Modeling Electromagnetic Wave Interactions with Nanostructures in Plasmonic Devices

August 28, 2011

Industrial Representatives: B. Pindzola (TIAX), M. Rona (TIAX)

Faculty: R. Braun (Delaware), C. Raymond (Delaware), G. Richardson (Southhampton), D. Schwendeman (RPI)

Students: D. Timmers (UBC), W. Woodruff (DSU)

Summary Report Prepared by: C. Raymond & G. Richardson

1 Introduction

The problem brought to the 2010 MPI Workshop by TIAX asked for the formulation of a mathematical model for the effect of interaction of light with metallic nanostructures, specifically the effect of surface features on the intensity of the electromagnetic field, when a metallic surface is illuminated by laser light of a given wavelength and the scale of the metallic structures is commensurate with the wavelength of the light. Devices based on this type of configuration are used in photovoltaics, catalysis, nonlinear optics, and sensors, just to name a few applications. The operation of many of these devices depends on effects which are proportional to the amplitude of the electromagnetic field in the vicinity of the metallic surface.

With a solution to this problem in hand, it then becomes possible to consider optimization and design. Two examples: If a given structure is imposed, what wavelength leads to the largest field in the vicinity of the

surface? Given a particular wavelength, is it possible to design a surface structure to maximize the intensity of the field in the vicinity of the surface?

We have partial answers to both of these questions. If we assume that a large field strength close to the metal surface is associated with the existence of plasmons whose properties are not strongly affected by the surface features, then a simple resonance argument gives a way to pick the wavelength of laser light which should most strongly couple to the plasmons, and (perhaps!) therefore lead to the greatest field enhancement near the metal surface.

The second question is more difficult. We have systematically derived a reduced problem for the field strength, but unfortunately this problem is still formidable. The asymptotic regime of interest to TIAX is one in which the wavelength of the incident light is of the same order as the size of the surface features, so simplifications based on the wavelength being large or small in comparison to the surface features are excluded. The main simplification available is based on the fact that the conductivity of the metal is large in comparison to that of the air: the leading order problem is essentially that of scattering from a perfect conductor. We suggest a few approaches that can be applied to this reduced problem, but which do not seem to be commonly used in the plasmonics literature.

The structure of the report is as follows:

In section 2, we carefully and systematically set up the governing equations for the problem, in the form most commonly encountered in the plasmonics literature, and contrast this with the form commonly presented in textbooks on electromagnetism (e.g. [10], [23]). While we eventually make many simplifying assumptions, these are pointed out as they are made. If any of these assumptions must be dropped in a later analysis, it will be clear where changes must be made.

In section 3, we solve the electromagnetic problem for surface waves, and construct the plasmon solutions. In the limit of large conductivity, these have a boundary layer structure near the metal-air interface which should be similar to that of the full electromagnetic problem. The problem of exciting plasmons with incident light is discussed, and a simple criterion for choosing the wavelength of the incident light to maximize the coupling to surface plasmons is suggested.

In section 4, we non-dimensionalize the governing equations and derive the leading order problem for the electric field in the air, in the asymptotic

limit of large conductivity of the metal. It is equivalent to the problem of scattering from a perfect conductor.

In section 5, we outline some potential numerical approaches to the scattering problem derived in section 4.

Finally, we conclude in section 6 with a summary and a few comments about further problems of interest.

2 Problem Set Up

2.1 Geometry of the Problem

The geometry of the problem is presented schematically in Figure 1. We assume that there are three regions: Ω_1 is the air above the device, Ω_2 is a thin nanostructured metal layer, and Ω_3 is a dielectric substrate. The metal layer is supposed to periodic in the lateral directions. The region Ω_2 might be connected, or it might consist of a periodic array of isolated metal islands on the dielectric substrate. The figure shows a cross section for the case when the metal layer is assumed continuous. If there is no variation laterally in the y direction (i.e. perpendicular to the page), this would correspond to a grating type geometry.

2.2 Microscopic Version of Maxwell's Equations

If we were willing to explicitly keep track of all charges in our system, it would in principle be possible to use the so-called microscopic version of Maxwell's equations:

$$\nabla \cdot \mathbf{e} = \frac{\rho}{\varepsilon_0}, \tag{1}$$

$$\nabla \cdot \mathbf{b} = 0, \tag{2}$$

$$\nabla \times \mathbf{e} = -\frac{\partial \mathbf{b}}{\partial t}, \tag{3}$$

$$\nabla \times \mathbf{b} = \mu_0 \mathbf{j} + \mu_0 \varepsilon_0 \frac{\partial \mathbf{e}}{\partial t}. \tag{4}$$

The unknowns in these equations are the electric field $\mathbf{e}$, the magnetic field $\mathbf{b}$, the charge density ρ, and the current density $\mathbf{j}$. The equations include two constant parameters as well: ε_0 is the permittivity of free space, and μ_0 is the permeability of free space.

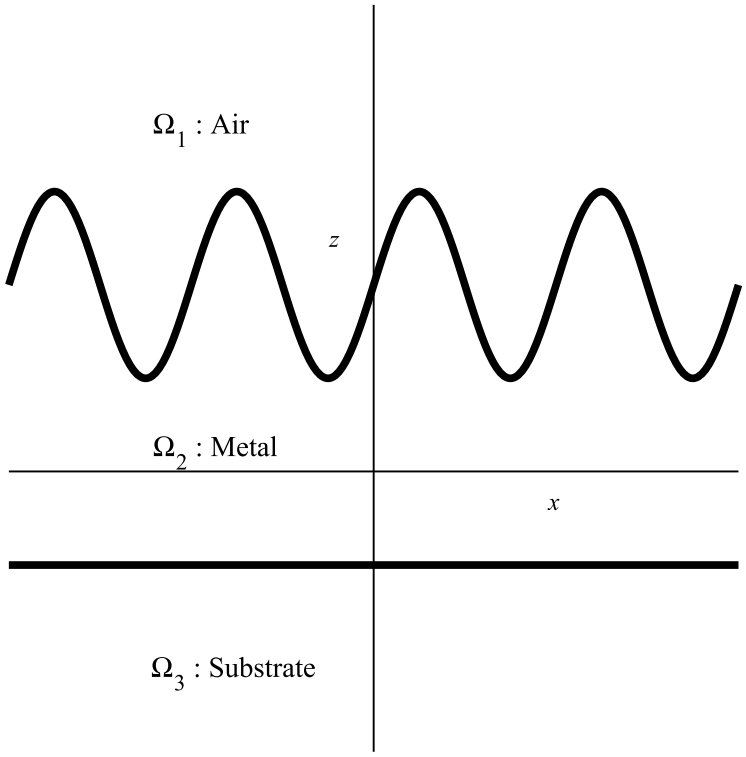

Figure 1: Geometry of the Problem. The x and z axes are shown. Since incident electromagnetic waves are supplied in region Ω_1, it is convenient to assume that z increases downward. The y axis (not shown) should point out of the page by the right hand rule.

