Surface Electromagnetic Field Radiated by a Subwavelength Hole in a Metal Film

A. Yu. Nikitin,[1,2,][*] F. J. García-Vidal,[3] and L. Martín-Moreno[1,][†]

[1]*Instituto de Ciencia de Materiales de Aragón and Departamento de Física de la Materia Condensada,
CSIC-Universidad de Zaragoza, E-50009, Zaragoza, Spain*
[2]*A. Ya. Usikov Institute for Radiophysics and Electronics, Ukrainian Academy of Sciences, 61085 Kharkov, Ukraine*
[3]*Departamento de Física Teórica de la Materia Condensada, Universidad Autónoma de Madrid, E-28049 Madrid, Spain*

We present an analytical expression for the electromagnetic field at the surface radiated by a hole in a metal film. This expression is valid for any metal, from the optical range to longer wavelengths, and for distances to the hole larger than a few tens of nanometers. The field pattern presents a rich behavior, showing three regions (a complex short distance, an intermediate range dominated by surface plasmon polaritons, and a long-distance one dominated by Norton waves). The crossover distances between these regimes depend strongly on both the wavelength and the angle with respect to the incident field.

Introduction.—The transmission of light through small apertures in a metal has been a subject of research for centuries [1]. Traditionally the interest has centered on the transmitted intensity and far-field radiation pattern. Recently, the discovery of optical transmission resonances in hole arrays [2,3] and the appearance of the field of plasmonics [4,5] has attracted attention to the properties of the electromagnetic (EM) field at the metal surface. However, knowing the spatial field distribution at the metal surface of the radiation transmitted through the hole still requires either the notoriously difficult numerical computation of Sommerfeld integrals, or the use of electromagnetic numeric solvers, which can only deal with a small volume around the hole [6,7].

Here we present an analytic expression for the field at the surface radiated by a hole in a metal film. This expression is valid for any metal and, despite originating from a long-distance asymptotic expansion, it turns out to be an excellent approximation for distances to the hole as small as $\lambda/10$ (λ being the wavelength).

The field at the surface and the Green's dyadic.—We consider a hole in an optically thick metal film, back-illuminated by a monochromatic plane wave with an electric field $\mathbf{E}_{\rm inc}$, see Fig. 1. The top metal interface is set at $z = 0$. The dielectric at $z > 0$ is assumed to be vacuum and the metal is characterized by a frequency-dependent dielectric constant ϵ (which here we take from [8]). Throughout this work, all distances expressed in lower case letters are dimensionless, such that $x = k_\omega X$, $y = k_\omega Y$, etc., where $k_\omega = 2\pi/\lambda$ and X, Y are usual coordinates.

The field emerging from the hole $\mathbf{E}(\mathbf{r}, z)$ [where $\mathbf{r} = (x, y)$ are coordinates along the surface] can be expressed in terms of the electric field in the hole at a distance of one skin depth $\delta = \lambda/[2\pi\,{\rm Im}(\sqrt{\epsilon})]$ from the opening, see the supplementary material (SM) [9].

$$\mathbf{E}\,(\mathbf{r}, z = 0^+) = C \int_S d\mathbf{r}'\hat{G}(\mathbf{r} - \mathbf{r}')\mathbf{E}_h(\mathbf{r}'), \quad (1)$$

where $\mathbf{E}_h(\mathbf{r}') = \mathbf{E}(\mathbf{r}', z = -\delta)$, S is the cross section of the aperture, and $C = \iota\sqrt{\epsilon - 1}/k_\omega$. The quantity $\hat{G}(\mathbf{r})$ is the 3×3 Green's dyadic; each of its vector columns provide the electric field at point $(\mathbf{r}, z = 0^+)$ created by a dipole placed at the origin, and aligned with one of the coordinate axes. The expressions for $\hat{G}(\mathbf{r})$ in terms of integrals over diffraction modes are well known (see [10] for a pedagogical derivation) and can be found in the SM. Physically, the integrals show that the field at the surface arises from the interference of all possible diffraction modes (both radiative and evanescent), which are excited by the opening with an amplitude that depends both on their density of states at the surface and their coupling with the field at the aperture.

