General mechanism involved in subwavelength optics of conducting microstructures: charge-oscillation-induced light emission and interference

Xian-Rong Huang[1,2,*] and Ru-Wen Peng[3]

[1]*National Synchrotron Light Source II, Brookhaven National Laboratory, Upton, New York 11973, USA*
[2]*Advanced Photon Source, Argonne National Laboratory, Argonne, Illinois 60439, USA*
[3]*National Laboratory of Solid State Microstructures, Nanjing University, Nanjing 210093, China*
**Corresponding author: xiahuang@aps.anl.gov*

Received August 7, 2009; revised January 10, 2010; accepted January 25, 2010;
posted February 1, 2010 (Doc. ID 115358); published March 15, 2010

Interactions between light and conducting microstructures or nanostructures can result in a variety of novel phenomena, but their underlying mechanisms have not been completely understood. From calculations of surface charge density waves on conducting gratings and by comparing them with classical surface plasmons, we revealed a general yet concrete picture regarding the coupling of light to free electron oscillation on structured conducting surfaces that can lead to oscillating subwavelength charge patterns (i.e., structured surface plasmons). New wavelets emitted from these light sources then destructively interfere to form evanescent waves. This principle, usually combined with other mechanisms, is mainly a geometrical effect that can be universally involved in light scattering from all periodic and non-periodic structures containing free electrons. This picture may provide clear guidelines for developing conductor-based nano-optical devices. © 2010 Optical Society of America

OCIS codes: 240.6680, 260.3910, 050.6624.

1. INTRODUCTION

The various novel and unusual optical properties of conducting microstructures or nanostructures, such as anomalous diffraction from metallic gratings, enhanced light transmission through subwavelength slits or holes, light polarizing through wire grid polarizers, surface-enhanced Raman scattering, etc., have attracted tremendous attention in recent years [1–4]. To date, the coupling of light with surface plasmons (SPs) has been widely adopted to explain these anomalous phenomena. However, the SP picture elaborated in numerous case studies in the literature actually corresponds to a very general concept about coupling of electromagnetic (EM) waves to free electron oscillation on conducting surfaces that can generate evanescent EM wave modes. This big picture is correct without doubt, but it is too general for one to obtain a clear and straightforward understanding of the essential underlying mechanism. Because of this uncertainty, the SP-like wave modes have been usually assumed to be the same as the classical SPs (CSPs) on planar metal surfaces [5], but this assumption is obviously challenged by the fact that (nearly) perfectly conducting structures that do not support CSPs still have *similar but stronger* anomalous light scattering properties [6,7]. Conductors with positive permittivity do not support CSPs either, but they can also exhibit light transmission anomalies [8–10]. (Extraordinary transmission through gratings can even occur for acoustic waves [11], which is completely irrelevant to SPs.) Because of these contradictions, the origin of anomalous light scattering from metallic microstructures is still being argued (e.g., see [12–16]).

Using modern computing techniques one may numerically solve Maxwell's equations for various complicated structures, but in the literature such computations have been largely focused on the EM fields. Surprisingly, the detailed mechanisms of free electron oscillation have been almost completely ignored, although they are known to play the fundamental role in the SP picture. Recently, we have briefly reported our computations of surface charge density waves (SCDWs) and the role they play in the process of enhanced light transmission through slit and hole arrays [9]. In this paper, we give a detailed and comprehensive illustration of the basic mechanism regarding light emission and interference from incident-wave-driven free electron oscillations, demonstrate that it is involved in light scattering from all periodic and non-periodic conducting structures (including perfect conductors), and thus establish a simple and universal *structured SP* picture. This picture provides a concrete view of the general SP concept and may bridge the gaps between the different mechanisms argued in the literature. It also provides solid guidelines for designing nano-optical devices by suggesting researchers concentrate on

the geometrical parameters of conducting nanostructures so as to control the locations, strengths, and interference of the charge-oscillation-induced light sources.

2. CHARGE-OSCILLATION-INDUCED LIGHT EMISSION AND INTERFERENCE

To illustrate the main picture, we start from the well-known principle of Thomson scattering of x rays by electrons [17,18], in which the incident x rays (EM waves with wavelengths ~0.1 nm) force the electrons in atoms (not necessarily free electrons) to oscillate with the same frequency. According to the fact that *accelerating charges radiate* (an elementary principle of electrodynamics), the oscillating electrons then emit new wavelets, which form the scattered waves. This effect also exists in the long wavelength range (say $\lambda > 0.1\ \mu$m), where electrons still oscillate with the incident wave (giving rise to oscillating polarization of the atoms). However, since now λ is much larger than the atoms (~0.1 nm), the net charge density averaged on the wavelength scale is zero in the bulk. Net polarization-induced charges do exist on surfaces (or interfaces), but for non-conducting materials, where the electrons are bound to the atoms and cannot move freely, the formation of net oscillating charges is very small even on rough surfaces.

A metal has free conduction electrons that move and oscillate easily on the surface in response to external EM waves and thus may emit new wavelets. But first note that a CSP corresponds to a surface-bound mode on the metal. If the oscillating charges emit light, how can the CSP be non-radiative? To clarify this ambiguity, let us use the Otto geometry in Fig. 1(a) as an example [5]. At a specific incident angle θ_{sp} [greater than the critical angle arcsin$(1/n_p)$ of the prism–vacuum interface], the incident wave can excite a CSP, which is a sinusoidal SCDW on the metal surface with a wavevector

$$k_{sp} = K(1 + 1/\varepsilon_c)^{-1/2}, \tag{1}$$

where ε_c is the permittivity of the metal and $K = 2\pi/\lambda$ (λ the incident wavelength in vacuum). Here θ_{sp} must satisfy $n_p K \sin\theta_{sp} = \mathrm{Re}(k_{sp})$, where n_p is the refractive index of the prism. Under this condition, the incident energy is largely transferred to the CSP, giving rise to a reflection dip, as can be proved by Fresnel theory [5].

Note that CSPs can be activated only on metals with $\mathrm{Re}(\varepsilon_c) < 0$ [and meanwhile $\mathrm{Im}(\varepsilon_c)$ being small] [5]. The reason is that under this condition, the spatial period of the CSP satisfies

$$d = 2\pi/\mathrm{Re}(k_{sp}) < \lambda \tag{2}$$

based on Eq. (1). Therefore, the CSP is a *subwavelength* charge pattern compared with the incident wavelength λ. Consider each period of the CSP in Fig. 1(a) as a scatter unit that emits new wavelets. Along any arbitrary direction $\theta \neq 0$, the wavelets emitted from two adjacent units have a path difference

$$\Delta S = d \sin\theta < d < \lambda, \tag{3}$$

i.e., the phase difference is less than 2π. This means that the oblique wavelets can never be *in phase*. Thus, they tend to cancel each other out in the far fields. The wavelets along the vertical direction $\theta = 0$, however, are in phase ($\Delta S = 0$), but viewed from a single period of the sinusoidal SCDW [Fig. 1(b)], each wavelet consists of two sub-wavelets with opposite electric fields $\mathbf{E}_1$ and $\mathbf{E}_2$ that also cancel each other out in the far fields. Therefore, all the emitted wavelets cannot escape the surface along any direction, so they form an *evanescent wave* above the surface [Fig. 1(c)] in the form $\mathbf{E}\exp(-ik_{sp}x)\exp(-\beta|z|)$, where $\beta = (k_{sp}^2 - K^2)^{1/2}$ is the evanescent factor. This gives a simple picture why a CSP corresponds to a surface-bound mode. The CSP can thus propagate outside the prism-covered region in Fig. 1(a) without radiation loss, and the propagation distance depends solely on the absorption of the metal (ohmic loss).

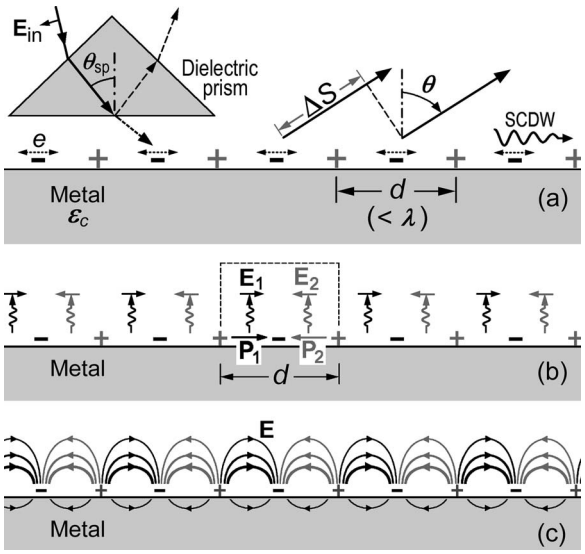

Fig. 1. CSP on a planar metal surface. "+" and "−" represent positive and negative charges, respectively. (a) Excitation of the CSP by Otto geometry. The SCDW is the result of electron oscillations (indicated by the dashed arrows) while the positive charges are fixed. (b) Symmetric sub-wavelets ($\mathbf{E}_1$ and $\mathbf{E}_2$) from a period of the charge wave (outlined by the dashed lines) along $\theta = 0$. $\mathbf{P}_1$ and $\mathbf{P}_2$ represent two oscillating dipoles with opposite directions, caused by the electron oscillation. (c) Near fields of the CSP (while the far fields tend to zero along any direction).

Here it is obvious that media with $\mathrm{Re}(\varepsilon_c) > 0$ do not support CSPs as the wavevector in Eq. (1) cannot satisfy Eq. (2). (Nearly) perfect conductors with $|\varepsilon_c| \to \infty$ do not support CSPs either because $k_{sp} \to K$, $d \to \lambda$, and $\beta \to 0$, i.e., the scattered wave above the conductor surface tends to be non-evanescent (not closely confined to the surface).

