

# Extraordinary transmission through 1, 2 and 3 holes in a perfect conductor, modelled by a mode expansion technique

**J.M. Brok**

*Department Imaging Science & Technology, Delft University of Technology, PO Box 5046,  
2600 GA Delft, The Netherlands*

[j.m.brok@tudelft.nl](mailto:j.m.brok@tudelft.nl)

**H.P. Urbach**

*Philips Research Laboratories, Professor Holstlaan 4, 5656 AA Eindhoven, The Netherlands*

**Abstract:** We discuss a mode expansion technique to rigorously model the diffraction from three-dimensional pits and holes in a perfectly conducting layer with finite thickness. On the basis of our simulations we predict extraordinary transmission through a single hole, caused by the Fabry-Perot effect inside the hole. Furthermore, we study the fundamental building block for extraordinary transmission through hole arrays: two and three holes. Coupled electromagnetic surface waves, the perfect conductor equivalent of a surface plasmon, are found to play a key role in the mutual interaction between two or three holes.

© 2006 Optical Society of America

**OCIS codes:** (050.1220) Apertures; (050.1960) Diffraction theory.

---

## References and links

1. H.A. Bethe, "Theory of diffraction by small holes," *Phys. Rev.* **66**, 163 (1944).
2. J. Meixner and W. Andrejewski, "Strenge Theorie der Beugung ebener elektromagnetischer Wellen and der vollkommen leitenden Kreisscheibe und an der kreisförmigen Öffnung im vollkommen leitenden ebenen Schirm," *Annalen der Physik* **7**, 157–168 (1950).
3. C. Flammer, "The vector wave function solution of the diffraction of electromagnetic waves by circular disks and apertures. I. Oblate spheroidal vector wave functions," *J. Appl. Phys.* **24**, 1218–1223 (1953).
4. C. Flammer, "The vector wave function solution of the diffraction of electromagnetic waves by circular disks and apertures. II. The diffraction problems," *J. Appl. Phys.* **24**, 1224–1231 (1953).
5. C.J. Bouwkamp, "Diffraction theory," *Reports on progress in physics* **17**, 35–100 (1954).
6. T.W. Ebbesen, H.J. Lezec, H.F. Ghaemi, T. Thio, and P.A. Wolff, "Extraordinary optical transmission through sub-wavelength hole arrays," *Nature* **391**, 667–669 (1998).
7. H.F. Schouten, T.D. Visser, D. Lenstra, and H. Blok, "Light transmission through a subwavelength slit: Waveguiding and optical vortices," *Phys. Rev. E* **67**, 036,608 (2003).
8. J. Bravo-Abad, L. Martin-Moreno, and F.J. Garcia-Vidal, "Transmission properties of a single metallic slit: From the subwavelength regime to the geometrical-optics limit," *Phys. Rev. E* **69**, 026,601 (2004).
9. H.F. Schouten, N. Kuzmin, G. Dubois, T.D. Visser, G. Gbur, P.F.A. Alkemade, H. Blok, G.W. Hooft, D. Lenstra, and E.R. Eliel, "Plasmon-assisted two-slit transmission: Young's experiment revisited," *Phys. Rev. Lett.* **94**, 053,901 (2005).
10. M.G. Moharam and T.K. Gaylord, "Rigorous coupled-wave analysis of metallic surface-relief gratings," *J. Opt. Soc. Am. A* **3**, 1780–1787 (1986).
11. J.A. Porto, F.J. Garcia-Vidal, and J.B. Pendry, "Transmission resonances on metallic gratings with very narrow slits," *Phys. Rev. Lett.* **83**, 2845–2848 (1999).
12. Q. Cao and P. Lalanne, "Negative role of surface plasmons in the transmission of metallic gratings with very narrow slits," *Phys. Rev. Lett.* **88**, 057,403 (2002).

13. L. Martin-Moreno and F.J. Garcia-Vidal, "Optical transmission through circular hole arrays in optically thick metal films," *Opt. Express* **12**, 3619–3628 (2004).
14. S.H. Chang, S.K. Gray, and G.C. Schatz, "Surface plasmon generation and light transmission by isolated nanoholes and arrays of nanoholes in thin metal films," *Opt. Express* **13**, 3150–3165 (2005).
15. F.J. Garcia-Vidal, E. Moreno, J.A. Porto, and L. Martin-Moreno, "Transmission of light through a single rectangular hole," *Phys. Rev. Lett.* **95**, 103,901 (2005).
16. F.J. Garcia de Abajo, "Light transmission through a single cylindrical hole in a metallic film," *Opt. Express* **10**, 1475–1484 (2002).
17. A. Roberts, "Electromagnetic theory of diffraction by a circular aperture in a thick, perfectly conducting screen," *J. Opt. Soc. Am. A* **4**, 1970–1983 (1987).
18. J.M. Brok and H.P. Urbach, "A mode expansion technique for rigorously calculating the scattering from 3D subwavelength structures in optical recording," *J. Mod. Opt.* **51**, 2059–2077 (2004).
19. J.B. Pendry, L. Martin-Moreno, and F.J. Garcia-Vidal, "Mimicking surface plasmons with structured surfaces," *Science* **305**, 847 (2004).
20. A.P. Hibbins, B.R. Evans, and J.R. Sambles, "Experimental verification of designer surface plasmons," *Science* **308**, 670 (2005).
21. Here and henceforth the square root of a complex number  $z$  is defined such that for real  $z > 0$  we have  $\sqrt{z} > 0$  and  $\sqrt{-z} = +i\sqrt{z}$ , with the branch cut along the negative real axis.
22. J.D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, New York, 1999).
23. Although  $\mathbf{v}_{\vec{\alpha}}(x, y)$  is real, we include the conjugation for consistency of the notation.
24. J. van Bladel, *Singular Electromagnetic Fields and Sources*, 1st ed. (Clarendon Press, Oxford, 1991).
25. H. Raether, *Surface Plasmons on Smooth and Rough Surfaces and on Gratings*, 1st ed. (Springer-Verlag, Berlin, 1988).
26. Although the  $\pm$  for  $\alpha_4$  now does not have anything to do with the propagation direction, for consistency, we stick to this notation.
27. The reader might wonder why we do not use the sine and cosine form always instead of using the exponential form only for  $|\gamma_z/k_p| < \epsilon$ . However, for large and purely imaginary  $\gamma_z$ , the functions  $\cos(\gamma_z z)$  and  $\sin(\gamma_z z)$  increase exponentially with  $|z|$ , which is not very convenient for numerical implementation.
28. The interaction integral also has some other properties that can save a lot of computation time. These properties are outside the scope of this paper.
29. This routine is a multi-dimensional adaptive quadrature over a hyper-rectangle. For a description, see the internet link: <http://www.nag.com/nagware/mt/doc/d01fcf.html>.

## 1. Introduction

The scattering from a single sub-wavelength hole in a conducting metal layer was already modelled in the 1940s and 1950s [1, 2, 3, 4, 5]. Apart from the assumption that the metal is perfectly conducting, the metal layer was also assumed to be infinitely thin; these assumptions made it possible to calculate the diffraction analytically. However, with respect to enhanced transmission through sub-wavelength hole arrays, as reported in Ref. [6], these assumptions are not valid. Furthermore, it is believed that the excitation of surface plasmons by neighboring holes plays an important role in the enhanced transmission.

To model this collective effect and to determine its importance, a lot of geometries have been studied and a lot of methods have been used. To name just a few: Infinitely long, single slits [7, 8] and double slits [9] in metals with finite conductivity were modelled with the Green's function approach. Coupled-wave analysis has been used to study gratings of slits (see, for example, [10, 11, 12]). Gratings that are periodic in two dimensions were studied with a modal method [13] and with the finite-difference time-domain method [14]. The modal method [15] and the boundary element method [16] were used for a single rectangular hole in a perfect conductor.

However, the specific case of two holes and their mutual interaction has not received much attention. This is remarkable, because it is a fundamental system and building block of a hole array. Yet, especially when the two holes are a wavelength or more apart, as the corresponding computational domain is then also large, this system is not easy to calculate.

For two-dimensional structures, such as infinite slits or circular symmetric setups, or for periodic systems in which the computational domain is only as large as one unit cell, current

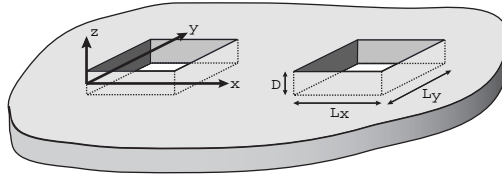


Fig. 1. Problem under consideration. Multiple rectangular holes in a perfectly conducting layer with finite thickness.

computer power suffices to obtain acceptable accuracy. For a finite number of three-dimensional holes that are far apart, however, a clever computational scheme is needed to prevent computation times of days or even weeks.

We present a rigorous modal method, similar to the methods in Refs. [17, 15, 18], that turns our three-dimensional diffraction problem into a two-dimensional numerical problem. Using this method, we calculate the transmission through one, two and three holes and we determine the influence that a second and third hole have on the transmission through the first. Parameters such as size of the hole, distance between two holes and thickness of the conducting layer are varied.

For the modal method that we use, the perfect conductor assumption is essential. Hence, our method is of quantitative value when the imaginary part of the index of refraction is so large, that the skin depth of the metal is small compared to the typical length scales of the geometry. This is for example the case in the terahertz and microwave frequency regimes, where also a lot of research is done in sub-wavelength hole arrays [19, 20]. However, we argue that, apart from absorption of energy in the metal, all relevant physics is present in our model. That also includes the surface plasmon equivalent of a perfect conductor: an electromagnetic surface wave. Thus, the understanding we gain from our results will also help us to understand extraordinary transmission in the optical regime.

We find that extraordinary transmission occurs for a single hole when the lowest order waveguide mode is just above cut-off. This is due to the Fabry-Perot resonance of this lowest order mode inside the waveguide. Furthermore, we calculate the transmission through two and three holes. In order to understand the mutual interaction between two and three holes, we normalize by the transmission through an identical but single hole. In this way we are able to isolate the effect that the presence of the second (and third) hole has on the transmission through the first. By changing the polarization of the incident light and by replacing one of the holes by a pit, we are able to show that coupled electromagnetic surface waves cause the enhanced and decreased transmission.