The continuity equation is a consequence of the equations above. If we take the partial derivative with respect to t of equation (1), then use equation (4) and a vector calculus identity ($\nabla \cdot (\nabla \times \mathbf{b}) = 0$), we obtain

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0. \tag{5}$$

Finally, we have the Lorentz force law, which gives the force $\mathbf{f}$ exerted on a particle with charge q and velocity $\mathbf{v}$ by the $\mathbf{e}$ and $\mathbf{b}$ fields:

$$\mathbf{f} = q\left(\mathbf{e} + \mathbf{v} \times \mathbf{b}\right). \tag{6}$$

2.3 Macroscopic version of Maxwell's Equations

Rather than keeping track of all the individual charges, and tracking the variation of the fields on very small space and time scales, it is customary to use a macroscopic version of Maxwell's equations, which is obtained by an averaging procedure [10]. The resulting equations as usually presented in electrodynamics textbooks (e.g. [10]) are:

$$\nabla \cdot \mathbf{D}_s = \rho_f, \tag{7}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{8}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{9}$$

$$\nabla \times \mathbf{H} = \mathbf{J}_f + \frac{\partial \mathbf{D}_s}{\partial t}. \tag{10}$$

Here $\mathbf{E}$ and $\mathbf{B}$ are the averaged versions of $\mathbf{e}$ and $\mathbf{b}$, and by definition

$$\mathbf{D}_s = \varepsilon_0 \mathbf{E} + \mathbf{P}_s, \tag{11}$$

and

$$\mathbf{H} = \frac{1}{\mu_0}\mathbf{B} - \mathbf{M}. \tag{12}$$

Here $\mathbf{D}_s$ is the so-called 'static' electric displacement, and $\mathbf{H}$ goes by a number of names. The other two fields that have been introduced in these definitions are the "static" polarization $\mathbf{P}_s$ and the magnetization $\mathbf{M}$. The term static, applied to the electric displacement and polarization, refers to the fact that these variables account for effects due to electrons which are strongly

bound to their atoms or molecules, and which are not free to move throughout the material. Finally, we have the so-called free charge density ρ_f and the free current density $\mathbf{J}_f$; here 'free' refers to the fact that these quantities are associated with relatively mobile conduction electrons which are not strongly bound to any particular atom or molecule. Essentially the same argument that yields (5) leads to the continuity equation for the free charges:

$$\frac{\partial \rho_f}{\partial t} + \nabla \cdot \mathbf{J}_f = 0. \tag{13}$$

Equations (7) – (13), while familiar, are not always the version of Maxwell's equations used in references studying plasmons and other surface modes [13], [14], [19], [21]. The distinction between 'free' and 'bound' charges is more or less artificial, and there are other possible formalisms [2], [15]. It is more common (though not universal!) in the plasmon literature to make a distinction between 'internal' charges and currents, whether they are associated with free or bound electrons, and 'external' charges and currents, due to externally imposed sources. The choice is entirely arbitrary, but it is necessary to know which convention is being used, since the meanings of ρ, $\mathbf{J}$, $\mathbf{D}$, and $\mathbf{P}$, and the proper form of the constitutive equations, all depend on the choice that is made.

We will use the following version of Maxwell's equations, which seems to be consistent with much of the literature on plasmons [13], [14], [19], [21]:

$$\nabla \cdot \mathbf{D} = \rho_e, \tag{14}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{15}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{16}$$

$$\nabla \times \mathbf{H} = \mathbf{J}_e + \frac{\partial \mathbf{D}}{\partial t}. \tag{17}$$

As above, $\mathbf{E}$ and $\mathbf{B}$ are the averaged versions of $\mathbf{e}$ and $\mathbf{b}$. Rather than (11), we use

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P}, \tag{18}$$

but we again use (12). Finally, there is again a version of the continuity equation, now for the 'external' (i.e. source) charge density and current:

$$\frac{\partial \rho_e}{\partial t} + \nabla \cdot \mathbf{J}_e = 0. \tag{19}$$

A major advantage of this version of Maxwell's equations is that we may immediately set $\rho_e = 0$ and $\mathbf{J}_e = \mathbf{0}$ if there are no externally imposed charges or currents, which is typically the case of interest for applications involving plasmons. This simplifies both the field equations and (see below) the boundary conditions.

2.4 Constitutive Equations

In each region, we in general have five constitutive equations:

$$\mathbf{D} = \varepsilon \mathbf{E}, \tag{20}$$

$$\mathbf{B} = \mu \mathbf{H}, \tag{21}$$

$$\mathbf{P} = \alpha \mathbf{E}, \tag{22}$$

$$\mathbf{M} = \chi \mathbf{H}, \tag{23}$$

$$\mathbf{J} = \sigma \mathbf{E}. \tag{24}$$

Here ε is known as the dielectric function, μ is the magnetic permeability, α is the dielectric polarizability, χ is the magnetic polarizability, and σ is the conductivity (it often called the optical conductivity in the current context). Notice that the internal current density $\mathbf{J}$ does not appear explicitly in (14) – (17).

The constitutive equations (20) – (24) all assume that the materials respond linearly; this assumption may not be reasonable at large field strength. For example, reference [16] devotes a chapter to surface modes at metal interfaces when the nonlinear response is important. Both [6] and [12] have articles related to nonlinear media, and almost every article in [22] is devoted to nonlinear effects.

Furthermore, we have for simplicity written these equations as if the material parameters are constants. In fact, these equations really should be written as convolutions, even for homogeneous and isotropic systems [10]; the simple product form used in (20) – (24) actually only holds once Fourier transforms in space and time are taken, or if we agree to work with only one normal mode, say proportional to $e^{-i\omega t + i\mathbf{k}\cdot\mathbf{r}}$, at a time. This is certainly reasonable for optical studies using monochromatic laser light if a linear response is expected. In this case, each material parameter will in general depend on both the frequency ω and the wave vector $\mathbf{k}$ of the mode under consideration, and will in general be complex. In many circumstances it is

reasonable to neglect the dependence of the material parameters on the wave vector and just retain the dependence on frequency; [10] briefly discusses certain cases when this approximation is not valid, such as when the mean free path of conduction electrons is on the order of or larger than the skin depth. As pointed out in [14], the dependence of the material parameters on the wave vector should also be retained when analyzing problems that involve surface modes, if the interface associated with these modes includes sufficiently small features. Another situation where the wave vector dependence should, strictly speaking, be retained is that of an emitter close to a metal surface. Even though this is exactly the situation of interest in surface enhanced raman scattering, this aspect of the problem is typically ignored by SERS researchers because it substantially increases the complexity of the analysis (see [9] for brief discussion). For simplicity, and since our interest here is in the general problem of the effect of a metal-dielectric interface on the electromagnetic field rather than in the details relevant to particular applications, we will ignore any dependence on the wave vector.

Magnetic effects are important for many plasmonic systems. Reference [16] devotes a chapter to the influence of magnetic responses on surface modes and almost the entirety of [19] is devoted to modeling systems where the magnetic response is important. We will however for simplicity ignore any magnetic response. We therefore assume that $\mu = \mu_0$ and that $\chi = 0$ in all regions. Also, any direct dependence of $\mathbf{J}$ on $\mathbf{B}$ has already been suppressed by our use of (24), even though such a dependence is suggested by the Lorentz force law (6). This assumption is usually quite reasonable for metallic conductors.

The remaining material parameters ε, α, and σ are not in general independent. In order to continue, we need a model for how electrons in a material respond to an electric field. The Drude model describes how free electrons respond to the electric field, while the Drude-Lorentz model explains how bound electrons respond to the electric field. While these models are based on classical mechanics, a more accurate quantum mechanical treatment produces similar results; [10] has a brief description, and [25] has a detailed and very clear discussion of the classical mechanics models as well as the correct quantum mechanical treatments; [13] also has a good short discussion.