Now we consider in Fig. 2(a) a plane wave incident on a conducting surface without the prism. For normal incidence, the incident electric field $\mathbf{E}_{in}$ drives free electrons on the surface to move homogenously. So there is no net charge, and the reflection obeys the Fresnel equations [19]. [The force exerted by the electric field $\mathbf{E}$ of an EM wave on an electron is $-|e|\mathbf{E}$ (e the charge per electron). Since the motion of free electrons (not the charge waves) in conductors is much slower than the speed of light (non-relativistic), the force caused by the magnetic field $\mathbf{H}$ of the wave is negligible.] In Fig. 2(b), a slit (or hole) is added. Obviously, the electron movement now can be impeded near the slit corner. Here some electrons may move

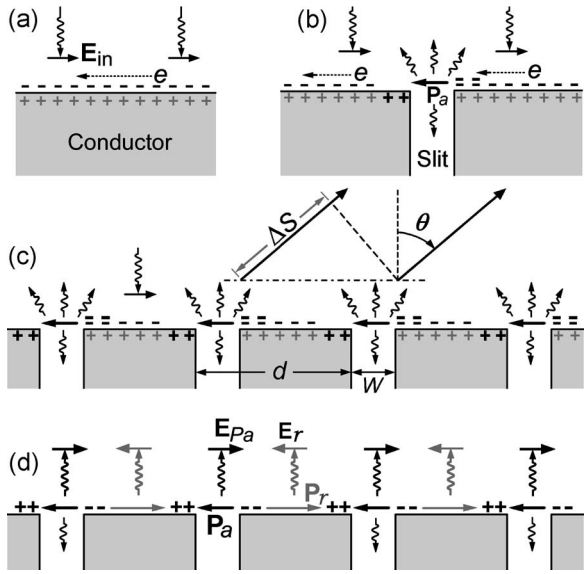

Fig. 2. Incident-wave-driven electron movement on (a) a flat conducting surface with no net charge, (b) a conducting surface with a single slit (net charges and their oscillation at the slit corners giving rise to a radiative light source), and (c) a conducting surface with a subwavelength slit array. (d) Periodic net charge pattern on the upper surface of (c). All the dipoles and electric fields have a common oscillating factor $e^{i\omega t}$ (t the time).

continuously to the vertical slit wall, but such movement corresponds to an abrupt 90° deflection (large angular acceleration), while the incident wave does not directly provide the necessary large driving force. So it is reasonable to assume that most of the moving electrons are stopped near one corner, while positive charges appear at the opposite corner because some electrons have moved away. This leads to the formation of an electric dipole $\mathbf{P}_a$ at the slit opening. $\mathbf{P}_a$ oscillates with the incident wave with a time factor $e^{i\omega t}$ (ω the angular frequency of the incident wave), thus acting as a new *light source* emitting wavelets [20]. Such a process is in fact a Thomson scattering process in the optical frequency range.

Next we apply this process to the one-dimensional (1D) periodic slit array in Fig. 2(c). For simplicity, we assume that the grating is semi-infinite so that there is no feedback from below. Similar to Fig. 2(b), now each slit becomes a light source, but along any oblique direction $\theta \neq 0$, the wavelets emitted from two adjacent sources have a path difference $\Delta S = d \sin\theta$, where d is the period of the slit array. For an incident wavelength $\lambda > d$, Eq. (3) is satisfied again. Then the oblique wavelets are canceled in the far fields (destructive interference, similar to the absence of x ray diffraction at non-Bragg angles), i.e., they also form evanescent waves near the surface. (This principle can also explain the fact that no light diffraction occurs from a single-crystal lattice, where the lattice constants are much smaller than the wavelength, although the electrons still oscillate with the incident wave.)

The charge pattern in Fig. 2(c) is similar to that of the CSP in Fig. 1, i.e., they are both *subwavelength charge patterns* ($d < \lambda$). However, there are two distinct differences. First, the CSP is a propagating charge wave with a specific wavevector determined by the metal's permittivity ε_c in Eq. (1), while the charge pattern in Fig. 2(c) is a standing wave (but not sinusoidal) with the period *always* equal to the grating period d. So the former is an intrinsic property of the metal (depending on ε_c), while the latter is a geometrical effect that can occur for any incident wavelength $\lambda > d$ and for any conducting materials containing free electrons (charges) [including perfect conductors and conductors with $\mathrm{Re}(\varepsilon_c) > 0$]. Second, as mentioned above, a CSP is a complete surface-bound mode. In contrast, the oscillating charge pattern in Fig. 2(c) is radiative along $\theta = 0$ ($\Delta S = 0$). This can be seen from Fig. 2(d), where we have discarded the oblique evanescent wavelets and added the dipoles $\mathbf{P}_r$ that are ignored in Fig. 2(c). In addition to the wavelet $\mathbf{E}_{Pa}$ emitted from $\mathbf{P}_a$, $\mathbf{P}_r$ also emits a wavelet $\mathbf{E}_r$ along $\theta = 0$ with a phase that is usually very close to that of the Fresnel reflected wave. So here we let $\mathbf{E}_r$ include Fresnel reflection for convenience in discussions. Then the wavelet emitted from a period consists of two sub-wavelets $\mathbf{E}_{Pa}$ and $\mathbf{E}_r$ along $\theta = 0$ with opposite directions (phases). But unlike in Fig. 1(b), $\mathbf{E}_{Pa}$ and $\mathbf{E}_r$ generally have different strengths, so they cannot completely offset each other. This leads to a propagating backward wave. Therefore, the charge pattern illustrated in Figs. 2(c) and 2(d) is not identical to a CSP. Previously we called such oscillating charge patterns *spoof SPs* in [9], but more precisely we may call them *structured SPs*, because the charge patterns have detailed structures and are mainly determined by the real patterns of the conducting media and because of their similarity (i.e., the subwavelength characteristic) to the true CSP in Fig. 1(a).

Note that the formation of oscillating charge patterns in 1D metallic slits or grooves has already been demonstrated in a number of theoretical and experimental results in the literature (e.g., [21–25]). In particular, Gay *et al.* [26] have directly shown interference in the far field from oscillating dipoles at the exit side of a finite-thickness slit-groove structure, but note that the distance between the slit and groove is greater than the incident wavelength ($d > \lambda$), such that the waves emitted from the two dipoles leads to far-field interference (non-evanescent waves). Here we provide a straightforward view of the formation mechanism of these charge patterns.

By studying perfect conductors perforated with periodic hole arrays, Pendry *et al.* [27] have found that the effective permittivity has the same plasmon form as CSPs and thus have proposed the concept of spoof SPs for perfectly conducting structures. Here we give a concrete picture (or at least a different view) of the fundamental mechanism underlying the formation of spoof or structured SPs. Most importantly, this mechanism is not limited to perfect conductors, but applicable to all conducting structures (including finite-conductivity structures). It should be noted that the original definition of spoof SPs by Pendry *et al.* is for surface modes that follow the same plasmon form as CSPs. This plasmon form may (roughly) apply to 2D hole arrays, but not necessarily to 1D or other structures (particularly with finite thickness). Here the structured SPs we illustrate have broader definitions, referring to any subwavelength charge patterns induced on structured conducting surfaces (including non-periodic structures discussed in Section 6 and perfect conductors). The common feature of structured SPs is that they lead to evanes-

cent modes near the surfaces. For periodic structures, the periodicity of the charge patterns is controlled by the structures rather than the "natural" wavelength of the CSP of real metals.

In addition to the CSP model (for finite-conductivity structures) and the spoof SP model (for perfect conductors), Lezec and Thio [8] have proposed the composite diffractive evanescent wave (CDEW) model to explain anomalous light scattering from metallic structures (also see [28]). From above one may see that these models do not contradict each other since the underlying mechanism for these models is the same, i.e., the incident wave induces subwavelength oscillating charge patterns (Thomson scattering), which further lead to SP-like evanescent wave modes near the surface. Thus, we believe our structured SP picture at least provides a close connection between the three models.

From Fig. 2 it is not difficult to obtain a general picture regarding light scattering from structured (or rough) conducting surfaces (either periodic or non-periodic). When light is incident on a non-planar conducting surface, it drives the free electrons to move, but the movement can be impeded by the rough parts (e.g., grooves, holes, bumps, particles) of the surface to form inhomogeneous oscillating charges, which become new light sources to emit wavelets [21]. It is the interference between these wavelets that may give rise to anomalous reflection or scattering. In the following, we will numerically prove this mechanism in the simple and well studied case of periodic 1D gratings using the rigorous coupled-wave analysis (RCWA) technique [29,30].

3. RCWA OF 1D LATTICE

For monochromatic waves in a nonmagnetic medium (permeability $\mu \equiv 1$), the electric and magnetic fields are coupled by Maxwell's equations (in c.g.s. units)

$$\nabla \times \mathbf{E} = -iK\mathbf{H}, \tag{4}$$

$$\nabla \times \mathbf{H} = iK\varepsilon\mathbf{E}, \tag{5}$$

where $K=2\pi/\lambda$ and ε is the *effective permittivity*. The effective permittivity of a conductor can be expressed as $\varepsilon_c=\varepsilon_c'-i4\pi\sigma/\omega$, where ε_c' is the regular permittivity and σ is the conductivity [19]. For perfect conductors, $\sigma\to\infty$ so that $\mathrm{Im}(\varepsilon_c)\to-\infty$ (which can also be derived from the Drude model of electrical conduction). The divergence of Eq. (5) gives $\nabla\cdot(\varepsilon\mathbf{E})=0$, or

$$\nabla\cdot\mathbf{E} = -[(\nabla\varepsilon)\cdot\mathbf{E}]/\varepsilon = 4\pi\rho, \tag{6}$$

where $\rho(\mathbf{r})$ is the *bulk charge density* (including both free and polarization-induced charges). In a modulated medium with varying $\varepsilon(\mathbf{r})$, $\nabla\varepsilon\neq 0$, which generally leads to inhomogeneous charge densities $\rho(\mathbf{r})$ according to Eq. (6). Mathematically, ε is discontinuous across a sharp interface (i.e., $\nabla\varepsilon\to\infty$), so one has to use the *surface charge density* $4\pi\tilde{\rho}_s=\delta E_z$ to describe the charge distribution on the interface, where δE_z is the jump of the perpendicular electric field component across the interface. As an exception, TE-polarization in a 1D structure satisfies $\rho\equiv 0$ since $(\nabla\varepsilon)\cdot\mathbf{E}=0$, so here we ignore it [9].

