PHYSICAL REVIEW B **83**, 035414 (2011)

Transmission and scattering properties of subwavelength slits in metals

M. Gorkunov,[1] E. Podivilov,[2] and B. Sturman[2]
[1]*Shubnikov Institute of Crystallography, Russian Academy of Sciences, RU-119333 Moscow, Russia*
[2]*Institute of Automation and Electrometry, Russian Academy of Sciences, RU-630090 Novosibirsk, Russia*

(Received 6 September 2010; revised manuscript received 3 November 2010; published 20 January 2011)

Using an advanced eigenmode-expansion method, we analyze the basic transmission and scattering properties of subwavelength slits in real metals characterized by the complex optical permittivity ε_m. This includes the slit-width, wavelength, and ε_m dependences of the efficiencies and cross sections of the main transformation processes: (i) transformation of the incident plane wave into the propagating mode, into the surface plasmons, and into diffracted waves in air and (ii) internal reflection of the propagating mode and its transformation to the surface plasmons and diffracted waves. In conjunction with the known perfect-metal-related efficiencies, the established dependences exhibit a wealth of important subwavelength features, including the nontrivial transmission peculiarities. Transition from the case of the periodic array of slits to the single-slit case when increasing the metal-wall width is considered as well. The established characteristics of a single interface between air and perforated metal are sufficient to describe the extraordinary light transmission through subwavelength slits in metal films.

DOI: 10.1103/PhysRevB.83.035414 PACS number(s): 42.25.Bs, 42.25.Fx, 42.70.Qs, 73.20.Mf

I. INTRODUCTION

Since the experimental discovery in 1998,[1] the phenomenon of the extraordinary light transmission through subwavelength holes in metals attracts a great research interest.[2–9] The subject is of great fundamental importance because the basic properties of metal-based light-transmitting nanostructures appeared to be under-investigated and full of fascinating effects. It is promising for applications because of the possibility of strong light confinement as well as in connection with the studies of metamaterials.

While the above research area is topical, it has deep historical roots in the works of Rayleigh, Wood, Sommerfeld, and Bethe.[10–15] It is well known that exact solutions of Maxwell's equations for metal-based nanostructures are available only for a few of the simplest cases in the perfect-metal limit. The direct (finite-difference) numerical methods often suffer from bad convergence, must deal with the so-called corner singularities, and can hardly be considered as reliable tools to achieve physical insights. The most adequate/physical tools of numerical analysis are based on the eigenmode expansions.

It became clear a few years ago that the basic concept of the eigenmodes in wave-guiding structures, which qualifies these modes into propagating and evanescent, is not fully applicable to metal-based structures.[16,17] Importantly, the eigenvalue problem for metal-based wave-guiding structures (single holes/slits, arrays of holes/slits, photonic crystals, etc.) is not Hermitian, even in the absence of light absorption, i.e., for real negative values of the optical permittivity of the metal ε_m. In addition to the propagating and evanescent modes, there is a sequence of anomalous eigenmodes with essentially complex eigenvalues, which is crucial for numerical simulations.[17,18] These modes can also be traced back to some early theoretical studies.[19]

The distinctive feature of the metal-based structures, which was recognized early, is the excitation of the surface plasmons (surface-plasmon polaritons) by the incident light.[5,20] The surface plasmon (SP) can propagate over long distances, coupling neighboring holes/slits and affecting the light-transmission properties. The presence of holes/slits on a metal surface modifies the SP properties, making them essentially different from the properties of flat-surface plasmons.[17,21,22] Furthermore, it was found recently[23–25] that a weak power-law decaying asymptotic wave (Norton wave) can become dominating over the SP for very long propagating distances.

Earlier, we applied the modified eigenmode approach to investigate the light-transmission and reflection properties of a periodic array of nanoslits in a metal slab.[18] This periodic structure [see Fig. 1(a)] is characterized by the slit half-width l, the wall half-width L, the half-period $w = l + L$, the slab thickness z_0, and the optical permittivity of metal $\varepsilon_m = \varepsilon_m' + i\varepsilon_m''$. The slab transmittance was expressed by the transmittance of a single interface. This fundamental characteristic possesses highly remarkable dependences on the key input parameters, the period-to-wavelength ratio $2w/\lambda$, the wall-to-slit ratio L/l, and $|\varepsilon_m|$. The sharp spectral features found have close links with the Rayleigh-Wood anomalies and surface-plasmon resonance.

Here we attack another fundamental problem of subwavelength optics of metals: the problem of light transmission, scattering, and SP excitation for a single slit [see Fig. 1(b)]. This differs substantially from the above periodic problem. First, we have a continuous angular distribution of scattered light waves in air. Second, the excited surface plasmons leave the slit area to decay far from it, causing no interference effects. Third, a large surface area (larger in size than the plasmon decay length) must be covered by our considerations.

Complexity of the single-slit problem for real metals is a serious issue. The spectrum of the eigenvalues includes (in contrast to the periodic problem) a continuous part, and the corresponding eigenfunctions do not decay for $|x| \to \infty$, even in the presence of light absorption.[17] This feature ensures the skin-depth penetration of light into the metal far from the slit and also the formation of the surface plasmons within the eigenmode-expansion method. On the other hand, the piecewise smooth structure of the eigenfunctions inevitably leads to the singularity of the integral equations for the eigenmode amplitudes: these equations include the principal

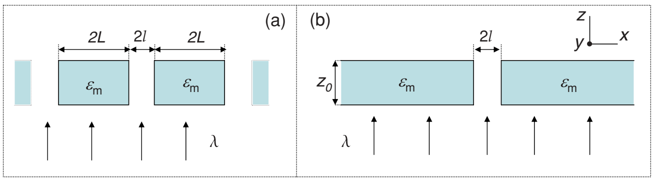

FIG. 1. (Color online) The periodic (a) and single-slit (b) structures; l and L are the slit and wall half-width, z_0 is the slab thickness, and ε_m is the optical permittivity of the metal.

integral value. Being mathematically correct, the singular integral equations are very difficult for numerical treatments. Any calculation method must deal, directly or indirectly, with this issue. In the perfect-metal case, the problem of singular integral equations is absent.[26]

Within our eigenmode-expansion method, we resolve the singularity by analyzing the transition from the periodic to the single-slit case when increasing the metal-wall width $2L$. This ensures correctness of the calculation procedure and clarifies the physics of this transition.

Our main findings concern the efficiencies (cross sections) of the elementary transformation processes (transmission, reflection, diffraction, and SP excitation) as functions of l/λ and ε_m in the subwavelength range. These basic interface characteristics give a true insight into the physics of metal-based subwavelength phenomena. The knowledge of these efficiencies also allows us to calculate easily the transmission and diffraction properties of a perforated metal slab.[3,18] This reductional approach has certain features in common with the microscopic theory of subwavelength phenomena.[9,27].

Despite a large number of publications on the single-slit properties,[26,28–41] surprisingly little is known about the elementary transformation processes. Two reasons can be indicated. First, there is a large variety of subwavelength phenomena. Second, there is a tendency to unify diverse physical factors into numerical examples, which typically cover a small part of the space of variable parameters of the system. The closest (in the subject matter) theoretical studies are the investigation of the SP-excitation efficiency in almost perfect conductors[36,37] and the analysis of the transformation processes in the perfect metal.[26]

II. BASIC RELATIONS

In accordance with Fig. 1, the normally incident plane wave is TM polarized, so that the total magnetic field has only the y component $H(x,z)\exp(-i\omega t) + \text{c.c.}$, where $\omega = 2\pi c/\lambda$, c is the speed of light, λ is the wavelength, and c.c. indicates complex conjugation. The light electric field has x and z components, which can readily be expressed by the amplitude H from Maxwell's equations.

A. Eigenmodes

The central point of our theory is an eigenmode expansion of $H(x,z)$ in the perforated-metal part. To specify a full set of eigenfunctions $\{h(x)\}$, we search for a solution of Maxwell's equations in the form $H(x,z) = h(x)\exp(i\beta z)$, where β is the propagation constant. The function $h(x)$ obeys the differential equation

$$\left(\varepsilon \frac{d}{dx}\frac{1}{\varepsilon}\frac{d}{dx} + \varepsilon k_0^2\right) h = \beta^2 h, \tag{1}$$

where $\varepsilon = \varepsilon(x)$ is $\varepsilon_d = 1$ and $\varepsilon_m \equiv \varepsilon'_m + i\varepsilon''_m$ in the dielectric and metal regions, respectively, and $k_0 = 2\pi/\lambda$ is the vacuum wave vector. The case $\varepsilon_d \neq 1$ can, if needed, be reduced to the case $\varepsilon_d = 1$ via a scaling procedure.[17]

Using the conventional boundary conditions for h and dh/dx at the metal walls, the condition of periodicity for $h(x)$, and the condition of parity $h(x) = h(-x)$ about a slit center, which is based on the symmetry grounds, we arrive at the dispersion relation for the eigenvalue β^2 (Ref. 17):

$$p_d \varepsilon_m \tan(p_d l) + p \tan(pL) = 0\,, \tag{2}$$

where $p = \sqrt{\varepsilon_m k_0^2 - \beta^2}$ and $p_d = \sqrt{p^2 + (1-\varepsilon_m)k_0^2}$. This algebraic equation gives a discrete infinite set of solutions for β^2 and, correspondingly, a double set for $\beta = \pm\sqrt{\beta^2}$. The sign symmetry is due to the presence of two equivalent propagation directions $\pm z$. The values of β, whose imaginary part β'' is positive/negative, correspond to solutions for $H(x,z)$ decaying for $z \to \pm\infty$.

Alternatively, one can use the quantity p as the eigenvalue and represent the propagating constant as $\beta(p) = (\varepsilon_m k_0^2 - p^2)^{1/2}$. This choice of the spectral parameter is convenient for what follows.