Lorentz suggested that a bound electron responding to an electric field would have the equation of motion

$$m\mathbf{x}_{tt} + m\gamma\mathbf{x}_t + m\omega_0^2\mathbf{x} = -e\mathbf{E}, \tag{25}$$

which is just Newton's second law, stating that mass times acceleration equals the force on a particle, with all terms except the external force due to the electric field collected on one side of the equation. The first term on the left hand side is just the mass times acceleration; the second is a phenomenological frictional force proportional to the velocity; the third is the restoring force which binds the electron to its nucleus. Here m is the mass of an electron and $-e$ is the charge on the electron. The spatial extent of an atom is miniscule, so we assume that any spatial dependence of $\mathbf{E}$ can be ignored. If we assume that the field varies harmonically, with time dependence $e^{-i\omega t}$, then we can easily solve for the steady state response,

$$\mathbf{x} = \frac{-e/m}{\omega_0^2 - \omega^2 - i\gamma\omega}\mathbf{E}, \tag{26}$$

and the resulting polarization due to this single electron,

$$\mathbf{p} = -e\mathbf{x} = \frac{e^2/m}{\omega_0^2 - \omega^2 - i\gamma\omega}\mathbf{E}. \tag{27}$$

Assuming that there is one such electron per atom, and N such atoms per unit volume, we have the total polarization density

$$\mathbf{P} = \frac{Ne^2/m}{\omega_0^2 - \omega^2 - i\gamma\omega}\mathbf{E}. \tag{28}$$

Therefore, in this case we have:

$$\alpha(\omega) = \frac{Ne^2/m}{\omega_0^2 - \omega^2 - i\gamma\omega}. \tag{29}$$

If there are instead several classes of electrons, with say f_j electrons per atom of class j, and that each class has its own frequency ω_j and damping constant γ_j, then we instead obtain

$$\alpha(\omega) = \frac{Ne^2}{m}\sum_j \frac{f_j}{\omega_j^2 - \omega^2 - i\gamma_j\omega}, \tag{30}$$

where $\sum_j = Z$, with Z the total number of electrons per atom.

Once $\alpha(\omega)$ is determined, we note that (18), (20), and (22) together imply that

$$\varepsilon(\omega) = \varepsilon_0 + \alpha(\omega). \tag{31}$$

However, there seems to be some disagreement in the literature as to how best to model $\alpha(\omega)$ and $\epsilon(\omega)$ for the metals of interest and the frequencies of interest in plasmonics. The electronic structure of the noble metals of interest in plasmonic applications is such that the conduction electrons experience a highly polarized background formed by the lattice of metal ions, leading to a dielectric function of the form

$$\varepsilon(\omega) = \varepsilon_\infty + \alpha(\omega), \tag{32}$$

but there seems to be some disagreement about the correct values of ε_∞. Also, if the frequency of the electric field is such that it can cause electrons to jump from one conduction band to another, each such transition will modify the dielectric function in ways that may or may not be well modeled by the inclusion of additional Lorentz terms in $\alpha(\omega)$ or adjustments to ε_∞. Furthermore, even the choice of the best model for $\varepsilon(\omega)$ is not agreed upon; [14] cites experiments to argue that a simple Drude model is unacceptable for silver, while [13] cites some of the same references, as well as a number of more recent ones, to argue that the simple Drude model actually does a very good job of modeling the dielectric function of silver for the frequencies of interest in plasmonic applications. In an appendix, [13] shows explicitly how to construct model dielectric functions for both silver and gold by fitting to experimental data collected from several references. This appendix also includes a discussion of some of the sources of experimental variability.

Lorentz's analysis was actually an extension of Drude's original model for conduction electrons. Drude used (25) with ω_0 set to zero (i.e. with no restoring force) to model the response of conduction electrons in a metal to an imposed electric field. One of the results of Drude's analysis is a derivation of Ohm's law, equation (24), and a representation of the dependence of σ on ω:

$$\sigma(\omega) = \frac{Ne^2/m}{-i\omega + \gamma}, \tag{33}$$

where here N is the number of conduction electrons per unit volume. The derivation is similar to that for $\alpha(\omega)$; where the polarization $\mathbf{P}$ is proportional to $\mathbf{x}$, the current $\mathbf{J}_f$ is proportional to $\mathbf{v} = \mathbf{x}_t$. Also note that (33) assumes a time dependence of the form $e^{-i\omega t}$. More generally, in the current context, all internal electrons should be treated on the same basis, which means that

$$\sigma(\omega) = -i\omega\alpha(\omega), \tag{34}$$

or equivalently,

$$\mathbf{J} = \frac{\partial \mathbf{P}}{\partial t}. \tag{35}$$

A simple model for a conducting material is to use (29) with $\omega_0 = 0$. A simple model for a nonconducting material is to use (29), with $\omega_0 \neq 0$. It is possible to model the response of a material with both free and bound electrons by using (30) with one of the ω_j's set to zero to account for the conduction electrons, and the other terms accounting for the bound electrons. This is an advantage of using (14) – (17) rather than (7) – (10): the response of both free and bound electrons in the material is treated on the same basis, and both are accounted for in the dielectric function ε. Even though $\mathbf{J}_f$ drops out of the governing equations, this does not mean that we are ignoring any conduction current. It gets lumped into $\mathbf{D}$ and $\mathbf{P}$ when we use these instead of of $\mathbf{D}_s$ and $\mathbf{P}_s$, and can be recovered if desired from (24).

2.5 Summary: Governing Equations and Constitutive Equations in the Various Regions

Assuming that there are no external charges or currents in the system, and assuming no magnetic effects, we have the following subproblems.

2.5.1 Dielectric Regions

In regions Ω_1 (air) and Ω_3 (substrate), we have:

$$\nabla \cdot \mathbf{D} = 0, \tag{36}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{37}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{38}$$

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t}. \tag{39}$$

We use constants for the material parameters ($i = 1, 3$):

$$\varepsilon = \varepsilon_i, \qquad \mu_i = \mu_0, \qquad \alpha_i = \varepsilon_i - \varepsilon_0, \qquad \chi_i = 0, \qquad \sigma_i = 0. \tag{40}$$

Therefore, we have

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P} \tag{41}$$

and

$$\mathbf{B} = \mu_0 \mathbf{H}, \tag{42}$$

as well as

$$\mathbf{D} = \varepsilon_i \mathbf{E}, \tag{43}$$

$$\mathbf{P} = (\varepsilon_i - \varepsilon_0)\mathbf{E}, \tag{44}$$

$$\mathbf{J} = \mathbf{0}. \tag{45}$$

2.5.2 Conductive Region

In region Ω_2 (metal) we have:

$$\nabla \cdot \mathbf{D} = 0, \tag{46}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{47}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{48}$$

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t}. \tag{49}$$

We again use $\mu_2 = \mu_0$ and $\chi_2 = 0$, and use a simple Drude model for the other parameters, without introducing ϵ_∞:

$$\varepsilon_2(\omega) = \varepsilon_0 - \frac{Ne^2/m}{\omega^2 + i\gamma\omega}. \tag{50}$$

$$\alpha_2(\omega) = -\frac{Ne^2/m}{\omega^2 + i\gamma\omega}, \tag{51}$$

$$\sigma_2(\omega) = \frac{Ne^2/m}{-i\omega + \gamma}, \tag{52}$$

We emphasize again that these three parameters are not independent:

$$\varepsilon_2(\omega) = \varepsilon_0 + \alpha_2(\omega) = \varepsilon_0 + \frac{i\sigma_2(\omega)}{\omega}. \tag{53}$$

For the same reason that $\mathbf{J} = \partial \mathbf{P}/\partial t$ above, rather than specifying $\alpha(\omega)$ directly, we could alternatively use

$$\frac{m}{Ne^2}\left(\frac{\partial^2 \mathbf{P}}{\partial t^2} + \gamma \frac{\partial \mathbf{P}}{\partial t}\right) = \mathbf{E}. \tag{54}$$