From Eqs. (4) and (5) one can obtain a second-order differential equation $\nabla\times(\varepsilon^{-1}\nabla\times\mathbf{H})=K^2\mathbf{H}$, of which the Fourier transformation form for the 1D lattice in Fig. 3 is

$$K^2\mathbf{H}_m = -\sum_n \zeta_{m-n}\mathbf{k}_m\times\mathbf{k}_n\times\mathbf{H}_n, \tag{7}$$

based on the Fourier expansions $\varepsilon^{-1}(x)=\Sigma\zeta_m e^{-iG_m x}$ and $\mathbf{H}(\mathbf{r})=\Sigma\mathbf{H}_m e^{-i\mathbf{k}_m\cdot\mathbf{r}}$, where $G_m=2\pi m/d$ (m being integers), $\zeta_m=d^{-1}\int_0^d\varepsilon^{-1}(x)e^{iG_m x}dx$, $\mathbf{k}_m=\mathbf{k}_0+G_m\hat{\mathbf{x}}$, and $\mathbf{k}_0$ is the forward wavevector [28]. Equation1 (7) can be numerically solved by RCWA. Here we briefly mention its main principles to make our calculations convincing.

In Fig. 3, the incident wave is $\mathbf{H}_{\mathrm{in}}\exp(-i\mathbf{K}_{\mathrm{in}}\cdot\mathbf{r})$ with $\mathbf{K}_{\mathrm{in}}=K(\sin\theta\hat{\mathbf{x}}+\cos\theta\hat{\mathbf{z}})$. The forward wavevector $\mathbf{k}_0$ can be written as $\mathbf{k}_0=k_{0x}\hat{\mathbf{x}}+q\hat{\mathbf{z}}$ with $k_{0x}=K\sin\theta$, where q is to be determined by the eigenequation. Then the internal diffracted wavevectors have the form $\mathbf{k}_m=k_{mx}\hat{\mathbf{x}}+q\mathbf{z}$ with $k_{mx}=k_{0x}+G_m$. Each diffraction order m corresponds to two diffracted waves $\mathbf{H}_m^R\exp(-i\mathbf{K}_m^R\cdot\mathbf{r})$ and $\mathbf{H}_m^T\exp(-i\mathbf{K}_m^T\cdot\mathbf{r})$ above and below the grating, respectively. Based on the conservation of the tangential wavevector components across the surfaces $z_s=0$ and τ, we have

$$K_{mx}^T = K_{mx}^R = k_{mx} = k_{0x}+G_m \text{ (always real)},$$

$$K_{mz}^T = -K_{mz}^R = \begin{cases} (K^2-k_{mx}^2)^{1/2} & \text{for } |k_{mx}|\le K, \\ -i(k_{mx}^2-K^2)^{1/2} & \text{otherwise.} \end{cases} \tag{8}$$

Here note that when $|k_{mx}|>K$, the corresponding external waves become evanescent waves $H_m^{R,T}e^{-ik_{mx}x}\exp[-(k_{mx}^2-K^2)^{1/2}|z|]$ along $\pm z$. In particular, for normal incidence ($k_{0x}=0$) and $\lambda>d$, all the external waves except for $m=0$ are evanescent, $H_m^{R,T}e^{-2\pi imx/d}\exp[-2\pi(m^2/d^2-1/\lambda^2)^{1/2}|z|]$. This is the mathematical description of the evanescent EM waves of the structured SP described in Fig. 2.

For TM polarization, all the magnetic fields are parallel to $\hat{\mathbf{y}}$. If we retain $2M+1$ diffraction orders $(0,\pm1,\cdots,\pm M)$, Eq. (7) can be written as a $(2M+1)\times(2M+1)$ matrix eigenequation. From this eigenequation and the boundary conditions (continuity of the tangential electric and magnetic fields) at the two surfaces $z_s=0$ and τ, one ob-

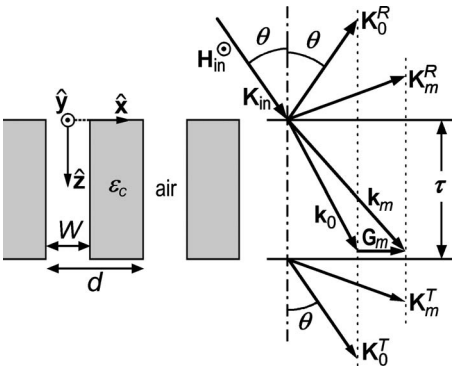

Fig. 3. Geometry of light diffraction from the 1D periodic grating. The two slit walls are located at $x=\pm W/2$ (plus any multiple of d). The vertical components of the internal wavevectors $\mathbf{k}_m$ are generally complex vectors. $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, $\hat{\mathbf{z}}$ are unit vectors along the x, y, z axes, respectively.

tains $4M+2$ sets of eigenvalues q_ι and eigenmodes $\{H_m^\iota\}$ ($m=-M, -M+1,\cdots,M$, and $\iota=1,2,\cdots,4M+2$) inside the grating and two sets of external fields $\{H_m^R\}$ and $\{H_m^T\}$ (see [28–30] for details). Then the zeroth-order reflectivity and transmissivity are $R_0=|H_0^R/H_{\text{in}}|^2$ and $T_0=|H_0^T/H_{\text{in}}|^2$, respectively. Meanwhile, the electric fields $\mathbf{E}_m^\iota$, $\mathbf{E}_m^R$, and $\mathbf{E}_m^T$ (parallel to the xz plane) are also obtained from (the Fourier transformation of) Eq. (5). Then the bulk charge density inside the grating ($0<z<\tau$) can be calculated from

$$4\pi\rho(\mathbf{r}) = -i\sum_{m=-M}^{M}\sum_{\iota=1}^{4M+2}(\mathbf{k}_m^\iota\cdot\mathbf{E}_m^\iota)\exp(-i\mathbf{k}_m^\iota\cdot\mathbf{r}), \tag{9}$$

where $\mathbf{k}_m^\iota=k_{mx}\hat{\mathbf{x}}+q_\iota\hat{\mathbf{z}}$. The surface charge density is

$$4\pi\tilde{\rho}_s(x) = -E_{\text{in},z}e^{-ik_{0x}x}+\sum_m e^{-ik_{mx}x}\left(-E_{mz}^R+\sum_\iota E_{mz}^\iota\right) \tag{10}$$

on the upper surface $z=0$ and

$$4\pi\tilde{\rho}_s(x) = \sum_m e^{-ik_{mx}x}\left(E_{mz}^{T\tau}-\sum_\iota E_{mz}^\iota\phi_\iota\right) \tag{11}$$

on the lower surface $z=\tau$, where $\phi_\iota=e^{-iq_\iota\tau}$ and $E_{mz}^{T\tau}=E_{mz}^T\exp(-iK_{mz}^T\tau)$. [For large τ, one may need to make the substitutions $H_m e^{-iq_\iota\tau}\to H_m^\tau$ and $H_m\to H_m^\tau e^{iq_\iota\tau}$ for $\text{Im}(q_\iota)>0$ to avoid numerical overflow in computing $e^{-iq_\iota\tau}$, where H_m^τ is the corresponding internal wave amplitude at the lower surface.] For a semi-infinite grating ($\tau\to\infty$), only half of the eigenmodes with $\text{Im}(q_\iota)<0$ are valid, so we need to use only the boundary conditions at the upper surface to compute the reflectivity and charge densities. Overall, RCWA is a first-principle method with the computation precision depending only on the number of diffraction orders $(2M+1)$ retained, but note that it usually converges more slowly in calculations of charges and near fields than in calculations of (far-field) reflectivity and transmissivity.

In calculating the bulk charge density $\rho(x,z)$ using Eq. (9), we find that when a large number of diffraction orders are retained, $\rho(x,z)$ approaches a delta function across the walls, which means that "bulk" charges exist only on the slit walls, i.e., they are also surface charges [9]. Mathematically, we let $\rho(x=\pm W/2,z)$ represent the surface charge densities on the slit walls (in arbitrary units). With sufficient orders retained, this approximation does not affect the shapes and phases of the real surface charge density curves on the slit walls.

4. SEMI-INFINITE GRATINGS

In the above RCWA descriptions of light scattering from a 1D (or 2D) lattice, the eigenmodes form pairs, each pair consisting of two eigenmodes with opposite (complex) vertical wavevectors q and $-q$. As will be demonstrated later, one mode propagating along $-z$ corresponds to reflection from the bottom surface for finite τ. This mode can resonate with the forward-propagating one (along $+z$). In order to verify the picture in Figs. 2(c) and 2(d) without the complication of the resonance, we first consider a semi-infinite grating ($\tau\to\infty$) where the backward eigenmodes do not exist.

Figure 4(a) shows the reflectivity curve calculated with RCWA from a semi-infinite gold grating (practically $\tau>200$ μm) under normal incidence (with the frequency-dependent permittivity data of gold taken from [31]). As a reference, the dashed curve is that of $(1-W/d)R_f$, with R_f being the Fresnel reflectivity from a flat gold surface ($R_f\simeq 0.98$ in the wavelength range 0.8–10 μm) and W the slit width. Compared with this reference curve, the anomalous reflection phenomenon from the grating is obvious. Generally the reflectivity R_0 is less than $(1-W/d)R_f$, except that near the Wood's anomalies $\lambda\simeq d/|m|$ ($m\neq 0$ being integers), R_0 is close to unity.