This paper is organized as follows. In the next section, we will define the problem under consideration. Then, we will describe the field above and below the layer, as well as inside the holes by mode expansions. Matching these expressions gives the system of equations that needs to be solved. After a few words on the numerical implementation, we will show results on single as well as multiple holes and pits.

## 2. Problem definition and system parameters

Let  $(x, y, z)$  be a rectangular Cartesian coordinate system. Perpendicular to the  $z$ -axis we have a perfectly conducting layer with finite thickness  $D$ . In this layer a finite number of rectangular holes and pits are present. See Fig. 1. A hole is a rectangular cylinder that is open on both sides and is as long as the thickness of the layer; a pit has an open end at one side, either at  $z = D/2$  or at  $z = -D/2$ , and a depth  $d_p < D$ . The sub- or superscript  $p$  denotes the number

of the pit or hole. The lengths in the  $x$ - and  $y$ -direction are  $L_x^p$  and  $L_y^p$ , respectively. The cross-section of a hole or a pit  $p$  is given by  $\Omega_p = \{ (x, y) \mid x_0^p < x < x_0^p + L_x^p, y_0^p < y < y_0^p + L_y^p \}$ . The halfspaces  $z > D/2$  and  $z < -D/2$  are filled with homogeneous dielectrics with index of refraction  $n_u$  and  $n_\ell$ , respectively. Every hole and pit is filled with a homogeneous dielectric with index of refraction  $n_p$ . The corresponding relative permeabilities are  $\epsilon_u = n_u^2$ ,  $\epsilon_\ell = n_\ell^2$  and  $\epsilon_p = n_p^2$ . The magnetic permeability is  $\mu_0$  everywhere.

A monochromatic incident field can originate from above and/or from below. The wavelength of the field in free space is given by  $\lambda$ . The local wavelengths are  $\lambda_u = \lambda/n_u$ ,  $\lambda_\ell = \lambda/n_\ell$  and  $\lambda_p = \lambda/n_p$ . The corresponding wave vectors are  $k_u = 2\pi/\lambda_u$ ,  $k_\ell = 2\pi/\lambda_\ell$  and  $k_p = 2\pi/\lambda_p$ . The harmonic time dependence of the electromagnetic field is given by the factor  $\exp(-i\omega t)$ , with  $\omega > 0$ , which will be omitted throughout.

### 3. Mode expansions

In each pit or hole, the electromagnetic field is expanded in a set of propagating or evanescent waveguide modes. In the next section, we describe these modes, that are characterized by the geometry of the pit or hole they live in, by their polarization, by their spatial frequency and by their direction of propagation.

In section 3.2, we describe the electromagnetic field above and below the conducting layer. This field is expanded in propagating and evanescent plane waves, that are characterized by their polarization and their direction of propagation.

#### 3.1. Inside the holes and pits

Solving Maxwell's equations inside the pits and holes means finding solutions of the scalar Helmholtz equation for every Cartesian component of the electromagnetic field. The boundary conditions imply that, at a perfect conductor, the tangential electric field as well as the normal magnetic field vanish. We then find solutions that are called waveguide modes. These are propagating or evanescent in the  $z$ -direction:

$$\begin{bmatrix} \mathbf{E}_\alpha(\mathbf{r}) \\ \mathbf{H}_\alpha(\mathbf{r}) \end{bmatrix} = \begin{bmatrix} \mathbf{E}_\alpha(x, y) \\ \mathbf{H}_\alpha(x, y) \end{bmatrix} e^{\pm i\gamma_z z}, \quad (1)$$

with propagation constant  $\gamma_z$  given by [21]:

$$\gamma_z = \sqrt{k_p^2 - \gamma_x^2 - \gamma_y^2}, \quad (2)$$

where  $\gamma_x$  and  $\gamma_y$  determine the spatial behaviour in  $x$  and  $y$ -direction:

$$\gamma_x = \frac{m_x \pi}{L_x^p}, \quad \gamma_y = \frac{m_y \pi}{L_y^p}, \quad (3)$$

with  $m_x$  and  $m_y$  integers. The bold subscript  $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$  is a multi-index that describes four discrete variables:  $\alpha_1$  (or  $p$ ) denotes the pit number,  $\alpha_2$  indicates the polarization (TE or TM),  $\alpha_3$  is determined by  $m_x$  and  $m_y$  and  $\alpha_4$  specifies whether the mode is travelling upwards or downwards.

Because the matching conditions at the interfaces  $z = \pm D/2$  only involve the  $x$ - and  $y$ -

component of the fields, it is convenient to introduce the following notation:

$$\mathbf{e}(x, y, z) \equiv \hat{\mathbf{i}}_z \times [\hat{\mathbf{i}}_z \times \mathbf{E}(x, y, z)] = \begin{pmatrix} -E_x \\ -E_y \\ 0 \end{pmatrix}, \quad (4a)$$

$$\mathbf{h}(x, y, z) \equiv \hat{\mathbf{i}}_z \times \mathbf{H}(x, y, z) = \begin{pmatrix} -H_y \\ H_x \\ 0 \end{pmatrix}, \quad (4b)$$

with  $\hat{\mathbf{i}}_z$  the unit vector in the  $z$ -direction. In this way, the lower case  $\mathbf{e}$  and  $\mathbf{h}$  are the rotated transverse components of the electric and magnetic field. Furthermore, we split the transverse components of the modes into a real part that depends on  $x$  and  $y$  and a complex part that depends on  $z$ :

$$\begin{bmatrix} \mathbf{e}_\alpha(x, y, z) \\ \mathbf{h}_\alpha(x, y, z) \end{bmatrix} = \mathbf{v}_{\bar{\alpha}}(x, y) \begin{bmatrix} \eta_\alpha(z) \\ \zeta_\alpha(z) \end{bmatrix}, \quad (5)$$

where the subscript  $\bar{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$  and thus the transverse vectorfield  $\mathbf{v}_{\bar{\alpha}}$  do not depend on the direction of propagation of the mode.

We normalize the real parts of the modes by [22]:

$$\langle \mathbf{v}_{\bar{\alpha}} | \mathbf{v}_{\bar{\alpha}} \rangle_{\Omega_p} \equiv \iint_{\Omega_p} [\mathbf{v}_{\bar{\alpha},x}(x, y) \mathbf{v}_{\bar{\alpha},x}(x, y)^* + \mathbf{v}_{\bar{\alpha},y}(x, y) \mathbf{v}_{\bar{\alpha},y}(x, y)^*] dx dy = 1, \quad (6)$$

where the superscript  $*$  denotes complex conjugation [23]. Furthermore, the modes are orthogonal such that for different modes  $\bar{\alpha}$  and  $\bar{\alpha}'$ :

$$\langle \mathbf{v}_{\bar{\alpha}} | \mathbf{v}_{\bar{\alpha}'} \rangle_{\Omega_p} = 0, \quad \text{if } \bar{\alpha} \neq \bar{\alpha}'. \quad (7)$$

Note that the time averaged Poynting vector in the  $z$ -direction of a mode is given by:

$$S_{\alpha,z} = \frac{1}{2} \text{Re} (E_{\alpha,x} H_{\alpha,y}^* - E_{\alpha,y} H_{\alpha,x}^*) = \frac{1}{2} \text{Re} (\eta_\alpha \zeta_\alpha^*) (\mathbf{v}_{\bar{\alpha},x} \mathbf{v}_{\bar{\alpha},x}^* + \mathbf{v}_{\bar{\alpha},y} \mathbf{v}_{\bar{\alpha},y}^*), \quad (8)$$

and, hence, the scalar product  $\langle \mathbf{v}_{\bar{\alpha}} | \mathbf{v}_{\bar{\alpha}} \rangle_{\Omega_p}$  of a mode  $\bar{\alpha}$  with itself is proportional to the flow of energy of this mode through a plane of constant  $z$ .

For a full listing of the waveguide modes, see Appendix A. The mode functions  $(\mathbf{E}_\alpha, \mathbf{H}_\alpha)$  are complete in the following sense: any time harmonic electromagnetic field with frequency  $\omega$  satisfying the source-free Maxwell equations inside the holes and pits can be expressed as a linear combination of these mode functions. Hence, for  $z$  between  $D/2$  and  $-D/2$ , we have:

$$\begin{bmatrix} \mathbf{E}^{\text{pit}}(\mathbf{r}) \\ \mathbf{H}^{\text{pit}}(\mathbf{r}) \end{bmatrix} = \sum_{\alpha} a_{\alpha} \begin{bmatrix} \mathbf{E}_{\alpha}(\mathbf{r}) \\ \mathbf{H}_{\alpha}(\mathbf{r}) \end{bmatrix}, \quad (9)$$

for some expansion coefficients  $a_{\alpha}$  that will be determined by matching the field inside the holes and pits to the field above and below the conducting layer.