Finally, we have

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P} \tag{55}$$

and

$$\mathbf{B} = \mu_0 \mathbf{H}, \tag{56}$$

as well as

$$\mathbf{D} = \varepsilon_2(\omega)\mathbf{E}, \tag{57}$$

$$\mathbf{P} = \alpha_2(\omega)\mathbf{E}, \tag{58}$$

$$\mathbf{J} = \sigma_2(\omega)\mathbf{E}. \tag{59}$$

2.6 Matching Conditions

The choice to use $\mathbf{D}$ and $\mathbf{P}$ rather than $\mathbf{D}_s$ and $\mathbf{P}_s$ means that we have the following matching conditions across any interface which separates region Ω_i from Ω_j, $i \neq j$:

$$[\mathbf{n} \cdot \mathbf{D}] = 0, \tag{60}$$

$$[\mathbf{n} \cdot \mathbf{B}] = 0, \tag{61}$$

$$[\mathbf{n} \times \mathbf{E}] = \mathbf{0}, \tag{62}$$

$$[\mathbf{n} \times \mathbf{H}] = \mathbf{0}. \tag{63}$$

In these equations, $\mathbf{n}$ is a normal vector to the interface. The square brackets represent the jump in the enclosed quantity across the interface.

At interfaces between the conductor and a dielectric region, where there is a discontinuity in the internal current density $\mathbf{J}$, we expect a surface charge density ρ_s to build up. If $\mathbf{D}_s$ had been used instead of $\mathbf{D}$, then this surface charge would show up explicitly in the jump conditions and the analogue of (60) would not be homogeneous. This is not an issue at a boundary separating dielectric regions, since we have assumed that the current vanishes in both.

3 Surface Plasmons at a Metal-Dielectric Interface

In this section, we will consider a simplified version of the problem. We will use the configuration in figure 1, but with only two layers and one interface: air above the interface, and metal below.

We will sketch the derivation of surface modes in this situation when the interface is planar. These modes are called surface plasmons or surface plasmon-polaritons in the literature. Reference [21] extensively discusses general features of such modes and constructs them in a number of cases; [14] and [17] give physically oriented discussions, with [17] in particular containing an extensive discussion of experiments; [13] gives a very detailed discussion both of the general features of such modes and of the silver-air and gold-air interfaces; [1] concentrates on similar modes in multilayer systems; and [19] briefly considers the simple set up discussed here, but also has extensive discussions of analogous modes in more complicated geometries, and for systems where magnetic effects are important.

We next consider the problem of exciting these modes with incident electromagnetic waves; [13] and [19] are good references for the theory, and [17] collects a great deal of experimental data related to this problem. Direct illumination of a flat interface is ineffective for this purpose; we mention some of the strategies that are actually used.

In the case of a single metal-dielectric surface with a grating type geometry, there is a relatively simple, basically geometrical theory for how to most efficiently pump energy into the surface plasmon modes. Figuring out the actual enhancement of the field close to the interface, and how the enhancement depends on the actual shape of the grating, is a different (and more difficult) problem.

3.1 The Electromagnetic Problem for Surface Modes

We first note that equations (36) and (46) imply that $\nabla \cdot \mathbf{E}$ in all regions. Therefore, using the vector identity

$$\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A},$$

we find that both $\mathbf{E}$ and $\mathbf{B}$ satisfy vector wave equations in the two domains. In domain Ω_1 (the air) we have

$$\nabla^2 \mathbf{E} = \frac{1}{c^2}\frac{\varepsilon_1}{\varepsilon_0}\frac{\partial^2 \mathbf{E}}{\partial t^2}, \qquad \nabla^2 \mathbf{B} = \frac{1}{c^2}\frac{\varepsilon_1}{\varepsilon_0}\frac{\partial^2 \mathbf{B}}{\partial t^2}, \tag{64}$$

and in domain Ω_2 (the metal) we have

$$\nabla^2 \mathbf{E} = \frac{1}{c^2}\frac{\varepsilon_2}{\varepsilon_0}\frac{\partial^2 \mathbf{E}}{\partial t^2}, \qquad \nabla^2 \mathbf{B} = \frac{1}{c^2}\frac{\varepsilon_2}{\varepsilon_0}\frac{\partial^2 \mathbf{B}}{\partial t^2}. \tag{65}$$

The jump conditions imply that all components of $\mathbf{B}$ are continuous across the interface; that the tangential components of $\mathbf{E}$ are continuous across the interface; and that there is in general a jump in the normal component of $\mathbf{E}$ across the interface, because there is in general a discontinuity in the dielectric function across the interface.

It is possible to show that there are no surface modes with TE polarization; only modes with TM polarization are possible unless the permeability μ varies between regions, which we have specifically excluded [2]. Since the problem is linear and homogeneous, the amplitude is arbitrary. Without any loss of generality, we find that the surface modes must take the form:

$$\begin{cases} \mathbf{E}_1 &= \dfrac{B_y}{\omega\mu_0\varepsilon_1}\,(k_{1z}, 0, -k_x)\exp(-i\omega t + ik_x x + ik_{1z} z), \\ \mathbf{E}_2 &= \dfrac{B_y}{\omega\mu_0\varepsilon_2}\,(k_{2z}, 0, -k_x)\exp(-i\omega t + ik_x x + ik_{2z} z), \\ \mathbf{B}_1 &= (0, B_y, 0)\exp(-i\omega t + ik_x x + ik_{1z} z), \\ \mathbf{B}_2 &= (0, B_y, 0)\exp(-i\omega t + ik_x x + ik_{2z} z), \end{cases} \tag{66}$$

where $\mathbf{E}_1$ means $\mathbf{E}$ in Ω_1, and so on. The form that we have assumed satisfies all the field equations as long as the following two conditions hold:

$$k_x^2 + k_{1z}^2 = \frac{\omega^2}{c^2}\frac{\varepsilon_1}{\varepsilon_0}, \qquad k_x^2 + k_{2z}^2 = \frac{\omega^2}{c^2}\frac{\varepsilon_2}{\varepsilon_0}. \tag{67}$$

The matching conditions on $\mathbf{B}$ and $\mathbf{H}$ are automatically satisfied, as is the matching condition on the normal component of $\mathbf{E}$. Continuity of the tangential components of $\mathbf{E}$ gives the final relationship that we need:

$$\frac{k_{1z}}{\varepsilon_1} = \frac{k_{2z}}{\varepsilon_2}. \tag{68}$$

Eliminating k_{1z} and k_{2z} among (67) and (68), we find the dispersion relation:

$$k_x^2 = \frac{\omega^2}{c^2}\frac{\varepsilon_1\varepsilon_2}{\varepsilon_0(\varepsilon_1 + \varepsilon_2)}. \tag{69}$$

The two possible solutions for k_x are physically equivalent, and by convention we can use

$$k_x = \frac{\omega}{c}\sqrt{\frac{\varepsilon_1\varepsilon_2}{\varepsilon_0(\varepsilon_1 + \varepsilon_2)}}, \tag{70}$$

meaning that we look for waves propagating in the positive x direction; the other choice of sign for k_x would correspond to waves propagating in the $-x$ direction. Once k_x is determined, we can work out k_{1z} and k_{2z}. Because $\varepsilon_2(\omega)$ is complex, it is somewhat involved to work out all the possibilities, but generically k_x will be complex for real metals.