Figure 4(b) shows the charge density function $\tilde{\rho}_s(x)$ on the upper surface ($z=0$) for an arbitrary wavelength in the $\lambda>d$ range. This function correctly shows that the incident wave indeed causes significant inhomogeneous charges on the grating surface with the charges strongly accumulating near the slit corners, which is excellently consistent with the charge distribution pattern predicted in Fig. 2(d). At a time when $\mathbf{E}_{\text{in}}$ is toward $+x$ at $z=0$, we have predicted in Fig. 2(d) that the phase of the charge pattern is constant, equal to $-\pi$ (negative charges) on the left half surface $W/2\leqslant x\leqslant d/2$, while for $d/2\leqslant x\leqslant d-W/2$, the phase is 0 (positive charges). Figure 4(b) shows that this prediction is largely correct, except that the calculated phases are slightly displaced from the predicted phases 0 and π by $\Delta\varphi\simeq 0.1\pi$ in most regions on the surface. The phases near the slit corners are closer to the predicted values. Our calculations show that *the charge patterns are nearly the same for any wavelength* $\lambda>1.1d$ with no resonance, and the phase shift $\Delta\varphi$ decreases with increasing λ, i.e., $\Delta\varphi\to 0$ for $\lambda\gg d$. Therefore, the calculations indeed confirm the picture of charge accumulation and oscillation on the subwavelength lattice in Figs. 2(c) and 2(d). Clearly, the period of the charge pattern in Fig. 4(b) is strictly equal to the lattice constant d and is irrelevant to the dispersion property of CSPs in Eq. (1). By performing RCWA calculations on gratings made of con-

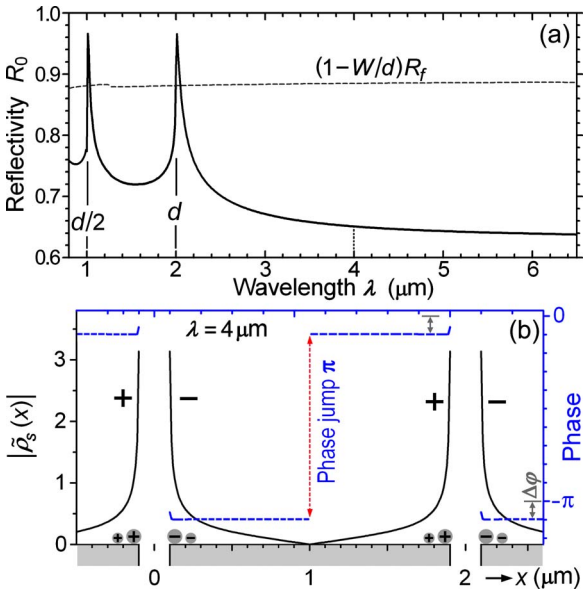

Fig. 4. (Color online) Reflectivity (a) and surface charge densities (b) of a semi-infinite gold grating with period $d=2$ μm and slit width $W=0.2$ μm. Normal incidence.

ductors with $\mathrm{Re}(\varepsilon_c)>0$ or perfect conductors with $\mathrm{Im}(\varepsilon_c)\to-\infty$, we find that the main features of the charge patterns remain unchanged.

To understand the anomalous reflection for $\lambda>d$ in Fig. 4(a), we may simply consider that, as we mentioned before, the $\mathbf{E}_r$ wavelet in Fig. 2(d) consists of two contributions, $\mathbf{E}_r=-(E_f+E_r')\hat{\mathbf{x}}$, with E_f corresponding to regular Fresnel reflection ($E_f=-\sqrt{R_f}E_{\mathrm{in}}$) and E_r' corresponding to the emission of light from dipole $\mathbf{P}_r$ along the backward direction $\theta=0$. Without charge accumulation, we have $E_{Pa}=E_r'=0$ and the reflection should obey the Fresnel theory $R_0=(1-W/d)R_f$. When charges appear at the slit corners, it can be verified by RCWA that $\mathbf{P}_a$ is strengthened much faster than $\mathbf{P}_r$ (for $W<d/2$). Then the effective strength of $\mathbf{E}_{Pa}$ becomes larger than E_r'. Consequently, $\mathbf{E}_{Pa}$ completely cancels E_r' and also partially offsets E_f. The net effect is that part of the photons incident on the metal surfaces are "funneled" into the slit. Thus, the overall reflectivity R_0 is smaller than $(1-W/d)R_f$.

When λ is reduced to less than (or close to) the grating period d, some of the external wavevectors K_{mz}^R in Eqs. (8) become (or tend to be) real, and the corresponding diffracted waves become non-evanescent. Then the diffraction effect appears, which can significantly change the reflectivity (particularly at the Wood's anomalies $\lambda\simeq d/|m|$). The details in the diffraction range $\lambda<1.1d$ are discussed in Appendix A, since they no longer belong to subwavelength optics. But, it is worth emphasizing again here that the diffraction effect is absent for the entire long wavelength range $\lambda>1.1d$, where the subwavelength charge patterns are always nearly the same as those in Fig. 4(b).

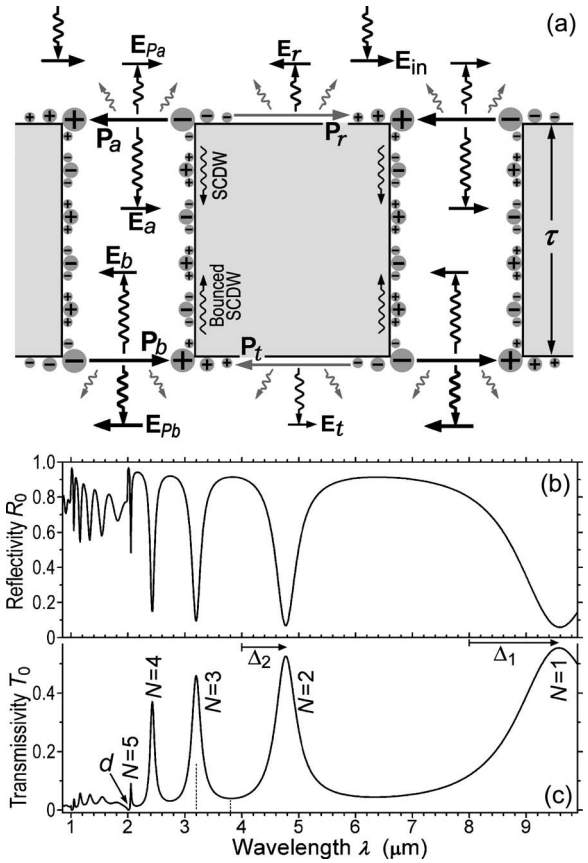

Fig. 5. (a) Process of charge-oscillation-induced light emission, resonance, and transmission through a conducting grating. Reflectivity (b) and transmissivity (c) of a gold grating with parameters $d=2$, $W=0.2$, and $\tau=4$ μm.

5. FINITE-THICKNESS GRATINGS

In Fig. 2 we have indicated that the dipole $\mathbf{P}_a$ also emits a wavelet in the slit toward $+z$ [see $\mathbf{E}_a$ in Fig. 5(a), which may also include a portion of the incident wave directly transmitted into the slit]. Because of the waveguide constraint, $\mathbf{E}_a$ tends to be a plane wave inside the slit, i.e., $\mathbf{E}_a\simeq E_a\exp(-ik_zz)\hat{\mathbf{x}}$, where $k_z\simeq 2\pi/\lambda$. Similarly, this wave drives electrons *on the slit walls* to oscillate, resulting in two SCDWs $\rho_a\exp(-ik_zz)$ and $-\rho_a\exp(-ik_zz)$ (with $\rho_a\propto E_a$) on the two opposite walls, respectively. The SCDWs and the $\mathbf{E}_a$ wave propagate along $+z$ and attenuate gradually because of the absorption of the conductor. If the grating is extremely thick, these waves can be completely absorbed before reaching the bottom surface, which corresponds to the semi-infinite case.

If the grating is thin enough, the SCDWs on the walls can reach the exit surface without significant absorption. Then in a similar way, the moving charges can be impeded at the lower slit corners, leading to another large oscillating dipole $\mathbf{P}_b$, as shown in Fig. 5(a) [9]. $\mathbf{P}_b$ can give a strong feedback to the upper surface by emitting a wavelet $\mathbf{E}_b\simeq E_b\exp(ik_zz)\hat{\mathbf{x}}$ propagating upward. (In other words, $\mathbf{E}_b$ can be considered to be the reflected wave of $\mathbf{E}_a$ from the bottom surface.) $\mathbf{E}_b$ also corresponds to two SCDWs, $\pm\rho_b\exp(ik_zz)$, on the two walls, respectively, which are in fact the back-bounced SCDWs of the $\pm\rho_a\exp(-ik_zz)$ waves by the bottom corners. If $\mathbf{E}_b$ is *in phase* with $\mathbf{E}_a$ at $z=0$, it enhances $\mathbf{P}_a$. The enhanced $\mathbf{P}_a$ subsequently strengthens $\mathbf{E}_{Pa}$, $\mathbf{E}_a$, $\mathbf{P}_b$, $\mathbf{E}_b$, and so on. Then a Fabry–Perot-like resonant state is formed, with $\mathbf{E}_a$ and $\mathbf{E}_b$ forming a standing wave $\mathbf{E}_a+\mathbf{E}_b\propto\cos(k_zz)e^{i\omega t}$ in the slit. Under this condition, $\mathbf{E}_r$ is largely offset by $\mathbf{E}_{Pa}$ in the far fields, leading to minimized backward reflection. Figure 5(b) shows the zeroth-order reflectivity curve of a gold grating with thickness $\tau=4$ μm. Compared with Fig. 4(a), one can see that Fig. 5(b) indeed shows a number of reflection dips corresponding to Fabry–Perot resonance.

