Comparative study of surface plasmon scattering by shallow ridges and grooves

Giovanni Brucoli[*] and L. Martín-Moreno

Instituto de Ciencia de Materiales de Aragón y Departamento de Física de la Materia Condensada,
CSIC-Universidad de Zaragoza, E-50009 Zaragoza, Spain

We revisit the scattering of surface plasmons by shallow surface defects for both protrusions and indentations of various lengths, which are deemed infinite in one dimension parallel to the surface. Subwavelength protrusions and indentations of equal shape present different scattering coefficients when their height and width are comparable. In this case, a protrusion scatters plasmons like a vertical point dipole on a metal surface, while an indentation scatters like a horizontal point dipole on a metal surface. We corroborate that long and shallow asymmetrically shaped surface defects have very similar scattering coefficients, as already found with approximate methods. In the transition from short shallow scatterers to long shallow scatterers, the radiation can be understood in terms of interference between a vertical and a horizontal dipole. The results attained numerically are exact and accounted for with analytical models.

PACS number(s): 42.82.Et, 41.20.Jb, 73.20.Mf, 78.67.−n

I. INTRODUCTION

Surface plasmon polaritons (SPPs) are electromagnetic bound modes responsible for the transport of light at the interface separating a metal from a dielectric. Their ability to confine light at an air-dielectric interface offers the prospect of developing a new technology consisting of photonic nanodevices.[1–4] Active research is currently focusing on the possibility of achieving control over the propagation of SPPs by means of optical elements that would couple or decouple light to them.[5–9] In order to conceive optical elements (lenses, mirrors, beam splitters) able to manipulate SPP propagation, we need to learn more about the interaction of surface plasmons with a subwavelength modification of the underlying dielectric metal interface. Indeed the interaction of SPPs with surface subwavelength defects on a metal surface is of great interest from a theoretical standpoint.[10,11]

In this article we shall study scattering of SPPs by a *shallow* surface defect. We will consider both indentations of the metal surface (grooves) and protrusions on it (ridges). We shall deal with bidimensional defects, which are deemed infinite in one dimension parallel to the interface (the *y* direction). Different aspects of this problem have been studied before with a variety of numerical techniques.[12–21] Here we present a systematic comparison between the different scattering coefficients and provide both analytical expressions and qualitative explanations.

It must be noted that in a previous work we presented such a comparison[22] but within an approximate numerical scheme. Within that framework it was found that ridges and grooves exhibited the same scattering, whenever they were shallow enough. In this paper we will revise that result. We solve the Maxwell equations through a discretization method whose accuracy depends only on the discretization mesh. We find that, as in the previous work, long asymmetric ridges or grooves (much wider than deep) do scatter very similarly. However, square shallow defects manifest a different scattering efficiency, mainly in the relative radiative loss and radiation pattern. The lack of distinction between these two cases did not emerge in the previous approximate treatment. The mistaken outcome of Ref. 22 for short defects may be traced back to the

breakdown of the assumption of small curvature in the defect geometry that was made there.

On the whole, the problem needs to be revisited so as to (i) substantiate why the approximate result does work in the case of elongated defects, (ii) point out what is the correct result in the case of shallow and short symmetric defects, and (iii) explain qualitatively how the scattering properties of short and shallow symmetric defects are gradually transformed into the scattering properties of elongated defects as the aspect ratio of the defect increases.

This paper is organized as follows. In Sec. II we state the basic assumptions on the scattering system as well as the solution method. In Sec. III we rearrange the asymptotic expansions of the far field to produce the scattering coefficients. Namely, we express the far field and the related Poynting vector in terms of the field inside the defect. Still in this section we look at an approximation for the scattering coefficients of shallow ridges. In Sec. IV we explain that, in general, we cannot quantitatively represent a scatterer (however small) by one mesh. We explain how we associate a small symmetric ridge or groove to a point dipole. In Sec. V we look at exact numerical results for the scattering of shallow defects of various horizontal lengths. We analyze these results and, in the case of square defects, show that a ridge can be associated with a vertical dipole, and a groove with a horizontal dipole. In Sec. VI we produce an analytical model that explains the radiation pattern of the surface plasmons scattered by small square ridges and grooves. In Sec. VII we look at the solutions for the case of shallow and long defects, and we present a clear-cut interpretation to support the results of the previous treatment.[22] Finally in Sec. VIII we explain qualitatively that the aspect ratio of the defect determines the orientation of the field induced in a shallow defect.

II. THE SCATTERING SYSTEMS CONSIDERED

The considered defects are infinite in the *y* dimension and shallow (with depth $h \ll \lambda$, where λ is the free-space wavelength). The defects are illuminated by a monochromatic surface plasmon propagating in the *x* direction, $\mathbf{e}_{spp+}(\mathbf{r})$, associated with an impinging energy flux S_{spp}, defined
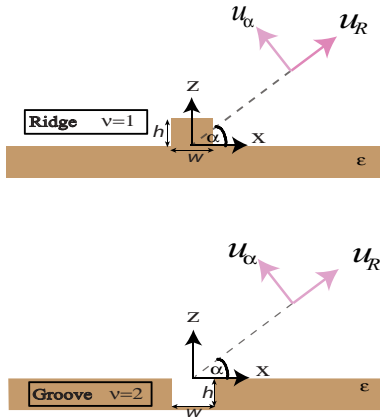
FIG. 1. (Color online) Schematic representation of the scattering systems considered. A ridge is a bidimensional defect located in air and labeled by the index $\nu = 1$. A groove is a bidimensional defect located in the metal and labeled by the index $\nu = 2$.

and derived in Appendix A. Therefore, only radiation into p-polarized (TM) waves needs to be considered. After we drop, out of symmetry, the y dependence on the whole problem, the field is expressed as $\mathbf{E}(\mathbf{r},t) = \mathbf{E}(x,z)e^{-i\omega t}$ where $\omega = cg$ and where $g = 2\pi/\lambda$ is the wave vector in vacuum. The material making the slab will be lossless silver, that is, $\varepsilon = \mathrm{Re}\{\varepsilon_{\mathrm{Ag}}(\lambda)\}$, where $\varepsilon_{\mathrm{Ag}}$ is taken from Ref. 23. Absorption is neglected as we consider nonresonant defects with widths much smaller than the SPP propagation length.

As represented in Fig. 1, we shall be expressing the source orientation in a Cartesian basis $(\mathbf{u}_x, \mathbf{u}_z)$, and the scattered fields in a right-handed orthogonal polar basis:

$$\mathbf{u}_R = \cos\alpha\, \mathbf{u}_x + \sin\alpha\, \mathbf{u}_z, \tag{1}$$

$$\mathbf{u}_\alpha = -\sin\alpha\, \mathbf{u}_x + \cos\alpha\, \mathbf{u}_z. \tag{2}$$

Finally a question of nomenclature: throughout the paper we shall refer to a bidimensional point dipole simply as a dipole. As represented in Fig. 1, each object lying in the vacuum semispace will be labeled by the superscript $\nu = 1$, while any object lying in the metal will be labeled by the superscript $\nu = 2$. In particular, scattering quantities related to ridges have the superscript $\nu = 1$, while the ones related to grooves have the superscript $\nu = 2$. The field within the cross-sectional area of the ridge is labeled $\mathbf{E}^r(\mathbf{r}')$ and the one within that of the groove is labeled $\mathbf{E}^g(\mathbf{r}')$.

III. SCATTERING COEFFICIENTS

The Green's tensor approach is a standard method to solve electromagnetic scattering problems.[11,24–30] Our first task in this section is to arrive at an explicit expression for the scattered electric far field. This is attained by propagating the field induced by a dipole density $\mathbf{P}^{(1)}(\mathbf{r}') = \Delta\varepsilon\, \mathbf{E}^r(\mathbf{r}')$ (where $\Delta\varepsilon = \varepsilon - 1$) inside the area of a ridge, to a point $\mathbf{R}$ very far from the source. For a groove we have the same relation between polarization and field (except for a change of sign) $\mathbf{P}^{(2)}(\mathbf{r}') = -\Delta\varepsilon\, \mathbf{E}^g(\mathbf{r}')$. To propagate the field from any of the

two, we use the standard formula[11]

$$\mathbf{E}_s(\mathbf{R}) = g^2 \int_A d\mathbf{r}'\, \hat{\mathbf{G}}(\mathbf{R},\mathbf{r}') \cdot \mathbf{P}(\mathbf{r}'). \tag{3}$$

Where $\hat{\mathbf{G}}(\mathbf{R},\mathbf{r}')$ is the Green's tensor for the air-metal background. The Green's tensor propagates the emission of a point source at $\mathbf{r}'$ to the distant point $\mathbf{R}$. One of the advantages of the Green's tensor technique is that once the fields inside the defects $\mathbf{E}(\mathbf{r}')$ [and thus $\mathbf{P}(\mathbf{r}')$] are computed numerically, the asymptotic expansions of scattered fields are analytic. This takes us to our second task, which is making a direct connection between the orientation of the induced polarization inside the defects and the far-field radiation pattern, and in so doing define the scattering coefficients.

First of all, finding the scattered electric far field $\mathbf{E}_s(\mathbf{R})$ requires the asymptotic expansions of the Green's tensor. The derivation is sketched in Appendix C. In what follows we give some simplifying rearrangements that will let us focus directly on the angular radiation pattern of surface defects.