It is helpful to introduce the non-dimensionalized versions of the dielectric functions,

$$\hat{\varepsilon}_i = \frac{\varepsilon_i}{\varepsilon_0}. \tag{71}$$

We assume for simplicity that $\hat{\varepsilon}_1 = 1$, and write

$$\hat{\varepsilon}_2 = (1 - \beta) + i\beta\Gamma, \tag{72}$$

where

$$\beta = \frac{\omega_p^2}{\omega^2 + \gamma^2}, \qquad \Gamma = \frac{\gamma}{\omega}, \qquad \omega_p^2 = \frac{Ne^2}{m\varepsilon_0}. \tag{73}$$

The plasma frequency ω_p is the natural frequency associated with oscillations of the plasma of free electrons in the metal. Recall that γ is a phenomenological damping constant introduced when modeling the response of electrons in the metal to an applied electric field. For the metals of interest in plasmonics applications, and the frequencies of interest, several references ([13], [21]) suggest that at least in certain situations, it is true that $\omega_p \gg \omega \gg \gamma$, so β is a large parameter and Γ is a small parameter. Some manipulation yields

$$\begin{aligned} k_x^2 &= \frac{\omega^2}{c^2} \frac{(1 - \beta^{-1})(1 - 2\beta^{-1}) + \Gamma^2 + i\Gamma\beta^{-1}}{(1 - 2\beta^{-1})^2 + \Gamma^2} \\ &\sim \frac{\omega^2}{c^2} \left((1 + \cdots) + i \left(\Gamma\beta^{-1} + \cdots \right) \right), \end{aligned} \tag{74}$$

which implies that for our simple Drude metal, k_x will have a small positive imaginary part. This means that the amplitude of the mode decays slowly in the propagation direction along the interface, i.e. the mode is a so-called pseudo-propagating wave. In the limit $\Gamma = 0$, it propagates without a decrease in amplitude. Further manipulation yields an expression for k_{1z}^2 :

$$k_{1z}^2 = -\frac{\omega^2}{c^2} \left(\frac{1 - 2\beta^{-1} + i\Gamma}{(1 - 2\beta^{-1})^2 + \Gamma^2} \right) \beta^{-1} \tag{75}$$

which means that k_{1z}^2 is very nearly real and negative. Since the small imaginary part is negative, if we assume the standard branch cut along the negative real axis for the square root function, then $\sqrt{k_{1z}^2}$ will be very nearly negative imaginary, with a small positive real part, and thus $\mathbf{E}_1$ and $\mathbf{B}_1$ will decay exponentially away from the interface (i.e. in the negative z direction). To leading order, assuming that β is large and Γ is small, we have

$$k_{1z} \sim -i\beta^{-1/2}\frac{\omega}{c}, \tag{76}$$

and so this decay is slow.

Similarly we can investigate k_{2z}; we find that

$$k_{2z}^2 = \frac{\omega^2}{c^2}\left((1-\beta) + i\beta\Gamma - \frac{(1-\beta^{-1})(1-2\beta^{-1}) + \Gamma^2 + i\Gamma\beta^{-1}}{(1-2\beta^{-1})^2 + \Gamma^2}\right). \tag{77}$$

For large β and small Γ, k_{2z}^2 is very nearly negative real, with a small positive imaginary part, so $k_2 z$ will be very nearly positive imaginary. To leading order, we have

$$k_{2z} \sim i\beta^{1/2}\frac{\omega}{c}, \tag{78}$$

and the solution decays exponentially away from the interface (i.e. in the positive z direction). Note that the solution decays at a much faster rate in the positive z direction (into the metal) than in the negative z direction (into the air). This is the so-called skin effect for a conductor.

We are interested in the possibility of using the air-metal interface to enhance the field strength locally, in the air but close to the interface. In principle, this means that we want the field in the air but close to the interface to be large in comparison to the incident field. This is not something we can really talk about for the surface mode problem, because there is no incident field in this situation, but we can say some other interesting things about the electric field in the vicinity of the interface. Let us restrict attention to the electric field. We can introduce the tangential and perpendicular components

of the electric field:

$$\begin{aligned}
\mathbf{E}_1^{\perp} &= \frac{B_y}{\omega\mu_0\varepsilon_1}(0,0,-k_x)\exp(-i\omega t + ik_x x + ik_{1z}z), \\
\mathbf{E}_1^{\parallel} &= \frac{B_y}{\omega\mu_0\varepsilon_1}(k_{1z},0,0)\exp(-i\omega t + ik_x x + ik_{1z}z), \\
\mathbf{E}_2^{\perp} &= \frac{B_y}{\omega\mu_0\varepsilon_2}(0,0,-k_x)\exp(-i\omega t + ik_x x + ik_{2z}z), \\
\mathbf{E}_2^{\parallel} &= \frac{B_y}{\omega\mu_0\varepsilon_2}(k_{2z},0,0)\exp(-i\omega t + ik_x x + ik_{2z}z).
\end{aligned} \tag{79}$$

At the interface (i.e. $z = 0$) in the air, we have:

$$\frac{|\mathbf{E}_1^{\perp}|^2}{|\mathbf{E}_1^{\parallel}|^2} = \frac{|k_x|^2}{|k_{1z}|^2} = \sqrt{(1+\Gamma^2)\beta^2 - 2\beta + 1}, \tag{80}$$

and so for large β, the electric field in the air near the interface is mainly perpendicular to the interface. Similarly, at the interface in the metal, we have

$$\frac{|\mathbf{E}_2^{\perp}|^2}{|\mathbf{E}_2^{\parallel}|^2} = \frac{|k_x|^2}{|k_{2z}|^2} = \frac{1}{\sqrt{(1+\Gamma^2)\beta^2 - 2\beta + 1}}, \tag{81}$$

and for large β the electric field in the metal near the interface is mainly parallel to the interface. Furthermore, we have the following expression for the ratio of the field amplitudes at the interface:

$$\frac{|\mathbf{E}_1|^2}{|\mathbf{E}_2|^2} = \frac{|k_x|^2 + |k_{1z}|^2}{|k_x|^2 + |k_{2z}|^2}\,|\hat{\varepsilon}_2|^2 = \sqrt{(1+\Gamma^2)\beta^2 - 2\beta + 1}, \tag{82}$$

indicating a large jump in the amplitude of the electric field from the metal to the air when β is large.

3.2 Coupling the Surface Modes to Incident Electromagnetic Waves

It is desirable for many applications to couple the surface modes found in the previous section to an incident electromagnetic field; this gives a way to

boost the amplitude of the electromagnetic field close to the interface. Since polarization is conserved at planar interfaces [2], in the current context only incident light which is TM polarized is useful for exciting the surface modes in which we are interested.

One of the difficulties that must be dealt with is that the wave vector for a photon in the dielectric region Ω_1 does not match the wave vector of the surface plasmon mode; therefore the momentum in the x direction for a photon does not match the momentum at the same frequency for the surface mode, and in fact the momentum of the photon is (slightly) smaller. Conservation of momentum therefore means that a photon can not (at least directly) excite the surface plasmon modes, and unless something special is done, almost no energy is transfered to the surface modes.

There are several strategies to get around the conservation of momentum restriction. This can be done by imposing some spatial structure on the surface along the x direction, for example by forming a surface grating (see figure 1). A periodic structure does not fully break the translational invariance, but it will allow couplings between the incident light and the surface modes for some specific frequencies and/or angles. If the translational invariance is totally broken by a surface with random roughness, then the photons will always couple to the surface modes, albeit somewhat weakly and in an uncontrolled way. See [17] for a presentation of many aspects of such problems. One aspect to keep in mind is that if we are going to continue to talk about the surface modes, then the perturbation from planarity must be small enough that the analysis in the previous section remains true to leading order; it is suggested in [13] that this is the case for a grating whose amplitude is small compared to its spatial period and to the decay length away from the surface in region Ω_1.