At the exit surface $z=\tau$ [Fig. 5(a)], dipoles $\mathbf{P}_b$ and $\mathbf{P}_t$ also emit wavelets toward the outside of the slit. For $\lambda>d$, only the wavelets $\mathbf{E}_{Pb}$ and $\mathbf{E}_t$ can propagate along $+z$ (while the oblique wavelets again form evanescent waves). Unlike the case above the upper surface where $\mathbf{E}_r$ contains specular reflection, here wavelet $\mathbf{E}_t$ is purely emitted from dipole $\mathbf{P}_t$. For $W<d/2$, the strength of $\mathbf{P}_b$ ($\mathbf{E}_{Pb}$) is much greater than that of $\mathbf{P}_t$ ($\mathbf{E}_t$), so the transmitted wave is dominated by $\mathbf{E}_{Pb}$. Consequently, the energy of the transmitted wave is highly localized near the exit opening. For long wavelengths $\lambda\gg d$, such a "near-field focusing" effect can achieve a focusing width W far smaller than λ, which has potential applications in nano focusing, beaming, lithography, imaging, etc. In the far-field region, however, this localization effect disappears as the transmitted beam becomes a plane wave for periodic slit arrays [32]. (For an isolated slit, the transmitted beam is divergent.) At resonant wavelengths, since the

strength of wavelet $\mathbf{E}_{Pb}$ is maximized, the zeroth-order transmissivity T_0 is also maximized [9], as can be seen in Fig. 5(c), where each reflection dip exactly corresponds to a transmission peak (also see similar results from finite-difference time-domain calculations in [15]).

If the waves $\mathbf{E}_a$ and $\mathbf{E}_b$ are ideal plane waves with wavevector $k_z=2\pi/\lambda$, Fabry–Perot resonance should occur at $\lambda_N=2\tau/N$ (N the resonance order), except that resonant peaks with $\lambda_N \leqslant d$ are suppressed by the diffraction effect [9]. However, the actual resonance wavelength is always redshifted, $\lambda_N=2\tau/N+\Delta_N$, where the redshift Δ_N may vary (slowly) with d, W, τ, and ε_c. One reason for the redshift is that the standing wave is distorted near the two ends of the slit [see Fig. 6(b)]. One may refer to [33] for discussions of other possible mechanisms. Here note that due to the redshift, the spatial period of the (sinusoidal) SCDWs on the wall is less than the incident wavelength by approximately Δ_N, i.e., they are also subwavelength charge waves.

Based on $\lambda_N \sim 2\tau/N$, thick (and highly conducting) gratings ($\tau \gg d$) have many resonance wavelengths in the non-diffraction range $\lambda > d$, as experimentally demonstrated in [34]. Not shown in Fig. 5(b) is that when λ increases above λ_1, the transmissivity T_0 first drops to a minimum, then monotonically increases with $\lambda \to \infty$. This is also true for very thin gratings with $\tau < d/2$, where the Fabry–Perot resonance condition cannot be satisfied in the $\lambda > d$ range. Then T_0 also increase monotonically from $\lambda = d$ toward the longer wavelength direction [see Fig. 1(a) in [35]]. The reason is that under these two conditions, the phase difference between $\mathbf{E}_b$ and $\mathbf{E}_a$ ($\mathbf{E}_{\text{in}}$) at $z=0$ is roughly $2\tau/\lambda$, which decreases toward zero (the "zeroth-order" Fabry–Perot resonance condition) when λ increases toward ∞.

In Figs. 6(a) and 6(b), the computed charge density distributions at resonant wavelength λ_3 are well consistent with the picture of Fig. 5(a). In Fig. 6(a), the two surface charge patterns $|\tilde{\rho}_s(x,z=0,\tau)|$ are very similar to those in Fig. 4(b), which confirms the existence of the large dipoles $\mathbf{P}_a$ and $\mathbf{P}_b$ in Fig. 5(a). Note that the charge densities in Figs. 6(a) and 4(b) are in the same (arbitrary) units. Therefore, the charge densities near the slit corners are much higher in Fig. 6(a) than in Fig. 4(b) because of the Fabry–Perot resonance/enhancement. As also shown in Fig. 6(a), in thin gratings where the attenuation of the charge density waves on the slit walls is negligible, the two SCDWs $\tilde{\rho}_s(x,z=0,\tau)$ are almost identical except that for odd resonant orders N, they have a phase difference π. For relatively thicker gratings, the strength of $\tilde{\rho}_s(x,z=\tau)$ drops with increasing τ. For $\tau > 200$ μm, $\tilde{\rho}_s(x,z=\tau)$ almost disappear while $\tilde{\rho}_s(x,z=0)$ tends to be the same as that in Fig. 4(b).

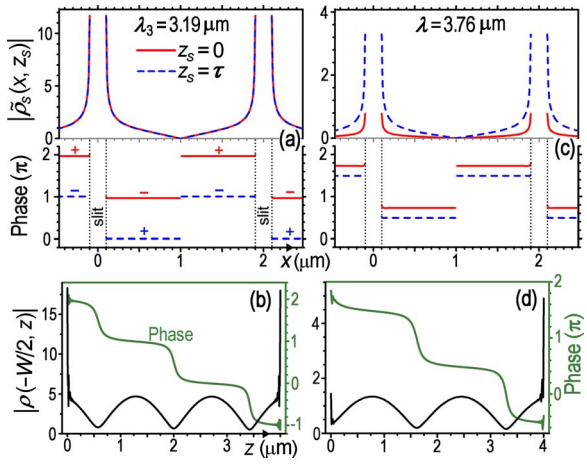

Fig. 6. (Color online) Charge patterns on the thin gold grating. (a) Charge densities on the two surfaces $z_s=0$ and τ for the resonant wavelength $\lambda_3=3.19$ μm of peak $N=3$ in Fig. 5. (b) Nearly standing charge wave on the slit wall $x=-W/2$ at λ_3. Note that the charge waves on the two walls are always identical except for a phase difference of π, i.e., $\rho(W/2,z) \equiv -\rho(-W/2,z)$. (c) Surface charge densities at a non-resonant wavelength $\lambda=3.76$ μm (corresponding to the valley between peaks $N=2$ and $N=3$ in Fig. 5). (d) Charge wave on the slit wall $x=-W/2$ for $\lambda=3.76$ μm, where $z=0$ is no longer an anti-node of the standing wave.

Figure 6(b) correctly reveals that on the slit walls, the charge density waves $\rho(x=\pm W/2,z)$ with approximately stepped phases are nearly standing waves. Here the ρ profile also shows high accumulation of charges at the slit corners ($x=\pm W/2$, $z=0,\tau$) that are (always) in phase with $\tilde{\rho}_s(x=\pm W/2,z=0,\tau)$ [9].

In Fig. 5(a), if $\mathbf{E}_b$ is not in phase with $\mathbf{E}_a$ (and $\mathbf{E}_{\text{in}}$) at $z=0$, it suppresses the strengths of $\mathbf{P}_a$ and $\mathbf{E}_{Pa}$ and influences their phases. Consequently the strengths of the charge waves on the slit walls are also reduced, leading to a weaker dipole $\mathbf{P}_b$ and weak transmissivity. This mechanism is clearly shown in Figs. 6(c) and 6(d) at a non-resonant wavelength. Compared with Figs. 6(a) and 6(b), the charge densities at the slit corners ($x=\pm W/2$, $z=0,\tau$) all drop significantly for both ρ and $\tilde{\rho}_s$, particularly at the upper corners. Meanwhile, the phases of the charge waves are also altered so that no resonance is formed. As stated above, without surface charges, the reflectivity from the upper surface should be the Fresnel reflectivity $R_0=(1-W/d)R_f$. Here one can see from Figs. 6(c) and 6(d) that at non-resonant wavelengths, the strengths of $\mathbf{P}_a$ and $\mathbf{P}_r$ at the upper surface are very small, and then the reflectivity R_0 in Fig. 5(b) is indeed very close to $(1-W/d)R_f \simeq 0.9$ in most of the non-resonant wavelength range. For the same wavelength, the non-resonant reflectivity in Fig. 5(b) is much stronger than that in Fig. 4(a), where charge oscillation is heavily involved. This further proves the essential role charge oscillation plays in extraordinary light scattering from conducting structures.

As mentioned above, perfect conductors with $\text{Im}(\varepsilon_c) \to -\infty$ and conductors with $\text{Re}(\varepsilon_c) > 0$ do not support CPSs. However, we have demonstrated in [9] that 1D gratings with $\text{Re}(\varepsilon_c) > 0$ may still show similar extraordinary light transmission, although the transmissivity is relatively lower (also see [8,10]). Here we use RCWA to simulate the transmission through a nearly perfectly conducting grating with a large constant imaginary permittivity $\varepsilon_c \equiv -i10^7$. Based on this value, the wavevector k_{sp} in Eq. (1) is almost identically equal to K, so CSPs should not exist. However, our calculations show that all the major properties in this case are almost identical to those of regular metallic gratings. For example, Fig. 7(a) shows the transmissivity curve calculated with the same geometrical parameters as in Fig. 5(c), while Fig. 7(b) shows the charge densities on the grating surfaces for the third-order resonance peak. Compared with the reflectivity curve in Fig.