A. Scattering into radiative modes

The asymptotic Green's tensor in the radiative zone for either a ridge or a groove can be written in a compact form as

$$\hat{\mathbf{G}}^{(\nu)}(R \to \infty, \alpha, \mathbf{r}') = \frac{e^{i(gR+\pi/4)}}{\sqrt{8\pi gR}}\, e^{-igx'\cos\alpha}$$
$$\times\, e^{-ik_z^{(\nu)}z'}\, \hat{\mathbf{G}}_\infty^{(\nu)}(\alpha, \mathbf{r}'). \tag{4}$$

In such form we can factor the asymptotic scalar Green's function out of the dyadic part of the Green's tensor. From Eq. (3), the direction of $\mathbf{E}_s(\mathbf{R})$ results from superposition of $\hat{\mathbf{G}}_\infty^{(\nu)}(\alpha, \mathbf{r}') \cdot \mathbf{P}^{(\nu)}(\mathbf{r}')$, the emission from all induced point polarization elements, or dipole density elements. Yet the direction of each contribution $\hat{\mathbf{G}}_\infty^{(\nu)}(\alpha, \mathbf{r}') \cdot \mathbf{P}^{(\nu)}(\mathbf{r}')$ must be independent of $\mathbf{r}'$. In other words, since electromagnetic waves are transverse waves in vacuum, far from their source, the field emitted by a dipole density element $\hat{\mathbf{G}}_\infty^{(\nu)}(\alpha, \mathbf{r}') \cdot \mathbf{P}^{(\nu)}(\mathbf{r}')$ must be proportional to $\mathbf{u}_\alpha$. In fact, using the standard asymptotic expansions (see Appendix C) we can write

$$\hat{\mathbf{G}}_\infty^{(\nu)}(\alpha, \mathbf{r}') \cdot \mathbf{P}^{(\nu)}(\mathbf{r}') = -[\boldsymbol{\Phi}^{(\nu)}(\alpha, \mathbf{r}') \cdot \mathbf{P}^{(\nu)}(\mathbf{r}')]\, \mathbf{u}_\alpha. \tag{5}$$

Where for a ridge,

$$\boldsymbol{\Phi}^{(1)}(\alpha, z') = \mathbf{k}_p^+(\alpha) + \mathbf{k}_p^-(\alpha)\, r_p(\alpha)\, e^{2igz'\sin\alpha}, \tag{6}$$

and for a groove,

$$\boldsymbol{\Phi}^{(2)}(\alpha) = t_p^{(1,2)}(\alpha)\, \mathbf{k}_p^{m+}(\alpha). \tag{7}$$

The vectors $\mathbf{k}_p(\alpha)$ are p waves defined in vacuum, while $\mathbf{k}_p^{m\pm}(\alpha)$ are defined in the metal. A reminder of their expressions, in terms of the angle α of Fig. 1, is reported in Appendix B, along with the expression for the Fresnel reflection and transmission coefficients: $r_p(\alpha), t_p^{(1,2)}(\alpha)$.

We are now in a position to write the expressions for the radiative fields. Plugging Eqs. (4) and (5) into Eq. (3), we can separate the electric far-field dependence into its radial and angular parts as

$$\mathbf{E}_s^{(\nu)}(R, \alpha) = -\frac{e^{i(gR-\pi/4)}}{\sqrt{8\pi gR}}\, E_s^{(\nu)}(\alpha)\, \mathbf{u}_\alpha. \tag{8}$$

Here the angular amplitude can be written as

$$E_s^{(\nu)}(\alpha) = g^2\, \Theta_{\text{rad}}^{(\nu)}(\alpha), \tag{9}$$

where $\Theta_{\text{rad}}^{(\nu)}(\alpha)$ is the scattering coefficient into radiative modes

$$\Theta_{\text{rad}}^{(\nu)}(\alpha) = \int_A d\mathbf{r}'\, e^{-igx'\cos\alpha}\, e^{-ik_z^{(\nu)}z'}\, \mathbf{\Phi}^{(\nu)}(\alpha,\mathbf{r}') \cdot \mathbf{P}^{(\nu)}(\mathbf{r}'). \tag{10}$$

In Eq. (8) the scattered field in the far zone consists of a cylindrical wave, transverse to the direction of propagation $\mathbf{u}_R$, and with a net angular amplitude determined by $g^2\Theta_{\text{rad}}^{(\nu)}(\alpha)$ resulting from an integral over the source region. The coefficient $\Theta_{\text{rad}}^{(\nu)}(\alpha)$ is actually the important bit in the formula, as its squared module determines the radiation pattern. As seen from Eq. (10), it represents the angular amplitude that results from the superposition of each scattering element taken with its own amplitude, phase, and optical path, analogous to how an antenna array determines its effective radiation pattern. The radiation is given by the intensity or Poynting vector in the far field. Accordingly the differential angular scattering cross section is

$$\frac{\partial \sigma_{\text{rad}}^{(\nu)}(\alpha)}{\partial \alpha} = \frac{|\mathbf{E}_s^{(\nu)}(R,\alpha)|^2 R}{S_{\text{spp}}} = \frac{g^3}{8\pi\, S_{\text{spp}}}\, |\Theta_{\text{rad}}^{(\nu)}(\alpha)|^2. \tag{11}$$

Finally, the net radiative loss σ_{rad} is defined as the integrated angular radiation

$$\sigma_{\text{rad}} = \int_0^\pi d\alpha\, \frac{\partial \sigma_{\text{rad}}^{(\nu)}(\alpha)}{\partial \alpha}. \tag{12}$$

B. Shallow defects and Green's tensor boundary conditions

Whenever the height of the defect is small enough, typically much smaller than the wavelength of the incident light, we can make the approximation $g|\mathbf{r}'| \ll 1$. That allows some simplification for the angular amplitude of a scattering element *above the surface*. Consider

$$\mathbf{\Phi}^{(1)}(\alpha,\mathbf{r}') = \mathbf{k}_p^+(\alpha) + \mathbf{k}_p^-(\alpha)\, r_p(\alpha)\, e^{2igz'\sin\alpha}$$
$$\simeq \mathbf{k}_p^+(\alpha) + \mathbf{k}_p^-(\alpha)\, r_p(\alpha) \equiv \mathbf{\Phi}^{(1)}(\alpha). \tag{13}$$

Hence, for shallow defects the Green's tensor dependence on $\mathbf{r}'$ of Eq. (4) is entirely given by the exponential factors $e^{-igx'\cos\alpha}\, e^{-ik_z^{(\nu)}z'}$, for both a source in the vacuum semispace and a source in the metal semispace. Indeed this turns out to simplify the comparison between $\mathbf{\Phi}^{(1)}(\alpha)$ and $\mathbf{\Phi}^{(2)}(\alpha)$, which we shall perform in detail in Sec. VI.

Before that, we need to highlight the relation between $\mathbf{\Phi}^{(1)}(\alpha)$ and $\mathbf{\Phi}^{(2)}(\alpha)$, under this approximation. Such relation emerges from the boundary conditions of the Green's tensor at the interface, which are

$$[\hat{\mathbf{G}}(\mathbf{R},x',z'=0^+) - \hat{\mathbf{G}}(\mathbf{R},x',z'=0^-)] \cdot \mathbf{u}_x = 0, \tag{14}$$

$$[\hat{\mathbf{G}}(\mathbf{R},x',z'=0^+) - \varepsilon\, \hat{\mathbf{G}}(\mathbf{R},x',z=0^-)] \cdot \mathbf{u}_z = 0. \tag{15}$$

Notice that, in the unperturbed system, space is translationally invariant in the horizontal direction x and this is reflected in

the x component of the vector in Eq. (13). Because of Eqs. (4) and (5), we can turn Eq. (14) into

$$\Phi_x^{(1)}(\alpha) = \Phi_x^{(2)}(\alpha) \equiv \Phi_x(\alpha). \tag{16}$$

The presence of surface charges at the interface implies, from Eq. (15), that the z components of the vector $\mathbf{\Phi}^{(\nu)}(\alpha)$ on either sides of the interface have the relation

$$\Phi_z^{(1)}(\alpha) = \varepsilon\, \Phi_z^{(2)}(\alpha). \tag{17}$$

C. Scattering into surface plasmons

Let us derive the scattering coefficient into surface plasmon modes. Note that in this one-dimensional problem, scattering will be into both the forward surface plasmon $\mathbf{e}_{\text{spp}+}(\mathbf{r})$, propagating in the positive x direction, and the backward plasmon $\mathbf{e}_{\text{spp}-}(\mathbf{r})$, propagating in the negative x direction, as defined in Appendix A. The emission by a point dipole or a point polarization element must result into a plasmon final state: $\hat{\mathbf{G}}_{p\pm}(\mathbf{R},\mathbf{r}') \cdot \mathbf{P}(\mathbf{r}') \propto \mathbf{e}_{\text{spp}\pm}(\mathbf{R})$, as shown in the derivation sketched in Appendix C. The asymptotic Green's tensor for a source *upon* ($\nu = 1$) or *in* ($\nu = 2$) the metal is

$$\hat{\mathbf{G}}_{p\pm}^{(\nu)}(\mathbf{R},\mathbf{r}') \cdot \mathbf{P}^{(\nu)}(\mathbf{r}') = \frac{-i}{2g\, S_{\text{spp}}}\left\{[\mathbf{e}_{\text{spp}\pm}^{(\nu)}(\mathbf{r}')]^* \cdot \mathbf{P}^{(\nu)}(\mathbf{r}')\right\}\mathbf{e}_{\text{spp}\pm}(\mathbf{R}). \tag{18}$$

Notice that $[\mathbf{e}_{\text{spp}\pm}^{(\nu)}(\mathbf{r}')]^*$ complies with Eqs. (14) and (15). Consequently the fields of the scattered plasmons are

$$\mathbf{E}_{p\pm}^{(\nu)}(\mathbf{R}) = -\frac{ig}{2\, S_{\text{spp}}}\Theta_{p\pm}^{(\nu)}\, \mathbf{e}_{\text{spp}\pm}(\mathbf{R}),$$
$$\Theta_{p\pm}^{(\nu)} = \int_{A^{(\nu)}} d\mathbf{r}'\, \mathbf{e}_{\text{spp}\pm}^*(\mathbf{r}') \cdot \mathbf{P}(\mathbf{r}'). \tag{19}$$

Furthermore the magnetic field related to the field scattered into SPPs is

$$\mathbf{H}_{p\pm}^{(\nu)}(\mathbf{R}) = -\frac{ig}{2\, S_{\text{spp}}}\Theta_{p\pm}^{(\nu)}\, \mathbf{h}_{\text{spp}\pm}(\mathbf{R}), \tag{20}$$

where $\mathbf{h}_{\text{spp}}$ is the magnetic field of a SPP, as proved in Appendix A.

