Theoretical analysis of the surface wave along a metal-dielectric interface

W. Dai
Ames Laboratory, U.S. DOE, and Department of Physics and Astronomy, Iowa State University, Ames, Iowa 50011, USA

C. M. Soukoulis*
Ames Laboratory, U.S. DOE, and Department of Physics and Astronomy, Iowa State University, Ames, Iowa 50011, USA and Institute of Electronic Structure and Lasers (IESL), FORTH, and Department of Material Science and Technology, University of Crete, Heraklion, 71110 Crete, Greece

(Received 16 July 2009; revised manuscript received 2 September 2009; published 2 October 2009)

The metal-dielectric interface supports surface plasmons. But the metal-dielectric interface with defects has not only surface plasmons but also residual waves. In this paper we calculate the fields along the metal-dielectric interface with defects from Maxwell's equations analytically using the surface impedance approximation and study the asymptotic behavior of the residual waves. These analytic results set up a solid foundation to understand various phenomena such as beaming and extraordinary optical transmission.

DOI: 10.1103/PhysRevB.80.155407 PACS number(s): 78.66.Bz, 42.25.Fx, 73.20.Mf, 42.25.Bs

In 1998, Ebbesen *et al.*[1] reported a phenomenon named as extraordinary optical transmission (EOT): the transmission through a metal film perforated with two-dimensional subwavelength hole arrays can be much larger than the standard aperture theory prediction.[2] Served as simpler models, transmission through one-dimensional subwavelength metallic slits is studied too, which include the cases of gratings,[3–5] one-slit structures,[6,7] double-slit structures,[8] and one-slit structures with grooves.[9] It is commonly accepted that the surface plasmons excited along the input and the output surfaces play an important role in the transmission.

Surface plasmons[10] are electromagnetic waves bounded along a metal-dielectric interface through their interaction with the free electrons in the metal. In the two-dimensional $x-y$ space, suppose $x=0$ is the metal-dielectric interface. Then the surface plasmon has H polarization (E_x, E_y, and H_z are the nonzero components of the electromagnetic fields) and the field along the interface is $H_z(y) \propto \exp(ik_{sp}y)$.

The surface-plasmon explanation was criticized in some aspects. First, it was found that perfect electric conductors (PECs) also support EOT (Refs. 11 and 12) and a flat PEC surface does not support surface plasmons. Second, surface plasmon describes the fields along a flat metal-dielectric interface precisely but not the fields along the interface with indentations (slits or grooves). To get better fitting results in the area close to an indentation, Gay *et al.*[13] gave up the surface-plasmon explanation and assumed that the surface wave excited by an indentation has the form $H_z(y) \propto (\kappa/y + \mu)\cos(k_{surf}y + \phi)$ (the indentation locates at $y=0$). Other groups also realized the surface plasmon alone is not enough.[14–17] They decomposed the surface wave into two parts: $H_z = H_{sp} + H_c$. H_{sp} is the surface plasmon and H_c is the residual quasicylindrical wave. It was verified numerically that $|H_c| \propto 1/y^{1/2}$ when y is small.[14,15]

In this paper, we calculate the surface wave along the metal-dielectric interface with indentations from Maxwell's equations analytically. We support the decomposition $H_z = H_{sp} + H_c$ and study the asymptotic behavior of the quasicylindrical wave. We explain the connection between the surface waves along the metal surface and PEC surface.

Figure 1 shows the structure we are studying: a dielectric-metal interface with grooves and slits. The left half-plane is metal and the right half-plane is dielectric. Surface impedance boundary condition (SIBC) (Ref. 18) is used to describe the interface here. SIBC is commonly used to replace the metal when $|\epsilon_m \mu_m| \gg |\epsilon_d \mu_d|$. Here ϵ_m and μ_m are relative permittivity and permeability of the metal; ϵ_d and μ_d are of the dielectric. Silver has relative permittivity $\epsilon_m = -8.5 + 0.76i$ at $\lambda = 500$ nm (λ is the wavelength in free space).[19] So the silver-air interface can be described by SIBC up to 500 nm. When λ becomes smaller, $|\epsilon_m|$ becomes smaller and we have to give up the approximation. When $|\epsilon_m|$ goes to infinity, the fields calculated from SIBC converge to the real fields.