In general, the field inside the hole must be computed self-consistently. However, for subwavelength holes the transmission process is governed by the fundamental mode inside the hole, and the pattern of the field can be readily obtained from a convolution of this mode and $\hat{G}(\mathbf{r})$ through Eq. (1).

Here we concentrate on computing $\hat{G}(\mathbf{r})$, which directly gives the field radiated by a pinhole through its effective electric dipole $\mathbf{p}$ as
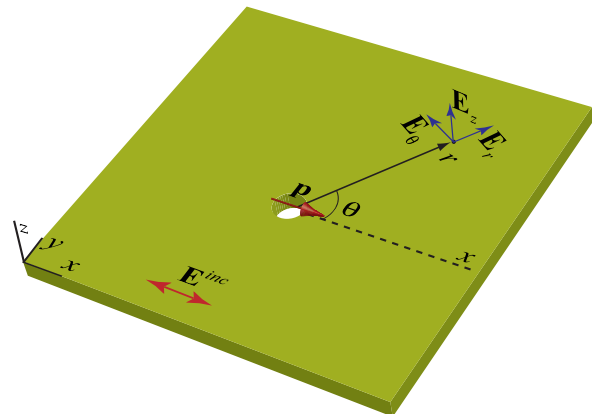

FIG. 1 (color online). The geometry of the studied system.

$$\mathbf{E}(\mathbf{r}, z = 0^+) = \hat{G}(\mathbf{r})\mathbf{p}, \qquad \mathbf{p} = C \int_S d\mathbf{r}' \mathbf{E}_h(\mathbf{r}'). \quad (2)$$

Because of the symmetry of the problem, it is convenient to compute both $\mathbf{E}(\mathbf{r}, z = 0^+)$ and $\hat{G}(\mathbf{r})$ in cylindrical coordinates (r, θ) (see Fig. 1). In this way, several elements of $\hat{G}$ are zero and the rest depend only on r. Moreover, as the field inside a small aperture has predominantly in-plane components, $p_z \approx 0$ and only a 2×3 subset of $\hat{G}(\mathbf{r})$ is needed. Assuming, without loss of generality, that the effective dipole p points along the x axis, the field at $\mathbf{r} = (r\cos\theta, r\sin\theta)$ is $\mathbf{E}(\mathbf{r}, z = 0^+) = \hat{G}(\mathbf{r}) \cdot (\cos\theta, \sin\theta)^T p$, with

$$\hat{G}(\mathbf{r}) = \hat{G}(r) = \begin{pmatrix} G_{rr}(r) & 0 \\ 0 & G_{\theta\theta}(r) \\ G_{zr}(r) & 0 \end{pmatrix}. \quad (3)$$

Asymptotic for the Green's dyadic.—The integrals defining $\hat{G}(\mathbf{r})$ do not have a closed analytical form and the difficulty of their numerical computation is legendary, due to the simultaneous presence of poles, branch cuts, and a strongly oscillatory integrand. To obtain an approximation to the Green's dyadic $\hat{G}^a(r)$, we have applied the steepest-descent method, modified in order to take into account the presence of both poles and branch points close to the path of integration [11]. This method is, in principle, valid for $r \gg 1$ but, as we will show later, in this particular problem it provides accurate results even for $r \approx 0.5$ (i.e., $R \approx 0.1\lambda$).

The details of the calculation are presented in the SM; here we give the final result:

$$\hat{G}^a(r) = f(r)\hat{g}(r), \qquad f(r) = \frac{\xi}{\lambda}\sqrt{\frac{1}{2\pi r}}e^{i\pi/4}, \quad (4)$$

where scalar term $f(r)$ reflects the 2D geometry of the waves in the plane. The dyadic $\hat{g}(r)$ can be decomposed as

$$\hat{g}(r) = \hat{g}^{\text{SPP}}(r) + \hat{g}^{\text{NW}}(r) + \widehat{\delta g}(r). \quad (5)$$