Now, if the source $\mathbf{P}(\mathbf{r}')$ is produced by an incident surface plasmon field (as in our case), we can define the cross section for scattering into SPPs as

$$\sigma_p^\pm = \frac{\mathbf{E}_{p\pm}(\mathbf{R}) \times \mathbf{H}_{p\pm}^*(\mathbf{R})\ \cdot \mathbf{u}_x}{\mathbf{e}_{\text{spp}\pm}(\mathbf{R}) \times \mathbf{h}_{\text{spp}\pm}^*(\mathbf{R})\ \cdot \mathbf{u}_x} = \left|\frac{g}{2\, S_{\text{spp}}}\Theta_{p\pm}\right|^2 \tag{21}$$

for either $\nu = 1$ or $\nu = 2$.

Finally, we can define the total scattering cross section, which in the lossless case is equivalent to the extinction cross section:

$$\sigma_{\text{xtn}} = \sigma_p^+ + \sigma_p^- + \sigma_{\text{rad}}. \tag{22}$$

IV. RAYLEIGH LIMIT: CAUTIONARY REMARKS

Next we are going to develop solutions to point sources in a metal-plane background. However, before we proceed, in this section we clarify the relation of either ridges or grooves to point dipoles.

When the field inside a defect is obtained by mesh discretization, we assume that the field inside a single mesh is uniform, and deviations from the field at its center are deemed negligible. Yet, in general, the field in a defect cannot be represented by the field at its center alone. Let us explain a little bit further this point.

For simplicity, let us consider a defect (with cross-sectional area A) in a homogenous medium with dielectric constant ε_b, but the argument is the same in other backgrounds. The field at every mesh is found by solving self-consistently a system of coupled equations:[31]

$$\mathbf{E}(\mathbf{r}_i) = \mathbf{E}_b(\mathbf{r}_i) + g^2 \sum_{j \neq i} \hat{\mathbf{G}}_b(\mathbf{r}_i - \mathbf{r}_j) \cdot \Delta\varepsilon \, \mathbf{E}(\mathbf{r}_j) \frac{A}{N}$$

$$+ g^2 \hat{\mathbf{M}} \cdot \Delta\varepsilon \, \mathbf{E}(\mathbf{r}_i) - \frac{\hat{\mathbf{L}}}{\varepsilon_b} \cdot \Delta\varepsilon \, \mathbf{E}(\mathbf{r}_i), \tag{23}$$

where N is the total number of mesh points used to represent the defect. The indices i and j go from 1 to N, and $\mathbf{E}(\mathbf{r}_i)$ is the field at the mesh center. $\hat{\mathbf{L}}$ is a term related to the depolarization of light and arises from the quasistatic contribution of the Green's tensor. $\hat{\mathbf{M}}$ is a correction term to the Green's tensor in the region of the scatterer, which is useful to improve the accuracy of the calculation when the inhomogeneity is discretized.[11,32]

In practice, the number of mesh points N is increased until the calculation converges to the required precision. Then variations of $\mathbf{E}(\mathbf{r})$ on the length scale $\sim\sqrt{A/N}$ are properly represented in the solution. In the Rayleigh limit, for a defect of area A so small that $g^2 A \ll 1$, the scatterer behaves like a point source or a point dipole and the background field (in this case the illumination) can be considered uniform over A: $\mathbf{E}_b(\mathbf{r}) = \mathbf{E}_b$. A circular defect in a homogenous medium with dielectric constant ε_b is an exceptional case in which the net field at any point $\mathbf{r}_i$ converges to

$$\mathbf{E} = \mathbf{E}_b - \frac{\hat{\mathbf{L}}}{\varepsilon_b} \cdot \Delta\varepsilon \, \mathbf{E}. \tag{24}$$

This is because the field inside an infinitesimal (very subwavelength) circular shape is actually uniform, and thus scattering by such circular defects can be described by one mesh. In fact the extinction coefficient[33,34] can be derived from the field at the center alone:

$$\sigma_{\mathrm{xtn}} = g \, \mathrm{Im}\left[\int_A d\mathbf{r}' \, \Delta\varepsilon \mathbf{E}_b^*(\mathbf{r}') \cdot \mathbf{E}(\mathbf{r}') \right] \tag{25}$$

$$= g \, A \, \mathrm{Im}[\Delta\varepsilon \mathbf{E}_b^* \cdot \mathbf{E}]. \tag{26}$$

To prove this numerically, we have calculated σ_{xtn} for a cylinder represented by a single mesh, as in Eq. (24), and illuminated by a plane wave. First, we have checked that the one-mesh cross section of Eq. (26) coincides with the Mie theory result. Second, we have subdiscretized the cylinder into square meshes as rendered in the inset of Fig. 2. As also rendered in the figure, applying Eq. (23) we found that as the number of meshes grows, the scattering cross section calculated by the collection of meshes in Eq. (25) converges to the initial value of one single mesh of Eq. (26). Notice also that since the geometrical representation of a circle with a few square meshes is inaccurate, when N is small
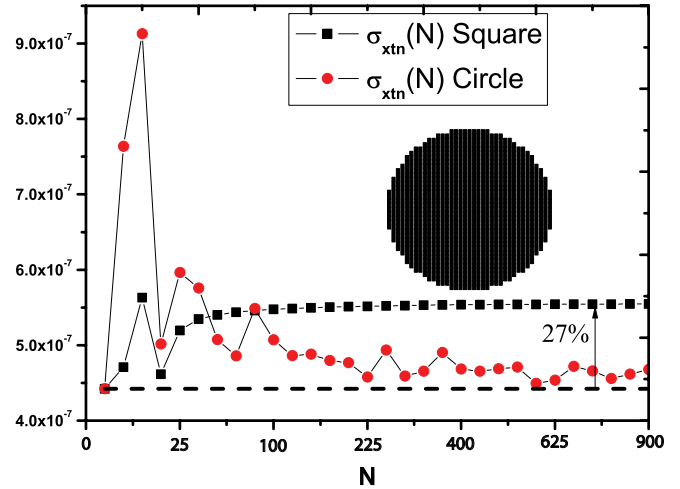

FIG. 2. (Color online) Dependence of the extinction coefficient on the number of meshes used in the calculation, for a square defect with $w = h = 1$ nm in vacuum, and a circular defect of the same area, illuminated by a plane wave. The dashed line represents extinction coefficient calculated with the Mie theory for the circle. The dielectric constant in the defect is $\varepsilon = -19.89$ at the wavelength of 700 nm. The inset represents the geometry of a discretized circle when tightly inscribed in a square represented by $N = 900$ mesh points.

the value of the extinction cross section deviates from the single-mesh value and the converged value (achieved when N is large).

As opposed to the case of a circle, the field inside a square scatterer can never be uniform if it is to satisfy real boundary conditions even in a homogenous medium or vacuum. Thus, it cannot be faithfully described by one mesh. This is illustrated in Fig. 2, which renders the extinction coefficient for a square defect of the same area as the circle. As it turns out, the converged value is $\sim$27% larger than that obtained by the one-mesh approximation. Remarkably this error is not reduced with the defect size: we obtained the same error for squares with sides 5 or 0.5 nm. This is just for reference in the optical range, since we found that the single-mesh approximation gives an error that actually depends on the type of defect and on the dielectric constant.

However, even if the field is not uniform, a small defect in the Rayleigh limit can be represented by a point source at the center of the mesh, with its field equal to the average field over the mesh $\overline{\mathbf{E}} = (1/A) \int_A d\mathbf{r}' \, \mathbf{E}(\mathbf{r}')$.

Indeed, if the variation of $\mathbf{E}_b(\mathbf{r})$ is negligible over the area of the defect, we have

$$\sigma_{\mathrm{xtn}} = g \, A \, \mathrm{Im}[\Delta\varepsilon \mathbf{E}_b^* \cdot \overline{\mathbf{E}}] \tag{27}$$

So the object behaves as a point dipole with dipole moment $\mathbf{p} = A \Delta\varepsilon \overline{\mathbf{E}}$.

The previous results are for a homogeneous background, but they also hold for the inhomogeneous one considered in this paper. We find that for a defect above the surface, the relative error is about 40% in the optical range, while it can reach 50% for a defects below the surface.

With very small nonelongated ridges and grooves, such that $w/\lambda \approx h/\lambda \ll 1$, the equivalent point dipoles are attained by averaging the fields over the area of the defects as follows:

$$\mathbf{p}^{(1)} = \Delta\varepsilon \overline{\mathbf{E}}^r A = \Delta\varepsilon \int_A d\mathbf{r}' \, \mathbf{E}^r(\mathbf{r}'), \qquad (28)$$

$$\mathbf{p}^{(2)} = -\Delta\varepsilon \overline{\mathbf{E}}^g A = -\Delta\varepsilon \int_A d\mathbf{r}' \, \mathbf{E}^g(\mathbf{r}') e^{-g|z'|\sqrt{|\varepsilon|}}. \qquad (29)$$

Accordingly, if we set $\mathbf{P}^{(\nu)}(\mathbf{r}') = \delta(\mathbf{r}')\mathbf{p}^{(\nu)}$, Eqs. (10) and (19) for small nonelongated defects become

$$\Theta_{\text{rad}}^{(\nu)}(\alpha) = \mathbf{\Phi}^{(\nu)}(\alpha) \cdot \mathbf{p}^{(\nu)}, \qquad (30)$$

$$\Theta_{p\pm}^{(\nu)} = \left[\mathbf{e}_{\text{spp}\pm}^{(\nu)}(0)\right]^* \cdot \mathbf{p}^{(\nu)}. \qquad (31)$$

V. NUMERICAL RESULTS

As an illustration, consider a square ridge and a groove of side $w = h = 10$ nm. We have calculated the scattering into radiative modes and SPPs without associating the defect to a point dipole but rather using Eqs. (11) and (21). In this case the major task is computing the Green's tensor for the plane metal surface required to attain the exact field within the surface defect. This can be achieved following the prescriptions of Refs. 35 and 36.