Figure 2(a) gives an example of the spectrum of $\beta = \beta' + i\beta''$ in the region $\beta'' > 0$ for a periodic nano-structure. The chosen value of ε_m is representative for silver at $\lambda \approx 500$ nm.[42,43]

First, we have a single propagating mode with an almost real β_0, a single evanescent mode with an almost imaginary β_1, and a discrete sequence of the anomalous modes with essentially complex eigenvalues. Additionally, we have a quasicontinuous spectrum, i.e., an infinite set of closely situated values of β with $\beta' \ll \beta''$ in the region $\beta''/k_0 > \sqrt{|\varepsilon_m|}$ (above the cutoff). Figure 2(b) shows the corresponding spectrum of p. The propagating and evanescent modes are characterized here by

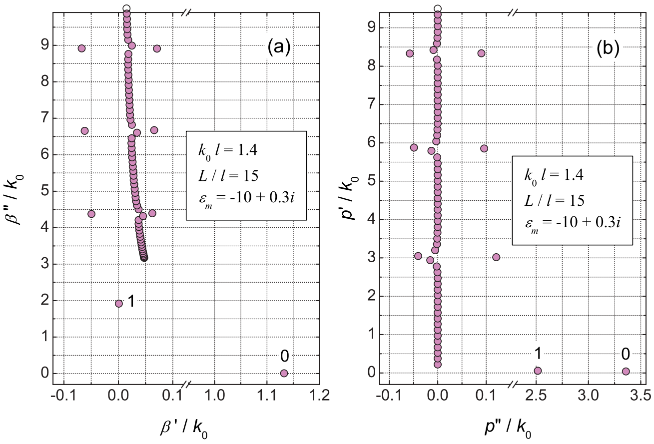

FIG. 2. (Color online) The eigenvalues for the periodic problem with $\varepsilon_m = -10 + 0.3i$, $k_0 l = 1.4$, and $L/l = 15$ in the β representation (a) and in the p representation (b). The numbers 0 and 1 refer to the propagating and evanescent modes. Note the breaks on the horizontal axes and different axes titles in (a) and (b).

almost imaginary values p_0 and p_1, while the quasicontinuous part of the spectrum consists of almost real roots. Three first pairs of the anomalous modes are clearly seen as well.

The discrete modes (propagating, evanescent, and anomalous) are greatly related to the single slit; the corresponding values of β (or p) remain practically unchanged by further increasing the wall half-width L. At the same time, the values of β (or p) for these modes essentially depend on the slit half-width l and also on ε_m. Let the slit width be decreasing. Then, the propagating root β_0 in Fig. 2(a) moves steadily to the right, while the evanescent root β_1 moves up, approaches the cutoff, and transforms into an anomalous mode. The anomalous roots in Figs. 2(a) and 2(b) move up and their vertical separation is increasing. If the slit width is increasing, we have a motion of the spectral dots in the opposite direction: the anomalous modes transform into the evanescent modes, the evanescent modes transform into the propagating modes, and the total number of the propagating modes increases. The dependence of the discrete spectrum on ε'_m is fairly weak. The second propagating mode appears, e.g., at $k_0 l \approx \pi/2$. For $\varepsilon''_m \to 0$, the values of β become purely real and imaginary for the propagating and evanescent modes, respectively. The anomalous roots remain complex, but they become strictly symmetric about the vertical axis in Figs. 2(a) and 2(b).

The quasicontinuous part of the spectrum is fairly simple. Most of the corresponding dots lie practically on the line $\beta'\beta''/k_0^2 = \varepsilon''_m/2$ with $\beta''/k_0 > |\varepsilon''_m|$, which represents the continuous spectrum of the single slit ($L \to \infty$). This line becomes strictly vertical for $\varepsilon''_m \to 0$. The vertical separation between the dots is $\pi/w \equiv \pi/(l+L)$ with a high accuracy. This separation decreases with increasing L, and we have a transition to the continuous spectrum. Within the p representation, the quasicontinuous part of the spectrum follows the line $p'' = 0$, $p' > 0$ even in the presence of light absorption.

The eigenfunctions can be found explicitly for our periodic structure.[17] Setting $x = 0$ at the slit center, we represent an even eigenmode within a half-period w as

$$h_p = \begin{cases} \cos(p_d x) & (0 \leqslant x \leqslant l), \\ c\,\cos[p(x-w)] & (l \leqslant x \leqslant w), \end{cases} \tag{3}$$

where $c = \cos(p_d l)/\cos(pL)$. Different eigenfunctions correspond to different values of p. According to Eq. (3), $h_p(0) = 1$; this normalization is chosen for convenience.

Numbering of the eigenmodes must be fixed. The spectral parameter p takes values p_ν with $\nu = 1, 2, \ldots$. The numeration order corresponds to a decrease of p'' (for the propagating modes) and then to an increase of p' [see Fig. 2(b)]. We set also $h_{p_\nu}(x) \equiv h_\nu(x)$ and $\beta_{p_\nu} \equiv \beta_\nu$.

Using Eq. (1) and the boundary conditions at the metal walls, it is not difficult to find that the following exact orthogonality relation takes place for $p \neq p'$:

$$\int_{-w}^{w} h_p(x)\, h_{p'}(x)\, \varepsilon^{-1}(x)\, dx = 0. \tag{4}$$

It is valid even for $\varepsilon''_m \neq 0$, does not include the sign of complex conjugation, and differs from the most well-known orthogonality relations. It can also be verified directly using Eqs. (2) and (3).

In the limit $L \to \infty$, we have a transition to the single-slit case. The quasicontinuous part of the spectrum becomes continuous with the real spectral parameter p ranging from 0 to ∞; the corresponding eigenfunctions are nonvanishing for $|x| \to \infty$. The discrete spectrum experiences practically no changes, and the corresponding eigenfunctions become localized.

III. TWO BASIC INTERFACE PROBLEMS

The following processes accompany the incidence of the light wave onto the periodic structure of Fig. 1(a): (i) reflection, i.e., the formation of a wave with the wave vector $-\boldsymbol{k}_0$; (ii) transmittance, i.e., the formation of a wave with the wave vector $\boldsymbol{k}_0$ at the output; (iii) diffraction, i.e., the generation of waves with the transversal components of the wave vectors $\pm\pi/w$, $\pm 2\pi/w, \ldots$ before and after the slab (the number of these diffraction orders grows with increasing L; for $L \to \infty$, the angular spectrum is continuous); (iv) SP excitation on the input and output faces (in the single-slit limit, the plasmons go the way of the slit to the right and left and decay); and (v) excitation of the propagating modes in the metal part (for subwavelength slits, we have only the modes with $\beta = \pm\beta_0$ propagating in the $\pm z$ directions).

Generally, these processes are mutually coupled. However, in the case of slabs, the thickness of which considerably exceeds the skin depth $k_0 z_0 \sqrt{|\varepsilon_m|} \gg 1$, the input and output interfaces are coupled only via a single propagating mode, and the general slab problem is reduced to the two basic interface problems depicted in Fig. 3. The first problem (*out*) deals with the incidence of a light wave from outside onto the semi-infinite structure [see Fig. 3(a)]. We are interested here in characterization of the reflectance, scattering, transmittance, and SP-excitation properties. The second problem (*in*) is about the processes that occur when the propagating mode reaches the interface from inside [see Fig. 3(b)]. These processes include the reflection back, the SP excitation, and the excitation of light waves at the output.

IV. COUPLED-MODE EQUATIONS

In the air region, where $z < 0$, we use the following exact Rayleigh expansion for $H(x,z)$:

$$H^< = a_0^+ e^{ik_0 z} + \sum_k a_k^-\, e^{-i\varkappa_k z} \cos(kx), \tag{5}$$

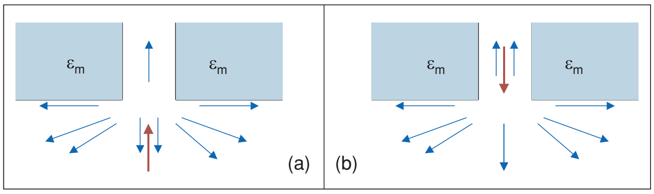

FIG. 3. (Color online) Two basic interface problems, (a) *out* and (b) *in*. (a) A plane wave, incident normally from outside onto the periodic structure (a small fragment is shown), reflects back, diffracts, and transforms into the propagation mode and SPs. (b) A propagating mode travels to the interface from inside, reflects back, and transforms into light waves in air and SPs.

where a_0^+ is the amplitude of the incident wave, $\varkappa_k = \sqrt{k_0^2 - k^2}$, and the transverse wave vector k takes the discrete values $0, \pi/w, 2\pi/w, \ldots$. The coefficients $\varkappa_k$ are real for $k \leqslant k_0 = 2\pi/\lambda$ and imaginary for $k > k_0$. The corresponding regions of k refer to the propagating and evanescent waves in air. The amplitude a_0^- characterizes the reflected wave [see Fig. 3(a)].

In the perforated-metal region, $z > 0$, and we employ the eigenmode expansion

$$H^> = b_0^- \, h_0(x) \, e^{-i\beta_0 z} + \sum_p b_p^+ \, h_p(x) \, e^{i\beta_p z}, \tag{6}$$

where $b_0^- \equiv b_{p_0}^-$ is the amplitude of the propagating mode incident onto the interface $z = 0$ from inside [see Fig. 3(b)], $\beta_p = (\varepsilon_m k_0^2 - p^2)^{1/2}$, and p takes the values p_ν, with $\nu = 0,1,\ldots$ in accordance with Eq. (2). All terms under the sum sign tend to zero for $z \to \infty$.

To consider the basic interface problems, *out* and *in*, it is sufficient to set $b_0^- = 0$, $a_0^+ = 1$ and $a_0^+ = 0$, $b_0^- = 1$, respectively, in the above relations (5) and (6).