SIBC says the fields along the metal-dielectric interface satisfy

$$E_{\parallel} = Z_m \hat{n} \times H_{\parallel}.$$

Here $\hat{n}$ is a unit vector normal to the interface pointing into the dielectric; $Z_m = \sqrt{\mu_m \mu_0 / \epsilon_m \epsilon_0}$. Along the metal-dielectric interface $x=0$ in Fig. 1, the boundary condition is $E_y + Z_m H_z = 0$.

Suppose no incident waves coming from the dielectric side in Fig. 1, using the mode expansion method,[20] we can prove the surface fields along $x=0$ satisfy

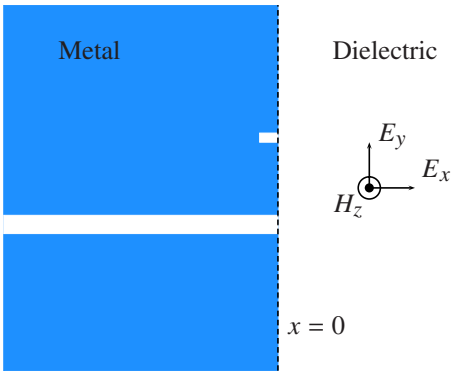

FIG. 1. (Color online) Schematic representation of a dielectric-metal interface with indentations in a two-dimensional space.

$$H_z(y) = \frac{1}{Z_d \lambda_d} \int_{-\infty}^{\infty} dy' (E_y(y') + Z_m H_z(y')) G(y,y'). \quad (1)$$

The Green's function is

$$G(y,y') = G(y-y') = \int_{-\infty}^{\infty} dk \frac{e^{ik(y-y')}}{k_x + Z_s k_d}. \quad (2)$$

Here $Z_s = Z_m / Z_d$ and $Z_d = \sqrt{\mu_d \mu_0 / \epsilon_d \epsilon_0}$, $\arg(Z_s) \in (-\pi/2, -\pi/4)$. $Z_s = 0$ when the metal is PEC. k_d and λ_d is the wave vector and the wavelength in the dielectric. k_x is the x component of the wave vector of the plane waves in the $x > 0$ half-plane. The plane waves propagate or decay along $+x$ direction, so

$$k_x = \begin{cases} \sqrt{k_d^2 - k^2} & \text{if } k_d \geq k \\ i\sqrt{k^2 - k_d^2} & \text{if } k_d < k \end{cases} \quad (3)$$

Since $E_y(y') + Z_m H_z(y') = 0$ at the metal surface, Eq. (1) is equivalent to the expression

$$H_z(y) = \frac{1}{Z_d \lambda_d} \int_{Indentation} dy' (E_y(y') + Z_m H_z(y')) G(y,y'). \quad (4)$$

So the surface wave is decided by the fields in the indentations. When the width of an indentation is much smaller than the wavelength, the fields at the exit of the indentation are close to a delta function. Then the Green's function describes the surface wave excited by the indentation. If the indentation is wide, we have to know the fields inside the indentation to calculate the surface wave. Though the surface wave around the indentation is complex, the Green's function still describes the surface wave in the region more than several widths away from the indentation.

Above we have shown the surface wave along the output surface of slits. It is easy to prove the fields along the input surface have the similar form. They are the summation of the surface wave excited by indentations plus some terms related to the incident and reflected waves. The surface wave excited by indentations can also be calculated by other methods,[21] but the Green's function method uses less assumptions and gets precise results. We have no need to assume the existence of surface plasmons beforehand. They will emerge from the Green's function automatically.

Define $y = s\lambda_d$, $k = hk_d$, then

$$G(y) = g(y/\lambda) = g(s) = \int_{-\infty}^{\infty} dh \frac{e^{i2\pi hs}}{\sqrt{1-h^2} + Z_s}.$$

Now let's focus on the nondimensional Green's function $g(s)$. The square-root function is defined as $\arg(\sqrt{1-h^2}) \in (-\pi/2, \pi/2]$ because of Eq. (3). $g(s)$ is an even function. We choose $s > 0$ here.