A related idea is to break the translational invariance by introducing a spatially localized perturbation to the surface, i.e. a bump or a valley, or by using a point source of illumination located sufficiently close to the surface. The localized illumination source or defect in the surface now acts as a point source for the surface plasmon modes. The translational symmetry is broken in these cases, but the problem is changed slightly as well; we must now consider cylindrical waves rather than plane waves.

There are several other methods of enhancing the coupling between the electromagnetic waves and the surface modes. One is based on using prisms [13], [14], [17], [19]. The idea is that if the prism exhibits total internal reflection of the electromagnetic wave at its surface, it must actually also

produce an evanescent wave near the surface. While the amplitude of this wave decreases rapidly away from the surface, it also produces photons with momentum larger than normal for propagating photons, and if the properties of the prism and are tuned properly, and the prism is held at the correct distance from the metal interface, it is possible to arrange things so that the photons couple with the surface modes. Related to this idea, if a thin metal layer is deposited on the prism, under the right conditions a surface mode can be excited at the metal-air boundary. Another technique is known as end-fire coupling or edge coupling [14], [19]. In this technique, a thin metal layer is sandwiched between dielectric slabs, and a laser is focused on the edge of the metal layer.

3.3 Resonance for Diffraction Gratings

Consider a plane wave incident on an air-metal interface, which is periodic in the x-direction, i.e. with a grating type geometry. Assuming that the amplitude of the surface corrugation is small enough that we can ignore its effect on the surface plasmon mode, we know from the analysis above that there is a mismatch between the wave vector in the air and the wave vector of the surface plasmon modes.

Assuming that plane wave is incident onto the grating with polar angle θ and azimuthal angle ϕ, we can project the wave vector of the incident photons onto the plane $z = 0$ to get that part of the wave vector of the incoming photons which is available to the surface plasmon modes:

$$\mathbf{k}_{in} = \frac{\omega}{c}\sqrt{\hat{\varepsilon}_1}(\sin(\theta)\cos(\phi), \sin(\theta)\sin(\phi), 0). \tag{83}$$

However, when a photon strikes a diffraction grating, integer multiples of the grating vector can be added to or subtracted from the wave vector of the photon via Bragg scattering (see [4]). If the grating has period d, and the grooves are parallel to the y direction, then the grating vector is

$$\mathbf{k}_g = \frac{2\pi}{d}(1, 0, 0). \tag{84}$$

If we define $\mathbf{k}_s$ to be the real part of the in-plane wave vector of the surface plasmon modes, then we need to match its magnitude with that of any one of the wave vectors accessible through Bragg scattering of the incoming wave:

$$|\mathbf{k}_s| = |\mathbf{k}_{in} \pm n\mathbf{k}_g| \tag{85}$$

where $n = 1, 2, 3, \ldots$, and typically only the $n = 1$ term generates a strong coupling. Note that the plasmons propagate along the surface in the direction given by the vector sum $\mathbf{k}_g \pm n\mathbf{k}_g$, not necessarily the x direction as chosen for convenience above when the surface modes were computed.

The predictions of this equation are compared to the results of a full numerical solution of Maxwell's equations for both TE and TM polarized waves, for several different values of ϕ, with parameters chosen for gold, and for gratings of several different shapes and amplitudes, in [19]. It seems to do a good job of predicting the best polar angle (i.e. the one which leads to the largest field strength in the vicinity of the metal surface), with all other parameters held fixed.

The theory for surfaces with random roughness is essentially built up by assuming that the roughness can be characterized as a superposition of many diffraction gratings of various periods and orientations on the surface. It seems that in the same way, a periodically structured surface could be characterized by a relatively small number of grating vectors, and the analogue of equation (85) used to look for potential resonances.

4 Nondimensionalization

We introduce dimensionless space and time variables as follows:

$$\mathbf{x} = l\mathbf{x}^*, \qquad t = \frac{1}{\omega}t^*, \tag{86}$$

where l is a typical length scale of the microstructure, and where as noted above, the frequency ω is a time scale associated with the laser illumination (we eventually assume that the dimensional variables all have time dependence $e^{-i\omega t}$). We scale the field variables as follows:

$$\mathbf{E} = E_0\mathbf{E}^*, \quad \mathbf{D} = \varepsilon_0 E_0 \mathbf{D}^*, \quad \mathbf{P} = \varepsilon_0 E_0 \mathbf{P}^*, \quad \mathbf{J} = \frac{E_0}{\mu_0 \omega l^2}\mathbf{J}^*, \tag{87}$$

and since we will not consider magnetic effects, we finally drop $\mathbf{H}$ in favor of $\mathbf{B}$:

$$\mathbf{B} = \frac{E_0}{\omega l}\mathbf{B}^*, \qquad \mathbf{H} = \frac{E_0}{\mu_0 \omega l}\mathbf{B}^*. \tag{88}$$

Note that we had to assume that the field variables had a time dependence $e^{-i\omega t}$ in order to get expressions for the material parameters, so while we write

the equations using partial derivatives with respect to t^* below, we should, strictly speaking, assume that the non-dimensionalized field variables have a time dependence of the form e^{-it^*}.

4.1 Nondimensional Equations in the Dielectric Regions

In regions Ω_1 (air) and Ω_3 (substrate), we have:

$$\nabla^* \cdot \mathbf{D}^* = 0, \tag{89}$$

$$\nabla^* \cdot \mathbf{B}^* = 0, \tag{90}$$

$$\nabla^* \times \mathbf{E}^* = -\frac{\partial \mathbf{B}^*}{\partial t^*}, \tag{91}$$

$$\nabla^* \times \mathbf{B}^* = \frac{1}{\Lambda}\frac{\partial \mathbf{D}^*}{\partial t^*}, \tag{92}$$

$$\mathbf{D}^* = \mathbf{E}^* + \mathbf{P}^*, \tag{93}$$

$$\mathbf{D}^* = \hat{\varepsilon}_i \mathbf{E}^*, \tag{94}$$

$$\mathbf{P}^* = (\hat{\varepsilon}_i - 1)\mathbf{E}^*, \tag{95}$$

$$\mathbf{J}^* = \mathbf{0}. \tag{96}$$

For simplicity, we will assume that $\hat{\varepsilon}_1 = 1$, as noted above.