Our nearest objective is to obtain closed equations for the amplitudes a_k^-. This can be achieved by a sequence of simple steps: First, we rewrite the boundary conditions $H_1(x,0) = H_2(x,0)$ and $H_{1,z}(x,0) = H_{2,z}(x,0)/\varepsilon(x)$ (the subscript z indicates the differentiation) in the terms of a_k and b_p using Eqs. (5) and (6). Second, we multiply the corresponding equalities by $h_p(x)/\varepsilon(x)$ and $h_p(x)$, respectively, integrate them over the half-period w, and obtain, using Eq. (4), two relations expressing the amplitude b_p^+ through the set $\{a_k^-\}$. Combining these relations, we arrive at the matrix equation

$$\sum_k T_{pk} \, a_k^- = A_p. \tag{7}$$

The interaction matrix T_{pk} and the effective driving force A_p are given generally by the relations

$$T_{pk} = \varkappa_k \, h_{pk} + \beta_p \, (h/\varepsilon)_{pk} \tag{8}$$

and

$$A_p = 2\beta_p \, \delta_{pp_0} \, (h^2/\varepsilon)_{p0} \, b_0^- + [k_0 \, h_{p0} - \beta_p \, (h/\varepsilon)_{p0}] \, a_0^+, \tag{9}$$

where $h_{pk} = \langle h_p \cos(kx) \rangle$, $(h/\varepsilon)_{pk} = \langle h_p \cos(kx)/\varepsilon(x) \rangle$, and $(h^2/\varepsilon)_{pk} = \langle h_p^2 \cos(kx)/\varepsilon(x) \rangle$ are the kth Fourier component of the periodic functions $h_p(x)$, $h_p(x)/\varepsilon(x)$, and $h_p^2(x)/\varepsilon(x)$; the brackets indicate averaging over the period. With the amplitudes a_k^- found, the amplitudes b_p^+ can be calculated from

$$b_p^+ = -b_0^- \, \delta_{pp_0} + \left(a_0^+ (h/\varepsilon)_{p0} + \sum_k a_k^- \, (h/\varepsilon)_{pk} \right) \Big/ (h^2/\varepsilon)_{p0}. \tag{10}$$

An equivalent way to find the amplitudes a_k and b_p is to deal with a matrix equation for b_p. This can be done in a similar manner. The only difference is multiplication of the initial equations (obtained from the boundary conditions) by $\cos(kx)$ instead of $h_p(x)/\varepsilon(x)$, and the use of the orthogonality relation for the cosine functions. The corresponding matrix equation for b_p^+ reads as

$$\sum_p T_{pk} \, b_p^+ = B_k, \tag{11}$$

where the interaction matrix T_{pk} is given again by Eq. (8) and the effective driving force is

$$B_k = 2k_0 \, a_0^+ \, \delta_{k0} + b_0^- \, [\beta_0 \, (h/\varepsilon)_{0k} - \varkappa_k \, h_{0k}]. \tag{12}$$

With the amplitudes b_p^+ found, the amplitudes a_k^- can be calculated from

$$a_k^- = -a_0^+ \delta_{k0} + q_k \left(b_0^- \, h_{0k} + \sum_p b_p^+ \, h_{pk} \right), \tag{13}$$

where $q_k = 1$ for $k = 0$ and 2 for $k \neq 0$.

In what follows, we refer to the above two calculation schemes, based on Eqs. (7) and (11), as to the *a* and *b* representations. In certain respects, the *a* representation is more physical: In the range $k < k_0$ and for $k \simeq k'_{\rm sp} > k_0$, the amplitudes a_k describe the diffraction orders in air and the surface plasmon, respectively. Within the *b* representation, only the amplitude of the propagating mode $b_0^+ \equiv b_{p_0}^+$ is directly linked to the observable characteristics. Getting the same final results with the *a* and *b* representations gives a strong evidence of correctness of the whole calculation procedure.

The Fourier components entering the expressions for T_{pk}, A_p, and B_k can be calculated exactly using Eq. (3). The corresponding expressions include trigonometric functions that become quickly oscillating with increasing L. This circumstance seriously complicates numerical calculations. However, the oscillating dependences on L can be completely excluded using Eq. (2). The final relation for h_{pk} is

$$h_{pk} = \frac{k \cos(p_d l) \sin(kl) - p_d \sin(p_d l) \cos(kl)}{w\left(k^2 - p_d^2\right)} + \frac{\varepsilon_m \, p_d \sin(p_d l) \cos(kl) - k \cos(p_d l) \sin(kl)}{w \, (k^2 - p^2)}. \tag{14}$$

The first and second terms originate from the integration over the half-slit and half-wall, respectively. To find the Fourier component $(h_p/\varepsilon)_k$, it is sufficient to divide the second term by ε_m.

After all, we obtain from Eqs. (8) and (14) the following explicit expression for the interaction matrix:

$$T_{pk} = i \, Q'_{pk} \frac{k \cos(p_d l) \sin(kl) - p_d \sin(p_d l) \cos(kl)}{w \left(k^2 - p_d^2\right)} + i \, Q_{pk} \frac{\varepsilon_m \, p_d \sin(p_d l) \cos(kl) - k \cos(p_d l) \sin(kl)}{w \, (k^2 - p^2)}, \tag{15}$$

where

$$\begin{aligned} Q'_{pk} &= \left(k^2 - k_0^2\right)^{1/2} + \left(p^2 - \varepsilon_m k_0^2\right)^{1/2}, \\ Q_{pk} &= \left(k^2 - k_0^2\right)^{1/2} + \varepsilon_m^{-1} \left(p^2 - \varepsilon_m k_0^2\right)^{1/2}. \end{aligned} \tag{16}$$

These relations possess important and general features: (a) The interaction matrix T_{pk} shows formally a singular behavior for $p \to k$. This limit is relevant to the single-slit case $L = \infty$, where k and p take independently continuous positive values.

Thus, Eqs. (7) and (11) transform into *singular integral equations* (including the principal-value integral) in this case. Consideration of the periodic case with a large value of L provides a true regularization of this singularity, and the denominator $w(k^2 - p^2)$ does not turn to zero for $L \to \infty$.

(b) The factor Q_{pk} accounts for the resonant SP properties. Solution of the equation $Q_{kk} = 0$ gives the SP wave vector $k_{\rm SP} = k'_{\rm SP} + i k''_{\rm SP} = k_0\sqrt{\varepsilon_m/(1+\varepsilon_m)}$. The value $1/k''_{\rm SP}$ is the SP decay length. For $\varepsilon''_m \ll |\varepsilon'_m|$ and $|\varepsilon'_m| \gg 1$, we have $k''_{\rm SP} \simeq k_0 \varepsilon''_m / 2|\varepsilon'_m|^2 \ll k_0$. For real values of k, the function Q^{-1}_{kk} shows a sharp plasmonic peak at $k = k'_{\rm SP} \simeq k_0\sqrt{|\varepsilon'_m|/(|\varepsilon'_m| - 1)}$. This k peak is also inherent in Q^{-1}_{pk} for $p \approx k'_{\rm SP}$. The factor Q'_{pk}, originating from the slit region, shows no sharp features.

Note finally two expected features of the transition from the periodic to the single-slit case: (i) One can expect that the wall width has to be considerably larger than the SP decay length in order to proceed to the single-slit case. This gives the requirement $L/\lambda \gtrsim |\varepsilon'_m|^2/\varepsilon''_m$, which is fulfilled only for pretty large values of L. (ii) There is a big difference between a^-_0 and the amplitudes a^-_k with $k \neq 0$. For $l/L \to 0$, we have $a^-_0/a^+_0 \to (\sqrt{\varepsilon_m} - 1)/(\sqrt{\varepsilon_m} + 1)$, which corresponds to the Fresnel reflection from the nonperforated metal. At the same time, we have $a^-_k/a^+_0 \to 0$ in this limit. This is why separation of the large contribution to a^-_0 is useful for calculation purposes within the a representation.

The infinite linear sets of algebraic equations (7) and (11) have been solved numerically for the basic interface problems *out* and *in*, depicted in Fig. 3, to calculate the observable characteristics, such as the efficiency of transmission, diffraction, and SP excitation (see Sec. VI). A numerical routine with the truncation of the sets at $n_{\max} = \nu_{\max} = N$ has been used. The truncation number N was chosen big enough to achieve the true (i.e., saturated in N) values of the observable characteristics. We have also made sure that the a and b representations give the same results. Inclusion of the anomalous modes in the calculation scheme is, as earlier,[18] crucial for the convergence of the calculation procedure.

Importantly, the achievement of the single-slit limit by increasing L imposes a harsh restriction on N, especially for subwavelength slits. If the half-widths of the slit and wall, l and L, are restricted by the inequalities $l \lesssim \lambda/2$ and $L \gtrsim |\varepsilon'_m|^2\lambda/\pi\varepsilon''_m$, then the physical condition $k_{\max} l \equiv \pi N l/(l + L) \gg \pi/2$ leads us to the inequality

$$N \gg \frac{1}{2\pi}\frac{|\varepsilon'_m|^2}{\varepsilon''_m}\frac{\lambda}{l}. \tag{17}$$

This restriction is in agreement with our numerical data. The maximum truncation number is $N_{\max} = 10^4$ in our calculations. Correspondingly, we are restricted to not very small values of l/λ and $\varepsilon''_m/|\varepsilon'_m|^2$ (see the following).

V. GENERAL PROPERTIES OF NUMERICAL SOLUTIONS

While the alternative sets of basic equations (7) and (11) can not be solved analytically, our numerical procedure reveals simple and general features of the solutions for a_k and b_p.