### 3.2. Above and below the layer

The total electric and magnetic field above and below the conducting layer consist of the (known) incident field, its corresponding reflected field (that results from the incident field when the conducting layer does not contain any pits or holes) and the scattered field (that results from the presence of the pits and holes):

$$\begin{bmatrix} \mathbf{E}(\mathbf{r}) \\ \mathbf{H}(\mathbf{r}) \end{bmatrix} = \begin{bmatrix} \mathbf{E}^i(\mathbf{r}) \\ \mathbf{H}^i(\mathbf{r}) \end{bmatrix} + \begin{bmatrix} \mathbf{E}^r(\mathbf{r}) \\ \mathbf{H}^r(\mathbf{r}) \end{bmatrix} + \begin{bmatrix} \mathbf{E}^s(\mathbf{r}) \\ \mathbf{H}^s(\mathbf{r}) \end{bmatrix}. \quad (10)$$

The reflected field can easily be calculated from the incident field and we therefore consider the sum of the incident and reflected field to be known. The scattered field can be written as:

$$\begin{bmatrix} \mathbf{E}^s(\mathbf{r}) \\ \mathbf{H}^s(\mathbf{r}) \end{bmatrix} = \sum_{\beta_1} \int_{-\infty}^{\infty} b_{\beta} \begin{bmatrix} \mathbf{E}_{\beta}(\mathbf{r}) \\ \mathbf{H}_{\beta}(\mathbf{r}) \end{bmatrix} d\beta_2, \quad (11)$$

where the coefficients  $b_{\beta}$  are still to be determined and  $(\mathbf{E}_{\beta}, \mathbf{H}_{\beta})$  are plane waves with wave vector  $\mathbf{k}_u$  above the layer and  $\mathbf{k}_l$  below the layer. The transverse components  $(k_x, k_y)$  of the wave vector are real and the  $z$ -component is given by:

$$k_z^u = +\sqrt{k_u^2 - k_x^2 - k_y^2}, \quad (12a)$$

$$k_z^l = -\sqrt{k_l^2 - k_x^2 - k_y^2}. \quad (12b)$$

The sign before the square root follows from the assumed time dependence  $\exp(-i\omega t)$  and from the fact that the scattered field propagates away from the conducting layer. The subscript  $\beta = (\beta_1, \beta_2)$  is a short notation for the polarization ( $\beta_1$ ) and the  $x$ - and  $y$ -component of the wave vector ( $\beta_2 = (k_x, k_y)$ ). The polarization can either be S or P. S-polarized means that the  $z$ -component of the electric field is zero (and thus corresponds to TE polarization inside the holes and pits), while for P-polarization the  $z$ -component of the magnetic field is zero (TM polarization). Note that the integral  $\int d\beta_2$  is a shorthand notation for  $\iint dk_x dk_y$ .

We use Eq. (4) to obtain the transverse components of the plane waves and, as before, we split these into a part that depends on  $x$  and  $y$  and a part that depends on  $z$ :

$$\begin{bmatrix} \mathbf{e}_{\beta}(x, y, z) \\ \mathbf{h}_{\beta}(x, y, z) \end{bmatrix} = \mathbf{v}_{\beta}(x, y) \begin{bmatrix} \eta_{\beta}(z) \\ \zeta_{\beta}(z) \end{bmatrix}. \quad (13)$$

These are given in Appendix B.

Analogous to the normalization of the waveguide mode functions, we normalize  $\mathbf{v}_{\beta}$  such that:

$$\langle \mathbf{v}_{\beta} | \mathbf{v}_{\beta'} \rangle_{\mathbb{R}^2} \equiv \iint_{\mathbb{R}^2} [\mathbf{v}_{\beta, x}(x, y) \mathbf{v}_{\beta', x}(x, y)^* + \mathbf{v}_{\beta, y}(x, y) \mathbf{v}_{\beta', y}(x, y)^*] dx dy = \delta_{\beta_1 \beta_1'} \delta(\beta_2 - \beta_2'). \quad (14)$$

Note that the integration is over an entire plane. The first  $\delta$  is the Kronecker delta and the second is the two-dimensional Dirac delta function:

$$\delta(\beta_2) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i(k_x x + k_y y)} dx dy, \quad \beta_2 = (k_x, k_y). \quad (15)$$

Analogous to the scalar product for the waveguide mode functions, the scalar product of two identical plane waves  $\langle \mathbf{v}_{\beta} | \mathbf{v}_{\beta} \rangle_{\mathbb{R}^2}$  is related to the flow of energy of the plane wave through a plane of constant  $z$ .

Because of the use of rotated transverse electric and magnetic fields, we have the following convenient relations between the transverse components of the electric and the magnetic field of the plane waves:

$$\mathbf{h}_{\beta}(x, y, z) = \frac{k_z}{\omega \mu_0} \mathbf{e}_{\beta}(x, y, z), \quad \beta_1 = \text{S}, \quad (16a)$$

$$\mathbf{h}_{\beta}(x, y, z) = \frac{\omega \epsilon \epsilon_0}{k_z} \mathbf{e}_{\beta}(x, y, z), \quad \beta_1 = \text{P}. \quad (16b)$$

Furthermore, every time-harmonic solution of Maxwell's equations with frequency  $\omega$  in the half spaces  $z > D/2$  and  $z < -D/2$ , that propagates away from the conducting layer can be expanded in terms of the plane waves  $(\mathbf{E}_\beta, \mathbf{H}_\beta)$ . In particular, we have, for the scattered transverse electric field:

$$\mathbf{e}^s(x, y, z) = \sum_{\beta_1=-\infty}^{\infty} \int_{\beta_1-\infty}^{\infty} b_\beta \mathbf{e}_\beta(x, y, z) d\beta_2, \quad (17)$$

where, as before, the integration over  $\beta_2$  is a short-hand notation for integrating over  $k_x$  and  $k_y$ .

By using Eq. (14) and (16) we define an operator  $\mathcal{A}$  that works on any two-dimensional vectorfield  $\mathbf{f}: \mathbb{R}^2 \rightarrow \mathbb{C}^2$ :

$$\mathcal{A}(\mathbf{f}) \equiv \int_{-\infty}^{\infty} \frac{k_z}{\omega\mu_0} \langle \mathbf{f} | \mathbf{v}_{\beta_2}^S \rangle_{\mathbb{R}^2} \mathbf{v}_{\beta_2}^S d\beta_2 + \int_{-\infty}^{\infty} \frac{\omega\epsilon\epsilon_0}{k_z} \langle \mathbf{f} | \mathbf{v}_{\beta_2}^P \rangle_{\mathbb{R}^2} \mathbf{v}_{\beta_2}^P d\beta_2, \quad (18)$$

where the superscript S or P naturally means that  $\beta_1 = S$  or  $\beta_1 = P$ . This operator is basically the integral version of the operator  $\frac{\mathbf{k}}{\omega\mu_0} \times$  that can be applied to the electric field of a plane wave to calculate the corresponding magnetic field. Please note that the factor  $\frac{\omega\epsilon\epsilon_0}{k_z}$  in the second integral is singular for  $k_x^2 + k_y^2 = \omega^2\epsilon\epsilon_0\mu_0$ . This is, of course, the fingerprint of the *coupled electromagnetic surface wave*. Although the integrand is integrable, in the numerical implementation prudence is necessary.

In any plane  $z$  is constant and in particular for  $z = \pm D/2$ , the scattered transverse magnetic field can now be expressed in terms of the electric field:

$$\mathbf{h}^s(x, y, \pm D/2) = \mathcal{A}[\mathbf{e}^s(x, y, \pm D/2)]. \quad (19)$$

This equation holds for all  $(x, y)$  with  $-\infty < x, y < \infty$ .

#### 4. Matching at the interfaces

At the interfaces  $z = \pm D/2$ , we have the following relations for the tangential electric and the tangential magnetic field:

$$\mathbf{e}^{\text{pit}} = \mathbf{e}^i + \mathbf{e}^r + \mathbf{e}^s, \quad \forall (x, y), z = \pm D/2, \quad (20a)$$

$$\mathbf{h}^{\text{pit}} = \mathbf{h}^i + \mathbf{h}^r + \mathbf{h}^s, \quad (x, y) \in \bigcup_{\alpha_1} \Omega_{\alpha_1}, z = \pm D/2, \quad (20b)$$

Here,  $\Omega_{\alpha_1}$  is, as before, the cross-section of the pit or hole that is denoted by index  $\alpha_1$ . In Eq. (20a), because the layer is perfectly conducting, the sum of the incident and reflected tangential electric field vanishes at  $z = \pm D/2$ , hence:

$$\mathbf{e}^{\text{pit}} = \mathbf{e}^s, \quad \forall (x, y), z = \pm D/2. \quad (21)$$

Using this together with Eq. (19), we have for Eq. (20b):

$$\mathbf{h}^{\text{pit}} = \mathbf{h}^i + \mathbf{h}^r + \mathcal{A}(\mathbf{e}^{\text{pit}}), \quad (22)$$

which is valid for all  $(x, y, \pm D/2)$  within the holes and pits. The waveguide modes that constitute  $\mathbf{e}^{\text{pit}}$  and  $\mathbf{h}^{\text{pit}}$  vanish outside the pits and holes, as indicated by the rectangle function  $\Pi$  in Eq. (31) and (32) in Appendix B.

In order to obtain a system of equations that is suitable for numerical implementation, we project this equation on the function  $\mathbf{v}_{\bar{\alpha}'}$  by using the scalar product defined in Eq. (34):

$$\sum_{\alpha_4} a_{\alpha'} \zeta_{\alpha'}(\pm D/2) - \sum_{\alpha} a_{\alpha} \eta_{\alpha}(\pm D/2) \langle \mathcal{A}(\mathbf{v}_{\bar{\alpha}}) | \mathbf{v}_{\bar{\alpha}'} \rangle_{\Omega_{p'}} = \langle \mathbf{h}^i + \mathbf{h}^r | \mathbf{v}_{\bar{\alpha}'} \rangle_{\Omega_{p'}}, \quad (23)$$

where the summation over  $\alpha_4$  is a summation over the two directions of propagation ( $\alpha_4$  is not contained in  $\bar{\alpha}$ ). This equation is valid for all  $\bar{\alpha}'$ , hence for all  $\alpha_1$  (counting the number of holes and pits), for all  $\alpha_2$  (TE and TM polarization) and for all  $\alpha_3$  (the mode numbers,  $m_x$  and  $m_y$ ). Consequently, solving the system of Eq. (23) for all  $\bar{\alpha}'$  and for  $z = \pm D/2$  gives the waveguide mode expansion coefficients  $a_{\alpha}$ . Note that the term on the right acts as a source term. The factor  $\langle \mathcal{A}(\mathbf{v}_{\bar{\alpha}}) | \mathbf{v}_{\bar{\alpha}'} \rangle_{\Omega_{p'}}$  is called the interaction integral. Physically speaking, it describes the interaction of a waveguide mode  $\bar{\alpha}$ , via the scattered plane waves through operator  $\mathcal{A}$ , with another mode  $\bar{\alpha}'$ . In Appendix C we will discuss some of its properties and a method to compute it numerically.