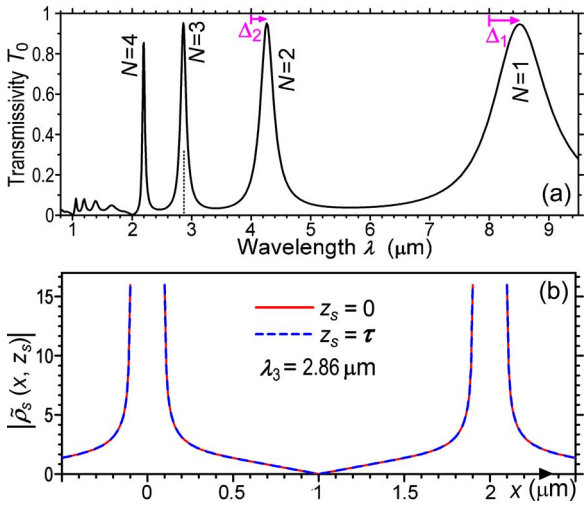

Fig. 7. (Color online) Simulated light scattering from a nearly perfectly conducting grating with $\varepsilon_c \equiv -i10\,000\,000$ (compared with $\varepsilon_c = -2834 - i1333$ for gold at $\lambda = 10$ μm). $d = 2$, $W = 0.2$, and $\tau = 4$ μm. Normal incidence. (a) Transmissivity curve. Compared with Fig. 5(c), here the redshifts (Δ_N) of the resonant wavelengths are smaller (so peak $N = 5$ below $\lambda = d$ is truncated). (b) Surface charge densities at the resonant wavelength λ_3 in (a).

5(c) and the charge density distribution in Fig. 6(a), Fig. 7 apparently shows that the light scattering mechanisms for the perfect-conductor case are the same, and thus are irrelevant to CSPs. The resonant transmissivity peaks and the charge densities in Fig. 7 are much higher than those for the gold grating, indicating that high conductivity can significantly enhance the extraordinary scattering effects. In fact, the transmission spectra and charge patterns calculated with $\varepsilon_c \equiv \pm 10^6 - i10^7$ are almost identical to those in Fig. 7 calculated with $\varepsilon_c \equiv -i10^7$, indicating that for highly conducting gratings, the extraordinary transmission is dominated by the large conductivity-related $\mathrm{Im}(\varepsilon_c)$ and nearly independent of $\mathrm{Re}(\varepsilon_c)$ (and its sign). [In particular, under the condition of $\varepsilon_c \equiv +10^6 - i10^7$, the CSP excitation condition in Eqs. (1) and (2) is completely destroyed.] Also note that the $R_0 + T_0$ values (not shown in Fig. 7) are always close to unity for $\lambda > d$, which means that the ohmic loss for the nearly perfectly conducting gratings is indeed negligible ($< 1\%$).

As demonstrated in [9,36], extraordinary transmission or scattering through 2D hole arrays involves the same mechanisms of light emission and interference, except that the tunneling of the SCDWs through the holes is different and that Fabry-Perot resonance is not involved. Recently we noted that the surface charge patterns of 2D hole arrays predicted in [9] had been experimentally demonstrated in [6] for microwaves. The details of oblique incidence geometry will be presented elsewhere, but the charge oscillation principle is similar.

6. NON-PERIODIC STRUCTURES

From the above demonstrations, it becomes obvious that charge-oscillation-induced light emission and interference are a fundamental and *universal* mechanism underlying various extraordinary light scattering processes from conducting structures, although these processes may also involve other mechanisms simultaneously (e.g., cavity resonance). The only basic requirement for this mechanism to work is that the structure have free electrons. So this mechanism applies to structures of metals, perfect conductors, conductors with $\mathrm{Re}(\varepsilon_c) > 0$, semiconductors, etc., but high conductivity can significantly enhance the anomalous scattering effects.

This mechanism also applies to non-periodic structures [37]. From Fig. 2(b) one can see that an isolated single slit also acts as a light source. If the conducting plate has a finite thickness, the exit opening of the slit at the lower surface becomes another strong light source at a Fabry–Perot resonant wavelength, emitting a transmitted beam below the plate [33]. Compared with the periodic slit array in Fig. 5(a), the waves emitted from a single slit have no interference and thus are completely divergent. If the slit is surrounded by periodic grooves on the entrance surface, as shown in Fig. 8(a), each groove now acts as a light source. Under the conditions that the groove period is less than the incident wavelength and that the Fabry–Perot-like resonance can be achieved simultaneously in both the grooves and the slit, the EM fields above the upper surface become similar to those in Fig. 5(a) with the oblique waves forming evanescent modes. Most importantly, the enhanced backward wavelets from the light sources [$\mathbf{E}_{Pa}$ in Fig. 5(a)] significantly reduce the Fresnel reflection ($\mathbf{E}_r$). Accordingly, the backward reflection is reduced while the transmission through the slit can be greatly enhanced (by up to two orders in [38]). But, the transmitted beam below the plate is still divergent.

Now if similar grooves are made on the exit surface, the wavelets emitted from the exit opening of the slit also drive free electrons to form oscillating dipoles at the groove openings. Thus, they also become light sources. Similarly, the oblique wavelets emitted from the slit opening and the grooves on the exit surface tend to form eva-

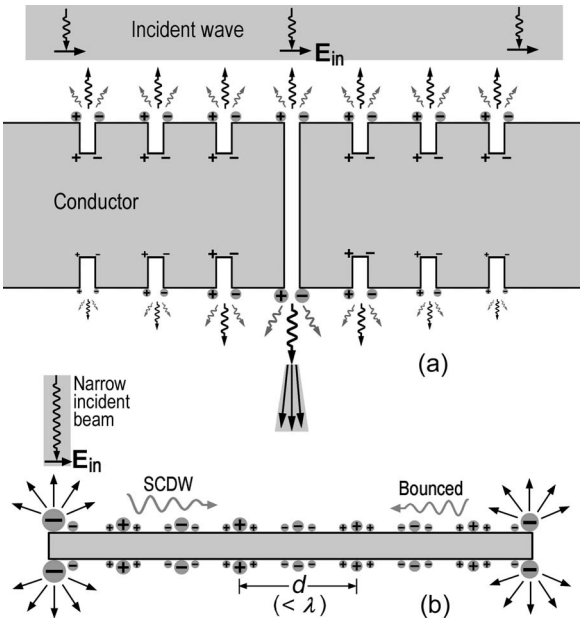

Fig. 8. (a) Schematic representation of enhanced light transmission and directed nanobeaming through a single slit surrounded by grooves. (b) Transfer of light on a conducting nanowire by subwavelength SCDWs.

nescent waves, giving rise to a narrow and directed transmitted beam below the slit opening. However, the dipoles on the upper entrance surface are formed and driven mainly by the wide incident wave, while those on the exit surface result only from the wavelets emitted from the dipole of the single slit. Therefore, the strengths of the light sources on the exit surface decrease quickly with increasing distance of the grooves from the slit. Because of this, the transmitted beam cannot be completely collimated, since the oblique wavelets cannot be completely suppressed. Meanwhile, the grooves on the exit surface have little influence on the overall transmission efficiency [38].

Obviously, the picture illustrated in Fig. 8(a) can also explain enhanced light transmission and directed nanobeaming through a single aperture surrounded by circular grooves in the 2D case, except that the cavity resonance mechanisms in the aperture and in the circular grooves may be different, and the directions and distributions of the light sources near the groove edges are more complicated [3].

As another non-peroidic structure example, it is known that when one end of a conducting nanowire is illuminated by a narrow-wavefront beam (with the electric field $\mathbf{E}_{\text{in}}$ being polarized along the wire), the other end that is not illuminated can emit light, as schematically illustrated in Fig. 8(b). The common explanation of this phenomenon is that light is transferred by CSPs on the wire surface [39]. However, it is found that this phenomenon is more pronounced in the low-frequency (e.g., terahertz) range, where most metals become nearly CSP-free perfect conductors. In particular, Wang and Mittleman [7] have experimentally demonstrated that in the terahertz range ($\lambda \sim 1$ mm), the wave modes on metallic nanowires have a dispersion trend that is opposite to that of CSPs.

In fact, according to our charge oscillation picture, the basic mechanism underlying light transfer on nanowires is very simple. In Fig. 8(b), the incident wave drives free electrons near the left input end to oscillate. The agitated electrons then propagate outside the illumination area toward the right side as a SCDW. At the other end the propagating charge wave is discontinued, giving rise to a strong charge accumulation there. The oscillation of these charges in turn emits new light near the exit end. Meanwhile, the charge wave is bounced back. When the bounced charge wave is in phase with the forward wave (and the incident wave) at the input end, Fabry–Perot resonance occurs. This is very similar to the charge movement on the slit walls in Fig. 5(a). In general, the resonant wavelength here also has a redshift. Accordingly, the standing charge wave on the wire surface is a subwavelength wave ($d < \lambda$), and based on Fig. 1, it generates little radiation loss when propagating on the wire. The Fabry–Perot resonance and the subwavelength charge patterns (proportional to the strengths of the near fields) indeed have been demonstrated both experimentally and theoretically [7,39,40]. Interestingly, our calculations show that a charge wave propagating in the unilluminated region of a flat/straight conducting surface (including the slit wall and the straight wire) is always a subwavelength SCDW, which indicates that the general SP picture elaborated in the literature, light→subwavelength SCDWs →light, is indeed correct except that the SCDWs are structured SPs and do not necessarily have the exact dispersion property of Eq. (1).

According to this picture, the conductivity of the nanowire is the dominant factor determining the propagating distance of the charge waves and the efficiency of light transfer. This explains the remarkably high transfer efficiency in the long-wavelength range where most metals are highly conducting. To further confine the near fields so as to reduce the radiation loss (caused by possible deviations of the actual charge waves from ideal subwavelength standing waves), one may activate charge waves in grooves and guide them to propagate inside the grooves. In these cases, the charge waves are channel structured SPs [41] that may have longer propagating distances.