For $k \approx k'_{\rm SP}$, i.e., nearby the SP resonance, we have with a good accuracy $a^-_k = ck_0/Q_{kk}$, where c is a complex constant. This fit describes the SP excitation. Moreover, for not too wide slits $k_0 l \lesssim 0.5$, the SP fit is valid within a much wider range of wave vectors $k \lesssim k'_{\rm sp}$. Variation of the input parameters results in changing c.

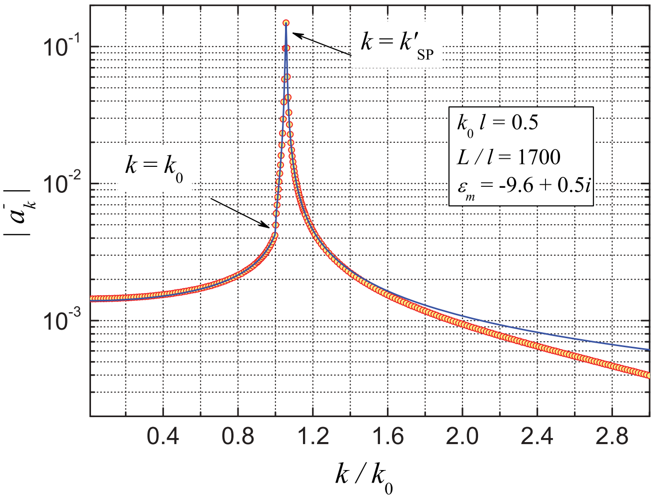

FIG. 4. (Color online) The case *out*: The functions $|a^-_k|$ (dots) and $0.00915\; k_0/|Q_{kk}|$ (solid line) for $\varepsilon_m = -9.6 + 0.5i$, $k_0 l = 0.5$, and $L/l = 1700$. The arrows indicate the critical value $k = k_0$ and the SP peak at $k = k'_{\rm SP}$. Outside the peak region, the dots are thinned out.

The mentioned features are illustrated by Fig. 4. The dots show $|a^-_k|$ versus k/k_0 on a semilogarithmic scale in a range of small and intermediate arguments, calculated numerically for the problem *out*, $\varepsilon_m = -9.6 + 0.5i$, $l = \lambda/4\pi$, and $r = L/l = 1700\,(L/\lambda \simeq 135)$. The chosen parameters correspond to silver at $\lambda \approx 500$ nm, fairly narrow slits, and very thick metal walls. The narrow plasmonic peak at $k = k'_{\rm SP}$ is surely the dominating feature here. The solid line is the function $|c|k_0/|Q_{kk}|$ with $|c| = 0.00915$, calculated with the same ε_m using Eq. (16). For $k < 1.6k_0$, this line describes perfectly well the numerical results.

Importantly, the constant c is directly linked to the SP amplitude a_p. Using Eqs. (5) and (16), and the residue theorem, one can find the following for $x > l$ and $\varepsilon''_m/2\varepsilon'^2_m \ll 1$:

$$\sum_k \frac{k_0 c}{Q_{kk}}\cos(kx) \simeq \frac{w k_0 c}{\pi}\int_{-\infty}^{\infty} Q^{-1}_{kk} e^{ikx}\, dk = a_p e^{ik_{\rm sp}x} \tag{18}$$

with $a_p = ik_0 wc/\sqrt{|\varepsilon'_m|}\,(1 - |\varepsilon'_m|^{-2})$.

As soon as the SP fit $a^-_k = ck_0/Q_{kk}$ is applicable to the range $k < k_0$, which corresponds to the propagating waves in air, the diffractive properties are linked to the SP characteristics. This takes place for $k_0 l \lesssim 0.5$. For $k_0 l \gtrsim 1$, when this fit is restricted to $k \approx k'_{\rm SP}$, the link between the diffraction and SP properties is relaxed.

In any case, the function $a^-_k Q_{kk}$ deviates strongly from a constant for $k \gg k_0$. In particular, the functions $|a^-_k||Q_{kk}|$ and $|a^-_k|$ possess deep minima at the points of minimum of the driving force $A_k \propto \sin(p_d l)$, which corresponds to the in-slit resonances. The behavior of a^-_k in this range is important for calculation of $b^+_{p_0}$ and for analysis of the light transmittance.

Consider now the x dependence of the magnetic field at the interface $z = 0$ using the numerical values of a^-_k and

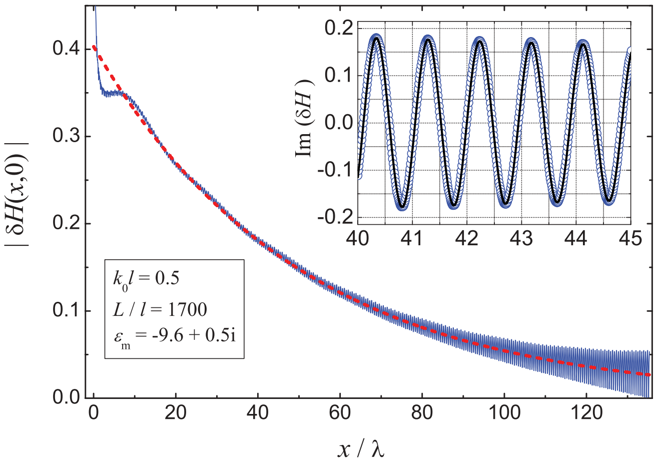

FIG. 5. (Color online) Calculated dependence $|\delta H(x,0)|$ (solid line) and the SP fit $|a_p|\,\exp(-k''_{\mathrm{SP}}x)$ (dashed line) for the SP amplitude $|a_p| = 0.4$. The inset shows the numerical results (dots) for $\mathrm{Im}(\delta H)$ and the fit $\mathrm{Im}[a_p\exp(ik_{\mathrm{SP}}x)]$ for $a_p = 0.4$ (solid line).

Eq. (5). Specifically, we analyze $\delta H(x) = H(x,0) - \langle H(x,0)\rangle_w$ by dropping the constant part. The solid line in Fig. 5 shows the corresponding numerical results for $|\delta H|$ within a half-period $0 < x < w$, and the dashed is the SP fit $|a_p|\exp(-k''_{\mathrm{sp}}x)$ with $|a_p| = 0.4$. The solid line exhibits very fast oscillations; their amplitude is very small near the slit and increasing slowly when approaching the wall center. For $x/\lambda \gtrsim 8$, the dashed line precisely follows the solid line averaged over the fast oscillations. For $x/\lambda \lesssim 0.4$, the field $|H|$ is rapidly increasing when approaching the slit and the SP fit fails.

This behavior becomes clear when taking into account the actual characteristic lengths. For the chosen input parameters, we have $k''_{\mathrm{SP}} \simeq 2.3/w \ll k_0$; this means that two plasmons irradiated in the vicinity of a slit propagate freely and symmetrically for long distances toward the neighboring slits and decay strongly, but not completely, after the half-way. This causes the plasmon interference (a standing wave), which is the strongest at $x \simeq w$.

The SP fit becomes even more impressive when we consider separately the real and imaginary parts of δH. This is exemplified by the inset of Fig. 5, which shows a fragment of the calculated dependence $\mathrm{Im}(\delta H)$ together with a fit for $a_p = 0.4\exp(i\varphi)$ with the SP phase $\varphi = -2.44$ rad. The fit accurately reproduces the fine quasiperiodic structure. The x dependence of $\mathrm{Re}(\delta H)$ is fitted simultaneously equally well.

It is worth mentioning that the dependence $a_k^- \propto Q_{kk}^{-1}$ includes a square-root-type inflection point at $k = k_0$. This means that the field $H(x,0)$ possesses a weak power-law decaying asymptotic wave (Norton wave), which should become dominating for very large $|x|$.[23–25] We have easily seen this asymptotic wave, especially for relatively large $\varepsilon''_m/|\varepsilon'_m|$ ratios. However, the powerlike tail and its influence on the observable characteristics, as seen in the following, were negligible in our calculations.

The last issue to consider for the periodic case is the L- dependence of the transmittance T_0, defined as the relative part of the energy transmitted into the propagating mode. This characteristic has been investigated earlier for subwavelength gratings $2w < \lambda$.[18] For sub-wavelength slits, we have $T_0 \propto |b_0^+|^2$. The amplitude of the propagating mode $b_0^+ \equiv b_{p_0}^+$, calculated indirectly from Eq. (10), coincides with a high accuracy with that found directly from Eq. (11); the a and b representations thus give the same result.

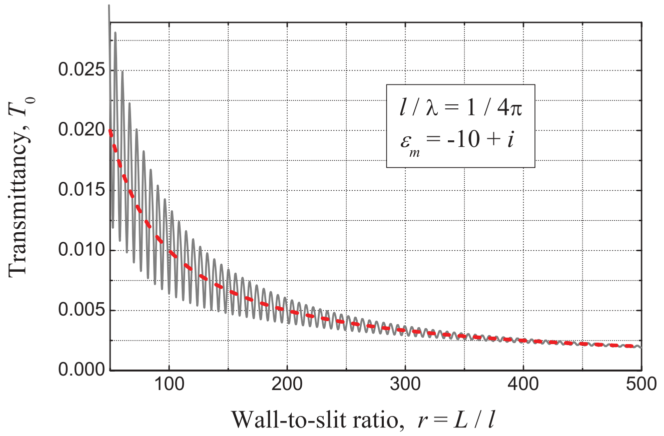

FIG. 6. (Color online) The transmittance T_0 versus the wall-to-slit ratio $r = L/l$ for $k_0 l = 1/2$ and $\varepsilon_m = -10 + i$ (gray line). The dashed line is the dependence $1/r$.