In this expression, $\hat{g}^{\text{SPP}}(r)$ is the contribution due to surface plasmon polaritons (SPP), which mathematically originates from the SPP pole:

$$\hat{g}^{\text{SPP}}(r) = e^{iq_p r}\hat{\Pi}(r), \qquad \hat{\Pi}(r) = 2\pi i \frac{q_{zp}}{\sqrt{q_p}}\left(\hat{P}_0 + \frac{1}{r}\hat{P}_1\right). \quad (6)$$

$$\hat{P}_0 = \begin{pmatrix} -q_{zp} & 0 \\ 0 & 0 \\ q_p & 0 \end{pmatrix}, \qquad \hat{P}_1 = \frac{iq_{zp}}{8q_p}\begin{pmatrix} -7 & 0 \\ 0 & 8 \\ 3q_p/q_{zp} & 0 \end{pmatrix}, \quad (7)$$

where $\xi = 1/\sqrt{\epsilon + 1}$, $q_p = \sqrt{1 - \xi^2}$, and $q_{zp} = -\xi$ are the in-plane and normal components of the SPP momentum, respectively, in units of k_ω. As it is well known, for a lossy metal $\text{Im} q_p \neq 0$ and the SPP fields decay exponentially with distance to the source. The SPP field is mainly perpendicular to the plane with a small in-plane component which is purely longitudinal in the $\hat{P}_0$ contribution ($G_{\theta\theta} = 0$), while the $\hat{P}_1$ term presents a transversal component. However, along any direction (different from $\theta = \pi/2$) the in-plane field tends to be radially polarized, as $E_\theta^{\text{SPP}}/E_r^{\text{SPP}} \sim \tan(\theta)/r$.

The term $\hat{g}^{\text{NW}}(r)$ is the contribution arising from the kink in the integrand related to the branch point at the light cone (see SM).

$$\hat{g}^{\text{NW}}(r) = \frac{e^{ir}}{r^{3/2}}\hat{N}, \qquad \hat{N} = \sqrt{2\pi}e^{-3i\pi/4}\begin{pmatrix} 1/\xi & 0 \\ 0 & -\xi \\ 1/\xi^2 & 0 \end{pmatrix}. \quad (8)$$

Its label NW reflects that it originates from the same part of the angular spectrum as the Norton wave found in the radiation of radio waves by a vertical dipole in a lossy dielectric [12]. As in that problem, the NW electric field amplitude decays algebraically with distance as $f(r)g^{\text{NW}}(r) \sim r^{-2}$ and points mainly along the z direction. However, while for the vertical dipole the in-plane component of the field is radially polarized and isotropic, in the case of radiation by a pinhole $E_\theta^{\text{NW}}/E_r^{\text{NW}} \sim \xi^2 \tan(\theta)$.

The close presence of the pole and the branch point modifies their respective density of EM states. As a result of this "interaction" there is a correction term $\widehat{\delta g}(r)$, given by

$$\widehat{\delta g}(r) = \frac{e^{-i\pi/4}}{2\sqrt{\pi\delta_p r}}e^{ir}\left[\hat{\Pi}(r) - \frac{\pi}{\delta_p r}\frac{q_{zp}}{\sqrt{q_p}}\hat{P}_0\right]$$
$$+ \left[\frac{1}{2}\text{erfc}(-\sqrt{i\delta_p r}) - 1\right]e^{iq_p r}\hat{\Pi}(r), \quad (9)$$

where $\text{erfc}(x) = (2/\sqrt{\pi})\int_x^\infty e^{-t^2}dt$ is the complementary error function, and $\delta_p = q_p - 1$. The expression for $\widehat{\delta g}(r)$ comes from a series expansion in inverse powers of r (see SM) where we have retained enough terms to represent both the leading long-distance and large ϵ dependencies. Actually, by performing the limit $\xi \to 0$ in Eqs. (2)–(9), we recover the exact (valid for all distances) field at the surface radiated by a hole in a perfect electrical conductor case (PEC, $|\epsilon| = \infty$): $E_z^{\text{PEC}}(r) = e^{ir}(ir^{-1} - r^{-2})/(2\pi) \times \int_S d\mathbf{r}' \mathbf{E}_h(\mathbf{r}')$. Evaluation of the previous expressions for different metals and wavelengths show that, at long distances, $\widehat{\delta g}(r)$ decays faster than both $\hat{g}^{\text{SPP}}(r)$ and $\hat{g}^{\text{NW}}(r)$.