For the ridge, and the chosen parameters, the fraction of energy scattered into SPPs is $\sigma_p^+ = 5.6 \times 10^{-5}$ and $\sigma_p^- = 6.2 \times 10^{-5}$, which is represented in Fig. 3 along with the radiation patterns of Eq. (11). The total energy scattered into radiative modes is obtained by integrating the radiation distribution over all angles as in Eq. (12). For the square ridge of side 10 nm, this equals $\sigma_{\text{rad}} = 7.9 \times 10^{-5}$. For a groove of the same size, the fraction of energy scattered into SPPs is $\sigma_p^+ = 3.8 \times 10^{-5}$ and $\sigma_p^- = 4.2 \times 10^{-5}$, while the total energy scattered into radiative modes is $\sigma_{\text{rad}} = 1.3 \times 10^{-4}$. These calculations show that, for symmetric defects, the net radiative loss is greater for a groove than for a ridge. This is so because, while both the scattering into SPPs and the radiation close to the surface (at $\alpha = 0, \pi$) are similar for both defects, their radiation patterns greatly differ normal to the
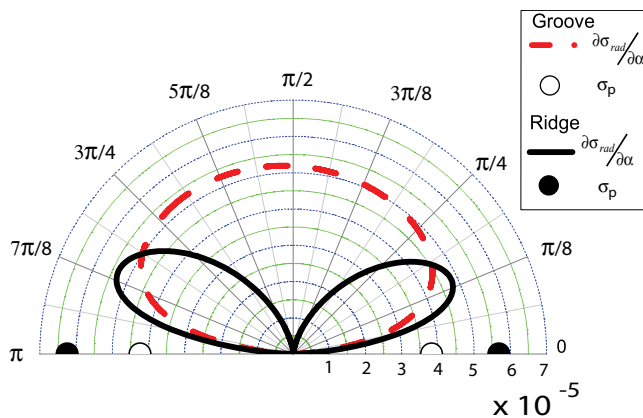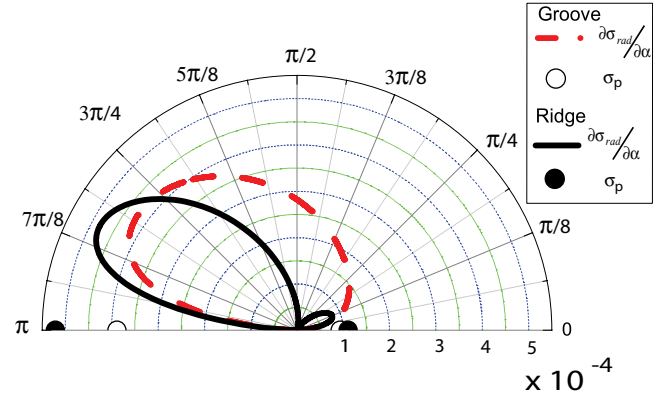
FIG. 3. (Color online) Angular radiative cross section [Eq. (11)] and surface plasmon cross section [Eq. (21) represented by the ● and ○ symbols at $\alpha = 0$ and $\alpha = \pi$], for square defects with 10 nm sides, illuminated by a SPP on silver at 700 nm.

FIG. 4. (Color online) Angular radiative cross section [Eq. (11)] and surface plasmon cross section [Eq. (21) represented by the ● and ○ symbols at $\alpha = 0$ and $\alpha = \pi$], for rectangular defects with 10 nm height and 50 nm width, illuminated by a SPP on silver at 700 nm.

surface ($\alpha = \pi/2$), where the groove radiation is maximum while the ridge radiation goes to zero. The ridge radiation pattern is distributed into two lobes on either side of $\alpha = \pi/2$, but the groove radiation pattern forms a single lobe. This is one of our main results and will be analyzed in detail in the next section. This result is not in agreement with those obtained in the approximate treatment of Ref. 22. We associate the discrepancy with the breakdown of the condition that the curvature of a short and shallow defect does not vary rapidly, used in that work.

Let us now keep the defect height at $h = 10$ nm and enlarge the width w. Figure 4 renders the radiation pattern for a rectangular defect of width 50 nm ($h = 10$ nm). The emergence of directivity in the out-of-plane radiation is part of a transitional behavior, in which the radiation patterns tend to align and, simultaneously, one of the lobes is shrunk while the other is blown up in the ridge radiation. Notice that the scattered energy into SPPs exhibits the same directivity, going mainly in reflection. Eventually, if we keep enlarging the defects until they are considerably asymmetric, the radiation patterns for both ridges and grooves are very similar, presenting single lobes (see Fig. 5). Noticeably, the scattering into SPPs is greatly reduced. Such similarity is explainable in the approximate framework presented in Ref. 22, which turns out to be quite acceptable in this limit of elongated defects, as we shall substantiate in Sec. VII. In Sec. VIII we shall account qualitatively for the transition observed in Fig. 4, explaining why the radiation pattern changes when the defects are enlarged.

A. Scattering by square ridges and grooves in the Rayleigh limit

The equivalence between nonelongated subwavelength defects and point dipoles gives us a chance to investigate in depth the individual radiation pattern by a symmetric defect as a single scattering element.

Figure 6 shows the averaged field inside the 10-nm ridges and grooves, prescribed in Eqs. (28) and (29). The field induced in a groove is mainly longitudinal, while the field inside the ridge is mainly transversal. This is due to both the illumination and the polarizability of the scatterers. When defects are almost
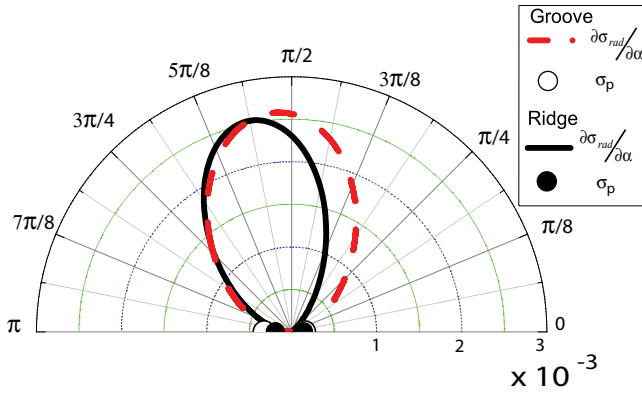
FIG. 5. (Color online) Angular radiative cross section [Eq. (11)] and surface plasmon cross section [Eq. (21) represented by the ● and ○ symbols at $\alpha = 0$ and $\alpha = \pi$], for rectangular defects with 10 nm height and 300 nm width, illuminated by a SPP on silver at 700 nm.

symmetric, their polarizabilities β_i are nearly isotropic and so the induced field and the incident field are virtually parallel. Hence the fields induced in a ridge and a groove are nearly parallel to the incident surface plasmon field $\mathbf{e}_{\text{spp}}$, which is mainly perpendicular to the plane in the vacuum semispace and is mainly parallel to the plane in the metal semispace. Therefore, in the Rayleigh limit, a ridge scatters SPPs into radiative modes like a vertical dipole on the plane, while the groove scatters SPPs into radiative modes like a horizontal dipole on a plane. The results for grooves is in agreement with Ref. 20.

Interestingly, we also have found numerically in Fig. 6 that

$$\left| \overline{E}_x^g \right| \sim \left| \sqrt{\varepsilon} \; \overline{E}_z^r \right|, \tag{32}$$

especially at short wavelengths. In reality we can see what happens by means of Eq. (24). Despite the fact that this equation is only exact for a circle in a homogenous background
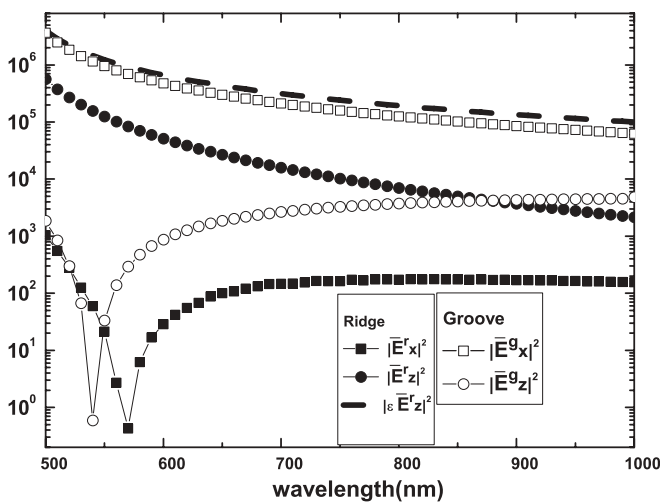

FIG. 6. (In a logarithmic scale) Intensity of various averaged field components defined in Eqs. (28) and (29) after having been normalized to the value of the SPP intensity in the air semispace $|\mathbf{e}_{\text{spp}}^{(\nu=1)}(0)|^2$. These values are calculated for a square groove and a square ridge of side 10 nm in lossless silver, as a function of the wavelength.

(as explained before), we can use it to show *qualitatively* the relation between the field inside the groove and the ridge, when their shapes are symmetric. Let us approximate the polarizability of a ridge for that of a circle of radius R in vacuum [whose polarizability is calculated through Eq. (24)], so $\beta_1 = 2/(\varepsilon + 1)$. Let us also approximate the groove polarizability by that of a circular hole in a homogenous metal medium, so we have $\beta_2 = 2\varepsilon/(\varepsilon + 1)$. Hence the field induced inside each object is estimated from the field at the center:

$$\overline{\mathbf{E}}^r \approx \beta_1 \mathbf{e}_{\text{spp}}(x = 0, z = R) = \beta_1 \mathbf{u}_z \, e^{-|k_{pz}|R}, \tag{33}$$

$$\overline{\mathbf{E}}^g \approx \beta_2 \mathbf{e}_{\text{spp}}(x = 0, z = -R) = \beta_2 \frac{\mathbf{u}_x}{\sqrt{\varepsilon}} \, e^{-|k_{pz}^m|R}. \tag{34}$$

Notice that for good metals like silver and gold, $e^{-|k_{pz}|R}$ grows monotonically with wavelength, from optical to infrared wavelengths, while $e^{-|k_{pz}^m|R}$ decays monotonically with wavelength in the same range.

Thus, these two factors are even more similar at short wavelengths than they are at large ones. Combining that for small defects, $e^{-|k_{pz}|R} \sim e^{-|k_{pz}^m|R}$ and that $\beta_2 = \varepsilon \beta_1$, Eqs. (33) and (34) lead to the result of Eq. (32).