The solid line in Fig. 6 shows T_0 versus the ratio $r = L/l$ for $\varepsilon_m = -10 + i$ and $l = \lambda/4\pi$. Initially we have very strong oscillations that are accompanied by an average decrease of T_0 approximately as $1/r$ (see the dashed line). These oscillations are caused by coupling with the SPs and the effects of SP interference caused by reflections from the corners. For $2k''_{\mathrm{SP}}L \gtrsim 1$, which corresponds to $r \gtrsim 200$, the interference is substantially suppressed by absorption and the oscillations become weak. For $r \gtrsim 400$, they become very small. At the same time, the decrease of T_0 as $1/r$ persists. It has a pure geometrical reason: Only a small fraction of the incident light ($\sim l/w$) interacts with the slits, while most of the light is reflected back. To characterize the light-slit interactions, we proceed to single-slit characteristics.

VI. SINGLE-SLIT CHARACTERISTICS

Our purpose here is to introduce single-slit characteristics by transition to large L and to analyze them on the basis of numerical data. The cases *out* and *in* (see Fig. 3) will be considered separately.

A. Incidence from the outside

In this case, we set $b_0^- = 0$ and $a_0^+ = 1$, which corresponds to a unit-amplitude wave incident from the outside. The set (7), (10) or (11), (13) then allows us to compute the amplitudes a_k^- and b_p^+. With these amplitudes known, we calculate the efficiencies of the actual transformation processes. An efficiency is defined here as the ratio of the energy coming to a certain channel to the incident energy for the slit.

Transmission efficiency. As we have seen, the transmittance T_0 for the periodic case tends to zero for $L \to \infty$. To characterize the slit transmittance, we multiply T_0 by the geometric factor w/l arriving at the efficiency

$$\eta_t^{\mathrm{out}} = (r + 1)\,T_0 \tag{19}$$

with $r = L/l$. The saturated value of $\eta_t^{\rm out}(l,r)$ with increasing r then characterizes the single slit. It is expected that the saturation occurs for $k''_{\rm SP}L \gtrsim 1$. The product $2l\,\eta_t^{\rm out}$ gives the transmission cross section.

Diffraction efficiency. Consider the diffracted (scattered) wave of the order n propagating at an angle $\theta_n = \arcsin(n\lambda/2w)$ [see also Fig. 3(a)]. According to Eq. (5), this wave is characterized by the amplitude $a_n^-/2 \equiv a_k^-/2$ with $k = \pi n/w$. To introduce the corresponding single-slit diffraction efficiency, we multiply $|a_k^-|^2/4$ by the geometric factor w/l and by the angular factor $\cos\theta_n$ (to pick out the energy flux coming from a period) and obtain

$$\eta_{d,n}^{\rm out} = (r+1)\,|a_n^-|^2\cos\theta_n/4. \tag{20}$$

The total diffraction efficiency $\eta_d^{\rm out}$ is given by the sum of $\eta_{d,n}^{\rm out}$ over all diffraction orders $n = \pm 1, \pm 2, \dots$.

For $L/\lambda \gg 1$, when $n_{\max} \gg 1$, it is useful to employ the continuous scattering angle θ instead of θ_n. The total diffraction efficiency can then be represented as

$$\eta_d^{\rm out} \simeq \int_{-\pi/2}^{\pi/2} \eta_d^{\rm out}(\theta)\,d\theta, \tag{21}$$

where

$$\eta_d^{\rm out}(\theta) = l\,(r+1)^2\,|a_\theta^-|^2\,\cos^2\theta/2\lambda \tag{22}$$

is the differential diffraction efficiency. It is even in θ for the case of normal incidence.

SP-excitation efficiency. Irradiation of surface plasmons propagating to the left and right occurs in a close vicinity of the slit. This feature is intuitively clear, and it is consistent with the results of Sec. V. Therefore, two SPs propagate over long distances and decay strongly before reaching two neighboring slits for $k''_{\rm SP}L \gtrsim 1$. The corresponding contribution to $H(x,0)$ is $a_p\,\exp(ik_{\rm SP}x)$ for $l \lesssim x \lesssim L$. Furthermore, the SP magnetic field $H_{\rm SP}(x,z)$ decays exponentially with increasing $|z|$. In the air and metal regions, the rates of decay are $k_0/\sqrt{|\varepsilon'_m|-1}$ and $k_0|\varepsilon'_m|/\sqrt{|\varepsilon'_m|-1}$ for $\varepsilon''_m \ll 2|\varepsilon'_m|^2$. This allows one to calculate the dominating x component of the pointing vector and to integrate it over z resulting in the total SP energy flux near the slit. Multiplying the flux by a factor of 2 to take into account two excited plasmons and dividing by the energy flux incident onto the slit, we obtain

$$\eta_{\rm SP}^{\rm out} \simeq \frac{\sqrt{|\varepsilon'_m|}\,|a_p|^2}{2k_0 l}\left(1 - \frac{1}{|\varepsilon'_m|^2}\right). \tag{23}$$

The second contribution in the bracket, originating from the metal part of the plasmonic energy flux, is typically very small. Multiplying $\eta_{\rm SP}^{\rm out}$ by $2l$, we obtain the SP-excitation cross section.

Furthermore, we find for $k_0 l \lesssim 0.5$, using Eqs. (22), (23), and the equality $a_k^- = ck_0/Q_{kk}$, that $\eta_d^{\rm out}/\eta_{\rm SP}^{\rm out} \simeq (\sqrt{|\varepsilon'_m|} - 1)/2$; this ratio does not depend on l.

Reflectance change. One more important characteristic is the reflectance $R^{\rm out} = |a_0^-|^2$ [see also Fig. 3(a)]. In the limit $l \to 0$, we have $a_0^- = (\sqrt{\varepsilon_m} - 1)/(\sqrt{\varepsilon_m} + 1)$ according to the Fresnel relations. The corresponding reflectance R_0 differs from 1 only because of the dissipative losses ($\varepsilon''_m \neq 0$). For $l/L \ll 1$, the reflectance $R^{\rm out}$ must be close to R_0 so that the efficiency

$$\eta_R^{\rm out} = (r+1)\,(R_0 - R^{\rm out}) \tag{24}$$

can be taken as a characteristic of the impact of the slit on the reflection properties.

Dissipative losses and energy balance. With the above efficiencies calculated, one can determine how much energy of the incident wave is transforming into the reflection, diffraction, transmission, and SPs. Correspondingly, one can evaluate the impact of the slit on the dissipative losses. The difference $\eta_R^{\rm out} - \eta_t^{\rm out} - \eta_d^{\rm out} - \eta_{\rm sp}^{\rm out}$ is nothing more than the energy ratio of the dissipation change for a period to the influx into the slit. If this difference is relatively small, the decrease of the reflectance R can be regarded as the consequence of the nondissipative losses as a result of the excitation of the propagating mode, diffraction orders, and SPs. The difference must be zero for $\varepsilon''_m \to 0$, and it is expected to be small for $\varepsilon''_m \ll |\varepsilon'_m|$.

B. Numerical results

Figure 7 shows the transition to the single-slit case by increasing the wall half-width L for $l = \lambda/4\pi$ and $\varepsilon'_m = 7.5$. At $\varepsilon''_m = 1$, it is practically completed for $r > 300$, i.e., for $k_0 L > 300$ or $k''_{\rm sp}L \gtrsim 1$. At $\varepsilon''_m = 0.3$, which corresponds to a SP decay length that is three times longer, this transition takes much longer (compare the solid and dotted lines for $\eta_t^{\rm out}$). At the same time, the single-slit values of $\eta_t^{\rm out}$ for $\varepsilon''_m = 1$ and 0.3 are almost the same. This is the general situation: for $\varepsilon''_m \ll |\varepsilon'_m|$, the effect of absorption on the single-slit characteristics is small.

In the single-slit limit, we have roughly $\eta_{t,d,sp}^{\rm out} \sim 1$ for the chosen $k_0 l$ and ε'_m. The corresponding cross sections are therefore of the order of the geometrical size $2l$. Furthermore, one can deduce from Fig. 7 that $\eta_r^{\rm out} \simeq \eta_t^{\rm out} + \eta_d^{\rm out} + \eta_{\rm SP}^{\rm out}$ in the single-slit case, i.e., the dissipation rate is not much changed as a result of these transformation processes.

The solid lines in Fig. 8 show the single-slit efficiencies versus $k_0 l$ in the subwavelength range (with a single propagating mode) for different values of ε'_m. Three dotted lines,

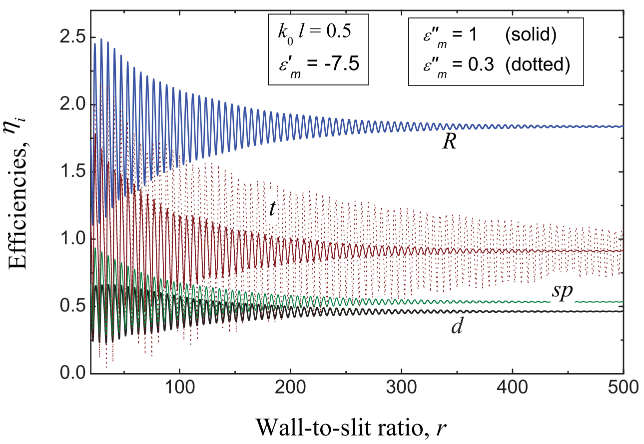

FIG. 7. (Color online) Dependence $\eta_i^{\rm out}(L)$ with $i = t, d$, SP, and R for $k_0 l = 0.5$ and $\varepsilon'_m = -7.5$. The solid lines correspond to $\varepsilon''_m = 1$, while the dotted line for $\eta_t^{\rm out}$ is plotted for $\varepsilon''_m = 0.3$.