When $s \to 0$, $g(s)$ diverges but

$$\int_{-\infty}^{\infty} dh \left(\frac{e^{i2\pi hs}}{\sqrt{1-h^2} + Z_s} - \frac{e^{i2\pi hs}}{\sqrt{1-h^2}} \right)$$

does not. So

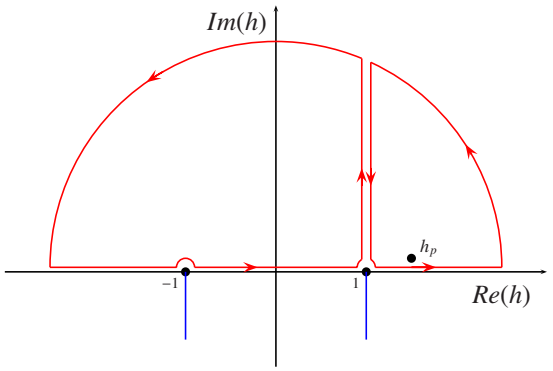

FIG. 2. (Color online) Integration path of the nondimensional Green's function in the complex space of h. The red curves are the integration path and the blue lines are the branch cut.

$$g(s) \to \pi H_0^{(1)}(2\pi s) \quad \text{when } s \to 0. \quad (5)$$

We moves to the complex space of h to do the integral in $g(s)$. The integration path is shown in Fig. 2 as the red curves with arrows. The blue lines are the branch cut to make $\sqrt{1-h^2}$ single valued. The function $\sqrt{1-h^2}$ is not continuous along the path $\mathrm{Im}(h) = 0$ at the point $h = 1$ under this branch cut. So the integration path has two loops.

Using Cauchy's integral formula to the left and right loops, we get

$$\int_{-\infty}^{1} dh \frac{e^{i2\pi hs}}{\sqrt{1-h^2} + Z_s} + \int_{0}^{\infty} dq \frac{ie^{i2\pi s} e^{-2\pi qs}}{\sqrt{q^2 - 2qi} + Z_s} = 0,$$

$$\int_{\infty}^{0} dq \frac{ie^{i2\pi s} e^{-2\pi qs}}{-\sqrt{q^2 - 2qi} + Z_s} + \int_{1}^{\infty} dh \frac{e^{i2\pi hs}}{\sqrt{1-h^2} + Z_s} = 2\pi i \frac{Z_s}{h_p} e^{i2\pi h_p s}.$$

Here h_p is a pole: $\sqrt{1-h_p^2} + Z_s = 0$. The integral variable is changed from h to q along the vertical parts of the path, $h = 1 + qi$ and $q \in [0, \infty)$. Again $\arg(\sqrt{1-h^2}) \in (-\pi/2, \pi/2]$, $\arg(\sqrt{q^2 - 2qi}) \in (-\pi/2, \pi/2]$.

Then

$$g(s) = 2\pi i \frac{Z_s}{h_p} e^{i2\pi h_p s} + 2ie^{i2\pi s} \int_0^{\infty} dq e^{-2\pi qs} \frac{\sqrt{q^2 - 2qi}}{Z_s^2 - q^2 + 2qi}.$$

In the second term, the integrand contributes significantly only when $q \in (0, 1/s)$. So we neglect q^2 when $s \gtrsim 1$. Then

$$g(s) \approx 2\pi i \frac{Z_s}{h_p} e^{i2\pi h_p s} + 2ie^{i2\pi s} \int_0^{\infty} dq e^{-2\pi qs} \frac{\sqrt{-2qi}}{Z_s^2 + 2qi}. \quad (6)$$

All the calculations below base on the approximation expression of Eq. (6), which is good when $s > 1$. But our numeric simulation results show the approximation works very well even when s is smaller than 1. We will also prove that Eq. (6) gives the same asymptotic result as Eq. (5) when s is very small. So the approximation in Eq. (6) is precise for almost any s.

The first term in Eq. (6) is the well-understood surface plasmon $g_{SP}(s)$. The second is the residual quasicylindrical wave $g_C(s)$,[17] which need further study. Define

$$I(s) = \int_0^\infty dq e^{-2\pi q s} \frac{\sqrt{q/Y}}{Y+q}; \quad Y = Z_s^2/(2i). \tag{7}$$

Then

$$I(s) = \sqrt{\frac{1}{2Ys}} - \pi e^{2\pi Ys} \mathrm{Erfc}(\sqrt{2\pi Ys}); \tag{8}$$

$$g_C(s) = \sqrt{Y}(1-i)e^{i2\pi s} I(s). \tag{9}$$

Here $\arg(\sqrt{Y}) \in (\pi/4, \pi/2)$; Erfc is the complementary error function; $|I(s)|$ describes the envelope of the quasicylindrical wave.