To justify the validity of the previous expressions we have compared, for the case of a gold surface, $G_{zr}(r)$, and the asymptotic expression $G_{zr}^a(r)$. Similar results to those shown here have been obtained for the other nonzero components of the Green's dyadic, and other metals, from the optical regime to longer wavelengths. Figure 2
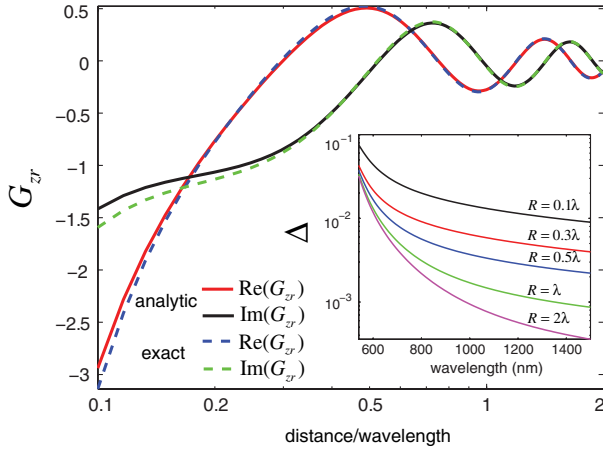
FIG. 2 (color online). The comparison between the exact numeric computation and the analytic approximation for $G_{zr}(r)$, in the case of a gold surface at $\lambda = 540$ nm. The inset shows the relative error Δ as a function of λ, for several distances to the hole.
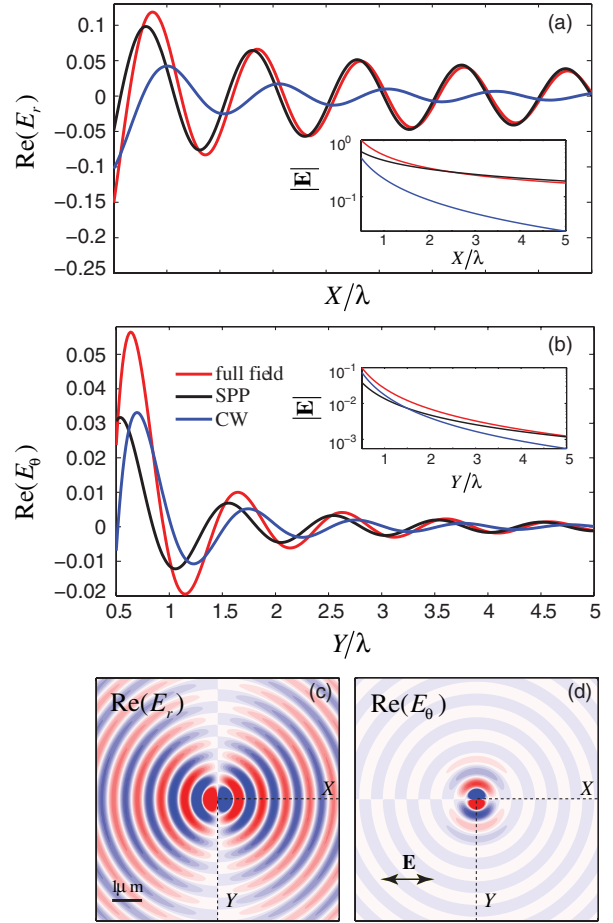

FIG. 3 (color online). The field along the gold surface radiated from a pinhole at $\lambda = 800$ nm. Panel (a) shows $\text{Re}[E_r(x, y = 0, z = 0)]$, together with the SPP and CW contributions, while panel (b) does the same for $\text{Re}[E_r(x = 0, y, z = 0)]$. Panels (c) and (d) render the spatial distribution of $\text{Re}[E_r(x, y, z = 0)]$ and $\text{Re}[E_\theta(x, y, z = 0)]$, respectively. All fields are normalized to $|E(x = \pi, y = 0, z = 0^+)|$.