To obtain an expression for the scattered field, we use Eq. (14) to project Eq. (21) for the tangential electric field on the plane wave mode function  $\mathbf{e}_{\beta'}$ :

$$b_{\beta'} = \sum_{\alpha} a_{\alpha} \langle \mathbf{e}_{\alpha} | \mathbf{e}_{\beta'} \rangle. \quad (24)$$

In this way, the scattered field can be expressed in the amplitudes of the modes of the pits and holes:

$$\begin{bmatrix} \mathbf{E}^s(\mathbf{r}) \\ \mathbf{H}^s(\mathbf{r}) \end{bmatrix} = \sum_{\beta_1} \sum_{\alpha=-\infty}^{\infty} a_{\alpha} \langle \mathbf{e}_{\alpha} | \mathbf{e}_{\beta} \rangle \begin{bmatrix} \mathbf{E}_{\beta}(\mathbf{r}) \\ \mathbf{H}_{\beta}(\mathbf{r}) \end{bmatrix} d\beta_2. \quad (25)$$

It can be shown that this integrand is integrable everywhere, except, possibly, at the edges of the holes and pits. The occurrence of infinite fields near infinitely sharp, conducting wedges is well-known [24]. For a protruding, right angle, perfectly conducting wedge, the field components perpendicular to the sharp edge may become infinite like  $r^{-1/3}$ , where  $r$  is the distance to the edge. The field components parallel to the edge remain finite. Furthermore, the charge density always remains finite. At an intruding wedge, like the inner part of a waveguide, all field components remain finite.

With Eq. (23) and (24) we have formulated our three-dimensional vectorial scattering problem as a linear system for the amplitudes of the waveguide modes only. Since these modes are parametrized by two parameters ( $\gamma_x$  and  $\gamma_y$ ), we have thus reduced the three-dimensional scattering problem to a two-dimensional numerical problem.

## 5. Numerical implementation

Of course, when implementing our diffraction problem into a computer code we will have to truncate the infinite series of waveguide modes. For large  $\gamma_x$  and  $\gamma_y$ , and depending on the size  $L_x^p$  and  $L_y^p$ , the mode of concern will be evanescent in the  $z$ -direction. For large imaginary  $\gamma_z$ , the mode will only penetrate into the hole or pit a very small distance. It is therefore reasonable to expect that only the modes with a small imaginary  $\gamma_z$  will contribute to the total result. Roberts [17] used 168 modes, while García-Vidal and coworkers [15] used only one. To compare the results of two calculations, one with a number  $N$  and the second with a smaller number  $\tilde{N}$ , we define the following measure:

$$F_{N\tilde{N}} = \frac{\iiint_{V_p} (\epsilon_0 \epsilon_p |\mathbf{E}_{\tilde{N}} - \mathbf{E}_N|^2 + \mu_0 |\mathbf{H}_{\tilde{N}} - \mathbf{H}_N|^2) dx dy dz}{\iiint_{V_p} (\epsilon_0 \epsilon_p |\mathbf{E}_{\tilde{N}}|^2 + \mu_0 |\mathbf{H}_{\tilde{N}}|^2) dx dy dz}, \quad (26)$$



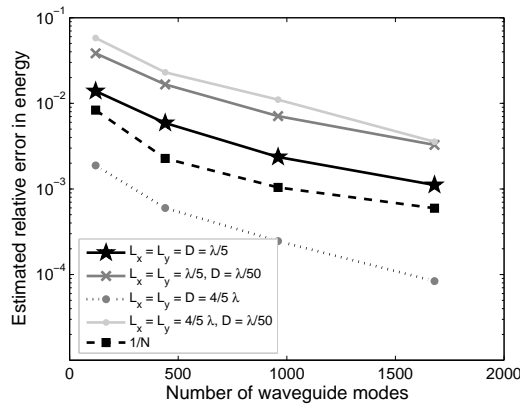


Fig. 2. Relative error  $F_{N\tilde{N}}$  with  $N = 2600$  as a function of the number of waveguide modes ( $\tilde{N} = 120, 440, 960, 1680$ ). Setup is a single hole with a perpendicular incident, linearly polarized plane wave. As a reference,  $1/N$  is also plotted.

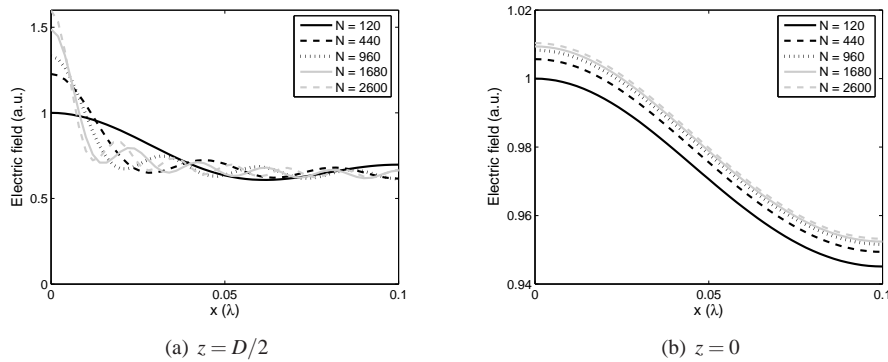


Fig. 3. Line scan of the absolute value of the  $x$ -component of the electric field in a plane with constant  $z$ , through the center of the hole, for various numbers of waveguide modes. Setup is a single hole,  $L_x = L_y = D = \lambda/5$ , with a perpendicular incident, linearly polarized plane wave.

where the integration is over the volume  $V_p$  of the hole (or pit). Here,  $(\mathbf{E}_{\tilde{N}}, \mathbf{H}_{\tilde{N}})$  is the electromagnetic field inside the hole for which the series are truncated after  $\tilde{N}$  waveguide modes and  $(\mathbf{E}_N, \mathbf{H}_N)$  is the electromagnetic field inside the hole obtained by truncation after  $N$  waveguide modes. Hence, this measure corresponds to the error in the energy.

Fig. 2 shows this error as a function of the number of unknowns for a few typical setups. It is clear that only several hundreds of unknowns per hole or pit are enough to model our three-dimensional problem accurately, provided that the thickness of the layer is not too small.

Fig. 3 shows a line scan of the electric field inside a single hole, for various numbers of waveguide modes. Fig. 3(a) shows the field at the entrance of the hole, whereas Fig. 3(b) shows the field in the middle of the hole. Large numbers of waveguide modes are needed to show the singular behaviour at the rim of the hole, as is clear from Fig. 3(a). Inside the hole, where the fields are smooth and regular, the difference with respect to the electric field between 120

waveguide modes and as much as 2600 waveguide modes is not much more than 1 percent. Hence, if the singular behaviour of the fields dominates the solution, as is the case when the conducting layer is thin as compared to the wavelength, our mode expansion technique is probably not the most suitable method.

A small system of equations with only several hundreds of unknowns is solved on a regular desk top computer in only a few seconds. Consequently, most computing time is spent on calculating the interaction integral, which takes a few hours. As discussed in Appendix C, these integrals contain an exponential factor that oscillates violently when mode  $\bar{\alpha}$  and  $\bar{\alpha}'$  live in pits or holes that are far apart. Moreover, the integral contains the factor  $1/k_z$  that is singular on the circle given by  $k_x^2 + k_y^2 = k^2$ . As stated before, this is the fingerprint of the *coupled electromagnetic surface wave*. The integrand is still integrable, but a careful implementation is required.

However, because the interaction integral only involves the plane  $z = \pm D/2$ , it does not depend on the following important parameters: the thickness  $D$  of the conducting layer; the index of refraction  $n_p$  inside the pit or hole; whether the scatterer is a pit or a hole and, in case of a pit, its depth  $d_p$ . Consequently, once the interaction integrals are calculated for a certain setup, we can vary these parameters with negligible computational effort. This is a great advantage of our method. The possibility to construct a library of calculated interaction integrals is also beneficial.

## 6. Extraordinary transmission

In this section we discuss our first results. Calculations were done for single as well as multiple holes and pits. For all calculations, we took into account a number of 440 waveguide modes, such that the error in the energy is less than 1 percent. All holes and pits are square ( $L_x = L_y = L$ ) and the index of refraction above and below the layer as well as inside the pits and holes is taken to be unity.

In the following we (among other things) calculate the energy flux through a hole for various setups. This energy flux through a plane with  $z = z_0$  is calculated directly from the coefficients of the waveguide modes in the following way:

$$\iint_{\Omega_p} S_z \, dx dy = \sum_{\bar{\alpha}} \sum_{\alpha_4} \sum_{\bar{\alpha}_4} \frac{1}{2} \operatorname{Re} \left[ a_{\bar{\alpha}}^{\alpha_4} a_{\bar{\alpha}_4}^{\alpha_4*} \eta_{\bar{\alpha}}^{\alpha_4}(z_0) \zeta_{\bar{\alpha}_4}^{\alpha_4}(z_0)^* \right], \quad (27)$$

where  $\alpha_4$  denotes the direction of propagation of the waveguide mode. Note that two waveguide modes that have an opposite direction of propagation but that are otherwise identical together produce a non-zero energy flux.

### 6.1. Extraordinary transmission through a single hole

Fig. 4 shows the energy flux through a single hole as a function of the layer thickness. The energy flux is normalized by the energy that is incident on the area of the hole (the scalar optics normalization). Results are shown for various sizes of the hole. For sizes where all waveguide modes are below or at cut-off, the amount of energy coming through the hole decreases exponentially with the layer thickness, as expected. When the lowest order modes are just above cut-off a strong modulation of the energy flux with layer thickness is seen. The period of this modulation is half of the effective wavelength ( $2\pi/\gamma_z$ ) of the propagating mode, indicating that the interference of this mode with its own reflection is responsible for the increased and decreased transmission. It follows from Fig. 4 that if the lowest order mode is just above cut-off, extraordinary transmission of a factor of about 1.5 seems possible. If the size of the hole is

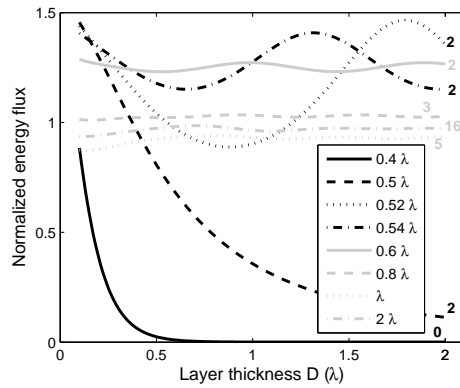


Fig. 4. Energy flux through single hole as a function of layer thickness, normalized by the energy that is incident on the area of the hole. Incident field is a perpendicular, linearly polarized plane wave. The hole is square:  $L_x = L_y = L$  with different values of  $L$  for every curve, as listed in the legend. The number at the right of each curve is the number of waveguide modes above cut-off.