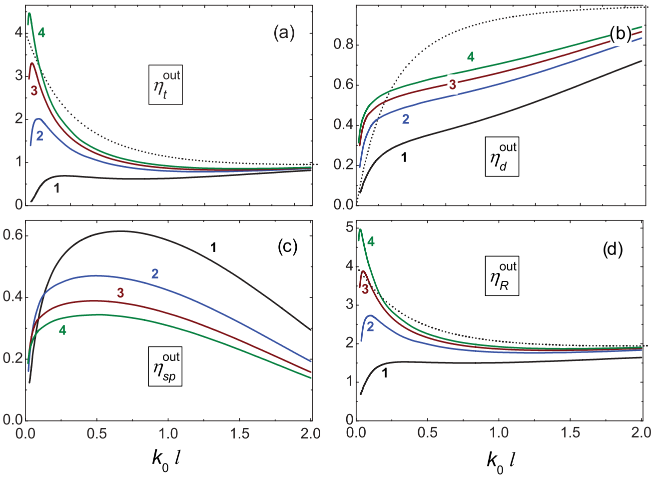

FIG. 8. (Color online) Efficiencies of the main processes (transmission, diffraction, SP excitation, and reflection) vs $k_0 l$ for $\varepsilon_m'' = 1$ and different values of ε_m'. Curves 1, 2, 3, and 4 correspond to $\varepsilon_m' = -5, -10, -15$, and -20, respectively.

given for comparison, show the functions $\eta_{t,d,R}^{\text{out}}(k_0 l)$ in the perfect-conductor limit.[26]

Discussion of the presented dependences is given in Sec. VII. Here we indicate some general features: The efficiencies η_t^{out}, η_d^{out}, and $\eta_{\text{sp}}^{\text{out}}$ behave substantially differently. The effect of ε_m'' is noticeable only for $\varepsilon_m' = -5$. There are similarities and also strong differences between the cases of real and perfect metals. For real metals, the efficiencies $\eta_{t,d,\text{sp}}^{\text{out}}$ tend to zero for $k_0 l \to 0$, while η_R^{out} tends to a small ε_m''-dependent constant. The presence of peaks of $\eta_{t,R}^{\text{out}}(k_0 l)$ for very small arguments is highly remarkable. Furthermore, we have with a good accuracy for $|\varepsilon_m| > 4$:

$$\eta_{t,\max}^{\text{out}} \simeq 0.28(|\varepsilon_m| - 3), \quad (k_0 l)_{\max}^{-1} \simeq 1.8(|\varepsilon_m| - 3). \tag{25}$$

Further increase of $k_0 l$ results in bifurcation changes to the described behavior. These changes correlate with the birth of the next propagation modes. Consideration of the corresponding features is beyond the scope of this paper, which is restricted to subwavelength slits.

We consider last the angular dependence $\eta_d^{\text{out}}(\theta)$. The solid lines in Fig. 9 show this function for $|\varepsilon_m'| = 10$ and for different values of $k_0 l$. The corresponding dotted lines are plotted for the perfect-metal case. For sufficiently narrow slits $k_0 l \lesssim 0.5$, the solid lines are close to those given by $\cos^2(\theta)/|Q_{kk}|$ with $k = k_0 \cos\theta$ (see Sec. V). Further increase of $k_0 l$ results in narrowing of the cental maximum. The value of $\eta_d^{\text{out}}(\theta = 0)$ grows monotonically with increasing $k_0 l$ in agreement with Fig. 8. For any $k_0 l$ and ε_m, we have $\eta_d^{\text{out}}(\theta) \to 0$ for $|\theta| \to \pi/2$, i.e., the grazing diffraction is absent. This contrasts to the perfect-metal case.

C. Incidence from the inside

In this case, we set $a_0^+ = 0$ and $b_0^- = 1$, which correspond to a unit-amplitude propagating mode traveling to the interface from the inside [see Fig. 3(b)], and compute again the amplitudes a_k^- and b_p^+. The simplest characteristic here is the internal reflection coefficient $R^{\text{in}} = |b_0^+|^2$. The diffraction efficiency of the order n is given by

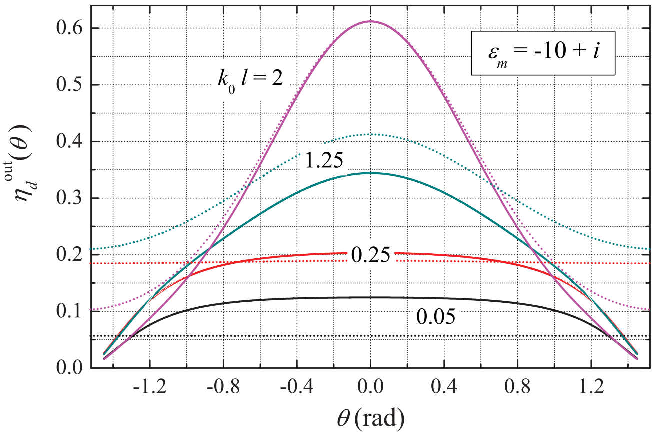

FIG. 9. (Color online) Differential diffraction efficiency $\eta_d^{\text{out}}(\theta)$ for $\varepsilon_m = -10 + i$ and for four values of $k_0 l$. The dotted lines correspond to the perfect-metal case.

$$\begin{aligned} \eta_{d,0}^{\text{in}} &= \frac{k_0 |a_0^-|^2}{\beta_0' \langle |h_0|^2/\varepsilon' \rangle} \quad (n = 0), \\ \eta_{d,n}^{\text{in}} &= \frac{k_0 |a_{k_n}^-|^2 \cos\theta_n}{4\beta_0' \langle |h_0|^2/\varepsilon' \rangle} \quad (n = \pm 1, \pm 2, \ldots). \end{aligned} \tag{26}$$

This definition accounts for the ratio of the corresponding energy fluxes and the geometric factor $\cos\theta_n$ [compare to Eq. (20)]. In contrast to the case *out*, the zero diffraction order is not dominating here, and there is no need to introduce the factor $r + 1 = w/l$.

The sum of $\eta_{d,n}^{\text{in}}$ over n gives the total diffraction efficiency η_d^{in}. For $L/\lambda \gg 1$, we employ again the continuous scattering angle θ and represent it as

$$\eta_d^{\text{in}} \simeq \int_{-\pi/2}^{\pi/2} \eta_d^{\text{in}}(\theta)\, d\theta\,, \tag{27}$$

where

$$\eta_d^{\text{in}}(\theta) = \frac{\pi w\, |a_\theta^-|^2 \cos^2\theta}{\beta_0' \lambda^2\, \langle |h_0|^2/\varepsilon' \rangle} \tag{28}$$

is the differential diffraction efficiency.

For the SP-excitation efficiency, we have

$$\eta_{\text{sp}}^{\text{in}} = \frac{\sqrt{|\varepsilon_m'|}\; |a_p|^2}{2w\beta_0'\, \langle |h_0|^2/\varepsilon' \rangle} \left(1 - \frac{1}{|\varepsilon_m'|^2}\right). \tag{29}$$

As in earlier discussion, the quantities introduced are generally functions of l and L. Their saturated values with increasing L give the corresponding single-slit characteristics.

According to our definitions, $R^{\text{in}} + \eta_d^{\text{in}} + \eta_{\text{sp}}^{\text{in}} = 1$ in the absence of dissipation $\varepsilon_m'' = 0$. For a weak dissipation, $\varepsilon_m'' \ll |\varepsilon_m'|$, this sum is expected to be approximately 1.

Figure 10 shows representative dependences of the introduced characteristics on $k_0 l$ for $\varepsilon_m = -10 + i$. The sum of these functions is close to 1. The most important feature here is the monotonic increase of R^{in} with decreasing slit width. Variation of ε_m does not result in strong changes of $R^{\text{in}}(k_0 l)$. The impact of ε_m on $\eta_{d,\text{sp}}^{\text{in}}(k_0 l)$ is similar to that on $\eta_{d,\text{sp}}^{\text{out}}(k_0 l)$.

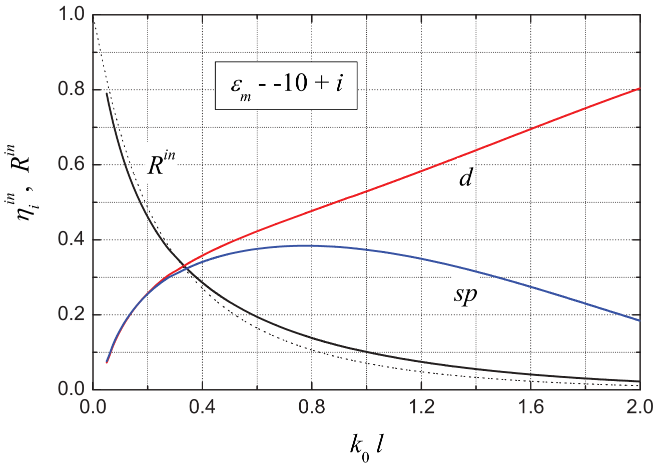

FIG. 10. (Color online) The efficiencies $\eta^{\text{in}}_{d,\text{SP}}$ and the reflection coefficient R^{in} versus $k_0 l$ for $\varepsilon_m = -10 + i$; the dotted line corresponds to the perfect-metal case.

VII. DISCUSSION

A. Methodical advances

Our eigenmode-expansion method reduces (without simplifying assumptions) the initial 2D problem of light transmission, diffraction, and SP excitation to a 1D interface problem; the latter is equivalent to the solution of linear algebraic equations. This reduction relies on the knowledge of the full set of eigenvalues and eigenfunctions: propagating, evanescent, and anomalous. Completeness of this set is verified within the method.

An essential feature of our approach is the consideration of the single-slit case as the limiting case of a periodic metal-dielectric structure when increasing the wall width $2L$. This allows one (i) to get a physically correct discrete regularization of the primary singular integral equations and (ii) to monitor the transition to the single-slit case. The latter occurs when the width $2L$ considerably exceeds the SP decay length L_{SP}.

Explicit dependences on the wall half-width L are excluded from the initial equations using the exact dispersion relation for the eigenvalues. Periodicity of the structure manifests itself only in the positions of the eigenvalues on the complex plane. This eliminates any sharp dependences when increasing L. This feature allows us to treat numerically the single-slit case.