The symmetry of the polarizations β_i and the property $\beta_2 \simeq \varepsilon \beta_1$ are strictly true for circular defects in homogeneous media. Our numerical calculations of Fig. 6 show that even though the field inside a ridge and a groove are quantitatively different from those of circular defects in homogenous media, the assumption that their mutual relation is preserved is in very good agreement with the exact result.

B. Reflection of surface plasmons by square shallow defects

As a corollary of the property of the fields in a ridge and a groove, $|\sqrt{\varepsilon} \overline{E}_z^r| \sim |\overline{E}_x^g|$, we can also substantiate that their reflection of surface plasmons is quite similar. In fact the reflection σ_p^- of Eq. (21) depends on the scattering coefficients of Eq. (31):

$$\left| \Theta_{p-}^{(1)} \right| \simeq |\Delta\varepsilon| A \left| \overline{E}_z^r \right|, \tag{35}$$

$$\left| \Theta_{p-}^{(2)} \right| \simeq |\Delta\varepsilon| A \left| \frac{\overline{E}_x^g}{\sqrt{\varepsilon}} \right| \simeq |\Delta\varepsilon| A \left| \overline{E}_z^r \right| \simeq \left| \Theta_{p-}^{(1)} \right|. \tag{36}$$

This property was obtained making the approximation $e^{-|k_{pz}|R} \sim e^{-|k_{pz}^m|R}$. As said, this tends to be more acceptable at short wavelengths, but even more so if the size of the particle (R) approaches zero.

To be precise, since strictly $e^{-|k_{pz}|R} > e^{-|k_{pz}^m|R}$, the reflection of SPPs by a ridge is always slightly larger than the one by a groove. Yet, the smaller the particle the smaller the difference between the reflection of SPPs by a symmetric ridge and a symmetric groove.

Once σ_{rad} from Eq. (12) and σ_p^- are determined, the value of the transmission of the surface plasmon is a constrained variable: $T = 1 - \sigma_p^- - \sigma_{\text{rad}}$, at least for the lossless case.[22] Since σ_{rad} is greater for grooves than for ridges, the transmission of a SPP after scattering is smaller in the case of a groove.
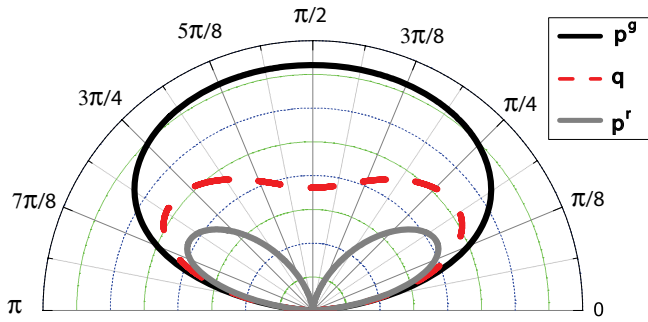
FIG. 7. (Color online) Radiation pattern $|\Theta_{\rm rad}(\alpha)|^2$ from the dipoles $\mathbf{p}^g = \sqrt{|\varepsilon|}\mathbf{u}_x$, $\mathbf{p}^r = \mathbf{u}_z$ and $\mathbf{q} = (\sqrt{|\varepsilon|}\mathbf{u}_x + \mathbf{u}_z)/\sqrt{2}$. The scale is linear with arbitrary units.

VI. RADIATION PATTERNS FOR HORIZONTAL AND VERTICAL POINT DIPOLES ON A REAL METAL INTERFACE

The first part of the expression (11) is a prefactor $g^3/(8\pi S_{\rm spp})$ whereas the second part is the *the radiation pattern* of a point dipole:

$$\left|\Theta_{\rm rad}^{(\nu)}\right|^2 = |\mathbf{\Phi}^{(\nu)}(\alpha) \cdot \mathbf{p}|^2. \quad (37)$$

A groove emits like a horizontal dipole. The angular amplitude of the field radiated by a horizontal unit dipole $\mathbf{p} = \mathbf{u}_x$, placed close to the interface $z = 0$, is $\Phi_x(\alpha)$, and it does not matter on which side of the interface it is placed. $\Phi_x(\alpha)$ can be derived using the relations in Appendix B 2 and the explicit result is

$$\Phi_x(\alpha) = \frac{2\sqrt{\varepsilon - \cos^2\alpha}\,\sin\alpha}{\sqrt{\varepsilon - \cos^2\alpha} + \varepsilon\sin\alpha}, \quad (38)$$

and the radiation pattern is $|\Phi_x(\alpha)|^2$. Notice $\Phi_x(\alpha)$ presents a mirror symmetry about the angle $\alpha = \pi/2$, the normal to the plane. Furthermore since $\Phi_x(\alpha)$ never changes sign between 0 and π (nor goes to zero), the field of a horizontal dipole has one single symmetric lobe, where the field always has the same sign.

To picture the form of the single-lobed radiation pattern emitted by a horizontal dipole, consider the radiation $|\sqrt{\varepsilon}\Phi_x(\alpha)|^2$ by a horizontal dipole $\mathbf{p}^g = \sqrt{\varepsilon}\,\mathbf{u}_x$ rendered in Fig. 7. Notice that such radiation pattern in Fig. 7 is in agreement with the one representing a groove in Ref. 19, obtained with a different numerical method.

If we consider $|\varepsilon| \gg 1$, then

$$\Phi_x \to 2\,\varepsilon^{-1/2}. \quad (39)$$

That is, when ε increases, this radiation pattern tends to simultaneously become isotropic and vanishing. In fact a horizontal dipole does not radiate on a perfect conductor.[37] For finite ε, the field $\Phi_x(\alpha)$ of a unit horizontal dipole within a real metal would not be completely screened, and while the pattern remains symmetric, its isotropy is disrupted parallel to the surface (i.e., $\alpha = 0,\pi$) to accommodate the emergence of surface plasmons in the density of electromagnetic states.[38]

For an individual vertical dipole $\mathbf{p} = \mathbf{u}_z$, which represents a ridge, the angular amplitude of the field is (see Appendix B 2)

$$\Phi_z^{(1)}(\alpha) = \frac{-2\varepsilon\sin\alpha}{\sqrt{\varepsilon - \cos^2\alpha} + \varepsilon\sin\alpha}\,\cos\alpha. \quad (40)$$

The field from a vertical dipole also goes to zero at $\alpha = 0,\pi$ for a finite ε but, since dipoles only radiate transversally, the field has a third zero at $\pi/2$. The field is antisymmetric with respect to the normal of the plane, while the intensity $|\Phi_z^{(1)}(\alpha)|^2$ is symmetric and is made up of the two lobes separated by a zero at $\pi/2$, see Fig. 7. Yet it is important to keep in mind that the field of one lobe is in antiphase with the field of the other.

Unlike what occurs for a unit horizontal dipole, the field radiated by a vertical dipole does not vanish for $|\varepsilon| \gg 1$; in fact,

$$\Phi_z^{(1)} \to -2\cos\alpha. \quad (41)$$

The total radiation from a unit vertical dipole has a larger weight than the radiation by a unit horizontal one, by a factor of $\sqrt{\varepsilon}$. This can be seen from Eq. (38) if we assume $\varepsilon \gg \cos^2\alpha$; this leads to

$$\Phi_z^{(1)}(\alpha) \simeq -\sqrt{\varepsilon}\;\Phi_x(\alpha)\cos\alpha. \quad (42)$$

In Fig. 7 we represent the radiation pattern of a horizontal dipole $\mathbf{p}^g = \sqrt{|\varepsilon|}\,\mathbf{u}_x$ and vertical dipole $\mathbf{p}^r = \mathbf{u}_z$ which corresponds to our analytic analog of the emission pattern of square ridges and grooves, respectively.

While we will consider a dipole with an intermediate orientation in the next section, we want to remark here that due to Eq. (42), the radiation patterns $|\mathbf{\Phi}^{(1)}(\alpha) \cdot \mathbf{p}^r|$ and $|\mathbf{\Phi}^{(2)}(\alpha) \cdot \mathbf{p}^g|$, by the vertical dipole $\mathbf{p}^r$ and the horizontal dipole $\mathbf{p}^g$, vanish parallel to the plane at $\alpha = 0,\pi$ in a similar manner, as illustrated in Fig. 7.

VII. SOLUTIONS FOR LONG AND SHALLOW RIDGES AND GROOVES

For shallow and long defects $w \gg h$ and $h/\lambda \ll 1$, we define the following height-averaged polarization densities and fields:

$$\widetilde{\mathbf{P}}^{(1)}(x') = \Delta\varepsilon \int_0^h dz'\mathbf{E}^r(x',z')$$
$$= \Delta\varepsilon\,\widetilde{\mathbf{E}}^{(1)}(x')\,h, \quad (43)$$

where the last equation defines $\widetilde{\mathbf{E}}^{(1)}(x')$. Likewise for a groove we can define $\widetilde{\mathbf{P}}^{(2)}(x')$ and $\widetilde{\mathbf{E}}^{(2)}(x')$ through the following equation:

$$\widetilde{\mathbf{P}}^{(2)}(x') = -\Delta\varepsilon \int_{-h}^0 dz'\mathbf{E}^g(x',z')\,e^{-g|z'|\sqrt{|\varepsilon|}}$$
$$= -\Delta\varepsilon\,\widetilde{\mathbf{E}}^{(2)}(x')\,h. \quad (44)$$

Notice that for $|\varepsilon| \gg 1$ we can make the approximation $k_{pz}^m \sim k_z^m \sim ig\sqrt{\varepsilon}$.

The benefit of using $\widetilde{\mathbf{P}}^{(\nu)}(x')$ is that the scattered-field coefficients for these defects in the far zone, $\Theta_{\rm rad}^{(\nu)}(\alpha)$ and $\Theta_{p\pm}^{(\nu)}$, are those emitted by a chain of point dipoles on the surface over the segment w, and set at 0^+ and 0^- for ridges and grooves, respectively.