From Fig. 8(b) it is obvious that to achieve high transfer efficiency, the diameter (vertical dimension) of the wire should be much smaller than the incident wavelength so that the agitated charge waves on the top and bottom of the wire have nearly the same phase. Otherwise, the charge waves with different phases will quickly mix together and thus offset each other outside the illuminated region, leading to a short propagation distance. This explains why light transfer is remarkable on *nanowires*. With respect to this effect, it is expected that a thin conducting slab would be more efficient since it can enhance the input coupling efficiency and reduce the electrical resistance without causing phase differences.

Note that in Fig. 2(c), when the moving electrons are stopped at the slit corners, they also have a tendency to be bounced back, similar to the moving electrons on the nanowires in Fig. 8(b) [and on the slit walls in Fig. 5(a)]. The difference in Fig. 2(c) is that the bounced back charges are suppressed by the incident electric field $\mathbf{E}_{\text{in}}$, since the driving force provided by $\mathbf{E}_{\text{in}}$ is always opposite to this tendency, while on the nanowire of Fig. 8(b), $\mathbf{E}_{\text{in}}$ is absent except at the input end.

7. SUMMARY

By numerically calculating the SCDWs on gratings, we have demonstrated that an incident wave can drive free electrons to accumulate and oscillate near the slit corners to form new light sources. These light sources then emit new wavelets. For periodic subwavelength structures ($d < \lambda$), the oscillating charges form subwavelength charge patterns (i.e., structured SPs), and the wavelets emitted from them destructively interfere with each other to form evanescent wave modes near the surfaces. Usually combined with other mechanisms (e.g., Fabry–Perot or cavity resonance, waveguiding), the structured SPs can lead to anomalous light reflection, transmission, or scattering. The structured SPs are mainly a geometrical effect and generally do not have the dispersion properties of CSPs. Note that in the literature, the SP-like modes on conductors with finite conductivity were widely assumed to be CSPs, while only those on perfectly conducting structures were believed to be spoof SPs. Here we have demonstrated that they are all structured SPs. (For transmission of acoustic waves through gratings [11,42], the counterpart of charge oscillation is the mechanical vibration of

the structured medium, particularly near the corners and edges, that emits acoustic wavelets.)

We also illustrated that the same mechanism of charge-oscillation-induced light emission and interference applies to *all structures with free electrons* (including perfect conductors and non-periodic structures). Thus, the structured SP picture represents a basic and universal mechanism of light scattering from conducting nanostructures. The guideline provided by this mechanism is that in designing novel nano-metamaterial devices, there is no CSP excitation constraint, but one needs to precisely design the geometrical parameters of the devices so as to accurately control the locations of the new light sources (including maximizing the strength of various resonance processes if involved) and their interference. Meanwhile, choosing highly conducting materials is generally another requirement for enhancing the anomalous scattering effects.

Finally, the picture presented in this paper also provides a bridge between light scattering and x ray scattering (which was already pursued in [28]) so that one can use the well-established x ray scattering mechanisms to easily and quickly understand many of the light scattering phenomena, including light scattering from non-periodic holes or slits and surface-enhanced Raman scattering.

APPENDIX A

Charge Patterns in the Diffraction Range for 1D Gratings

One might intuitively think that for normal incidence with the electric fields $\mathbf{E}_{\text{in}}$ parallel to the surface in Fig. 2(c), the charge pattern on the surface should remain the same even if $\lambda \leqslant d$, as $\mathbf{E}_{\text{in}}$ provides the same driving force. However, the free electrons on the grating surface are driven by both the incident wave and the wavelets emitted from the charges/dipoles. The latter may significantly influence the charge distribution (to achieve self-consistency of the system) when $\lambda \leqslant d$.

Figure 9 shows the charge patterns in the short-wavelength range (λ close to or below d) for the semi-infinite grating of Fig. 4(a). To understand these patterns, recall in Section 3 that the $\pm m$th-order wave components ($m>0$) above the grating can be written as $\mathbf{E}^R_m e^{-2\pi i m x/d}\exp(-iK^R_{mz}z)$ and $\mathbf{E}^R_{-m} e^{2\pi i m x/d}\exp(-iK^R_{mz}z)$, respectively, for normal incidence, where $K^R_{mz}=-2\pi(1/\lambda^2-m^2/d^2)^{1/2}$. By symmetry, $|\mathbf{E}^R_m|=|\mathbf{E}^R_{-m}|$. So these two waves form a standing wave $\boldsymbol{\mathcal{E}}_m \sin(2\pi m x/d)\exp(-iK^R_{mz}z)$. Accordingly, there exists a standing SCDW $\tilde{\rho}_{s,m}\sin(2\pi m x/d)$ corresponding to this wave on the metal surface (but discontinued in the slit gap). For $\lambda > d/m$, the propagating factor $\exp(-iK^R_{mz}z)$ of the $\boldsymbol{\mathcal{E}}_m$ wave becomes a decaying factor $\exp(-\beta_m|z|)$ along $-z$, where $\beta_m=2\pi(m^2/d^2-1/\lambda^2)^{1/2}$ is the decaying coefficient. Therefore, the $\boldsymbol{\mathcal{E}}_m$ wave is a standing evanescent wave under this condition. For $\lambda > d$ (so that $\lambda > d/m$), all the $\boldsymbol{\mathcal{E}}_m$ waves except for $m=0$ are evanescent, so they all have weak strengths. The charge pattern in Fig. 4(b) is the collective contribution from a large number of charge wave components $\tilde{\rho}_{s,m}\sin(2\pi m x/d)$, and this pattern (as well as the collective near fields of the wavelets emitted from the dipoles) coincides with the grating lattice.

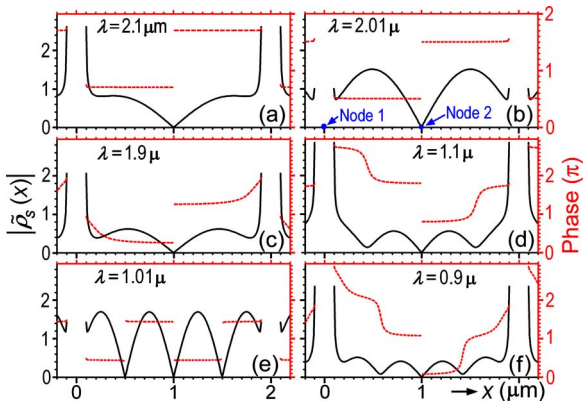

Fig. 9. (Color online) Diffraction-affected charge densities on the surface $z=0$ of a semi-infinite gold grating. $d=2$ and $W=0.2$ μm. Dashed lines are the phases. Normal incidence.

When λ decreases toward d from above, the first-order $\boldsymbol{\mathcal{E}}_1$ wave tends to be strengthened as its decaying coefficient β_1 decreases. Compared with Fig. 4(b), one can see in Fig. 9(a) the appearance of the standing SCDW component $\tilde{\rho}_{s,1}\sin(2\pi x/d)$ superimposed on the overall charge pattern that is similar to that in Fig. 4(b). (Here the charges are influenced by the near fields even if the $\boldsymbol{\mathcal{E}}_1$ wave is evanescent in the far fields.) Meanwhile, the phase shift $\Delta\varphi$ indicated in Fig. 4(b) increases to -0.28π in Fig. 9(a) because of the phase of the complex constant $\tilde{\rho}_{s,1}$.

The wavelength $\lambda=d$ corresponds to the first-order Wood's anomaly [9]. Under this condition, the wavelets emitted from two adjacent units of the grating along the horizontal direction $\theta=\pi/2$ have a phase difference of 2π according to Fig. 2(c), which corresponds to a special *Bragg diffraction condition*. Then the strength of the $\boldsymbol{\mathcal{E}}_1$ wave is maximized and becomes outstanding among other components, as can be verified by RCWA. Therefore, the charge pattern in Fig. 9(b) is dominated by the corresponding charge wave $\tilde{\rho}_{s,1}\sin(2\pi x/d)$. Here one (virtual) node (Node 1) of this standing SCDW is located at ($x=0$, $z=0$), the middle of the slit [while the other node (Node 2) is at ($x=d/2$, $z=0$)]. Node 1 significantly suppresses the charge densities at the two slit corners (because the nodes of a standing wave have zero amplitudes), so the strength of the dipole $\mathbf{P}_a$ [see Figs. 2(c) and 2(d)] is minimized. As discussed with respect to Figs. 6(c) and 6(d), when the strength of $\mathbf{P}_a$ is very small, the reflectivity should approach the Fresnel reflectivity $(1-W/d)R_f$. But note that in Fig. 9(b), the charge pattern $\tilde{\rho}_{s,1}\sin(2\pi x/d)$ also contributes *positively* to the reflectivity (equivalent to the increase of E_r' in Fig. 2(d) although it has a phase difference $\simeq\pi/2$ from the incident wave). This is the reason why in Fig. 4(a) the reflectivity at Wood's anomaly wavelengths is even higher than $(1-W/d)R_f$. Note that the first-order Wood's anomaly in Fig. 4(a) is slightly redshifted from $d=2$ to 2.01 μm.

When λ decreases to 1.9 μm in Fig. 9(c), the $\boldsymbol{\mathcal{E}}_1$ wave is still a propagating mode, but its strength decreases as the diffraction condition deviates from the Bragg condition. Meanwhile, the $\tilde{\rho}_{s,2}\sin(4\pi x/d)$ wave begins to gain

strength, which modifies the stepped phase profile. When λ further decreases toward $d/2$ in Fig. 9(d), the $\tilde{\rho}_{s,2}\sin(4\pi x/d)$ wave becomes appreciable. At the second-order Wood's anomaly $\lambda = 1.01$ μm (also slightly redshifted from $d/2$), this wave is maximized [Fig. 9(e)]. For $\lambda < d/2$ [Fig. 9(f)], the third-order wave $\tilde{\rho}_{s,3}\sin(6\pi x/d)$ begins to show strength, and so on. Thin gratings have similar properties, but these effects are meanwhile mixed with the resonance in the slits.