The integrand of $I(s)$ in Eq. (7) contributes significantly only when $0 < q \lesssim 1/s$. When $|Y|s \gg 1$, q in the interval $(0, 1/s)$ satisfies

$$\frac{1}{Y+q} \approx \frac{1}{Y}\left(1 - \frac{q}{Y}\right)$$

$$\Rightarrow I(s) \approx \frac{1}{4\sqrt{2}\pi} \frac{1}{(Ys)^{3/2}} - \frac{1}{16\sqrt{2}\pi^2} \frac{1}{(Ys)^{5/2}}.$$

When $|Y|s$ is small, $I(s)$ can be simplified by the Taylor expansion of $e^{u^2}\mathrm{Erfc}(u)$,

$$e^{u^2}\mathrm{Erfc}(u) = 1 - \frac{2}{\sqrt{\pi}}u + u^2 + \cdots.$$

We can call $|Y|s$ as the surface distance because it describes how the envelope of the quasicylindrical wave evolves during the propagation. The piecewise function below gives good estimation about $I(s)$.

$$I(s) \approx \begin{cases} \sqrt{1/2Ys} & \text{if } |Y|s < 0.002 \\ \sqrt{1/2Ys} - \pi & \text{if } |Y|s < 0.02 \\ 1/[4\sqrt{2}\pi(Ys)^{3/2}] & \text{if } |Y|s > 1 \end{cases}.$$

When $0.02 < |Y|s < 1$, we have to include more terms of the Taylor expansion or we can calculate Erfc function directly.

We now compare the amplitude of g_{SP} and g_C. Since g_{SP} decays exponentially, g_C will be stronger for sure when s is very big. But in practice, h_p is close to 1 with a very small imaginary part for good noble metals. So $|g_{SP}(s)/g_C(s)| \approx |I(s)/2\pi|$. Figure 3 shows $|g_{SP}(s)/g_C(s)|$ along a metal-air interface when $\epsilon_m = -31.39 + 2.22i$, which is the relative permittivity of silver at $\lambda = 852$ nm.[19] If the metal is lossless, $I(s)$ is a function of $|Y|s$, which is the horizontal axis in Fig. 3. We can find the quasicylindrical wave is negligible when $|Y|s > 0.1$.

Put the results together, we have

$$g_{SP}(s) = 2\pi i \frac{Z_s}{h_p} e^{i2\pi h_p s},$$

$$g_{C0}(s) = (1-i)e^{i2\pi s} \frac{1}{\sqrt{2s}},$$

$$g_{C1}(s) = -i\pi Z_s e^{i2\pi s},$$

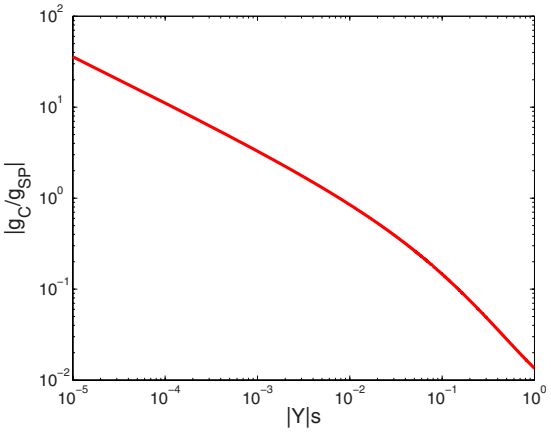

FIG. 3. (Color online) $|g_C(s)/g_{SP}(s)|$ along a metal-air interface when the metal permittivity is $\epsilon_m = -31.39 + 2.22i$.

$$g_{CB}(s) = Z_s \frac{i}{4\sqrt{2}\pi} \frac{e^{i2\pi s}}{(Ys)^{3/2}}.$$

The piecewise function below gives good approximation about the Green's function,

$$g(s) \approx \begin{cases} g_{SP}(s) + g_{C0}(s) & \text{if } |Y|s < 0.002 \\ g_{SP}(s) + g_{C0}(s) + g_{C1}(s) & \text{if } |Y|s < 0.02 \\ g_{SP}(s) + g_{CB}(s) & \text{if } |Y|s > 0.1 \end{cases}.$$

The metal-dielectric interface is divided into three regions based on the surface distance: the near region ($|Y|s < 0.02$), the intermediate region ($0.02 < |Y|s < 0.1$) and the far region ($|Y|s > 0.1$).