The scattered field angular amplitude $\Theta_{\rm rad}^{(\nu)}(\alpha)$ from Eqs. (10) and (13) is obtained as

$$\Theta_{\rm rad}^{(\nu)}(\alpha) \simeq \mathbf{\Phi}^{(\nu)}(\alpha) \cdot \int_0^w dx'\,\widetilde{\mathbf{P}}^{(\nu)}(x')\,e^{-igx'\cos\alpha}. \quad (45)$$

This holds for the scattering into surface plasmon modes as well, since we have

$$\Theta_{p\pm}^{(\nu)} = \left[\mathbf{e}_{\text{spp}\pm}^{(\nu)}(0)\right]^* \cdot \int_0^w dx' \, e^{\mp ik_{px}x'} \, \widetilde{\mathbf{P}}^{(\nu)}(x'). \quad (46)$$

When we illuminate a shallow and long defect, with a SPP, an equivalent linear density of dipole sources $\widetilde{\mathbf{P}}(x')$ stems from how the induced fields are distorted inside the scatterer, namely, by its polarizability. When the defect was larger in the horizontal direction than in the vertical one, ridges and grooves were found to give the same scattering by an approximated Rayleigh expansion.[22] We have an alternative first principles argument to justify the Rayleigh expansion result, which is based entirely on the assumption that these defects are needle shaped. The field induced in the ridge tends to be that induced in a needle-shaped protrusion placed horizontally on the surface 0^+. For a groove, we have a horizontal needle-shaped cavity at 0^-. In such an idealistic simplification it is clear-cut to deduce the fields inside the defects from the boundary conditions. Namely, the parallel component of the incident field is always continuous and equal, as in Eqs. (A1) and (A2):

$$\widetilde{E}_x^{(1)}(x') = \mathbf{e}_{\text{spp}}(x',0) \cdot \mathbf{u}_x = \widetilde{E}_x^{(2)}(x'), \quad (47)$$

which preserves the continuity of Eq. (14). However, we are generating fields which, normal to the surface, make up for the discontinuity perpendicular to the metal surface of Eq. (15). In fact, for a horizontal needle-like ridge, the boundary conditions imposed by the continuity of the displacement vector are

$$\widetilde{E}_z^{(1)}(x') = \mathbf{e}_{\text{spp}}(x',0^+) \cdot \mathbf{u}_z/\varepsilon, \quad (48)$$

while for a needle-like slit,

$$\widetilde{E}_z^{(2)}(x') = \varepsilon \, \mathbf{e}_{\text{spp}}(x',0^-) \cdot \mathbf{u}_z. \quad (49)$$

Ultimately, using the expressions in Appendix A,

$$\widetilde{E}_x^{(1)}(x') = \widetilde{E}_x^{(2)}(x'), \quad (50)$$

$$\varepsilon \widetilde{E}_z^{(1)}(x') = \widetilde{E}_z^{(2)}(x'), \quad (51)$$

which, matched with Eqs. (14) and (15), yield

$$|\hat{\mathbf{G}}(\mathbf{R},x',z'=0^+) \cdot \widetilde{\mathbf{E}}^{(1)}(x')| = |\hat{\mathbf{G}}(\mathbf{R},x',z'=0^-) \cdot \widetilde{\mathbf{E}}^{(2)}(x')|, \quad (52)$$

and thus the property of producing the same scattering coefficients, previously found in Ref. 22. Of course this approximation considers the polarizability of the defects uniform and neglects deviations at the end points. Nonetheless it explains why elongated defects have similar scattering properties. The plasmon scattering by protrusions and indentations is similar because, far from the edges, a shallow but elongated defect behaves as an infinitely elongated one, as confirmed by numerical calculations. As an example, we report in Fig. 8 a numerical calculation of the fields averaged over the height for defects of $w = 300$ nm and $h = 10$ nm. This shows that Eqs. (50) and (51) are quite accurate at the center of the defect, and they deviate from the needle model prediction due to fringe effects at the edges.

It is worth mentioning that this equivalence is valid in the Rayleigh limit when the defect size is much smaller than the wavelength, and it may be altered at resonant wavelengths.
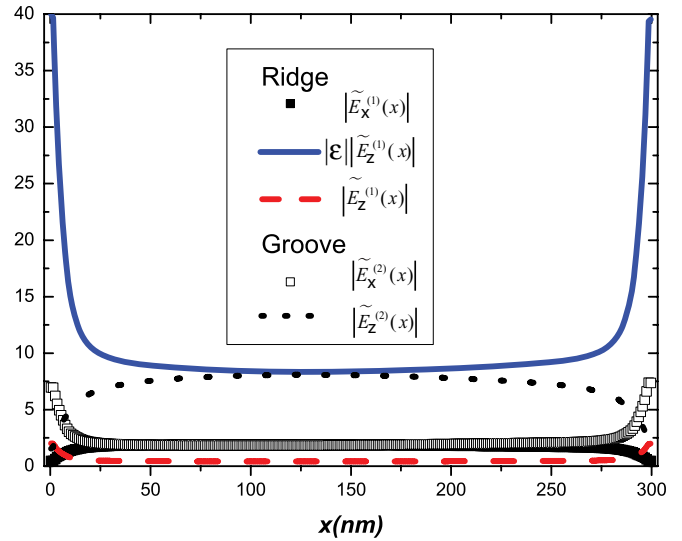

FIG. 8. (Color online) Modules of the averaged field components inside a ridge and a groove, $|\widetilde{E}_x^{(1)}|$, the quantity $|\varepsilon \widetilde{E}_z^{(1)}|$, and the rest of the components $|\widetilde{E}_x^{(1)}|$, $|\widetilde{E}_x^{(2)}|$, $|\widetilde{E}_z^{(2)}|$, for rectangular defects of $w = 300$ nm and $h = 10$ nm, normalized to the value of the SPP intensity in the air semispace $|\mathbf{e}_{\text{spp}}^{(\nu=1)}(0)|$. The system is illuminated by a SPP in lossless silver at $\lambda = 700$ nm.

VIII. THE TRANSITION FROM SHORT AND SHALLOW DEFECTS TO LONG AND SHALLOW DEFECTS: OBLIQUE DIPOLES ON A REAL METAL PLANE

The models presented in Secs. V and VI for symmetric surface defects are based on the fact that their polarizability is isotropic. As the defect width is increased, the aspect ratio becomes larger and this leads, progressively, to an asymmetric polarizability tensor. The first effect is that the induced field is gradually less and less parallel to the incident field. Therefore a ridge would develop a non-negligible horizontal electric field component, thus ceasing to be equivalent to a vertical dipole. Likewise the groove, which in the symmetric case behaves as a horizontal dipole, gradually starts having a non-negligible vertical component as its shape is elongated. The process goes on until we recover the case of a needle-shaped defect of Sec. VII. The fields inside a defect having intermediate width, as in Fig. 4, are intermediate between those for the needle case and the square symmetric case. Therefore in these cases, defects emit qualitatively like *oblique dipoles*, with the orthogonal components out of phase.

In order to better understand the radiation pattern by ridges and grooves, we decompose the oblique dipole in its horizontal and vertical components. First of all, we focus on the radiation pattern emitted by a ridge $\nu = 1$. From Eq. (30) a dipole with arbitrary orientation emits close to the surface, with a field angular amplitude:

$$\Theta_{\text{rad}}^{(1)}(\alpha) = \mathbf{\Phi}^{(1)}(\alpha) \cdot \mathbf{p}^{(1)} = \Phi_x(\alpha)\Delta^{(1)}(\alpha), \quad (53)$$

where $\Delta^{(1)}(\alpha) = p_x^{(1)} + [\Phi_z^{(1)}(\alpha)/\Phi_x(\alpha)] \, p_z^{(1)}$ and equals

$$\Delta^{(1)}(\alpha) = p_x^{(1)} + i \, \frac{\varepsilon \cos\alpha}{\sqrt{\cos^2\alpha + |\varepsilon|}} \, p_z^{(1)}. \quad (54)$$

$\Delta^{(1)}(\alpha)$ shows that the contribution to the radiative field coming from the vertical and horizontal dipoles on a metal

plane have a phase difference factor of $\pi/2$. This was already evident from Eq. (42), when $\varepsilon < 0$. Such phase difference arises from the impedance of a metal plane[22] $Z_s = -i/\sqrt{|\varepsilon|}$.

The radiation pattern for a dipole with arbitrary orientation and lying above the metal is written in our formalism as $|\Phi_x(\alpha)\Delta^{(1)}(\alpha)|^2$. The net angular amplitude for an oblique dipole is resolved into the superposition of the angular envelope of the horizontal dipole, with the other radiation factor $|\Delta^{(1)}(\alpha)|^2$. This last factor contains both the orientation and phase of the field. To envisage how these combine, we may develop $|\Delta^{(1)}(\alpha)|^2$ into three terms. These consist of the individual emission from the horizontal and vertical dipoles plus an interference term:

$$|\Delta^{(1)}(\alpha)|^2 = |p_x|^2 + \frac{|\varepsilon|^2 \cos^2\alpha}{|\varepsilon| + \cos^2\alpha}|p_z|^2$$
$$+ (-2)\frac{|\varepsilon|}{\sqrt{\cos^2\alpha + |\varepsilon|}}\mathrm{Im}[p_{1x}p_z^*]\cos\alpha. \quad (55)$$

In the presence of the plane metal background, we have that horizontal and vertical dipoles behave as individual sources, but their interference presents an intrinsic added phase difference of $\pi/2$, which is due to the different interaction of a horizontal and a vertical dipole with the plane. As a result, when in phase they do not interfere, and their radiation pattern is always symmetric regardless of the orientation of the dipole. This is the case for $\mathbf{q} = (\sqrt{|\varepsilon|}\mathbf{u}_x + \mathbf{u}_z)/\sqrt{2}$ where, as in Fig. 7, the radiation pattern is the sum of the angular intensity of a vertical and a horizontal dipole, so that at $\pi/2$ there is a minimum due to the vanishing of the vertical dipole contribution, and yet it never goes to zero because of the horizontal dipole contribution. Nevertheless, when the dipole components are not in phase, we can get asymmetric radiation patterns and additional zeros (to those at 0 and π), because the interference term can be negative. In such a case, the interference of the horizontal radiative field (with only one lobe) with the vertical radiative (with two lobes of different sign) is responsible for an asymmetric radiation pattern and exhibits directionality. This is illustrated in Fig. 9 for a dipole emission whose main contribution comes from the vertical dipole. In Fig. 10 we show the radiation pattern for a dipole whose main contribution comes from the horizontal dipole radiation.
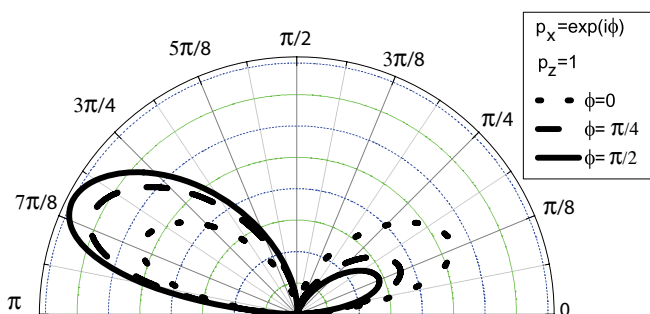