A disadvantage of the method is the necessity to deal with very large values of the wall-to-slit ratio L/l and, correspondingly, with inversion of high-rank matrices. For realistic values of ε_m, the rank of the matrix is $\sim 10^4$, which brings us to the limit of the conventional computational resources.

B. Physical insights

Very little is known about the efficiencies and cross sections of the elementary single-slit transformation processes in the subwavelength range for real metals. Only recently, these efficiencies were systematically investigated for the perfect-metal case.[26] The SP excitation is absent in this case. Earlier, the SP excitation was analyzed for almost perfect metals.[36,37] We consider in the following the status of our findings versus the known results.

The most striking new result is perhaps the changing behavior of the transmission efficiency deeply inside the subwavelength range $k_0 l \lesssim |\varepsilon_m|^{-1}$: η_t^{out} possesses here a peak, the height $\eta^{\text{out}}_{t,\max}$ of which grows linearly with increasing $|\varepsilon_m|$ [see Fig. 8(a)]. For $|\varepsilon_m| \gtrsim 16$, the peak value $\eta^{\text{out}}_{t,\max}$ exceeds 4, i.e., the perfect-metal maximum. Furthermore, the efficiency η_t^{out} tends to zero for $l \to 0$ and remains noticeably smaller than its perfect value for $k_0 l \sim 1$. Transition to the perfect-metal case with increasing $|\varepsilon_m|$ takes place only for $\sqrt{|\varepsilon_m|}k_0 l \gg 1$.

The $k_0 l$ dependence of the total diffraction efficiency is similar to that in the perfect-metal case in the sense that $\eta_d(k_0 l)$ is a monotonically increasing function [see Fig. 8(b)]. However, there are real-metal specific features: Compared to the perfect case, the diffraction losses are essentially increased for $k_0 l \lesssim |\varepsilon_m|^{-1}$ and decreased for $k_0 l \gg |\varepsilon_m|^{-1}$. Within the whole subwavelength range, these losses significantly grow with increasing $|\varepsilon_m|$. It is remarkable that the differential diffraction efficiency $\eta_d^{\text{out}}(\theta)$ tends to zero for $\theta \to \pm\pi/2$ (see Fig. 9). This feature, which is absent in the perfect-metal case,[26] can be naturally viewed as capture of the grazing diffracted waves into the SP mode.

Let us turn to the SP-excitation efficiency η_{SP} relevant to real metals. As shown, this characteristic is coupled to the diffraction efficiency for $k_0 l \lesssim 0.5$: $\eta_{\text{SP}} \simeq 2\eta_d/(\sqrt{|\varepsilon_m|} - 1)$. This feature is in line with Figs. 8(b) and 8(c). For $k_0 l \gtrsim 1$, the functions $\eta_d(k_0 l)$ and $\eta_{\text{SP}}(k_0 l)$ are different, i.e, increasing and decreasing. While η_{SP} decreases with increasing $|\varepsilon_m|$, it remains comparable with η_d within a wide range of parameters. The maximum value $\eta^{\max}_{\text{SP}} \approx 0.62$ occurs in the middle of the subwavelength range $2l/\lambda \approx 0.4$. The presence of a broad maximum of $\eta_{\text{SP}}(k_0 l)$ in the middle of the range can be linked to the behavior of $\eta_d^{\text{out}}(\theta)$ for the perfect metal: The amount of grazing diffracted waves, to be captured to SPs, is also maximal here (see the dotted lines at $\theta \simeq \pm\pi/2$ in Fig. 9) because of redistribution of the diffracted energy from large angles to the central peak with further increasing $k_0 l$.

Earlier, the SP-excitation efficiency was analyzed for almost perfect metals using an approximate model and direct numerical methods.[36,37] This model works well for $|\varepsilon_m| \gg 10$ and $k_0 l \gtrsim |\varepsilon_m|^{-1}$, when $\eta_{\text{SP}} \ll 1$. Comparison with numerical data of Ref. 37 is possible for $|\varepsilon_m| \approx 12$, and it shows a good agreement.

The last efficiency to be considered for the case *out* is η_R^{out}. On one hand, it characterizes the decrease of the reflectivity caused by the slit structure. However, on the other hand, it gives the total losses by transmission, diffraction, and SP excitation. The approximate equality $\eta_R \simeq \eta_t + \eta_d + \eta_{\text{SP}}$ would be exact and represent the energy conservation law for $\varepsilon''_m = 0$. For $\varepsilon''_m \ll |\varepsilon_m|$, it is fulfilled with good accuracy.

One sees from Fig. 8(d) that the total losses in real metals are noticeably, but not much, smaller compared to those in the perfect-metal case for $k_0 l \gtrsim |\varepsilon_m|^{-1}$; with increasing $|\varepsilon_m|$, they are approaching from below the perfect-metal losses. It is evident that the SP excitation takes away some of the energy from the transmission and diffraction channels. In particular, the shallow minimum of $\eta_t(k_0 l)$ in the middle of the interval is caused by this effect. No SP-assisted funneling of light into the slit takes place. For $k_0 l \lesssim |\varepsilon_m|^{-1}$ and sufficiently large $|\varepsilon_m|$, the total losses become larger compared to those in the perfect-metal case, primarily because of the transmission enhancement. The peak of η_R^{out} is thus linked to the peak of η_t^{out}.

The case *in* is similar to the case *out* in the sense that $\eta_{d,\text{SP}}^{\text{in}}(k_0 l,|\varepsilon_m|) \approx \eta_{d,\text{SP}}^{\text{out}}(k_0 l,|\varepsilon_m|)$ (see also Figs. 8 and 10). The equality $\eta_d^{\text{in}}(\theta) = \eta_d^{\text{out}}(\theta)$, valid for the perfect metal, also becomes approximate for real metals. The efficiency $\eta_{\text{SP}}^{\text{in}}$ is slightly larger than $\eta_{\text{SP}}^{\text{out}}$; this feature can also be extracted from the numerical data of Ref. 37. The energy conservation law for the case *in*, $R^{\text{in}} \simeq 1 - \eta_d^{\text{in}} - \eta_{\text{SP}}^{\text{in}}$, is also approximate for $\varepsilon_m'' \neq 0$. In accordance with this relation, R^{in} tends to 1 for $k_0 l \to 0$, which is not different from the perfect-metal case. This feature is important for the Fabry-Perot enhancement of the light transmission (see also below).

The oscillatory-type transition from the periodic to the single-slit case, considered in Sec. VI B, possesses an important general feature: It is caused by the SP-assisted influence of neighboring slits. This transition is completed when the wall half-width L exceeds the decay length $L_{\text{SP}} \simeq |\varepsilon_m'|^2/k_0\varepsilon_m''$; for real metals, it typically corresponds to very large values of L/λ. In the perfect-metal case, where the interaction between neighboring slits is caused by scattering of the grazing waves, the transition to the single-slit case is essentially different.

C. Interpretation of the transmission-efficiency peak

Let us set $\varepsilon_m'' = 0$ and represent the transmission efficiency as $\eta_t^{\text{out}} = (\beta_0 w\langle h_0^2/\varepsilon\rangle/2k_0 l)|b_0^+|^2$, where the product $w\langle h_0^2/\varepsilon\rangle$ does not depend on w for $w \gg |\varepsilon_m| l$. The first factor merely expresses the properties of the propagating mode for the single slit. It tends to infinity as $1/l$ for $l \to 0$ (because $\beta_0 \propto 1/l$) and affects the right wing of the peak of $\eta_t^{\text{out}}(k_0 l)$ [see Fig. 8(a)]. The abrupt left wing is determined by the l dependence of the amplitude $b_0^+ \equiv \langle H h_0/\varepsilon\rangle/\langle h_0^2/\varepsilon\rangle$. The magnetic-field distribution $H(x,0)$ is thus the main issue.

The function $H(x,0)$ must be continuous at $x = \pm l$ and tending to $H_0 = 2\sqrt{|\varepsilon_m|}/(\sqrt{|\varepsilon_m|} - i) \simeq 2$ far from the slit, as prescribed by Fresnel's relations. The difference $\delta H(x) = H(x,0) - H_0$ must therefore tend to 0 for $l \to 0$. This is valid for both perfect and real metals.

By setting $H(x,0) = H_0$, one can make sure that $b_0^+ \propto l^2$ and $\eta_t^{\text{out}} \propto l^3$ for $l \to 0$. This result comes from the property of the eigenfunction h_0: $\langle h_0/\varepsilon\rangle \propto l^3$ for $l \to 0$, i.e., from a strong cancellation of the contributions to $\langle h_0/\varepsilon\rangle$ from the metal and air parts of the slit. Qualitatively, it gives the necessary peak dependence of $\eta_t^{\text{out}}(k_0 l)$ for $|\varepsilon_m| k_0 l \lesssim 1$.

Unfortunately, neglect of the small correction δH can not be justified. Moreover, taking δH into account, one is able to modify the law of vanishing of η_t^{out} for $k_0 l \to 0$. The point is that the well-known corner singularities of the electric field[15,44,45] strongly affect the behavior of $\delta H(x)$: this function possesses inflection points $\pm l$ with infinite derivatives. Cancellation of the contributions to $\langle \delta H h_0/\varepsilon\rangle$ from the metal and air parts cannot be strong.

We suppose a strong involvement of the corner singularities in the behavior of η_t^{out} for $k_0 l \lesssim |\varepsilon_m|^{-1}$. Apart from these general arguments, proportionality of $\eta_{t,\max}^{\text{out}}$ and $1/(k_0 l)_{\max}$ to $|\varepsilon_m'| - 3$ points to this origin; the 90° corner singularities become increasingly pronounced (nonintegrable) for $\varepsilon_m \to -3$.[45–47] The exact law of vanishing of η_t^{out} for $k_0 l \to 0$ remains unknown.

D. Impact on slab-transmission

A thorough analysis of the transmission-scattering properties of a single subwavelength slit in a metallic film and a comparison with a few relevant experiments go beyond the scope of this paper. We are confident, however, that our results provide a firm basis for such an analysis. Furthermore, they already allow us to understand the many expected observable properties.

If the film thickness z_0 considerably exceeds the decay length of the propagating mode $\beta_0'' z_0 \gtrsim 1$, the multiple reflections for the propagating mode are negligible, and the observable properties can be easily described in terms of the calculated efficiencies. In particular, the reflected and diffracted signals are described by the cross sections $2l\eta_r^{\text{out}}$ and $2l\eta_d^{\text{out}}(\theta)$, while the transmitted signal is given approximately by the product $2l\eta_t^{\text{out}}\,(1 - R^{\text{in}})\exp(-2\beta_0'' z_0)$.

If the propagating losses are weak, $\beta_0'' z_0 \ll 1$, and the film thickness is much larger than the skin depth $k_0 z_0\,|\varepsilon_m|^{1/2} \gg 1$, the multiple reflections and the Fabry-Perot resonances strongly affect the observable properties. Expressions for the slab transmittance in this case are known.[3,18,28] The resonant transmittance is expressed by the above interface characteristics, while the positions of the Fabry-Perot resonances depend additionally on the phase change during an internal reflection. Large values of η_t^{out} and R^{in} (see Figs. 8 and 10) facilitate the the slab transmittance for sufficiently narrow slits.

VIII. CONCLUSIONS

For single subwavelength slits $2l \lesssim \lambda$, efficiencies of the main interface transformation processes (transmission, diffraction, SP excitation, and reflection) are calculated for real metals as functions of $k_0 l$ and $|\varepsilon_m|$ and compared with the relevant single-slit characteristics of the perfect metal.

For very narrow slits $k_0 l \lesssim 1/|\varepsilon_m|$, these fundamental characteristics show a special behavior, which is absent in the perfect-metal case. In particular, the transmission efficiency η_t shows here a pronounced peak with an amplitude that grows with increasing $|\varepsilon_m|$ and exceeds the maximum perfect-metal value of 4. This peculiarity is interpreted in terms of the fundamental eigenmode characteristics. For modest subwavelength slits $\pi \gtrsim k_0 l \gg 1/\sqrt{|\varepsilon_m|}$, the transformation efficiencies show strong $|\varepsilon_m|$ dependences with a slow transition to the perfect-metal case.

For $k_0 l \lesssim 0.5$, the total diffraction efficiency η_d and SP-excitation efficiency η_{SP} are linked with each other; opening of the SP-excitation channel in real metals results in a redistribution of the scattering losses between diffraction and SP excitation. This feature correlates with the intrinsic property of the differential diffraction efficiency $\eta_d(\theta)$ in real metals: it vanishes for $|\theta| \to \pi/2$, which corresponds to capture of the grazing diffracted waves into the localized SP mode.

In real and perfect metals, the internal-reflection coefficient R^{in} for the fundamental mode approaches 1 for $k_0 l \to 0$. This feature strongly facilitates the slab transmittance via the familiar Fabry-Perot resonances.

Methodically, an advanced version of the eigenmode method of Ref. 18 has been used. It allows us to treat the single-slit case as the limit of the periodic-array case with increasing wall width. Correspondingly, we avoid any mathematical assumptions on the regularization of the primary singular single-slit integral equation; at every instant, we are dealing with a genuine physical problem. Furthermore, this scheme allows us to determine the separation between the neighboring slits, which is necessary to realize the single-slit limit.

ACKNOWLEDGMENTS

Financial support from the Russian Academy of Sciences (Program of the Presidium 21.2 and Program of the Branch of Physical Sciences “Physics of new materials and structures”) is gratefully acknowledged.

---

[1] T. W. Ebbesen, H. J. Lezec, H. F. Ghaemi, T. Thio, and P. A. Wolff, Nature (London) **391**, 667 (1998).

[2] J. A. Porto, F. J. Garcia-Vidal, and J. B. Pendry, Phys. Rev. Lett. **83**, 2845 (1999).

[3] F. J. Garcia-Vidal and L. Martin-Moreno, Phys. Rev. B **66**, 155412 (2002).

[4] Q. Cao and Ph. Lalanne, Phys. Rev. Lett. **88**, 057403 (2002).

[5] W. L. Barnes, A. Dereux, and T. W. Ebbesen, Nature (London) **424**, 824 (2003).

[6] J. Bravo-Abad, A. Degiron, F. Przybilla, C. Genet, F. J. Garcia-Vidal, L. Martin-Moreno, and T. W. Ebbesen, Nat. Phys. **2**, 120 (2006).

[7] Y. Xie, A. R. Zakharian, J. V. Moloney, and M. Mansuripur, Opt. Express **14**, 6400 (2006).

[8] C. Genet and T. W. Ebbensen, Nature (London) **445**, 39 (2007).

[9] H. Liu and Ph. Lalanne, Nature (London) **452**, 728 (2008).

[10] Lord Rayleigh, Proc. R. Soc. London Ser. A **89**, 194 (1913); Philos. Mag. Ser. 6 **14**, 60 (1907).

[11] R. W. Wood, Philos. Mag. Ser. 6 **4**, 396 (1902).

[12] H. A. Bethe, Phys. Rev. **66**, 163 (1944).

[13] A. Sommerfeld, *Optics* (Academic Press, New York, 1954), pp. 249–272.

[14] W. Pauli, Phys. Rev. **54**, 924 (1938).

[15] L. D. Landau and E. M. Lifshits, *Electrodynamics of Continuous Media* (Pergamon Press, Oxford, 1984).

[16] B. Sturman, E. Podivilov, and M. Gorkunov, Europhys. Lett. **80**, 24002 (2007).

[17] B. Sturman, E. Podivilov, and M. Gorkunov, Phys. Rev. B **76**, 125104 (2007).

[18] B. Sturman, E. Podivilov, and M. Gorkunov, Phys. Rev. B **77**, 075106 (2008).

[19] A. Roberts and R. C. McPhedran, J. Mod. Opt. **34**, 511 (1987).

[20] H. Raether, *Surface Plasmons* (Springer-Verlag, Berlin, 1998).

[21] S. Collin, F. Pardo, R. Teissier, and J.-L. Pelouard, Phys. Rev. B **63**, 033107 (2001).

[22] P. Lalanne, C. Sauvan, J. P. Hugonin, J. C. Rodier, and P. Chavel, Phys. Rev. B **68**, 125404 (2003).

[23] A. Yu. Nikitin, S. G. Rodrigo, F. J. García-Vidal, and L. Martín-Moreno, New J. Phys. **11**, 123020 (2009).

[24] X. Y. Yang, H. T. Liu, and P. Lalanne, Phys. Rev. Lett. **102**, 153903 (2009).

[25] A. Yu. Nikitin, F. J. García-Vidal, and L. Martín-Moreno, Phys. Rev. Lett. **105**, 073902 (2010).

[26] B. Sturman, E. Podivilov, and M. Gorkunov, Phys. Rev. B **82**, 115419 (2010).

[27] H. Liu and P. Lalanne, Phys. Rev. B **82**, 115418 (2010).

[28] Y. Takakura, Phys. Rev. Lett. **86**, 5601 (2001).

[29] F. Yang and J. R. Sambles, Phys. Rev. Lett. **89**, 063901 (2002).

[30] J. Lindberg, K. Lindfors, T. Setälä, M. Kaivola, and A. T. Friberg, Opt. Express **12**, 623 (2004).

[31] H. F. Schouten, T. D. Visser, G. Gbur, D. Lenstra, and H. Blok, Phys. Rev. E **67**, 036608 (2003).

[32] H. F. Schouten, T. D. Visser, G. Gbur, D. Lenstra, and H. Blok, Phys. Rev. Lett. **93**, 173901 (2004).

[33] J. Bravo-Abad, L. Martin-Moreno, and F. J. Garcia-Vidal, Phys. Rev. E **69**, 026601 (2004).

[34] R. Gordon, Phys. Rev. B **73**, 153405 (2006).

[35] R. Gordon, Phys. Rev. B **75**, 193401 (2007).

[36] Ph. Lalanne, J. P. Hugonin, and J. C. Rodier, Phys. Rev. Lett. **95**, 263902 (2005).

[37] Ph. Lalanne, J. P. Hugonin, and J. C. Rodier, J. Opt. Soc. Am. B **23**, 1608 (2006).

[38] G. Gay, O. Alloschery, B. Viaris de Lesegno, J. Weiner, and H. J. Lezec, Phys. Rev. Lett. **96**, 213901 (2006).

[39] A. I. Fernandez-Dominguez, F. J. Garcia-Vidal, and L. Martin-Moreno, Phys. Rev. B **76**, 235430 (2007).

[40] M. Mechler, O. Samek, and S. V. Kukhlevsky, Phys. Rev. Lett. **98**, 163901 (2007).

[41] H. I. Perez, C. I. Valencia, E. R. Mendez, and J. A. Sanchez-Gil, J. Opt. Soc. Am. A **26**, 909 (2009).

[42] P. B. Jonson and R. W. Christy, Phys. Rev. B **6**, 4370 (1972).

[43] *Handbook of Optics*, edited by M. Bass (McGraw-Hill, New York, 1995), Vol. 2.

[44] R. E. Collin, *Field Theory of Guided Waves* (Wiley Interscience, New York, 1991).

[45] J. Meixner, IEEE Trans. Antennas Propag. **20**, 442 (1972).

[46] J. H. Hetherington and M. F. Thorpe, Proc. R. Soc. London A **438**, 591 (1992).

[47] G. R. Hardley, J. Lightwave Technol. **20**, 1219 (2002).