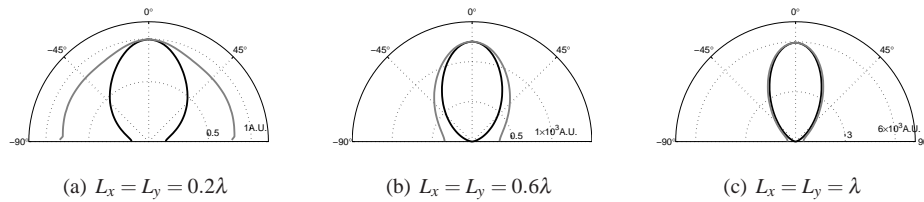


Fig. 5. Polar plot of the near field scattering from a single square pit with a depth  $d = \lambda/4$ . For different sizes, the Poynting vector in the radial direction is shown, at a half circle with radius  $\lambda$ , with its center coinciding with the center of the pit, at  $z = D/2$ . Black line is for the  $(y, z)$ -plane, gray for the  $(x, z)$ -plane. Incident field is a perpendicular plane wave, with its electric field linearly polarized along the  $x$ -direction. The radial scale is arbitrary, but equal for all three figures.

increased further, more and more modes are propagating and the normalized energy flux decreases below unity. Going from  $L = \lambda$  to  $L = 2\lambda$ , the energy flux increases to just below unity. For large holes, one expects an energy flux of unity, of course. The energy flux that is shown, is calculated directly from the coefficients found for the waveguide modes and hence, it is not necessarily the energy that will travel along the  $z$ -axis and, possibly, arrive at a far field detector. However, because of the perfect conductor assumption, none of this energy is absorbed.

Fig. 5 gives information on the direction along which most energy is scattered. It shows polar plots of the energy, scattered from a single pit with depth  $\lambda/4$ . Shown is the Poynting vector in the radial direction along a half circle with radius of one wavelength, hence the scattering in the near field. The scattering in the plane in which the incident electric field is polarized can be non-zero along the interface, because the corresponding electric field then points in the  $z$ -direction. The scattering in the perpendicular plane can not be along the  $z$ -direction, for the tangential field at a perfect conductor must be zero. For a pit with size  $L = \lambda/5$  the scattering is like that of a dipole, whereas for larger pits the scattering is mainly along the optical axis (the  $z$ -axis).

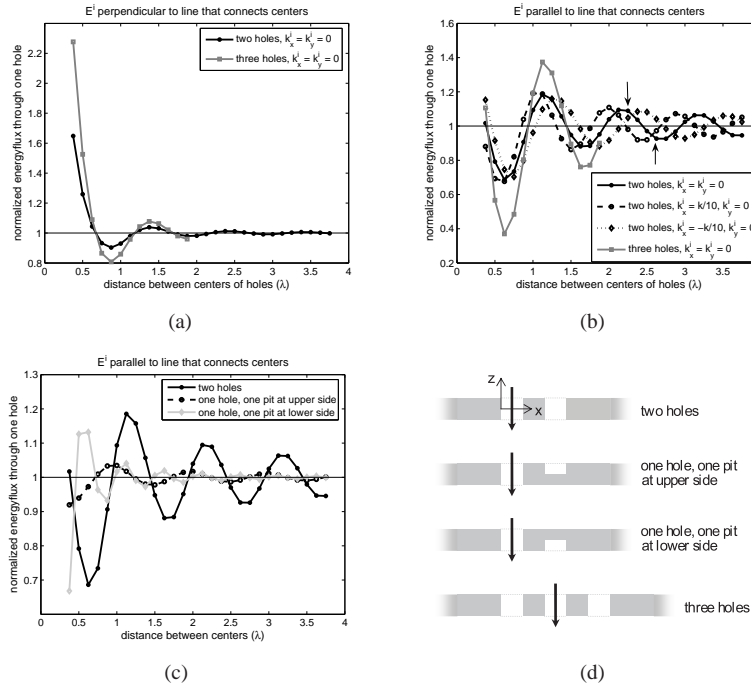


Fig. 6. Energy flux through a hole, normalized by the energy flux through an identical single hole, as a function of the distance between two holes (or one hole and one pit). Pits and holes are all square, with  $L_x = L_y = L = \lambda/4$ . Incident field is a linearly polarized plane wave, with polarization as stated at the top of the figures. Incidence is always perpendicular, except for the dashed and dotted lines in Fig. 6(b). Here,  $k_x^i \neq 0$  means that the plane of incidence is the  $(x, z)$ -plane. Fig. 6(d) shows cross-sections of the four used geometries. The bold arrows denote where the energy flux is calculated. The thickness of the layer is  $\lambda/2$  and the depth of the pits is  $\lambda/4$ .

### 6.2. Extraordinary transmission through multiple holes

In Fig. 6, we show the effect that the presence of a second and third hole or pit have on the energy flux, as calculated with equation (27), through the first hole. As a function of the distance between the centers of the two or three scatterers, the energy flux through one hole is calculated. This energy flux is normalized by the energy flux through an identical, but single hole, also calculated with our method. With this special normalization, we are able to isolate the effect that the presence of the second and third scatterer have on the transmission through the first. The incident field is again a linearly polarized plane wave. To distinguish the two basic directions of polarization we define a reference plane through the line that connects the centers of the scatterers and the  $z$ -axis. The electric field of the incident plane wave is either directed perpendicular to this plane or else parallel to this plane. Fig. 6(a) shows the normalized energy flux through one of two holes and through one of three holes for perpendicular polarization. A modulation of the energy flux as a function of distance between the centers of the holes only occurs for holes that are less than two wavelengths apart. We believe that enhanced or decreased transmission in this case is caused by the coupling of evanescent fields scattered from one hole to the other and/or by polarization rotation at the corners of the hole. The solid lines in Fig. 6(b) show the same calculations, but now for parallel polarization. The modulation

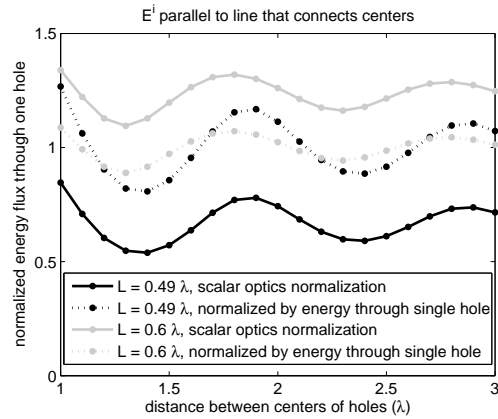


Fig. 7. Comparison between the scalar optics normalization (solid line) and the single hole normalization (dashed line). The vertical axis shows the energy flux through one of two holes, the horizontal axis the distance between the centers of the two holes. The holes are square, with  $L = 0.49\lambda$  (black) and  $L = 0.6\lambda$  (gray). The thickness of the conducting layer is  $\lambda/2$ .

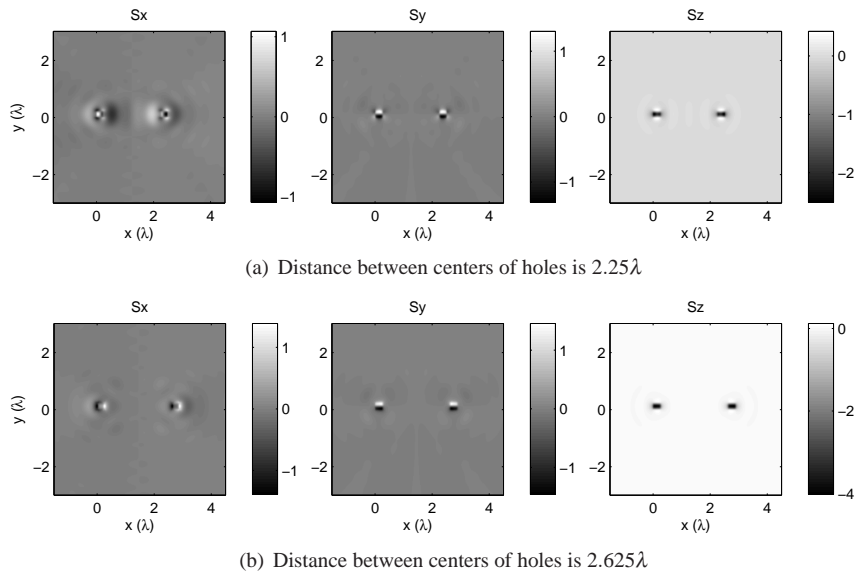


Fig. 8. Cartesian components of the Poynting vector in a plane of constant  $z$ , at a distance of  $\lambda/20$  below the metal layer. Gray scale is in arbitrary units, the same for all figures. The layer contains two holes with  $L_x = L_y = \lambda/4$  and the thickness of the layer is  $\lambda/2$ . A parallel polarized plane wave is incident from above, perpendicular incidence. Top figure corresponds to a maximum in energy throughput, lower figure to a minimum. See arrows in Fig. 6(b).

of the energy flux is now also present for large distances between the holes. Its period is equal to the wavelength of the incident field. Its amplitude for three holes is twice the amplitude for two holes. Furthermore, the amplitude is proportional to the inverse of the distance between the centers of the holes, as expected for a cylindrical wave. If the propagation direction of the incident plane wave is slightly tilted (dashed and dotted line in Fig. 6(b)) then a phase shift occurs that is equal to the delay that the incident field experiences in reaching the farthest hole as compared to the nearest hole. Fig. 6(c) shows, for the incident field polarized parallel, the normalized energy flux through a hole in the presence of a pit. This pit has its open end at the upper side (dashed, black line) or at the lower side (solid, gray line). The field is incident from above. The modulation period for the latter is half that of the first. Furthermore, the amplitude for both cases is much smaller than for the case of two holes (solid, black line).