FIG. 9. (Color online) Radiation pattern $|\Theta_{\mathrm{rad}}(\alpha)|^2$ for a point dipole: $\mathbf{p} = \mathbf{u}_x e^{i\phi} + \mathbf{u}_z$, lying on top of a metal surface. The scale is linear with arbitrary units.
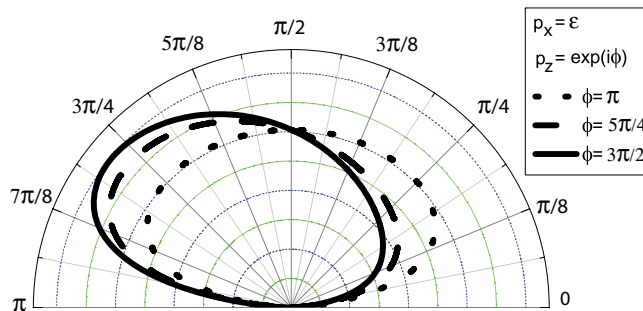

FIG. 10. (Color online) Radiation pattern $|\Theta_{\mathrm{rad}}(\alpha)|^2$ for a point dipole: $\mathbf{p} = \varepsilon\mathbf{u}_x + e^{i\phi}\mathbf{u}_z$, lying below a metal surface. The scale is linear with arbitrary units.

For the case of a groove ($\nu = 2$), the radiative angular field amplitude is, from Eq. (17),

$$\Theta_{\mathrm{rad}}^{(2)}(\alpha) = \Phi_x(\alpha)\Delta^{(2)}(\alpha)\, e^{-g|z'|\sqrt{|\varepsilon|}}, \quad (56)$$

$$\Delta^{(2)}(\alpha) = p_x^{(2)} + i\,\frac{\cos\alpha}{\sqrt{\cos^2\alpha + |\varepsilon|}}\, p_z^{(2)}, \quad (57)$$

where we recall that we have also added the approximation $k_z^m \simeq g\sqrt{\varepsilon}$ for $|\varepsilon| \gg 1$.

Remarkably, as opposed to the dipole emission over the surface, in the net emission from a dipole under the surface, the horizontal dipole contribution has a greater weight than the vertical dipole contribution. Apart from this, all the arguments used for a dipole over the surface apply.

The interference between the vertical and horizontal components of the field induced in the field generates the directional patterns of Fig. 4. For a ridge with length slightly larger than its height, the directional radiation is dominated by its vertical component. Figure 9 exemplifies the effect of the interference of a dominant vertical component with a smaller but non-negligible horizontal component. For even larger aspect ratios, the contribution from the other component may be comparable.

Likewise when a groove has a small aspect ratio, it is predominantly a horizontal source interfering with a smaller vertical source. The result is an interference pattern that looks like the one rendered in Fig. 10. Yet again this can be modified by increasing the aspect ratio. This transition is in good agreement with Fig. 11 of Ref. 19 where, using a different numerical method, the radiation pattern of a groove was computed for different aspect ratios.

IX. CONCLUSIONS

Our analysis of the surface plasmon scattering by square shallow defects into radiative modes and plasmon modes reveals that a groove scatters a larger fraction of incident energy than a ridge. The reflection by a symmetric ridge and a symmetric groove of the same size is similar and so is the radiative emission close to the horizontal direction. Indeed their scattering essentially differs in the vertical direction, where a groove radiates while a ridge does not. When defects become longer the polarizability gets more asymmetric. Correspondingly, since both components of the incident plasmon are out of phase, defects must be represented

by interfering horizontal and vertical dipoles on a plane, which interfere constructively in some direction, thus producing directionality in the radiation pattern. Finally when ridges and grooves are shallow and long, they tend to produce the same scattering because, apart for fringe effects, their polarizability exactly counterbalances the discontinuity of the incident surface plasmon field at the air-metal interface.

APPENDIX A: SURFACE PLASMON POLARITON MODE

The incident illumination is the field of a surface plasmon wave mode propagating in the positive x direction $(+)$ or negative x direction $(-)$, that is,

$$\mathbf{e}_{\mathrm{spp}\pm}^{(\nu=1)}(\mathbf{r}) = \left(\frac{\pm\mathbf{u}_x}{\sqrt{\varepsilon}} + \mathbf{u}_z\right) e^{i\pm k_p x} e^{-|k_{pz}|z}, \quad z > 0, \quad \text{(A1)}$$

$$\mathbf{e}_{\mathrm{spp}\pm}^{(\nu=2)}(\mathbf{r}) = \left(\frac{\pm\mathbf{u}_x}{\sqrt{\varepsilon}} + \frac{\mathbf{u}_z}{\varepsilon}\right) e^{i\pm k_p x} e^{-|k_{pz}^m|z}, \quad z < 0, \quad \text{(A2)}$$

where $k_p = g[\varepsilon/(\varepsilon+1)]^{1/2}$, $k_{pz} = ig/\sqrt{-\varepsilon-1}$, and $k_{pz}^m = -\varepsilon k_{pz}$. This can, alternatively, be written as $\mathbf{e}_{\mathrm{spp}\pm}(\mathbf{r}) = \mathbf{e}_{\mathrm{spp}\pm}^{(\nu=1)}$ for $z > 0$; and $\mathbf{e}_{\mathrm{spp}\pm}(\mathbf{r}) = \mathbf{e}_{\mathrm{spp}\pm}^{(\nu=2)}$ for $z < 0$.

The associated magnetic field is continuous at the interface and equal to

$$\mathbf{h}_{\mathrm{spp}\pm}(\mathbf{r}) = \frac{-i}{g}\nabla\times\mathbf{e}_{\mathrm{spp}\pm}(\mathbf{r}). \quad \text{(A3)}$$

Now consider a lossless metal (characterized by a real and negative dielectric constant ε) and a plasmon moving in the forward direction (the subscript $+$ will be omitted). The incident Poynting vector of the plasmon in the air side is

$$S_{\mathrm{spp}}^{(\nu=1)} = \int_0^\infty dz\, \mathbf{e}_{\mathrm{spp}}(\mathbf{r})\times\mathbf{h}_{\mathrm{spp}}^*(\mathbf{r})\cdot\mathbf{u}_x = \frac{k_p}{g}\frac{Z_s^2+1}{2|k_{pz}|}, \quad \text{(A4)}$$

while in the metal it is

$$S_{\mathrm{spp}}^{(\nu=2)} = \int_{-\infty}^0 dz\, \mathbf{e}_{\mathrm{spp}}(\mathbf{r})\times\mathbf{h}_{\mathrm{spp}}^*(\mathbf{r})\cdot\mathbf{u}_x = Z_s^4 S_{\mathrm{spp}}^{(\nu=1)}, \quad \text{(A5)}$$

where $Z_s = -i/\sqrt{|\varepsilon|}$. The total Poynting vector energy flux associated with a plasmon mode in a lossless metal is

$$S_{\mathrm{spp}} = S_{\mathrm{spp}}^{(\nu=1)} + S_{\mathrm{spp}}^{(\nu=2)} = \frac{\sqrt{-\varepsilon}}{2g}\frac{(\varepsilon+1)(\varepsilon^2-1)}{\varepsilon^3} \geqslant 0. \quad \text{(A6)}$$

APPENDIX B: P MODES

We shall repeat, for completeness, the explicit expression for p waves, particularly in the far field when $\mathbf{k}/g = \mathbf{u}_r$. In this case these modes are expressed in terms of the direct space polar angle α by noticing that $k_x = g\cos\alpha$ and $k_z = k_z^{(\nu=1)} =$ $g\sin\alpha$ in the air semispace, and $k_z^m = k_z^{(\nu=2)} = g\sqrt{\varepsilon-\cos\alpha^2}$ in the metal. Hence

$$\mathbf{k}_p^\pm(\alpha) = \frac{1}{g}(k_z\mathbf{u}_x \mp k_x\mathbf{u}_z) = \sin\alpha\,\mathbf{u}_x \mp \cos\alpha\,\mathbf{u}_z,$$

$$\mathbf{k}_p^{m\pm}(\alpha) = \frac{1}{\sqrt{\varepsilon}g}\left(k_z^m\mathbf{u}_x \mp k_x\mathbf{u}_z\right) \quad \text{(B1)}$$

$$= \sqrt{\frac{|\varepsilon|+\cos^2\alpha}{|\varepsilon|}}\mathbf{u}_x \mp \frac{\cos\alpha}{\sqrt{\varepsilon}}\mathbf{u}_z.$$

1. Reflection and transmission coefficients for a plane surface

For reference, we give here the Fresnel coefficients for an air-metal interface. In the present treatment we only deal with the reflection coefficient for a p wave propagating from air to metal, and this is

$$r_p = r_p^{(1,1)} = \frac{k_z^m - \varepsilon k_z}{k_z^m + \varepsilon k_z}, \quad \text{(B2)}$$

where notice that, for the sake of tidiness, we omit the superscript throughout.