Overall, we may call the short-wavelength range $\lambda < 1.1d$ the *diffraction range*, where the diffraction effects of the 1D lattice appear, particularly at the Wood's anomalies. In the diffraction process, it is interesting that the incident energy is largely back-reflected rather than diffracted, although the corresponding diffracted waves become non-evanescent, as can be seen from Fig. 4(a). This is also true in Figs. 5(b) and 5(c) for the thin grating, where the reflectivity is on average very high in the entire diffraction range while the transmissivity (particularly at the Fabry-Perot resonant transmission peaks) is much lower than that in the non-diffraction range $\lambda > 1.1d$.

ACKNOWLEDGMENTS

This work was supported by grants from the National Science Foundation of China (NSFC) (Grant Nos. 10625417, 50672035, and 10874068), the Ministry of Science and Technology of China (MOST) (Grant Nos. 2004CB619005 and 2006CB921804), and partly by the Ministry of Education of China (ME) and Jiangsu Province (Grant Nos. NCET-05-0440 and BK2008012). X. R. H was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract Nos. DE-AC-02-98CH10886 and DE-AC-02-06CH11357. We thank Yong Q. Cai for helpful discussions.

REFERENCES

1. T. W. Ebbesen, H. J. Lezec, H. F. Ghaemi, T. Thio, and P. A. Wolff, "Extraordinary optical transmission through subwavelength hole arrays," Nature (London) **391**, 667–669 (1998).
2. W. L. Barnes, A. Dereux, and T. W. Ebbesen, "Surface plasmon subwavelength optics," Nature (London) **424**, 824–830 (2003).
3. C. Genet and T. W. Ebbesen, "Light in tiny holes," Nature (London) **445**, 39–46 (2007).
4. X. J. Yu and H. S. Kwok, "Optical wire-grid polarizers at oblique angles of incidence," J. Appl. Phys. **93**, 4407–4412 (2003).
5. J. R. Sambles, G. W. Bradbery, and F. Yang, "Optical excitation of surface plasmons: an introduction," Comtemp. Phys. **32**, 173–183 (1991).
6. B. Hou, H. Hang, W. Wen, C. T. Chan, and P. Sheng, "Microwave transmission through metallic hole arrays: Surface electric field measurements," Appl. Phys. Lett. **89**, 131917 (2006).
7. K. Wang and D. M. Mittleman, "Dispersion of surface plasmon polaritons on metal wires in the terahertz frequency range," Phys. Rev. Lett. **96**, 157401 (2006).
8. H. J. Lezec and T. Thio, "Diffracted evanescent wave model for enhanced and suppressed optical transmission through subwavelength hole arrays," Opt. Express **12**, 3629–3651 (2004).
9. X. R. Huang, R. W. Peng, Z. Wang, F. Gao, and S. S. Jiang, "Charge-oscillation-induced light transmission through subwavelength slits and holes," Phys. Rev. A **76**, 035802 (2007).
10. M. Sarrazin and J. P. Vigneron, "Optical properties of tungsten thin films perforated with a bidimensional array of subwavelength holes," Phys. Rev. E **68**, 016603 (2003).
11. M.-H. Lu, X.-K. Liu, L. Feng, J. Li, C.-P. Huang, Y.-F. Chen, Y.-Y. Zhu, S.-N. Zhu, and N.-B. Ming, "Extraordinary acoustic transmission through a 1D grating with very narrow apertures," Phys. Rev. Lett. **99**, 174301 (2007).
12. Cao and Lalanne, "Negative role of surface plasmons in the transmission of metallic gratings with very narrow slits," Phys. Rev. Lett. **88**, 057403 (2002).
13. G. Gay, O. Alloschery, B. Viaris de Lesegno, C. O'Dwyer, J. Weiner, and H. J. Lezec, "The optical response of nanostructured surfaces and the composite diffracted evanescent wave model," Nat. Phys. **2**, 262–267 (2006).
14. P. Lalanne and J. P. Hugonin, "Interaction between optical nano-objects at metallo-dielectric interfaces," Nat. Phys. **2**, 551–556 (2006).
15. N. Garcia and M. Nieto-Vesperinas, "Theory of electromagnetic wave transmission through metallic gratings of subwavelength slits," J. Opt. A, Pure Appl. Opt. **9**, 490–495 (2007).
16. N. Garcia and M. Bai, "Theory of transmission of light by sub-wavelength cylindrical holes in metallic films," Opt. Express **14**, 10028–10042 (2006).
17. B. C. Cullity and S. R. Stock, *Elements of X-Ray Diffraction*, 3rd ed. (Prentice Hall, 2001), p. 123.
18. http://en.wikipedia.org/wiki/Thomson_scattering.
19. M. Born and E. Wolf, *Principles of Optics* (Cambridge Univ. Press, 1999).
20. http://en.wikipedia.org/wiki/Dipole.
21. Y. Xie, A. R. Zakharian, J. V. Moloney, and M. Mansuripur, "Transmission of light through periodic arrays of subwavelength slits in metallic hosts," Opt. Express **14**, 6400–6413 (2006).
22. B. Ung and Y. Sheng, "Interference of surface waves in a metallic nanoslit," Opt. Express **15**, 1182–1190 (2007).
23. G. Lévêque, O. J. F. Martin, and J. Weiner, "Transient behavior of surface plasmon polaritons scattered at a subwavelength groove," Phys. Rev. B **76**, 155418 (2007).
24. D. Pacifici, H. J. Lezec, H. A. Atwater, and J. Weiner, "Quantitative determination of optical transmission through subwavelength slit arrays in Ag films: Role of surface wave interference and local coupling between adjacent slits," Phys. Rev. B **77**, 115411 (2008).
25. J. Weiner and F. D. Nunes, "High-frequency response of subwavelength-structured metals in the petahertz domain," Opt. Express **16**, 21256–21270 (2008).
26. G. Gay, O. Alloschery, B. V. de Lesegno, J. Weiner, and H. J. Lezec, "Surface wave generation and propagation on metallic subwavelength structures measured by far-field interferometry," Phys. Rev. Lett. **96**, 213901 (2006).
27. J. B. Pendry, L. Martín-Moreno, and F. J. Garcia-Vidal, "Mimicking surface plasmons with structured surfaces," Science **305**, 847–848 (2004).
28. M. M. J. Treacy, "Dynamical diffraction explanation of the anomalous transmission of light through metallic gratings," Phys. Rev. B **66**, 195105 (2002).
29. M. G. Moharam, E. B. Grann, D. A. Pommet, and T. K. Gaylord, "Formulation for stable and efficient implementation of the rigorous coupled-wave analysis of binary gratings," J. Opt. Soc. Am. A **12**, 1068–1076 (1995).
30. P. Lalanne and G. M. Morris, "Highly improved convergence of the coupled-wave method for TM polarization," J. Opt. Soc. Am. A **13**, 779–784 (1996).
31. *Handbook of Optical Constants and Solids*, E. D. Palik, ed. (Academic, 1985).
32. S. J. McNab and R. J. Blaikie, "Contrast in the evanescent near field of $\lambda/20$ period gratings for photolithography," Appl. Opt. **39**, 20–25 (2000).
33. J. R. Suckling, A. P. Hibbins, M. J. Lockyear, T. W. Preist, J. R. Sambles, and C. R. Lawrence, "Finite conductance

governs the resonance transmission of thin metal slits at microwave frequencies," Phys. Rev. Lett. **92**, 147401 (2004).
34. H. E. Went, A. P. Hibbins, J. R. Sambles, C. R. Lawrence, and A. P. Crick, "Selective transmission through very deep zero-order metallic gratings at microwave frequencies," Appl. Phys. Lett. **77**, 2789–2791 (2000).
35. J. A. Porto, F. J. García-Vidal, and J. B. Pendry, "Transmission resonances on metallic gratings with very narrow slits," Phys. Rev. Lett. **83**, 2845–2848 (1999).
36. Z. J. Zhang, R. W. Peng, Z. Wang, F. Gao, X. R. Huang, W. H. Sun, Q. J. Wang, and M. Wang, "Plasmonic antenna array at optical frequency made by nanoapertures," Appl. Phys. Lett. **93**, 171110 (2008).
37. T. Matsui, A. Agrawal, A. Nahata, and Z. V. Vardeny, "Transmission resonances through aperiodic arrays of subwavelength apertures," Nature (London) **446**, 517 (2007).
38. F. J. García-Vidal, H. J. Lezec, T. W. Ebbesen, and L. Martín-Moreno, "Multiple paths to enhance optical transmission through a single subwavelength slit," Phys. Rev. Lett. **90**, 213901 (2003).
39. H. Ditlbacher, A. Hohenau, D. Wagner, U. Kreibig, M. Rogers, F. Hofer, F. R. Aussenegg, and J. R. Krenn, "Silver nanowires as surface plasmon resonators," Phys. Rev. Lett. **95**, 257403 (2005).
40. T. Laroche and C. Girard, "Near-field optical properties of single plasmonic nanowires," Appl. Phys. Lett. **89**, 233229 (2006).
41. S. I. Bozhevolnyi, V. S. Volkov, E. Devaux, J.-Y. Laluet, and T. W. Ebbesen, "Channel plasmon subwavelength waveguide components including interferometers and ring resonators," Nature (London) **440**, 508–511 (2006).
42. J. Christensen, L. Martin-Moreno, and F. J. Garcia-Vidal, "Theory of resonant acoustic transmission through subwavelength apertures," Phys. Rev. Lett. **101**, 014301 (2008).