4.2 Nondimensional Equations in the Conductive Region

In region Ω_2 (the metal) we have:

$$\nabla^* \cdot \mathbf{D}^* = 0, \tag{97}$$

$$\nabla^* \cdot \mathbf{B}^* = 0, \tag{98}$$

$$\nabla^* \times \mathbf{E}^* = -\frac{\partial \mathbf{B}^*}{\partial t^*}, \tag{99}$$

$$\nabla^* \times \mathbf{B}^* = \frac{1}{\Lambda}\frac{\partial \mathbf{D}^*}{\partial t^*}, \tag{100}$$

$$\mathbf{D}^* = \mathbf{E}^* + \mathbf{P}^*, \tag{101}$$

$$\mathbf{D}^* = \hat{\varepsilon}_2 \mathbf{E}^* = ((1-\beta) + i\beta\Gamma)\,\mathbf{E}^*, \tag{102}$$

$$\mathbf{P}^* = \hat{\alpha}_2 \mathbf{E}^* = (-\beta + i\beta\Gamma)\,\mathbf{E}^*, \tag{103}$$

$$\mathbf{J}^* = \hat{\sigma}_2 \mathbf{E}^* = \frac{1}{\Lambda}(\beta\Gamma + i\beta)\,\mathbf{E}^*. \tag{104}$$

4.3 Nondimensional Jump Conditions

The jump conditions become:

$$[\mathbf{n} \cdot \mathbf{D}^*] = 0, \tag{105}$$

$$[\mathbf{n} \cdot \mathbf{B}^*] = 0, \tag{106}$$

$$[\mathbf{n} \times \mathbf{E}^*] = \mathbf{0}, \tag{107}$$

$$[\mathbf{n} \times \mathbf{B}^*] = \mathbf{0}. \tag{108}$$

4.4 Nondimensional Parameters

The only non-dimensional group that we have not seen yet is:

$$\Lambda = \frac{1}{\varepsilon_0 \mu_0 \omega^2 l^2} = \frac{c^2}{\omega^2 l^2} = \frac{1}{4\pi^2}\frac{\lambda^2}{l^2}, \tag{109}$$

where we have used the fact that $\varepsilon_0 \mu_0 = c^2$, where c is the speed of light in vacuum, and the fact that $\omega = 2\pi c\lambda$, where λ is the wavelength of a light wave with angular frequency ω. Therefore, Λ is essentially the square of a ratio of the wavelength of the incident light to a representative length scale for the nano-structured metal surface. It was suggested during the MPI workshop that the assumption $\Lambda = O(1)$ be made, even though Λ could in principle vary over several orders of magnitude and might also be small.

We also have the relative dielectric functions $\hat{\varepsilon}_1$ and $\hat{\varepsilon}_3$ in Ω_1 and Ω_3 respectively, as well as β and Γ, all of which were introduced above.

4.5 Reduced Problem

From now on, we will drop the *s from the non-dimensional variables. Furthermore, we will assume that the electric and magnetic fields have time dependence e^{-it}, as noted above. If we restrict consideration to a problem

with one interface, air above and metal below, we can consider the following problem.

$$\begin{aligned} \mathbf{x} \in \Omega_1 : \quad & \nabla \cdot \mathbf{E}_1 = 0, \;\; \nabla^2 \mathbf{E}_1 + \Lambda^{-1} \mathbf{E}_1 = 0, \\ & \nabla \cdot \mathbf{B}_1 = 0, \;\; \nabla^2 \mathbf{B}_1 + \Lambda^{-1} \mathbf{B}_1 = 0, \\ & \mathbf{B}_1 = i \nabla \times \mathbf{E}_1, \end{aligned} \tag{110}$$

$$\begin{aligned} \mathbf{x} \in \Omega_2 : \quad & \nabla \cdot \mathbf{E}_2 = 0, \;\; \nabla^2 \mathbf{E}_2 + \Lambda^{-1} \left((1-\beta) + i\beta\Gamma \right) \mathbf{E}_2 = 0, \\ & \nabla \cdot \mathbf{B}_2 = 0, \;\; \nabla^2 \mathbf{B}_2 + \Lambda^{-1} \left((1-\beta) + i\beta\Gamma \right) \mathbf{B}_2 = 0, \\ & \mathbf{B}_2 = i \nabla \times \mathbf{E}_2. \end{aligned} \tag{111}$$

If we assume that the interface is given by $z = h(x, y)$, with z increasing downward (into the metal) for consistency with our derivation of the surface plasmon wave problem above, we have the outer normal vector for Ω_1:

$$\mathbf{n} = \frac{(-h_x, -h_y, 1)}{\sqrt{1 + h_x^2 + h_y^2}}. \tag{112}$$

The matching conditions at the interface are:

$$\begin{aligned} z = h(x,y) : \quad & \mathbf{n} \cdot \mathbf{E}_1 = \mathbf{n} \cdot \left((1-\beta) + i\beta\Gamma \right) \mathbf{E}_2, \\ & \mathbf{n} \cdot \mathbf{B}_1 = \mathbf{n} \cdot \mathbf{B}_2, \\ & \mathbf{n} \times \mathbf{E}_1 = \mathbf{n} \times \mathbf{E}_2, \\ & \mathbf{n} \times \mathbf{B}_1 = \mathbf{n} \times \mathbf{B}_2. \end{aligned} \tag{113}$$

Finally, we should impose appropriate conditions on the far field.

4.6 Asymptotics for β Large, Λ $O(1)$: Leading Order Problem

Assuming that β is large, and that Λ is $O(1)$, suggests the expansions

$$\begin{aligned} \mathbf{E}_1 &= \mathbf{E}_{1,0} + \beta^{-1} \mathbf{E}_{1,1} + \cdots, \\ \mathbf{B}_1 &= \mathbf{B}_{1,0} + \beta^{-1} \mathbf{E}_{1,1} + \cdots, \\ \mathbf{E}_2 &= \qquad \beta^{-1} \mathbf{E}_{2,1} + \cdots, \\ \mathbf{B}_2 &= \qquad \beta^{-1} \mathbf{B}_{2,1} + \cdots, \end{aligned} \tag{114}$$

which produces a leading order problem in which the metal is treated as a perfect conductor. The problem for the leading order electric field in the air is:

$$\begin{aligned} \mathbf{x} \in \Omega_1 : \quad & \nabla \cdot \mathbf{E}_{1,0} = 0, \quad \nabla^2 \mathbf{E}_{1,0} + \Lambda^{-1} \mathbf{E}_{1,0} = 0, \\ \mathbf{x} \in \partial\Omega_1 : \quad & \mathbf{n} \times \mathbf{E}_{1,0} = 0, \end{aligned} \tag{115}$$

which must be supplemented by the appropriate radiation condition. Based on the discussion above, if we want to excite surface plasmons, the incoming wave will need to hit the boundary surface obliquely. Furthermore, since the surface is not planar, no matter how the incident wave is chosen, the scattered wave will in general contain both possible polarizations. The fact that we have scaled the spatial coordinates so as to make $h(x, y)$ of $O(1)$ means that even at leading order we have to deal with a nontrivial scattering problem.

Further analytical progress is possible if we assume that the amplitude of the surface corrugation is small. This effectively gives a linearized boundary condition at a flat interface, which couples the two possible polarizations and depends on the shape of the interface, but is imposed on the unperturbed planar interface. If Λ^{-1} is small, we get back to the electrostatic approximation, which is commonly used as a first approximation, even in cases where it is well understood not to be formally valid [13]. If Λ^{-1} is large, then the geometric theory of diffraction ([3], [5], [11], [26]) might be helpful. If none of these further approximations is acceptable, then problem (115) must in general be solved numerically.

5 Potential Numerical Approaches

There has been a great deal of work done on the numerical modeling of plasmonic devices. Quite a few different methods have been proposed and implemented, as pointed out in the introduction to the problem provided by TIAX. We mention here a few techniques which can be applied to the reduced problem derived above, but which do not appear common in the plasmonics literature.

5.1 Boundary Element Method

One possibility is to use the Boundary Element Method. The idea is to reduce the problem to an integral equation involving the values of the unknown on the boundary. Once this equation is solved, the same equation provides an integral representation of the solution over the whole domain. The required reformulation of the problem in terms of a boundary integral equation is based on the Stratton-Chu formulas ([24], see also [23] and [8]).