Fig. 7 shows a comparison of the usual scalar optics normalization, where the transmission is normalized by the energy that is incident on the area of the hole, with the normalization used in Fig. 6. The setup is two square holes, at varying distance of each other. The incident field is a parallel incident plane wave. Shown is, again, the transmission through one of these two holes. The transmission, when normalized by the transmission through an identical but single hole, always varies around unity. Hence, the presence of the second hole can lead to an increase as well as a decrease in transmission. However, when the scalar optics normalization is used, we already know from Fig. 4 that the layer thickness has an influence on the transmission. This influence is of another nature. For holes that are so small that all modes are evanescent (shown in black), the transmission decreases exponentially with layer thickness. When the size of the holes is such that at least one mode is propagating (shown in gray) the Fabry-Perot effect can cause enhanced transmission for a single hole.

We think that, for the cases in which the incident electric field is polarized perpendicular to the reference plane, the enhanced and reduced transmission that occurs when the single hole normalization is used, is mainly caused by waves with  $k_z = 0$  that are scattered along the metal surface. These scattered waves cause a periodicity of a wavelength when two scatterers at the same side of the metal layer contribute and a periodicity of half the wavelength when there is only one source, as is the case when one hole is accompanied by a pit with its open end at the non-illuminated side. In this case, the surface wave that is excited at the exit of the hole travels to the pit. It there excites another surface wave that travels back to the hole and interferes. This results in a phase shift that corresponds to twice the distance between the hole and the pit. See also Ref. [9]. The doubling of the amplitude for three holes as compared to two holes also fits nicely in the picture of the interfering surface waves.

For a perfect conductor, a wave propagating along the surface with its wave vector equal to the wave vector of the incident light, is the analogue of a surface plasmon. It is often stated that perfect metals do not support surface plasmons. However, we believe that the expression *surface plasmon* (or *surface plasmon polariton*) is confusing. The ending *-on* suggests a sort of localization. For a surface plasmon, as described by, for example, Raether [25], this means that it is bound to the interface between the metal and the dielectric. The derivation by Raether of the surface plasmon wave vector (which is only valid for a flat interface) is also valid for an interface between a perfect conductor and a dielectric. Then, the plasmon wave vector is just the wave vector of the light. The penetration depth inside the metal is zero and the charges inside the metal oscillate only in the plane of the interface between the metal and the dielectric. The plasmon - or, better, *coupled electromagnetic surface wave* - then has a constant field strength in the half space above the metal and is not bound to the surface. The existence and the physical nature of the phenomenon, however, are the same for both a conductor with finite conductivity and an idealized perfect conductor. The only difference is the finite decay length and absorption caused by finite conductivity.

Fig. 8 shows the three cartesian components of the Poynting vector in a plane that is  $\lambda/20$  below the metal layer. The metal layer contains two square holes. The small arrows in Fig. 6(b) indicate the data points for which we calculated the near field as shown in the upper and lower figure. Hence, the incident electric field is polarized in the  $x$ -direction. This means that the wave along the surface is mainly propagating energy in the  $x$ -direction. For the two setups, the difference is indeed largest for the  $x$ -component of the Poynting vector. Note that, near the holes, especially in the upper figure, the  $z$ -component of the Poynting vector points towards the metal layer.

## 7. Conclusion

We have described a mode expansion method to rigorously calculate the diffraction by a perfectly conducting layer with finite thickness, that contains rectangular pits and/or holes. The electromagnetic field above and below the conducting layer is written as an integral over plane waves. These plane waves can be S- or P-polarized and they can be propagating or evanescent. The field inside the pits and holes is expanded into waveguide modes, that can have TE or TM polarization and can also be propagating or evanescent. A system of equations is derived by matching the tangential field components at the interfaces. As unknowns, this system only contains the expansion coefficients of the waveguide modes and is therefore very small. For each pit or hole with a size that is of the order of the wavelength of the incident light, about 400 waveguide modes are sufficient for an accuracy in the energy of less than one percent. Once the system of equations of a particular diffraction geometry is composed, important parameters such as the thickness of the conducting layer and the index of refraction inside the pit or hole can be varied with negligible computational effort.

We have shown results on extraordinary transmission through single as well as multiple holes. For a single hole, with transverse sizes such that the lowest order waveguide mode is just above cut-off, a strong modulation of the transmitted energy as a function of layer thickness was obtained. This is due to the interference of this lowest order mode with its own reflection (the Fabry-Perot effect). For two or three holes, a strong modulation in the transmitted energy was found as a function of the distance between neighbouring holes, but only for one polarization of the incident field. Since we normalize the transmitted energy by the energy transmitted by a single, identical hole, the displayed increase and decrease is only due to the presence of the second (and third) hole. The fact that this modulation is hardly present when the incident field is polarized in the perpendicular direction clearly points to a plasmon mechanism. Actually, the term *coupled electromagnetic surface wave* would be more appropriate, as the boundedness of this kind of wave depends on the conductivity of the metal.

Except for absorption, all relevant physics is present in the model. This means that, apart from the microwave and terahertz frequency regimes, the results found can also be of use in the optical domain, as long as the penetration depth inside the metal is not too large as compared to the thickness of the metal layer and the distances between the holes or pits.

## Acknowledgements

This research was supported by the Dutch Technology Foundation STW.

### A. The waveguide modes

We here give a complete listing of the waveguide modes. The rotated transverse components of the modes are split into a transverse part and a part that depends on  $z$ :

$$\begin{bmatrix} \mathbf{e}_\alpha(x, y, z) \\ \mathbf{h}_\alpha(x, y, z) \end{bmatrix} = \mathbf{v}_\alpha(x, y) \begin{bmatrix} \eta_\alpha(z) \\ \zeta_\alpha(z) \end{bmatrix}. \quad (28)$$

With:

$$\Gamma_{\alpha} = \sqrt{\gamma_x^2 + \gamma_y^2}, \quad (29)$$

and

$$\Lambda_{\alpha} = \begin{cases} 2(L_x^p L_y^p)^{-1/2}, & \text{if } m_x \neq 0 \text{ and } m_y \neq 0, \\ \sqrt{2}(L_x^p L_y^p)^{-1/2}, & \text{if } m_x = 0 \text{ or } m_y = 0, \end{cases} \quad (30)$$

we first define the following auxiliary functions:

$$\mathbf{v}_{\bar{\alpha}}(\bar{x}_p, \bar{y}_p) = \begin{cases} \frac{\Lambda_{\alpha}}{\Gamma_{\alpha}} \Pi(\bar{x}_p, \bar{y}_p) \begin{bmatrix} \gamma_y \cos(\gamma_x \bar{x}_p) \sin(\gamma_y \bar{y}_p) \\ -\gamma_x \sin(\gamma_x \bar{x}_p) \cos(\gamma_y \bar{y}_p) \end{bmatrix}, & \alpha_2 = \text{TE}, \\ \frac{\Lambda_{\alpha}}{\Gamma_{\alpha}} \Pi(\bar{x}_p, \bar{y}_p) \begin{bmatrix} \gamma_x \cos(\gamma_x \bar{x}_p) \sin(\gamma_y \bar{y}_p) \\ \gamma_y \sin(\gamma_x \bar{x}_p) \cos(\gamma_y \bar{y}_p) \end{bmatrix}, & \alpha_2 = \text{TM}, \end{cases} \quad (31)$$

$$\vartheta_{\bar{\alpha}}(\bar{x}_p, \bar{y}_p) = \begin{cases} -i\Lambda_{\alpha} \Gamma_{\alpha} \Pi(\bar{x}_p, \bar{y}_p) \cos(\gamma_x \bar{x}_p) \cos(\gamma_y \bar{y}_p), & \alpha_2 = \text{TE}, \\ i\Lambda_{\alpha} \Gamma_{\alpha} \Pi(\bar{x}_p, \bar{y}_p) \sin(\gamma_x \bar{x}_p) \sin(\gamma_y \bar{y}_p), & \alpha_2 = \text{TM}, \end{cases} \quad (32)$$

where we have introduced local coordinates for every pit or hole:  $\bar{x}_p \equiv x - x_0^p$ ,  $\bar{y}_p \equiv y - y_0^p$ . Furthermore, the function  $\Pi(\bar{x}_p, \bar{y}_p)$  is a rectangle function that indicates that the mode functions are identical to zero outside the cross-sectional area of the  $p$ -th hole:

$$\Pi \equiv [\text{H}(\bar{x}_p) - \text{H}(\bar{x}_p - L_x^p)] [\text{H}(\bar{y}_p) - \text{H}(\bar{y}_p - L_y^p)], \quad (33)$$

where  $\text{H}(x)$  is the Heaviside step function. By using (31) it is easy to see that two different modes are orthogonal in the sense:

$$\langle \mathbf{v}_{\bar{\alpha}} | \mathbf{v}_{\bar{\alpha}'} \rangle_{\Omega_{p'}} \equiv \iint_{\Omega_{p'}} [\mathbf{v}_{\bar{\alpha},x}(x,y) \mathbf{v}_{\bar{\alpha}',x}(x,y)^* + \mathbf{v}_{\bar{\alpha},y}(x,y) \mathbf{v}_{\bar{\alpha}',y}(x,y)^*] dx dy = 0, \quad \bar{\alpha} \neq \bar{\alpha}'. \quad (34)$$

Hence,  $\{\mathbf{v}_{\bar{\alpha}}\}$  is an orthonormal system with respect to the scalar product (34).