If the metal is nearly lossless, g_{SP} dominates in the far region. If the metal is a PEC, $Z_s = 0$ and $Y = 0$. The whole surface belongs to the near region and

$$g(s) \approx g_{C0} = (1-i)e^{i2\pi s} \frac{1}{\sqrt{2s}}.$$

It agrees with the PEC's Green's function with Hankel function form.[11] When s is very small, $|g_{C0}| \gg |g_{SP}|$, we can neglect g_{SP} and return to the result in Eq. (5).

The fields along a metal-air interface are calculated numerically using the commercial finite element method software COMSOL Multiphysics. The simulation area is similar to Fig. 1 but without the groove. The incident wave comes through the slit in $+x$ direction. The fields along the output surface are recorded and compared with analytic results in Figs. 4 and 5. The slit width is 0.05λ. The relative error of the simulations is around 2%. The analytic results in the figures are calculated by Eq. (4) by replacing $G(y, y')$ with the different simplified forms of the Green's function shown in the legends. We assume that $E_y(y') + ZH_z(y')$ is constant at the exit of the slit and the value is taken from the numerical simulations. There are no fitting parameters in the analytic calculations.

In Figs. 4 and 5, the real and imaginary parts of H_z are plotted in the range $y \in (0.1\lambda, 40\lambda)$ at two incident wavelengths. We always plot the simulation results along with the analytic results from $g_{SP} + g_C$ and the two curves always

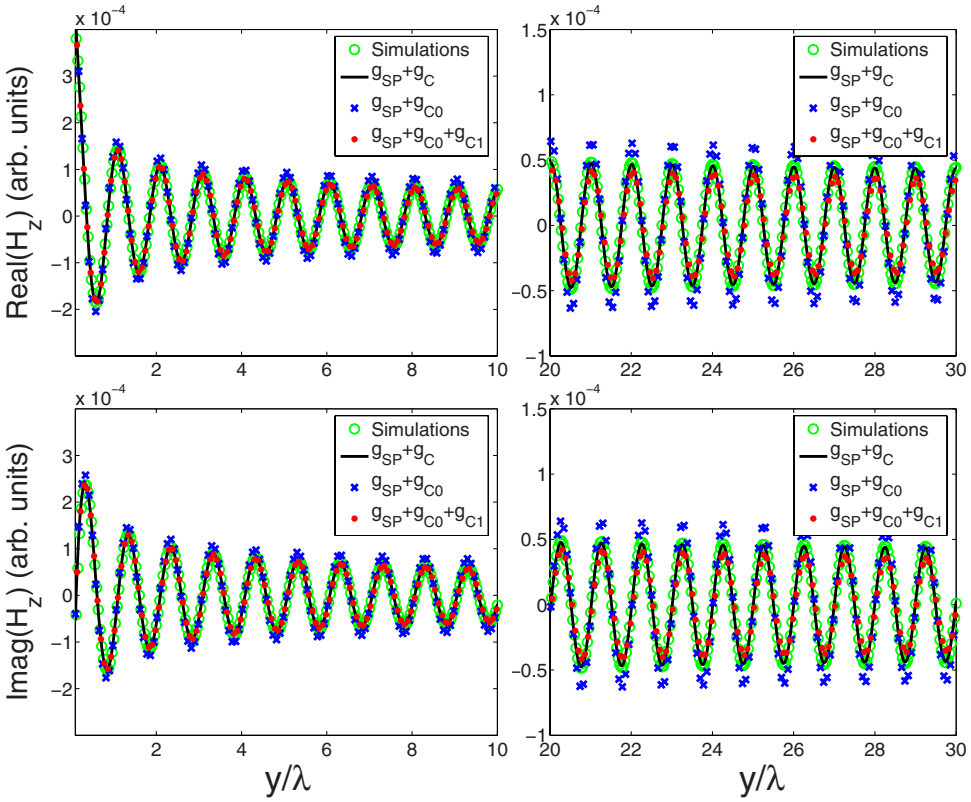

FIG. 4. (Color online) Real and imaginary parts of H_z along a silver-air interface in the regions $y \in (0.1\lambda, 10\lambda)$ and $y \in (20\lambda, 30\lambda)$ when $\lambda = 3000$ nm and $\epsilon_m = -329 + 47.5i$. The simulation results are calculated by COMSOL Multiphysics, a commercial finite element method software. The other curves are calculated analytically using different simplified forms of the Green's function shown in the legends.

overlap. It shows the approximation in Eq. (6) is good in the region $y > 0.1\lambda$.