renders the results for $\lambda = 540$ nm ($\epsilon = -5.26 + 2.06i$), which has been chosen to illustrate the validity of our result even in the unfavorable case of a small $|\epsilon|$. As can be seen, the asymptotic expression is an excellent approximation to the exact result, both in amplitude and phase, even at distances smaller than the wavelength. This is further illustrated in the inset to Fig. 2, which shows the dependence, with both wavelength and distance to the source, of the relative error $\Delta = |(G_{zr}^a - G_{zr})/G_{zr}|$. The relative error is larger for short distances and short wavelengths. Still, even at $\lambda = 540$ nm the maximum relative error is less than 10% for $R/\lambda \sim 0.1$ and decays rapidly to 1% already at the telecom wavelength $\lambda \simeq 1500$ nm.

The field at the surface radiated by a pinhole.—Equations (2)–(9) represent the main result of this Letter. In what follows we use them to get additional insight into the fields radiated by a pinhole. According to Eq. (3), the field along x has radial and z components, while it has only transversal θ component along the y direction. Figure 3 shows the dependence of the in-plane components of the electric field with distance to the pinhole along the x axis [panel (a)] and y axis [panel(b)], for the representative case of a gold surface at $\lambda = 800$ nm. These panels show that the field is strongly suppressed along the y direction and that, as in the 2D case [13–15], the field is well represented by a SPP already at 2–3 wavelengths away from the pinhole.

At shorter distances, the field pattern is more complex and, as was done for the 2D case [16], it is useful to define a "wave" substructing the SPP contribution from the total field, as $\mathbf{E}^{CW} \equiv \mathbf{E} - \mathbf{E}^{SPP}$ (denoted as CW, from its usual names "creeping wave" or "quasicylindrical wave").

Notice, however, that, in the intermediate region ($0.1\lambda < R < 2$–3λ), the CW contribution is larger relative to the SPP one along the y direction than along the

x direction. This is so because the SPP originates from p-polarized waves (which, as discussed before, are mainly longitudinal in the plane), while both p and s waves contribute to the total field (and thus to the CW). This angular dependence of the relative importance of SPPs and CWs is further illustrated in the insets to panels Figs. 3(a) and 3(b), where the amplitude of the total field at the surface is plotted along the two coordinate axes.

At even larger distances the exponentially decaying SPP is overcome by the algebraically decaying part in the total field corresponding to the NW contribution (in this distance regime $\widehat{\delta g} \ll \hat{g}^{NW}$). This transition has been recently studied in the simplified 2D geometry (radiation by a slit) [17–19], where it was found that the crossover distance depends upon the material and wavelength but, as a rule of thumb, it is located at 6–9 SPP propagation lengths, for most of the metals in the visible and infrared spectral regions [19].
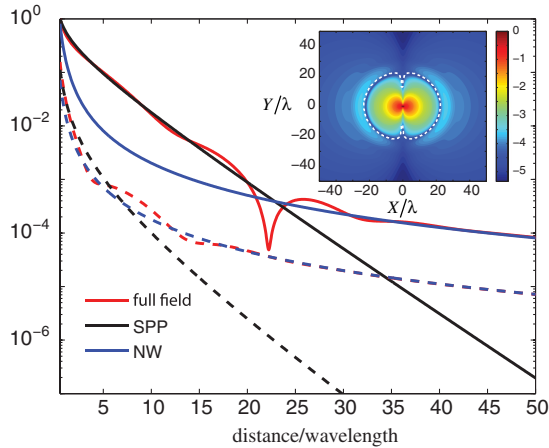
FIG. 4 (color online).   The electric field amplitude $|E(x, y, z = 0)|$, at the Au surface for $\lambda = 540$ nm, along with the SPP and Norton wave contributions. The continue curves correspond to the dependencies along Ox, while the dashed ones are along Oy. The inset shows the spatial distribution for the $\log_{10}[|E(x, y, z = 0)|]$. The crossover distance between SPP and NW is shown by the white discontinuous curve. All fields are normalized to $|E(x = \pi, y = 0, z = 0^{+})|$.