The following auxiliary functions are needed for the  $z$ -dependent parts  $\eta_{\alpha}$  and  $\zeta_{\alpha}$  in Eq. (5):

$$f_{\alpha}(z) = \begin{cases} -\exp[-i\gamma_z(z - z_1^p)], & |\gamma_z/k_p| \geq \varepsilon, \alpha_4 = -, \\ \exp[i\gamma_z(z - z_2^p)], & |\gamma_z/k_p| \geq \varepsilon, \alpha_4 = +, \\ -ik_p \gamma_z^{-1} \cos(\gamma_z z), & |\gamma_z/k_p| < \varepsilon, \alpha_4 = -, \\ i \sin(\gamma_z z), & |\gamma_z/k_p| < \varepsilon, \alpha_4 = +, \end{cases} \quad (35)$$

$$g_{\alpha}(z) = \begin{cases} \exp[-i\gamma_z(z - z_1^p)], & |\gamma_z/k_p| \geq \varepsilon, \alpha_4 = -, \\ \exp[i\gamma_z(z - z_2^p)], & |\gamma_z/k_p| \geq \varepsilon, \alpha_4 = +, \\ k_p \gamma_z^{-1} \sin(\gamma_z z), & |\gamma_z/k_p| < \varepsilon, \alpha_4 = -, \\ \cos(\gamma_z z), & |\gamma_z/k_p| < \varepsilon, \alpha_4 = +, \end{cases} \quad (36)$$

The constants  $z_1^p$  and  $z_2^p$  are the  $z$ -coordinates of the upper and lower end of the pit or hole, respectively. Hence, for holes we have  $z_1^p = D/2$  and  $z_2^p = -D/2$ , for pits we either have  $z_1^p = D/2$  and  $z_2^p = D/2 - d_p$  or we have  $z_1^p = -D/2 + d_p$  and  $z_2^p = -D/2$ . We have introduced



the constant factors like  $\exp(i\gamma_z z_1^p)$  to make sure that, for imaginary  $\gamma_z$ , the moduli of the exponents are always equal to or smaller than unity. In Eq. (35) and (36) we use the sine and cosine functions instead of the exponential function if the propagation constant  $\gamma_z$  is very small as compared to  $k_p$ , with smallness parameter  $\varepsilon$  [26]. If we would not do this, for  $\gamma_z = 0$  for some mode, we would miss the mode function that is linear in  $z$  and our set of modes would not be complete [27]. In our implementation, we take  $\varepsilon = 10^{-5}$ . For the  $z$ -dependent parts we now have:

$$\eta_{\alpha}(z) = \begin{cases} \omega\mu_0 g_{\alpha}(z), & \alpha_2 = \text{TE}, \\ \gamma_z \sqrt{\frac{\mu_0}{\varepsilon_p \varepsilon_0}} f_{\alpha}(z), & \alpha_2 = \text{TM}, \end{cases} \quad (37)$$

$$\zeta_{\alpha}(z) = \begin{cases} \gamma_z f_{\alpha}(z), & \alpha_2 = \text{TE}, \\ k_p g_{\alpha}(z), & \alpha_2 = \text{TM}. \end{cases} \quad (38)$$

Note that the modes are propagating in the  $z$ -direction if  $\Gamma_{\alpha} < k_p$ , while for  $\Gamma_{\alpha} > k_p$  the modes are evanescent. For a square pit or hole ( $L_x = L_y$ ) with  $L_x < \lambda_p/2$  all modes are evanescent. Finally, the (rotated) transverse components of the modes are then given by:

$$\begin{bmatrix} \mathbf{e}_{\alpha}(x, y, z) \\ \mathbf{h}_{\alpha}(x, y, z) \end{bmatrix} = \mathbf{v}_{\alpha}(x, y) \begin{bmatrix} \eta_{\alpha}(z) \\ \zeta_{\alpha}(z) \end{bmatrix}, \quad (39)$$

and the longitudinal components:

$$E_{\alpha,z}(x, y, z) = \begin{cases} 0, & \alpha_2 = \text{TE}, \\ \sqrt{\frac{\mu_0}{\varepsilon_p \varepsilon_0}} \vartheta_{\alpha}(x, y) g_{\alpha}(z), & \alpha_2 = \text{TM}, \end{cases} \quad (40a)$$

$$H_{\alpha,z}(x, y, z) = \begin{cases} \vartheta_{\alpha}(x, y) g_{\alpha}(z), & \alpha_2 = \text{TE}, \\ 0, & \alpha_2 = \text{TM}. \end{cases} \quad (40b)$$

We note here that the normalization of the waveguide modes involves only  $\mathbf{v}_{\alpha}$ , which is the part of the transverse field that does not depend on  $z$ . This means that the above  $z$ -dependent part is only defined up to a constant. We have chosen this constant such that for both TE and TM polarization the waveguide modes have the same order of magnitude.

## B. The plane waves above and below the layer

The transverse components of the plane waves  $\mathbf{E}_{\beta}, \mathbf{H}_{\beta}$  are divided into a part that depends on  $x$  and  $y$  and a part that depends on  $z$ :

$$\begin{bmatrix} \mathbf{e}_{\beta}(x, y, z) \\ \mathbf{h}_{\beta}(x, y, z) \end{bmatrix} = \mathbf{v}_{\beta}(x, y) \begin{bmatrix} \eta_{\beta}(z) \\ \zeta_{\beta}(z) \end{bmatrix}. \quad (41)$$

We will give a listing of these functions here. With:

$$\Gamma_{\beta} = \sqrt{k_x^2 + k_y^2}, \quad (42)$$

and

$$\Lambda_{\beta} = -\frac{1}{2\pi}, \quad (43)$$

we define the following functions that are independent of  $z$ :

$$\mathbf{v}_{\beta}(x, y) = \begin{cases} \frac{\Lambda_{\beta}}{\Gamma_{\beta}} e^{i(k_x x + k_y y)} \begin{pmatrix} k_y \\ -k_x \end{pmatrix}, & \beta_1 = \text{S}, k_x^2 + k_y^2 > 0, \\ \Lambda_{\beta} \begin{pmatrix} 0 \\ -1 \end{pmatrix}, & \beta_1 = \text{S}, k_x^2 + k_y^2 = 0, \\ \frac{\Lambda_{\beta}}{\Gamma_{\beta}} e^{i(k_x x + k_y y)} \begin{pmatrix} k_x \\ k_y \end{pmatrix}, & \beta_1 = \text{P}, k_x^2 + k_y^2 > 0, \\ \Lambda_{\beta} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, & \beta_1 = \text{P}, k_x^2 + k_y^2 = 0, \end{cases} \quad (44)$$

$$\vartheta_{\beta}(x, y) = \Lambda_{\beta} \Gamma_{\beta} e^{i(k_x x + k_y y)}, \quad (45)$$

For the  $z$ -dependent part we have the following auxiliary function:

$$f_{\beta}(z) = \begin{cases} \exp[ik_z^u(z - D/2)], & z > D/2, \\ \exp[ik_z^{\ell}(z + D/2)], & z < -D/2, \end{cases} \quad (46)$$

and the actual  $z$ -dependent parts:

$$\eta_{\beta}(z) = \begin{cases} \omega \mu_0 f_{\beta}(z), & \beta_1 = \text{S}, \\ k_z \sqrt{\frac{\mu_0}{\varepsilon \varepsilon_0}} f_{\beta}(z), & \beta_1 = \text{P}, \end{cases} \quad (47)$$

$$\zeta_{\beta}(z) = \begin{cases} k_z f_{\beta}(z), & \beta_1 = \text{S}, \\ k f_{\beta}(z), & \beta_1 = \text{P}. \end{cases} \quad (48)$$

Note that  $k$ ,  $k_z$  and  $\varepsilon$  in the above equations are either  $k_u$ ,  $k_z^u$  and  $\varepsilon_u$  or  $k_{\ell}$ ,  $k_z^{\ell}$  and  $\varepsilon_{\ell}$ , depending on  $z$ . The  $z$ -components of the plane waves are given by:

$$E_{\beta,z}(x, y, z) = \begin{cases} 0, & \beta_1 = \text{S}, \\ \sqrt{\frac{\mu_0}{\varepsilon \varepsilon_0}} \vartheta_{\beta}(x, y) f_{\beta}(z), & \beta_1 = \text{P}, \end{cases} \quad (49a)$$

$$H_{\beta,z}(x, y, z) = \begin{cases} \vartheta_{\beta}(x, y) f_{\beta}(z), & \beta_1 = \text{S}, \\ 0, & \beta_1 = \text{P}. \end{cases} \quad (49b)$$

Note that we did not take special precautions for the case that  $k_z = 0$ . For the waveguide modes in the previous section, we made sure that, when it happens that  $\gamma_z = 0$ , the set of mode functions is still complete. The plane waves in the upper and lower half spaces, however, form a continuous set, parametrized by  $-\infty < k_x, k_y < \infty$ . The plane waves with  $k_z = 0$  are only a set of measure zero in the space of all plane waves and are therefore irrelevant for the completeness.

### C. The interaction integral

In this appendix we elaborate on the interaction integral that describes the interaction of waveguide mode  $\bar{\alpha}$ , via the scattered field through operator  $\mathcal{A}$ , with another waveguide mode  $\bar{\alpha}'$ . We will first write out the integral and then we will discuss the numerical implementation.