The boundary integral equation equivalent to (115) is:

$$\mathbf{E}_{1,0}(\mathbf{x}) = -\Lambda\nabla\times\nabla\times\int_{\partial\Omega_1}\mathbf{n}(\mathbf{y})\times\nabla\times\mathbf{E}_{1,0}(\mathbf{y})\Phi(\mathbf{x},\mathbf{y})\,dS(\mathbf{y}), \qquad (116)$$

where Φ is the free space Green's Function for the Helmholtz equation, which for our problem is given by

$$\Phi(\mathbf{x},\mathbf{y}) = \frac{1}{4\pi}\frac{e^{i\Lambda^{-1/2}|\mathbf{x}-\mathbf{y}|}}{|\mathbf{x}-\mathbf{y}|}, \qquad (117)$$

with $\mathbf{x} = (x, y, z)$, and $\mathbf{y} = (\xi, \eta, \zeta)$, and $\partial\Omega_1$ represents the air-metal boundary. The differential of surface area is written $dS(\mathbf{y})$ to indicate that the parameterization of the surface is to be done in terms of $\mathbf{y} = (\xi, \eta, \zeta)$. The surface integral is improper for two reasons: the singularity of Φ when $\mathbf{x} = \mathbf{y}$, and the fact that $\partial\Omega_1$ is an infinite two dimensional surface. This still appears to be a nontrivial problem. (Note that while it is possible to reduce the problem to a single integral equation, this problem is typically reduced to two coupled boundary integral equations, which involve $\mathbf{E}$ and $\mathbf{H}$ in dual roles.)

Some related problems are treated in [18] by a similar method. In this paper, the case of finite conductivity is treated, which means that the full problem (i.e. involving both half spaces Ω_1 and Ω_2) must be tackled. Metal-air interfaces which are rough (i.e. random) in one direction along the interface and do not vary in the orthogonal direction along the surface are treated. The problem is reduced to a single integral equation, and a number of practical aspects (i.e. how to regularize the singularity of the Green's function) are discussed. Since a number of physical assumptions are used along the way to simplify the form of the integral equation, a direct comparison to our problem is not possible.

5.2 Other Approaches

The boundary element method is mentioned above, but there are a number of other alternatives. One which is fairly easy to understand conceptually is based on the idea of building a solution to the scattering problem of interest, which involves a corrugated interface, in terms of a superposition of solutions to a simpler scattering problem. There are a number of variations on this idea.

In [7], the author considers scattering from a perfectly conducting metallic surface which is rough in two dimensions. The problem is reduced from three dimensional to two dimensional by expanding in terms of local eigenfunctions which depend on the normal coordinate (z for us); the problem is then reduced to a set of coupled integral equations for the amplitudes of the various modes. These are then solved iteratively.

In [20], the authors treat the problem of scattering of a TE-polarized electromagnetic wave from a surface which is corrugated in one dimension, and does not vary in the orthogonal dimension along the interface. The corrugations can be deterministic or random. Solutions are built from modified plane waves which by design have the correct behavior both at the corrugated interface and at infinity, but are not necessarily solutions to Maxwell's equations. The solution is constructed as an incident plane wave (which is a solution of Maxwell's equations), plus a reflected modified plane wave, plus corrections. For the case of a planar interface, the reflected plane wave actually solves Maxwell's equations and there are no additional corrections; for the case of a non-planar interface, the modified reflected plane wave is just the reflected plane wave from the planar case, multiplied by a factor which enforces the boundary conditions at the interface. Since the reflected modified plane wave is generally not a solution of Maxwell's equations, additional correction terms are necessary. These can be thought of as current distributions which compensate for the errors in the field variables. These corrections are computed iteratively. At least conceptually, this seems to be the easiest approach to understand, although its implementation is somewhat involved. The authors are able to show that the iteration scheme converges in certain circumstances. The authors limit themselves to explicit calculation of the first correction (i.e. implementing one iteration of the method), but are able to produce some impressive results none the less. When only the first correction is computed, the results can be simplified quite a bit, and the answer written in terms of relatively simple formulas depending on the corrugation profile.

6 Conclusions

We have set up and non-dimensionalized the governing equations for the interaction of a plane wave with a structured metal surface. In section 3.3 we suggest a very simple criterion for choosing the wavelength of incident

light in order to maximize the coupling to surface plasmons and potentially maximize the field strength in the vicinity of the interface, if the structured surface can be represented as a superposition of gratings.

In order to validate this suggestion, a solution of the full electromagnetic problem that can be studied parametrically is required. We have taken the first step in this direction by deriving in section 5 the leading order problem for the electric field in the air. While this is a difficult problem, its solution it is only the first step in a boundary layer analysis which is required to construct the electromagnetic field everywhere. It is clear from the structure of the expansions for large β, and from the structure of the surface plasmon solution found in section 3.1, that the rapidly varying variable in the boundary layer will be $\beta^{1/2}\Lambda^{-1/2}z$. (Here z is the non-dimensionalized variable that was originally denoted z^*). Furthermore, the solution for the surface plasmon problem suggests that the solution in the air will also depend on a slowly varying variable $\beta^{-1/2}\Lambda^{-1/2}z$. If the strategy suggested in [20] can be extended to produce a reasonably accurate analytic representation for the leading order problem in the air, then it might be possible to work out the boundary layer analysis explicitly, and in so doing, obtain formulas for both the local field and the extent to which plasmons are excited. Otherwise, it seems that a full numerical solution will be necessary.

References

[1] E. L. Albuquerque and M. G. Cottam. *Polaritons in Periodic and Quasiperiodic Structures*. Elsevier, 2004.

[2] C. A. Balanis. *Advanced Engineering Electromagnetics*. Wiley, 1989.

[3] N. Bleistein. *Mathematical Methods for Wave Phenomena*. Academic Press, 1984.

[4] M. Born and E. Wolf. *Principles of Optics* (seventh edition). Cambridge, 1999.

[5] D. Bouche, F. Molinet, and R. Mittra. *Asymptotic Methods in Electromagnetics*. Springer, 1997.

[6] M. L. Brongersma and P. G. Kik (eds.) *Surface Plasmon Nanophotonics*. Springer Series in Optical Sciences 131, Springer, 2007.

[7] R. E. Collin. Electromagnetic scattering from perfectly conducting rough surfaces (a new full wave method). *IEEE Transactions on Antennas and Propagation*, **40**: 1466 – 1477, 1992.

[8] D. Colton and R. Kress. *Inverse Acoustic and Electromagnetic Scattering*. Springer, 1992.

[9] G.W. Ford and W.H. Weber. Electromagnetic Interactions of molecules with metal surfaces. *Physics Reports*, **113**: 195–287, 1984.

[10] J. D. Jackson. *Classical Electrodynamics*, 3rd. Ed., Wiley, 1999.

[11] G. L. James. *Geometrical Theory of Diffraction for Electromagnetic Waves* (revised edition). IEE, 1980.

[12] K. Kneipp, M. Moskovits, and H. Kneiff (eds.) *Surface-Enhanced Raman Scattering: Physics and Applications*. Topics in Applied Physics 103, Springer 2006.

[13] E. C. Le Ru and P. G. Etchegoin. *Principles of Surface-Enhanced Raman Spectroscopy*. Elsevier, 2009.

[14] S. A. Maier. *Plasmonics: Fundamentals and Applications*. Springer, 2007.

[15] M.P. Marder. *Condensed Matter Physics*. John Wiley and Sons, 2010.

[16] D.L. Mills. *Nonlinear Optics: Basic Concepts* (second enlarged edition). Springer, 1998.

[17] H. Raether. *Surface Plasmons on Smooth and Rough Surfaces and on Gratings*. