that of the Bethe–Bouwkamp theory. These results are important for near-field optics technology. Numerical results show that the surface plasmon polaritons can play an important role in the diffraction of optical waves by a subwavelength aperture. So, the Bethe–Bouwkamp theory is insufficient to explain the basic characteristics of the near field around the subwavelength aperture in a practical metallic screen.

Appendix A

The scattered field $\mathbf{E}^s(\mathbf{x})$ can be obtained by subtracting transmitted plane wave $\mathbf{E}^t(\mathbf{x})$ from total field Eq. (9). We show only how to obtain the scattered field expression of only the first term on the RHS of Eq. (9) as follows:

$$\mathbf{E}^s(\mathbf{x}) = {k_0}^2 \iiint_V [\varepsilon_r(\mathbf{x}') - 1]\bar{G}(\mathbf{x}|\mathbf{x}') \cdot \mathbf{E}^c(\mathbf{x}')\mathrm{d}v'. \tag{A1}$$

The dyadic Green's function can be expressed as

$$\bar{G}(\mathbf{x}|\mathbf{x}') = (\bar{I} + 1/k_0{}^2\nabla\nabla)g(\mathbf{x}, \mathbf{x}'), \quad \text{(A2)}$$

where

$$\nabla^2 g(\mathbf{x}, \mathbf{x}') + k_0{}^2 g(\mathbf{x}, \mathbf{x}') = -\delta(\mathbf{x}, \mathbf{x}'), \quad \text{(A3)}$$

$$g(\mathbf{x}, \mathbf{x}') = \exp(-k_0|\mathbf{x} - \mathbf{x}'|)/(4\pi|\mathbf{x} - \mathbf{x}'|), \quad \text{(A4)}$$

and $\bar{I}$ is a unit dyad. Substituting Eqs. (A2) and (A3) into Eq. (A1), we can rewrite Eq. (A1) as

$$\begin{aligned}\mathbf{E}^s(\mathbf{x}) &= k_0{}^2 \iiint_V [\varepsilon_r(\mathbf{x}') - 1][-(1/k_0{}^2)\nabla^2 g(\mathbf{x}, \mathbf{x}')\bar{I} \\ &\quad + (1/k_0{}^2)\nabla\nabla g(\mathbf{x}, \mathbf{x}')] \cdot \mathbf{E}^e(\mathbf{x}')\mathrm{d}v' \\ &= \iiint_v [\varepsilon_r(\mathbf{x}') - 1][-\nabla^2 g(\mathbf{x}, \mathbf{x}')\bar{I} \\ &\quad + \nabla\nabla g(\mathbf{x}, \mathbf{x}')] \cdot \mathbf{E}^e(\mathbf{x}')\mathrm{d}v' \\ &= \iiint_v [\varepsilon_r(\mathbf{x}') - 1][\nabla \times \nabla \times g(\mathbf{x}, \mathbf{x}')\bar{I}] \\ &\quad \cdot \mathbf{E}^e(\mathbf{x}')\mathrm{d}v' \\ &= \nabla \times \nabla \times \iiint_v [\varepsilon_r(\mathbf{x}') - 1]g(\mathbf{x}, \mathbf{x}') \\ &\quad \mathbf{E}^e(\mathbf{x}')\mathrm{d}v'. \end{aligned} \quad \text{(A5)}$$

When observation point $\mathbf{x}$ is far from source point $\mathbf{x}'$, we can use the asymptotic expression of the free-space scalar Green's function as

$$\begin{aligned} g(\mathbf{x}, \mathbf{x}') &\approx \exp(-jk_0 r)/(4\pi r) \\ &\quad \times \exp(jk_0\mathbf{x}' \cdot \mathbf{i}_r), \quad k_0 r \gg 1, \end{aligned} \quad \text{(A6)}$$

where r is the distance between the observation point $\mathbf{x}$ and the origin. So we can show that

$$\begin{aligned}\nabla \times g(\mathbf{x}, \mathbf{x}')\mathbf{E}(\mathbf{x}') \approx\ & -jk_0\mathbf{i}_r \exp(-jk_0 r)/(4\pi r) \\ & \times \mathbf{E}(\mathbf{x}')\exp(jk_0\mathbf{x}' \cdot \mathbf{i}_r), \quad k_0 r \gg 1. \end{aligned} \quad \text{(A7)}$$

We finally obtain the expression for the scattered far field:

$$\begin{aligned}\mathbf{E}^s(\mathbf{x}) = \exp(-jk_0 r)/(4\pi k_0 r)\Big[& -\mathbf{i}_r \times \mathbf{i}_r \\ & \times k_0{}^3 \iiint_v [\varepsilon_r(\mathbf{x}') - 1]\mathbf{E}^e(\mathbf{x}')\exp(jk_0\mathbf{x}' \cdot i_r)\mathrm{d}v'\Big]. \end{aligned} \quad \text{(A8)}$$

We can obtain a similar expression for the second term in Eq. (9) and can obtain relation (10).

Appendix B

We consider the geometry shown in Fig. 12. We define total electric and magnetic fields by $(\mathbf{E}_1, \mathbf{H}_1)$ for incident angles of (θ_1, ϕ_1) and by $(\mathbf{E}_2, \mathbf{H}_2)$ for incident angles of (θ_2, ϕ_2) in the space whose permittivity and permeability are given by ε and μ, respectively. Let us start with the well-known reciprocity relation between total fields $(\mathbf{E}_1, \mathbf{H}_1)$ and $(\mathbf{E}_2, \mathbf{H}_2)$ derived from Maxwell's equations without source terms as

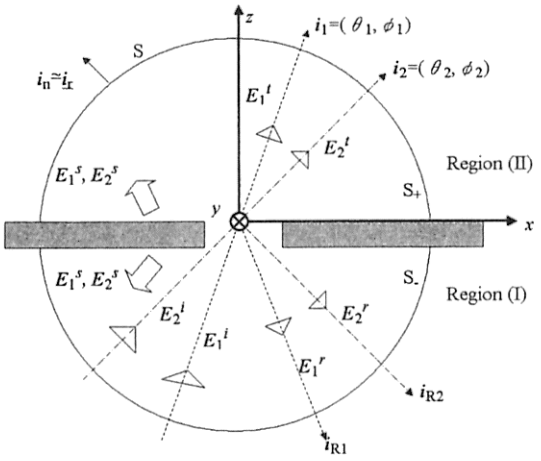

Fig. 12. Geometry of the problem used for derivation of reciprocity of the problem.

$$\int_S (\mathbf{E}_1 \times \mathbf{H}_2 - \mathbf{E}_2 \times \mathbf{H}_1) \cdot \mathbf{i}_n \mathrm{d}s = 0, \quad \text{(B1)}$$

where S is an infinitely large sphere enclosing a metallic screen with an aperture that intersects the screen far from the aperture, as shown in Fig. 12, and $\mathbf{i}_n$ is an outward unit vector perpendicular to surface S. We decompose the total fields into components of scattered fields $\mathbf{E}_1{}^s$, $\mathbf{E}_2{}^s$ and plane wave terms $\mathbf{E}_1{}^p$, $\mathbf{E}_2{}^p$ as

$$\mathbf{E}_k = \mathbf{E}_k{}^p + \mathbf{E}_k{}^s, \quad \mathbf{H}_k = \mathbf{H}_k{}^p + \mathbf{H}_k{}^s, \ (k = 1, 2), \quad \text{(B2)}$$

where

$$\begin{aligned}\mathbf{E}_k{}^p &= \mathbf{E}_k{}^i + \mathbf{E}_k{}^r, \quad \mathbf{x} \text{ exists in region (I)}, \\ &= \mathbf{E}_k{}^t, \quad (k = 1, 2) \quad \mathbf{x} \text{ exists in region (II)}, \end{aligned} \quad \text{(B3)}$$

$$\begin{aligned}\mathbf{H}_k{}^p &= \mathbf{H}_k{}^i + \mathbf{H}_k{}^r, \quad \mathbf{x} \text{ exists in region (I)}, \\ &= \mathbf{H}_k{}^t, \quad (k = 1, 2) \quad \mathbf{x} \text{ exists in region (II)}, \end{aligned} \quad \text{(B4)}$$

$$\mathbf{E}_k{}^i = \boldsymbol{\epsilon}_k \exp(-jk_0\mathbf{x} \cdot \mathbf{i}_k),$$

$$\mathbf{H}_k{}^i = \frac{1}{\zeta}(\mathbf{i}_k \times \boldsymbol{\epsilon}_k)\exp(-jk_0\mathbf{x} \cdot \mathbf{i}_k), \quad (k = 1, 2), \quad \text{(B5)}$$

$$\mathbf{E}_k{}^r = \mathbf{R}_k \exp(-jk_0\mathbf{x} \cdot \mathbf{i}_{Rk}),$$

$$\mathbf{H}_k{}^r = \frac{1}{\zeta}(\mathbf{i}_{Rk} \times \mathbf{R}_k)\exp(-jk_0\mathbf{x} \cdot \mathbf{i}_{Rk}), \quad (k = 1, 2), \quad \text{(B6)}$$

$$\mathbf{E}_k{}^t = \mathbf{T}_k \exp(-jk_0\mathbf{x} \cdot \mathbf{i}_k),$$

$$\mathbf{H}_k{}^t = \frac{1}{\zeta}(\mathbf{i}_k \times \mathbf{T}_k)\exp(-jk_0\mathbf{x} \cdot \mathbf{i}_k), \quad (k = 1, 2), \quad \text{(B7)}$$

where $\zeta = (\varepsilon/\mu)^{1/2}$. Since we consider the field on surface S of an infinitely large sphere that encloses the metallic screen, we can use asymptotic expressions of the scattered far field (see Appendix A) as

$$\mathbf{E}_k^s(r, \theta, \phi) \approx \frac{\exp(-jk_0 r)}{k_0 r} \mathbf{F}_k(\theta, \phi), \quad (k = 1, 2), \tag{B8}$$

$$\mathbf{H}_k^s(r, \theta, \phi) \approx \frac{1}{\zeta} [\mathbf{i}_r \times \mathbf{E}_k^S(r, \theta, \phi)], \quad (k = 1, 2). \tag{B9}$$

In Eqs. (B5)–(B7) and approximations (B8) and (B9), $\mathbf{i}_k$ represents the unit vectors of incident direction of incident angle (θ_1, ϕ_1) for $k = 1$ and (θ_2, ϕ_2) for $k = 2$, and $\mathbf{i}_k$ is equal to those of the transmitted direction as given by Eqs. (B7). Unit vectors $\boldsymbol{\varepsilon}_k$ $(k = 1, 2)$ represent the polarization of the incident wave. Unit vectors $\mathbf{i}_{Rk}$ $(k = 1, 2)$ represent the unit vector of the reflected direction, as shown in Fig. 12. The expressions $\mathbf{R}_k$ $(k = 1, 2)$ represent the vector reflection coefficients defined by Eqs. (B6) for incident angle (θ_1, ϕ_1) for $k = 1$ and (θ_2, ϕ_2) for $k = 2$. Similarly, $\mathbf{T}_k$ $(k = 1, 2)$ represents the vector transmission coefficients defined by Eqs. (B7) for incident angle (θ_1, ϕ_1) for $k = 1$ and (θ_2, ϕ_2) for $k = 2$. A unit vector $\mathbf{i}_r$ represents the radial direction shown in Fig. 12.

Substituting Eqs. (B2) into Eq. (B1), we obtain

$$\int_S (\mathbf{E}_1^p + \mathbf{E}_1^s) \times (\mathbf{H}_2^p + \mathbf{H}_2^s) - (\mathbf{E}_2^p + \mathbf{E}_2^s) \times (\mathbf{H}_1^p + \mathbf{H}_1^s)] \cdot \mathbf{i}_n ds = 0. \tag{B10}$$

Equation (B10) yields

$$\int_S (\mathbf{E}_1^p \times \mathbf{H}_2^p - \mathbf{E}_2^p \times \mathbf{H}_1^p) \cdot \mathbf{i}_n ds + \int_S (\mathbf{E}_1^s \times \mathbf{H}_2^s - \mathbf{E}_2^s \times \mathbf{H}_1^s) \cdot \mathbf{i}_n ds + \int_S (\mathbf{E}_1^p \times \mathbf{H}_2^s - \mathbf{E}_2^p \times \mathbf{H}_1^s) \cdot \mathbf{i}_n ds + \int_S (\mathbf{E}_1^s \times \mathbf{H}_2^p - \mathbf{E}_2^s \times \mathbf{H}_1^p) \cdot \mathbf{i}_n ds = 0. \tag{B11}$$

Considering the relations

$$\int_S (\mathbf{E}_1^p \times \mathbf{H}_2^p - \mathbf{E}_2^p \times \mathbf{H}_1^p) \cdot \mathbf{i}_n ds = 0, \tag{B12}$$

$$\int_S (\mathbf{E}_1^s \times \mathbf{H}_2^s - \mathbf{E}_2^s \times \mathbf{H}_1^s) \cdot \mathbf{i}_n ds = 0, \tag{B13}$$

we obtain the following relation from Eq. (B11):

$$\int_S (\mathbf{E}_1^p \times \mathbf{H}_2^s - \mathbf{E}_2^p \times \mathbf{H}_1^s) \cdot \mathbf{i}_n ds + \int_S (\mathbf{E}_1^s \times \mathbf{H}_2^p - \mathbf{E}_2^s \times \mathbf{H}_1^p) \cdot \mathbf{i}_n ds = 0. \tag{B14}$$

Substituting Eqs. (B12)–(B14) into Eq. (B11), we derive the following equation:

$$\int_{S_-} (\mathbf{E}_1^i \times \mathbf{H}_2^s - \mathbf{E}_2^i \times \mathbf{H}_1^s) \cdot \mathbf{i}_n ds + \int_{S_-} (\mathbf{E}_1^s \times \mathbf{H}_2^i - \mathbf{E}_2^s \times \mathbf{H}_1^i) \cdot \mathbf{i}_n ds + \int_{S_-} (\mathbf{E}_1^r \times \mathbf{H}_2^S - \mathbf{E}_2^r \times \mathbf{H}_1^s) \cdot \mathbf{i}_n ds + \int_{S_-} (\mathbf{E}_1^s \times \mathbf{H}_2^r - \mathbf{E}_2^s \times \mathbf{H}_1^r) \cdot \mathbf{i}_n ds + \int_{S_+} (\mathbf{E}_1^t \times \mathbf{H}_2^s - \mathbf{E}_2^t \times \mathbf{H}_1^s) \cdot \mathbf{i}_n ds + \int_{S_+} (\mathbf{E}_1^s \times \mathbf{H}_2^t - \mathbf{E}_2^s \times \mathbf{H}_1^t) \cdot \mathbf{i}_n ds = 0, \tag{B15}$$

where S_- and S_+ represent, respectively, the lower surface of S in region (I) and that of the upper surface in region (II) in Fig. 12. It is not difficult to show the following relations:

$$\int_{S_-} (\mathbf{E}_1^i \times \mathbf{H}_2^s - \mathbf{E}_2^i \times \mathbf{H}_1^s) \cdot \mathbf{i}_n ds_- = 0,$$

$$\int_{S_-} (\mathbf{E}_1^s \times \mathbf{H}_2^i - \mathbf{E}_2^s \times \mathbf{H}_1^i) \cdot \mathbf{i}_n ds_- = 0. \tag{B16}$$

We now calculate one of the surviving terms in Eq. (B15). For example, we can perform the following calculation:

$$\begin{aligned}\int_S (\mathbf{E}_1^r \times \mathbf{H}_2^s - \mathbf{E}_2^r \times \mathbf{H}_1^s) \cdot \mathbf{i}_n ds &= \int_S \left\{ \mathbf{R}_1 \exp(-jk_0 \mathbf{x} \cdot \mathbf{i}_{R1}) \times \exp \frac{(-jk_0 r)}{(k_0 r)} \left(\frac{1}{\zeta}\right) [\mathbf{i}_r \times \mathbf{F}_2(\theta, \phi)] \right. \\ &\quad \left. - \mathbf{R}_2 \exp(-jk_0 \mathbf{x} \cdot \mathbf{i}_{R2}) \times \exp \frac{(-jk_0 r)}{(k_0 r)} \left(\frac{1}{\zeta}\right) [\mathbf{i}_r \times \mathbf{F}_1(\theta, \phi)] \right\} \cdot \mathbf{i}_n ds \\ &= \left(\frac{1}{\zeta}\right) \int_S \left\{ \exp \frac{(-jk_0 r)}{(k_0 r)} \exp(-jk_0 \mathbf{r} \cdot \mathbf{i}_{R1}) [\mathbf{R}_1 \times \mathbf{i}_r \times \mathbf{F}_2(\theta, \phi)] \right. \\ &\quad \left. - \exp \frac{(-jk_0 r)}{(k_0 r)} \exp(-jk_0 \mathbf{x} \cdot \mathbf{i}_{R2}) [\mathbf{R}_2 \times \mathbf{i}_r \times \mathbf{F}_1(\theta, \phi)] \right\} \cdot \mathbf{i}_n ds. \end{aligned} \tag{B17}$$

$$\int_{S_-} (\mathbf{E}_1^s \times \mathbf{H}_2^r - \mathbf{E}_2^s \times \mathbf{H}_1^r) \cdot \mathbf{i}_n ds = \int_{S_-} \left\{ \exp \frac{(-jk_0 r)}{(k_0 r)} \mathbf{F}_1(\theta, \phi) \times \left(\frac{1}{\zeta}\right)(\mathbf{i}_{R2} \times \mathbf{R}_2)\exp(-jk_0 \mathbf{x} \cdot \mathbf{i}_{R2}) \right.$$
$$\left. - \exp \frac{(-jk_0 r)}{(k_0 r)} \mathbf{F}_2(\theta, \phi) \times \left(\frac{1}{\zeta}\right)(\mathbf{i}_{R1} \times \mathbf{R}_1)\exp(-jk_0 \mathbf{x} \cdot \mathbf{i}_{R1}) \right\} \cdot \mathbf{i}_n ds$$
$$= \left(\frac{1}{\zeta}\right) \int_{S_-} \left\{ \exp \frac{(-jk_0 r)}{(k_0 r)} \exp(-jk_0 \mathbf{x} \cdot \mathbf{i}_{R2})[\mathbf{F}_1(\theta, \phi) \times \mathbf{i}_{R2} \times \mathbf{R}_2] \right.$$
$$\left. - \exp \frac{(-jk_0 r)}{(k_0 r)} \exp(-jk_0 \mathbf{x} \cdot \mathbf{i}_{R1})[\mathbf{F}_2(\theta, \phi) \times \mathbf{i}_{R1} \times \mathbf{R}_1] \right\} \cdot \mathbf{i}_n ds. \quad \text{(B22)}$$

When surface S is sufficiently large, we can set $\mathbf{i}_n = \mathbf{i}_r$. Using Eqs. (A8), we derive the relation as

$$\mathbf{i}_n \cdot \{\mathbf{R}_1 \times [\mathbf{i}_r \times \mathbf{F}_2(\theta, \phi)]\}$$
$$= \mathbf{i}_n \cdot [\mathbf{R}_1 \cdot \mathbf{F}_2(\theta, \phi)]\mathbf{i}_r - \mathbf{i}_n \cdot (\mathbf{R}_1 \cdot \mathbf{i}_r)\mathbf{F}_2(\theta, \phi)$$
$$= \mathbf{i}_n \cdot [\mathbf{R}_1 \cdot \mathbf{F}_2(\theta, \phi)]\mathbf{i}_r = \mathbf{R}_1 \cdot \mathbf{F}_2(\theta, \phi). \quad \text{(B18)}$$

Therefore, we can calculate Eq. (B17) as

Using the relation

$$\mathbf{i}_n \cdot [\mathbf{F}_1(\theta, \phi) \times \mathbf{i}_{R2} \times \mathbf{R}_2] = \mathbf{i}_n \cdot [\mathbf{F}_1(\theta, \phi) \cdot \mathbf{R}_2]\mathbf{i}_{R2} - \mathbf{i}_n \cdot [\mathbf{F}_1(\theta, \phi) \cdot \mathbf{i}_{R2}]\mathbf{R}_2$$
$$= [\mathbf{F}_1(\theta, \phi) \cdot \mathbf{R}_2](\mathbf{i}_n \cdot \mathbf{i}_{R2}) - [\mathbf{F}_1(\theta, \phi) \cdot \mathbf{i}_{R2}](\mathbf{i}_n \cdot \mathbf{i}_{R2}), \quad \text{(B23)}$$

$$\left(\frac{1}{\zeta}\right) \int_{S_-} \left\{ \exp \frac{(-jk_0 r)}{(k_0 r)} \exp(-jk_0 \mathbf{x} \cdot \mathbf{i}_{R1})[\mathbf{R}_1 \cdot \mathbf{F}_2(\theta, \phi)] - \exp \frac{(-jk_0 r)}{(k_0 r)} \exp(-jk_0 \mathbf{x} \cdot \mathbf{i}_{R2})[\mathbf{R}_2 \cdot \mathbf{F}_1(\theta, \phi)] \right\} ds$$
$$= \left(\frac{1}{\zeta}\right) \iint \left\{ \exp \frac{(-jk_0 r)}{(k_0 r)} \exp(-jk_0 \mathbf{x} \cdot \mathbf{i}_{R1})[\mathbf{R}_1 \cdot \mathbf{F}_2(\theta, \phi)] r^2 \sin\theta d\theta d\phi \right\}$$
$$- \exp \frac{(-jk_0 r)}{(k_0 r)} \exp(-jk_0 \mathbf{x} \cdot \mathbf{i}_{R2})[\mathbf{R}_2 \cdot \mathbf{F}_1(\theta, \phi)] r^2 \sin\theta d\theta d\phi$$
$$\approx \frac{-2j\pi}{(\zeta k_0^2)} \exp(-j2k_0 r)[\mathbf{R}_1 \cdot \mathbf{F}_2(\theta_{R1}, \phi_{R1})] + \frac{2j\pi}{(\zeta k_0^2)} \exp(-j2k_0 r)[\mathbf{R}_2 \cdot \mathbf{F}_1(\theta_{R2}, \phi_{R2})]. \quad \text{(B19)}$$

When calculating Eq. (B19), we used the following saddle-point integration formula:

$$I = \int_S g(\theta, \phi)\exp[jkf(\theta, \phi)]d\theta d\phi$$
$$\approx j2\pi\sigma(\alpha\beta - \gamma^2)^{-1/2} g(\theta_0, \phi_0)$$
$$\times \exp \frac{[jkf(\theta_0, \phi_0)]}{k} \quad \text{for } k \to \infty, \quad \text{(B20)}$$

where

$$f_\theta(\theta, \phi)|_{\theta=\theta_0, \phi=\phi_0} = 0,\ f_\phi(\theta, \phi)|_{\theta=\theta_0, \gamma=\phi_0} = 0,$$
$$\alpha = f_{\theta\theta}(\theta_0, \phi_0), \quad \beta = f_{\phi\phi}(\theta_0, \phi_0), \quad \gamma = f_{\theta\phi}(\theta_0, \phi_0),$$
$$\sigma = +1(\alpha\beta > \gamma^2, \alpha > 0),\ -1(\alpha\beta > \gamma^2, \alpha < 0),$$
$$-j(\alpha\beta < \gamma^2). \quad \text{(B21)}$$

Similarly, we can calculate

we can derive Eq. (B24) [see next page]. Here

$$\theta_{R1} = \pi - \theta_1, \quad \theta_{R2} = \pi - \theta_2, \quad \phi_{R1} = \phi_1, \quad \phi_{R2} = \phi_2. \quad \text{(B25)}$$

The remaining terms of Eq. (B15) can be calculated similarly. Finally, the surviving terms of Eq. (B15) can be written as

$$\mathbf{R}_1 \cdot \mathbf{F}_2(\theta_{R1}, \phi_{R1}) + \mathbf{T}_1 \cdot \mathbf{F}_2(\theta_1, \phi_1)$$
$$= \mathbf{R}_2 \cdot \mathbf{F}_1(\theta_{R2}, \phi_{R2}) + \mathbf{T}_2 \cdot \mathbf{F}_1(\theta_2, \phi_2), \quad \text{(B26)}$$

and Eq. (B26) can be expressed as Eq. (13).

This study was financially supported in part by a Grant in Aid (2000–2002) for Scientific Research from the Ministry of Education, Science, Sports, and Culture of Japan. The authors thank D. Molenda and colleagues for the useful information about their experimental results at the University of Munster, Germany.

$$\left(\frac{1}{\zeta}\right)\int_{S_-}\left(\exp\frac{(-jk_0r)}{(k_0r)}\exp(-jk_0\mathbf{x}\cdot\mathbf{i}_{R2})\{[\mathbf{F}_1(\theta,\phi)\cdot\mathbf{R}_2](\mathbf{i}_n\cdot\mathbf{i}_{R2})-[\mathbf{F}_1(\theta,\phi)\cdot\mathbf{i}_{R2}](\mathbf{i}_n\cdot\mathbf{R}_2)\}-\exp\frac{(-jk_0r)}{(k_0r)}\right.$$
$$\left.\exp(-jk_0\mathbf{x}\cdot\mathbf{i}_{R1})\{[\mathbf{F}_2(\theta,\phi)\cdot\mathbf{R}_1](\mathbf{i}_n\cdot\mathbf{i}_{R1})-[\mathbf{F}_2(\theta,\phi)\cdot\mathbf{i}_{R1}](\mathbf{i}_n\cdot\mathbf{R}_1)\}\right)ds\approx\frac{-2j\pi}{(\zeta k_0{}^2)}\exp(-2jk_0r)[\mathbf{R}_2\cdot\mathbf{F}_1(\theta_{R2},\phi_{R2})]$$
$$+\frac{2j\pi}{(\zeta k_0{}^2)}\exp(-2jk_0r)[\mathbf{R}_1\cdot\mathbf{F}_2(\theta_{R1},\phi_{R1})]. \quad \text{(B24)}$$

References

1. E. Betzig and R. J. Chichester, "Single molecules observed by near-field scanning optical microscopy," Science **262,** 1422–1425 (1993).
2. E. H. Synge, "A suggested method for extending microscopic resolution into the ultra-microscopic region," Philos. Mag. **6,** 356–362 (1928).
3. D. W. Pohl and D. Courjon, eds., *Near-Field Optics* (Kluwer Academic, Dordrecht, The Netherlands, 1993).
4. M. Ohtsu and H. Hori, *Near-Field Nano-Optics* (Kluwer Academic, Dordrecht, The Netherlands, 1999).
5. H. A. Bethe, "Theory of diffraction by small holes," Phys. Rev. **66,** 163–182 (1944).
6. C. J. Bouwkamp, "On the diffraction of electromagnetic waves by small circular disks and holes," Philips Res. Rep. **5,** 401–422 (1950).
7. H. Levine and J. Schwinger, "On the theory of diffraction by an aperture in an infinite plane screen. I," Phys. Rev. **74,** 958–974 (1948).
8. Y. Leviatan, R. F. Harrington, and J. R. Mautz, "Electromagnetic transmission through apertures in a cavity in a thick conductor," IEEE Trans. Antennas Propag. **AP-30,** 1153–1165 (1982).
9. Y. Leviatan, "Study of near-zone fields of a small aperture," J. Appl. Phys. **60,** 1577–1583 (1986).
10. C. M. Butler, Y. Rahmat-Samii, and R. Mittra, "Electromagnetic penetration through apertures in conducting surfaces," IEEE Trans. Antennas Propag. **AP-26,** 82–93 (1978).
11. R. E. English, Jr. and N. George, "Diffraction from a small square aperture: approximate aperture fields," J. Opt. Soc. Am. A **5,** 192–199 (1988).
12. A. Roberts, "Near-zone fields behind circular apertures in thick, perfectly conducting screens," J. Appl. Phys. **65,** 2896–2899 (1989).
13. A. Roberts, "Small-hole coupling of radiation into a near-field probe," J. Appl. Phys. **70,** 4045–4049 (1991).
14. H. Furukawa and S. Kawata, "Analysis of image formation in a near-field scanning optical microscope: effects of multiple scattering," Opt. Commun. **132,** 170–178 (1996).
15. R. Chang, P-K. Wei, W. S. Fann, M. Hayashi, and S. H. Lin, "Theoretical investigation of near-field optical properties of tapered fiber tips and single molecule fluorescence," J. Appl. Phys. **81,** 3369–3376 (1997).
16. A. Chavez-Pirson and S. K. Chu, "A full vector analysis of near-field luminescence probing of a single quantum dot," Appl. Phys. Lett. **74,** 1507–1509 (1999).
17. O. J. F. Martin, "3D simulations of the experimental signal measured in near-field optical microscopy," J. Microsc. (Oxford) **194,** 235–239 (1999).
18. D. J. Shin, A. Chavez-Pirson, S. H. Kim, S. T. Jung, and Y. H. Lee, "Diffraction by a subwavelength-sized aperture in a metal plane," J. Opt. Soc. Am. A **18,** 1477–1486 (2001).
19. C. Obermüller and K. Karrai, "Far field characterization of diffracting circular apertures," Appl. Phys. Lett. **67,** 3408–3410 (1995).
20. D. J. Shin, A. Chavez-Pirson, and Y. H. Lee, "Diffraction of circularly polarized light from near-field optical probes," J. Microsc. (Oxford) **194,** 353–359 (1999).
21. A. Liu, A. Rahmani, G. W. Bryant, L. J. Richter, and S. Stranick, "Modeling illumination-mode near-field optical microscopy of Au nanoparticles," J. Opt. Soc. Am. A **18,** 704–716 (2001).
22. O. J. F. Martin, C. Girard, and A. Dereux, "Generalized field propagator for electromagnetic scattering and light confinement," Phys. Rev. Lett. **74,** 526–529 (1995).
23. D. Barchiesi, C. Girard, O. J. F. Martin, D. Van Labeke, and D. Courjon, "Computing the optical near-field distributions around complex subwavelength surface structures: a comparative study of different methods," Phys. Rev. E **54,** 4285–4292 (1996).
24. O. J. F. Martin, C. Girard, and A. Dereux, "Dielectric versus topographic contrast in near-field microscopy," J. Opt. Soc. Am. A **13,** 1801–1808 (1996).
25. O. J. F. Martin and C. Girard, "Controlling and tuning strong optical field gradients at a local probe microscope tip apex," Appl. Phys. Lett. **70,** 705–707 (1997).
26. C. Girard, J.-C. Weeber, A. Dereux, O. J. F. Martin, and J.-P. Goudonnet, "Optical magnetic near-field intensities around nanometer-scale surface structures," Phys. Rev. B **55,** 16487–16497 (1997).
27. K. Kobayashi and O. Watanuki, "Characteristics of photon scanning tunneling microscope read-out," J. Vac. Sci. Technol. B **14,** 804–808 (1996).
28. K. Kobayashi and O. Watanuki, "Polarization-dependent contrast in near-field optical microscopy," J. Vac. Sci. Technol. B **15,** 1966–1970 (1997).
29. K. Tanaka, M. Yan, and M. Tanaka, "A simulation of near-field optics by three-dimensional volume integral equation of classical electromagnetic theory," Opt. Rev. **8,** 43–53 (2001).
30. K. Tanaka, M. Yan, and M. Tanaka, "Simulated output images of near-field optics by volume integral equation: object placed on the dielectric substrate," Opt. Rev. **9,** 213–221 (2002).
31. E. K. Miller, L. Medgyesi-Mitschang, and E. H. Newman, *Computational Electromagnetics: Frequency-Domain Method of Moments* (Institute of Electrical and Electronics Engineers, Piscataway, N.J., 1992).
32. J. H. Wang, *Generalized Moment Method in Electromagnetics: Formulation and Computer Solution of Integral Equations* (Wiley, New York, 1991).
33. J. Van Bladel, "Some remarks on Green's dyadic for infinite space," IEEE Trans. Antennas Propag. **AP-9,** 563–566 (1961).
34. A. D. Yaghjian, "Electric dyadic Green's functions in the source region," Proc. IEEE **68,** 248–263 (1980).
35. J. J. H. Wang and J. R. Dubberley, "Computation of fields in an arbitrary shaped heterogeneous dielectric or biological body by an iterative conjugate gradient method," IEEE Trans. Microwave Theory Tech. **37,** 1119–1125 (1989).
36. C.-C. Su, "Electromagnetic scattering by a dielectric body with arbitrary inhomogeneity and anisotropy," IEEE Trans. Antennas Propag. **AP-37,** 384–389 (1989).
37. C.-C. Su, "The three-dimensional algorithm of solving the elec-

tric field integral equation using face-centered node points, conjugate gradient method, and FFT," IEEE Trans. Microwave Theory Tech. **41,** 510–515 (1993).
38. R. Barrett, M. Berry, T. F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. van der Vorst, *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods* (Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1994).
39. A. B. Samokhin, "Integral equations of the electrodynamics for three-dimensional structure and iterative method of solving them," J. Commun. Technol. Electron. **38,** 15–34 (1993).
40. D. Molenda, C. Höppener, H. Fuchs, and A. Naber, Physics Institute, University of Münster, Wilhelm-Klemen Strasse 10, D-48149 Münster, Germany (personal communication, 2002).
41. C. Höppener, D. Molenda, H. Fuchs, and A. Naber, "Simultaneous topographical and optical characterization of near-field optical aperture probes by way of imaging fluorescent nanospheres," Appl. Phys. Lett. **80,** 1331–1333 (2002).
42. A. Naber, D. Molenda, U. C. Fischer, H.-J. Maas, C. Höppener, N. Lu, and H. Fuchs, "Enhanced light confinement in a near-field optical probe with a triangular aperture," Phys. Rev. Lett. **89,** 210801 (2002).
43. A. D. Boardman, ed., *Electromagnetic Surface Modes* (Wiley, New York, 1982).
44. K. Tanaka and M. Tanaka, "Simulation of an aperture in the thick metallic screen that gives high intensity and small spot size using surface plasmon polariton," J. Microsc. (Oxford) **210,** 294–300 (2003).