In the case of radiation by a pinhole we find that the crossover distance, from a SPP-dominated to a NW-dominated in-plane field, depends strongly on θ. This occurs because the SPP and the NW present the same dependence of E_θ/E_r with θ [$\sim \tan(\theta)$], but a different dependence along r. This is illustrated in Fig. 4 which, for a gold surface and $\lambda = 540$ nm, shows the full field, and the SPP and NW contributions along both the x axis (continuous lines) and y axis (broken lines). Notice that the good agreement with the full field confirms that $\widehat{\delta g}(r)$ is already negligible in this range of distances. In general, the crossover distance is maximum along the x direction and minimum along the y direction (23λ and 7λ, respectively, for the represented case). The crossover distances strongly depend on wavelength and metal; however, the calculations show that the crossover distance along the x direction is approximately 3 times larger than along the y direction in the optical regime.

To conclude, in this Letter we have derived a simple analytical expression for the field at the surface radiated by a pinhole in a metal film. The comparison with the exact numerical results shows that this analytical expression is valid for all metals, from the optical regime to longer wavelengths and, despite its asymptotic character, even at distances as small as $\lambda/10$. Three spatial regions have been identified: a short-distance region where the field differs from a SPP, an intermediate region where the SPP domi-

nates, and a long-distance region, where the SPP has died away and a different type of wave (the Norton wave) dominates. All crossover distances between these regions are strongly angular dependent. We expect our results to be of fundamental value in the research on problems such as extraordinary optical transmission, launching of SPPs by apertures, the shaping of EM fields by nanopatterning of metal surfaces, and the radiation of localized sources, such as molecules and quantum dots, placed on a metal surface.

*alexeynik@rambler.ru
†lmm@unizar.es

[1] C. Genet and T. Ebbesen, Nature (London) **445**, 39 (2007).
[2] T. W. Ebbesen, H. J. Lezec, H. F. Ghaemi, T. Thio, and P. A. Wolff, Nature (London) **391**, 667 (1998).
[3] F. J. Garcia-Vidal, L. Martin-Moreno, T. W. Ebbesen, and L. Kuipers, Rev. Mod. Phys. **82**, 729 (2010).
[4] A. V. Zayats, I. I. Smolyaninov, and A. A. Maradudin, Phys. Rep. **408**, 131 (2005).
[5] S. A. Maier, *Plasmonics: Fundamentals and Applications* (Springer-Verlag, New York, 2006).
[6] S.-H. Chang, S. K. Gray, and G. C. Schatz, Opt. Express **13**, 3150 (2005).
[7] E. Popov, N. Bonod, M. Neviere, H. Rigneault, P.-F. Lenne, and P. Chaumet, Appl. Opt. **44**, 2332 (2005).
[8] E. D. Palik, *Handbook of Optical Constants of Solids* (Academic Press, New York, 1985).
[9] See supplementary material at http://link.aps.org/supplemental/10.1103/PhysRevLett.105.073902 for mathematical details.
[10] L. Novotny and B. Hetch, *Principles of Nano-Optics* (Cambridge University Press, New York, 2006).
[11] L. P. Felsen and N. Marcuvitz, *Radiation and Scattering of Waves* (IEEE, Piscataway, NJ, 1994).
[12] K. A. Norton, Proc. IRE **24**, 1367 (1936).
[13] F. López-Tejeira, F. J. Garcia-Vidal, and L. Martin-Moreno, Phys. Rev. B **72**, 161405 (2005).
[14] P. Lalanne, J. P. Hugonin, and J. C. Rodier, Phys. Rev. Lett. **95**, 263902 (2005).
[15] B. Ung and Y. Sheng, Opt. Express **16**, 9073 (2008).
[16] P. Lalanne and J. P. Hugonin, Nature Phys. **2**, 551 (2006).
[17] P. Lalanne, J. P. Hugonin, H. T. Liu, and B. Wang, Surf. Sci. Rep. **64**, 453 (2009).
[18] W. Dai and C. M. Soukoulis, Phys. Rev. B **80**, 155407 (2009).
[19] A. Y. Nikitin, S. G. Rodrigo, F. J. Garcia-Vidal, and L. Martin-Moreno, New J. Phys. **11**, 123020 (2009).