Figure 4 shows the fields when $\lambda = 3000$ nm and $\epsilon_m = -329 + 47.5i$. The near region is the area $y \lesssim 12\lambda$. The analytic results from $g_{SP} + g_{C0} + g_{C1}$ agree with the simulation results very well in this region and g_{C1} component do improve the analytic results. When s enters the intermediate region, results from $g_{SP} + g_{C0}$ and $g_{SP} + g_{C0} + g_{C1}$ have visible difference from the simulation results. The difference looks small here because g_{SP} is strong. In Fig. 5, the incident

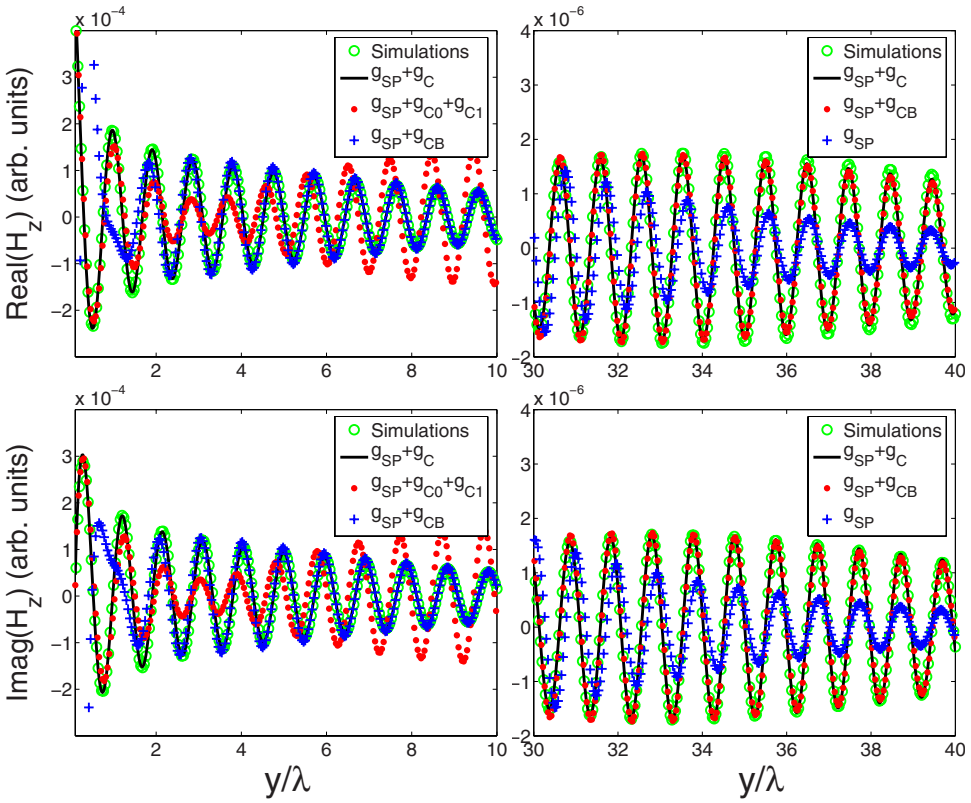

FIG. 5. (Color online) Real and imaginary parts of H_z along a metal-air interface in the regions $y \in (0.1\lambda, 10\lambda)$ and $y \in (30\lambda, 40\lambda)$ when $\lambda = 500$ nm and $\epsilon_m = -8.50 + 6.00i$. The simulation results are calculated by COMSOL Multiphysics. The other curves are calculated analytically using different simplified forms of the Green's function shown in the legends.

wavelength is $\lambda=500$ nm. The permittivity of the metal is set as $\epsilon_m=-8.50+6.00i$ with a big imaginary part to suppress the surface plasmon in the far region. Since $|Y|$ is big here, the near and intermediate regions are short. $g_{SP}+g_{CB}$ gives good analytic results when $y>2\lambda$. Figure 5 shows clearly that the quasicylindrical wave decays as $y^{-3/2}$ in the far region.

Our analytic calculation also proves $g_{SP}(s)+g_{C0}(s)$ is a very good approximation of $g(s)$ for a good noble metal at visible wavelengths.[15–17] It is the approximation we used in the near region. In the far region, we have proven $g_C(s)$ is negligible; so is $g_{C0}(s)$. In the intermediate region, this approximation is poor. But the intermediate region is short and it is difficult to notice the fitting error.

In EOT research, the interesting surface area is normally between $\lambda/2$ and 100λ. For a good noble metal at visible wavelengths, the area belongs to the far region and surface plasmon dominates. When wavelength increases and ϵ_m is more negative, the metal surface converges to the PEC surface in two ways: (i) the near region becomes longer and (ii)g_{C0} becomes stronger comparing with g_{SP} in the near region. The whole surface works like a mixture of PEC and metal surfaces and $|Y|$ servers as a good index to describe the mixture state. Reference 14 shows the converging process graphically. Both metal and PEC surfaces support the strong slow-decaying surface waves; so they have similar phenomena such as EOT.

In this paper, we have analyzed the surface wave along a metal-dielectric interface excited by an indentation. We get the asymptotic forms of the wave far away from and close to the indentation. Based on the surface distance $|Y|s$, the interface is divided into several regions and the surface wave behaves differently in every region. The complete description of the surface wave would give us deeper understanding about light transmission mechanics through metallic apertures.

Work at Ames Laboratory was supported by the Department of Energy (Basic Energy Sciences) under Contract No. DE-AC02-07CH11358. This work was partially supported by the AFOSR under MURI Grant No. FA9550-06-1-0337.

*soukoulis@ameslab.gov

[1] T. W. Ebbesen, H. J. Lezec, H. F. Chaemi, T. Thio, and P. A. Wolff, Nature (London) **391**, 667 (1998).
[2] H. A. Bethe, Phys. Rev. **66**, 163 (1944).
[3] J. A. Porto, F. J. García-Vidal, and J. B. Pendry, Phys. Rev. Lett. **83**, 2845 (1999).
[4] Q. Cao and P. Lalanne, Phys. Rev. Lett. **88**, 057403 (2002).
[5] K. G. Lee and Q. Han Park, Phys. Rev. Lett. **95**, 103902 (2005).
[6] Y. Takakura, Phys. Rev. Lett. **86**, 5601 (2001).
[7] F. Yang and J. R. Sambles, Phys. Rev. Lett. **89**, 063901 (2002).
[8] H. F. Schouten, N. Kuzmin, G. Dubois, T. D. Visser, G. Gbur, P. F. A. Alkemade, H. Blok, G. W. 't Hooft, D. Lenstra, and E. R. Eliel, Phys. Rev. Lett. **94**, 053901 (2005).
[9] O. T. A. Janssen, H. P. Urbach, and G. W. 't Hooft, Phys. Rev. Lett. **99**, 043902 (2007).
[10] H. Raether, *Surface Plasmons on Smooth and Rough Surfaces and on Gratings* (Springer-Verlag, Berlin, 1988).
[11] F. J. García-Vidal, H. J. Lezec, T. W. Ebbesen, and L. Martín-Moreno, Phys. Rev. Lett. **90**, 213901 (2003).
[12] J. W. Lee, M. A. Seo, D. S. Kim, S. C. Jeoung, Ch. Lienau, J. H. Kang, and Q.-Han Park, Appl. Phys. Lett. **88**, 071114 (2006).
[13] G. Gay, O. Alloschery, B Viaris De Lesegno, C. O'Dwyer, J. Weiner, and H. J. Lezec, Nat. Phys. **2**, 262 (2006).
[14] P. Lalanne and J. P. Hugonin, Nat. Phys. **2**, 551 (2006).
[15] L. Chen, J. Robinson, and M. Lipson, Opt. Express **14**, 12629 (2006).
[16] L. Aigouy, P. Lalanne, J. P. Hugonin, G. Julíe, V. Mathet, and M. Mortier, Phys. Rev. Lett. **98**, 153902 (2007).
[17] X. Y. Yang, H. T. Liu, and P. Lalanne, Phys. Rev. Lett. **102**, 153903 (2009).
[18] J. D. Jackson, *Classical Electrodynamics*, 2nd ed. (Wiley, New York, 1975).
[19] E. D. Palik, *Handbook of Optical Constants of Solids* (Academic, New York, 1985).
[20] F. J. García-Vidal and L. Martín-Moreno, Phys. Rev. B **66**, 155412 (2002).
[21] P. Lalanne, J. P. Hugonin, and J. C. Rodier, Phys. Rev. Lett. **95**, 263902 (2005).