As to the transmission coefficients, the one for a wave $(2,1)$ propagating from the metal to air is $t_p^{(2,1)}$, while the one for a p wave transmitted from the air medium to the metal is $t_p^{(1,2)}$.

$$t_p^{(2,1)} = \frac{2k_z^m\sqrt{\varepsilon}}{k_z^m + \varepsilon k_z}, \quad t_p^{(1,2)} = \frac{2k_z\sqrt{\varepsilon}}{k_z^m + \varepsilon k_z}. \quad \text{(B3)}$$

Notice that the transmission coefficients are related as follows:

$$\frac{t_p^{(1,2)}}{k_z^{(1)}} = \frac{t_p^{(2,1)}}{k_z^{(2)}}. \quad \text{(B4)}$$

2. Key identities

The following expressions for the reflection and transmission coefficients are essential to deriving Eqs. (38) and (40):

$$t_p^{(1,2)}(\alpha) = \frac{2\sqrt{\varepsilon}\sin\alpha}{\sqrt{\varepsilon-\cos^2\alpha} + \varepsilon\sin\alpha}, \quad \text{(B5)}$$

$$1 + r_p(\alpha) = \frac{2\sqrt{\varepsilon-\cos^2\alpha}}{\sqrt{\varepsilon-\cos^2\alpha} + \varepsilon\sin\alpha}, \quad \text{(B6)}$$

$$1 - r_p(\alpha) = \frac{2\varepsilon\sin\alpha}{\sqrt{\varepsilon-\cos^2\alpha} + \varepsilon\sin\alpha}. \quad \text{(B7)}$$

APPENDIX C: ASYMPTOTIC GREEN's TENSORS

The asymptotic expressions for the Green's tensor for three-dimensional scatterers are found in references.[11,33,39] For bidimensional grooves, we have already presented the derivation scheme in Appendix B of Ref. [30]. As explained therein the surface plasmon Green's tensor and the far-field Green's tensor are obtained from its angular spectrum. From the relevant Sommerfeld integral, the surface plasmon contribution is obtained by applying the residue theorem, and the far-field Green's tensor instead is obtained by applying the method of the steepest descent.

For the case of a ridge, we use the total Green's tensor of the background in the vacuum semispace. This can be written as the sum of the direct Green's tensor (the free-space Green's tensor) and the indirect Green's tensor (which gives the contribution due to the reflections at the metal-plane interface). Hence

$$\hat{\mathbf{G}}^{(1)}(\mathbf{R},\mathbf{r}') = \hat{\mathbf{G}}_0(\mathbf{R},\mathbf{r}') + \hat{\mathbf{G}}_s(\mathbf{R},\mathbf{r}'), \qquad \text{(C1)}$$

where the spectral representation for the direct Green's tensor is

$$\hat{\mathbf{G}}_0(\mathbf{R},\mathbf{r}') = \frac{i}{4\pi} \int_{-\infty}^{\infty} \frac{dk_x}{k_z} e^{ik_z(Z-z')} e^{ik_x(X-x')} \mathbf{k}_p^+ \mathbf{k}_p^+, \qquad \text{(C2)}$$

while for the indirect Green's tensor,

$$\hat{\mathbf{G}}_s(\mathbf{R},\mathbf{r}') = \frac{i}{4\pi} \int_{-\infty}^{\infty} \frac{dk_x}{k_z} e^{ik_x(X-x')} e^{ik_z(Z+z')} r_p \mathbf{k}_p^- \mathbf{k}_p^+. \qquad \text{(C3)}$$

Applying the residue theorem and the steepest descent method to $\hat{\mathbf{G}}^{(1)}(\mathbf{R},\mathbf{r}')$ we end up with Eqs. (4) and (18) for ($\nu = 1$).

For the groove case, we need to expand the Green's tensor connecting a point in the metal to a point in air. This is just

$$\hat{\mathbf{G}}^{(2)}(\mathbf{R},\mathbf{r}') = \frac{i}{4\pi} \int_{-\infty}^{\infty} \frac{dk_x}{k_z^m} e^{ik_x(X-x')} e^{i(k_z Z - k_z^m z')} t_p^{(2,1)} \mathbf{k}_p^{m+} \mathbf{k}_p^+. \qquad \text{(C4)}$$

Applying the residue theorem and the steepest descent method to $\hat{\mathbf{G}}^{(2)}(\mathbf{R},\mathbf{r}')$ we end up with Eqs. (4) and (18) for ($\nu = 2$). Notice that the form of $\hat{\mathbf{G}}^{\infty}(\alpha,\mathbf{r}') = -\mathbf{\Phi}^{(\nu)}\mathbf{u}_{\alpha}$ given in Sec. III is obtained by recognizing $\mathbf{k}_p^+ = -\mathbf{u}_{\alpha}$.

One more subtlety that might be confusing is how we pass from the transmission coefficient $t_p^{(2,1)}$ in the integral to the transmission coefficient $t_p^{(1,2)}$ in the asymptotic form $\mathbf{\Phi}^{(2)}$. This comes about because when we apply the method of the steepest descent to the integral, we get Eq. (4) with

$$\hat{\mathbf{G}}^{\infty}(\mathbf{r}') = \frac{k_z^m}{k_z} t_p^{(2,1)} \mathbf{k}_p^m \mathbf{u}_{\alpha} = t_p^{(1,2)} \mathbf{k}_p^m \mathbf{u}_{\alpha}, \qquad \text{(C5)}$$

where, in the last equation, we have used the identity (B4).

---

*gianni@unizar.es

[1] R. Zia, J. Sculler, A. Chandran, and M. Brongersman, Mater. Today **9**, 20 (2006).

[2] T. Ebbessen, C. Genet, and S. Bozhevolny, Phys. Today **61**, 44 (2008).

[3] S. A. Maier, *Plasmonics: Fundamentals and Applications* (Springer-Verlag, New York, 2006).

[4] E. Ozbay, Science **311**, 189 (2006).

[5] J. Krenn, H. Ditlbacher, G. Schider, A. Hohenau, A. Leitner, and F. R. Aussenegg, J. Microsc. **209**, 167 (2003).

[6] J.-C. Weeber, Y. Lacroute, A. Dereux, E. Devaux, T. Ebbesen, C. Girard, M. González, and A. Baudrion, Phys. Rev. B **70**, 235406 (2004).

[7] M. U. González, J.-C. Weeber, A.-L. Baudrion, A. Dereux, A. L. Stepanov, J. R. Krenn, E. Devaux, and T. W. Ebbesen, Phys. Rev. B **73**, 155416 (2006).

[8] I. P. Radko, S. I. Bozhevolnyi, G. Brucoli, L. Martín-Moreno, F. J. García-Vidal, and A. Boltasseva, Phys. Rev. B **78**, 115115 (2008).

[9] I. P. Radko, S. I. Bozhevolnyi, G. Brucoli, L. Martin-Moreno, F. J. Garcia-Vidal, and A. Boltasseva, Opt. Express **17**, 7228 (2009).

[10] A. V. Zayats, I. Smolyaninov, and A. Maradudin, Phys. Rep. **408**, 131 (2005).

[11] L. Novotny and B. Hecht, *Principles of Nano-Optics* (Cambridge University Press, Cambridge, England, 2006).

[12] F. Pincemin, A. A. Maradudin, A. D. Boardman, and J.-J. Greffet, Phys. Rev. B **50**, 15261 (1994).

[13] P. J. Valle, F. Moreno, J. M. Saiz, and F. González, Phys. Rev. B **51**, 13681 (1995).

[14] A. V. Shchegrov, I. V. Novikov, and A. A. Maradudin, Phys. Rev. Lett. **78**, 4269 (1997).

[15] J. A. Sánchez-Gil, Appl. Phys. Lett. **73**, 3509 (1998).

[16] J. A. Sánchez-Gil and A. A. Maradudin, Phys. Rev. B **60**, 8359 (1999).

[17] J. A. Sanchez-Gil and A. A Maradudin, Opt. Express **12**, 883 (2004).

[18] J. A. Sanchez-Gil and A. A. Maradudin, Appl. Phys. Lett. **86**, 251106 (2005).

[19] I. Chremmos, J. Opt. Soc. Am. A **27**, 85 (2010).

[20] G. Lévêque, O. J. F. Martin, and J. Weiner, Phys. Rev. B **76**, 155418 (2007).

[21] F. López-Tejeira, F. J. García-Vidal, and L. Martín-Moreno, Phys. Rev. B **72**, 161405 (2005).

[22] A. Y. Nikitin, F. López-Tejeira, and L. Martín-Moreno, Phys. Rev. B **75**, 035129 (2007).

[23] E. D. Palik, *Handbook of Optical Constants of Solids* (Academic, New York, 1985).

[24] H. W. Hohmann, Geophysics **40**, 309 (1975).

[25] G. Protásio, D. Rogers, and A. Giarola, Radio Sci. **17**, 503 (1982).

[26] O. Keller, Phys. Rev. B **34**, 3883 (1986).

[27] L. W. Li, J. Bennet, and P. Dyson, Int. J. Electron. **70**, 803 (1991).

[28] O. J. F. Martin, C. Girard, and A. Dereux, Phys. Rev. Lett. **74**, 526 (1995).

[29] L. B. Felsen and N. Marcuvitz, *Radiation and Scattering of Waves* (IEEE Press, New York, 2003).

[30] A. Y. Nikitin, G. Brucoli, F. J. García-Vidal, and L. Martín-Moreno, Phys. Rev. B **77**, 195441 (2008).

[31] M. Paulus and O. J. F. Martin, Phys. Rev. E **63**, 066615 (2001).

[32] O. J. F. Martin and N. B. Piller, Phys. Rev. E **58**, 3909 (1998).

[33] A. B. Evlyukhin, G. Brucoli, L. Martín-Moreno, S. I. Bozhevolnyi, and F. J. García-Vidal, Phys. Rev. B **76**, 075426 (2007).

[34] T. Søndergaard and S. I. Bozhevolnyi, Phys. Rev. B **67**, 165405 (2003).

[35] M. Paulus, P. Gay-Balmaz, and O. J. F. Martin, Phys. Rev. E **62**, 5797 (2000).

[36] T. Søndergaard and S. I. Bozhevolnyi, Phys. Rev. B **69**, 045422 (2004).

[37] A. Sommerfeld, *Partial Differential Equations in Physics* (Academic, New York, 1964).

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

[39] L. Novotny, J. Opt. Soc. Am. A **14**, 105 (1997).