### C.1. The interaction integral

The interaction integral is given by:

$$\begin{aligned} \langle \mathcal{A}(\mathbf{v}_{\alpha}) | \mathbf{v}_{\alpha'} \rangle_{\Omega_{p'}} &= \int_{-\infty}^{\infty} \frac{k_z}{\omega \mu_0} \langle \mathbf{v}_{\alpha} | \mathbf{v}_{\beta_2}^S \rangle_{\Omega_p} \langle \mathbf{v}_{\alpha'} | \mathbf{v}_{\beta_2}^S \rangle_{\Omega_{p'}}^* d\beta_2 \\ &+ \int_{-\infty}^{\infty} \frac{\omega \epsilon \epsilon_0}{k_z} \langle \mathbf{v}_{\alpha} | \mathbf{v}_{\beta_2}^P \rangle_{\Omega_p} \langle \mathbf{v}_{\alpha'} | \mathbf{v}_{\beta_2}^P \rangle_{\Omega_{p'}}^* d\beta_2. \end{aligned} \quad (50)$$

Please recall that  $\beta_2 = (k_x, k_y)$  and that  $\int d\beta_2 = \iint dk_x dk_y$ . We consider one of the scalar products, with  $\beta = (\beta_1, \beta_2)$  and  $\beta_1 = S, P$ :

$$\begin{aligned} \langle \mathbf{v}_{\alpha} | \mathbf{v}_{\beta} \rangle_{\Omega_p} &= \int_{x_0^p}^{x_0^p + L_x^p} \int_{y_0^p}^{y_0^p + L_y^p} \mathbf{v}_{\alpha}(x, y) \cdot \mathbf{v}_{\beta}(x, y)^* dx dy, \\ &= e^{-i(k_x x_0^p + k_y y_0^p)} \int_0^{L_x^p} \int_0^{L_y^p} \mathbf{v}_{\alpha}(\bar{x}_p, \bar{y}_p) \cdot \mathbf{v}_{\beta}(\bar{x}_p, \bar{y}_p)^* d\bar{x}_p d\bar{y}_p, \end{aligned} \quad (51)$$

where we have changed to local coordinates. This final double integral is in fact a Fourier integral that can be calculated analytically:

$$\begin{aligned} F_{\alpha}^{\beta_1}(k_x, k_y) &\equiv \int_0^{L_x^p} \int_0^{L_y^p} \mathbf{v}_{\alpha}(\bar{x}_p, \bar{y}_p) \cdot \mathbf{v}_{\beta}(\bar{x}_p, \bar{y}_p)^* d\bar{x}_p d\bar{y}_p, \\ &= \frac{\Lambda_{\alpha} \Lambda_{\beta}}{\Gamma_{\alpha} \Gamma_{\beta}} \begin{cases} \gamma_y k_y c_{m_x}^p(k_x) s_{m_y}^p(k_y) + \gamma_x k_x s_{m_x}^p(k_x) c_{m_y}^p(k_y), & \alpha_2 = \text{TE}, \beta_1 = S, \\ \gamma_x k_y c_{m_x}^p(k_x) s_{m_y}^p(k_y) - \gamma_y k_x s_{m_x}^p(k_x) c_{m_y}^p(k_y) = 0, & \alpha_2 = \text{TM}, \beta_1 = S, \\ \gamma_y k_x c_{m_x}^p(k_x) s_{m_y}^p(k_y) - \gamma_x k_y s_{m_x}^p(k_x) c_{m_y}^p(k_y), & \alpha_2 = \text{TE}, \beta_1 = P, \\ \gamma_x k_x c_{m_x}^p(k_x) s_{m_y}^p(k_y) + \gamma_y k_y s_{m_x}^p(k_x) c_{m_y}^p(k_y), & \alpha_2 = \text{TM}, \beta_1 = P, \end{cases} \end{aligned} \quad (52)$$

where the functions  $c_{m_j}^p$  and  $s_{m_j}^p$  with subscript  $j = x, y$  are given by:

$$c_{m_j}^p(k_j) \equiv \int_0^{L_j^p} \cos(\gamma_j z) e^{-ik_j z} dz = \begin{cases} \frac{ik_j}{\gamma_j^2 - k_j^2} [1 - (-1)^{m_j} e^{-ik_j L_j}], & k_j \neq \pm \gamma_j, \\ \frac{1}{2} L_j, & k_j = \pm \gamma_j, k_j \neq 0, \\ L_j, & k_j = \gamma_j = 0, \end{cases} \quad (53a)$$

$$s_{m_j}^p(k_j) \equiv \int_0^{L_j^p} \sin(\gamma_j z) e^{-ik_j z} dz = \begin{cases} \frac{\gamma_j}{\gamma_j^2 - k_j^2} [1 - (-1)^{m_j} e^{-ik_j L_j}], & k_j \neq \pm \gamma_j, \\ \mp \frac{1}{2} i L_j, & k_j = \pm \gamma_j, k_j \neq 0, \\ 0, & k_j = \gamma_j = 0. \end{cases} \quad (53b)$$

Note that  $F_{\alpha}^{\beta_1}(k_x, k_y)$  is zero for all  $(k_x, k_y)$  when the waveguide mode is TM polarized and the plane wave is S-polarized. Hence, these polarizations do not interact. For the interaction

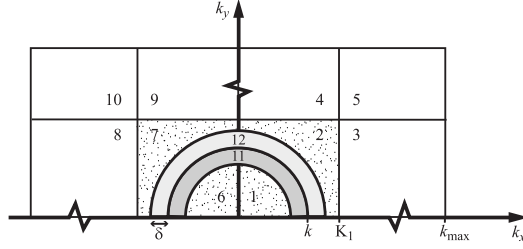


Fig. 9. The division of the integration area into 12 domains. Not on scale.

integral we now have:

$$\begin{aligned} \langle \mathcal{A}(\mathbf{v}_{\bar{\alpha}}) | \mathbf{v}_{\bar{\alpha}'} \rangle_{\Omega_{p'}} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{k_z}{\omega \mu_0} e^{i(k_x \Delta_x + k_y \Delta_y)} F_{\bar{\alpha}}^S(k_x, k_y) F_{\bar{\alpha}'}^S(k_x, k_y)^* dk_x dk_y, \\ &+ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\omega \epsilon \epsilon_0}{k_z} e^{i(k_x \Delta_x + k_y \Delta_y)} F_{\bar{\alpha}}^P(k_x, k_y) F_{\bar{\alpha}'}^P(k_x, k_y)^* dk_x dk_y, \quad (54) \end{aligned}$$

with  $\Delta_x = x_0^{p'} - x_0^p$  and  $\Delta_y = y_0^{p'} - y_0^p$ , the distance between pit  $p'$  and  $p$ . From the above equation it is clear that this double integral is difficult for two reasons. First, the factor  $\exp[i(k_x \Delta_x + k_y \Delta_y)]$  oscillates violently when pit  $p'$  and  $p$  are far apart. Second, the factor  $k_z^{-1}$  is singular for  $k_x^2 + k_y^2 = k^2$ . The integrand is still integrable, but care has to be taken.

### C.2. Numerical computation of the interaction integrals

Since we have to calculate a lot of interaction integrals, we have to find a method that is both fast and accurate. To select the best numerical integration routine we concentrate on the two most difficult parts of the interaction integral: the circle in the  $(k_x, k_y)$ -plane where the square root  $k_z = (k^2 - k_x^2 - k_y^2)^{1/2}$  is equal to zero and the exponential factor  $\exp(ik_x \Delta_x + ik_y \Delta_y)$  that oscillates violently when the two holes or pits under consideration are far apart. The square root term would be best tackled with polar coordinates, whereas the exponential term would be easier to integrate with cartesian coordinates. To overcome this problem, we split the integration area in 12 domains. See Fig. 9. From symmetry properties of the integrand, it follows that it suffices to integrate over half the  $(k_x, k_y)$ -plane [28]. Furthermore, in the domains 1, 6 and 11 that are situated within the circle  $k_x^2 + k_y^2 = k^2$ , only the real parts need to be calculated. For the other domains, we only need the imaginary parts:

$$\langle \mathcal{A}(\mathbf{v}_{\bar{\alpha}}) | \mathbf{v}_{\bar{\alpha}'} \rangle_{\Omega_{p'}} = \sum_{N \in \mathcal{N}_{\text{in}}} m_f \text{Re}(I_N) + \sum_{N \in \mathcal{N}_{\text{out}}} m_f i \text{Im}(I_N), \quad (55)$$

with  $m_f$  a multiplication factor and  $\mathcal{N}$  a set of domain numbers:

$$\text{for } \alpha_1 = \alpha_1' : \quad m_f = 4, \mathcal{N}_{\text{in}} = \{1, 11\}, \mathcal{N}_{\text{out}} = \{2, 3, 4, 5, 12\}, \quad (56a)$$

$$\text{for } \alpha_1 \neq \alpha_1' : \quad m_f = 2, \mathcal{N}_{\text{in}} = \{1, 6, 11\}, \mathcal{N}_{\text{out}} = \{2, 3, 4, 5, 7, 8, 9, 10, 12\}, \quad (56b)$$

and with

$$I_N = \iint_{\text{domain } N} I_{\bar{\alpha}\bar{\alpha}'}(k_x, k_y) dk_x dk_y. \quad (57)$$

the double integral over one of the domains. Here,  $I_{\vec{\alpha}\vec{\alpha}'}$  is shorthand notation for the integrand of the interaction integral.

The two half rings, domains 11 and 12, are called the inner  $\delta$ -ring  $[(k - \delta)^2 \leq k_x^2 + k_y^2 \leq k^2]$  and the outer  $\delta$ -ring  $[(k^2 < k_x^2 + k_y^2 \leq (k + \delta)^2]$ . The value of  $\delta$  is chosen such, that the number of oscillations of the exponential factor inside the rings is small. However, the  $\delta$ -rings should be wide enough to contain the steepest part of the square root factor (typically larger than or equal to  $k/10$ ). Within the two  $\delta$ -rings, we choose a polar coordinate system. Furthermore, we apply a substitution to get rid of the square root singularity. We then use a standard, adaptive quadrature routine from the NAG foundation toolbox for Matlab, D01FCF [29].

Regarding the domains 1 to 10, we assume that the square root factor is sufficiently flat. In these areas, we split the integrand in a slowly varying part and the (possibly) quickly oscillating exponential factor. Domains 1, 2, 6 and 7 are the areas that are bounded by the inner and outer  $\delta$ -ring. We approximate this boundary linearly on a cartesian grid. The slowly varying part is also approximated linearly, on a cartesian grid. This linear approximation times the exponential factor can now be integrated exactly. The domains 3, 4, 5 and 8, 9, 10 are rectangular areas, bounded by  $K_1$  and  $k_{\max}$ . Here, we approximate the slowly varying part parabolically and we integrate exactly this parabolic approximation times the exponent.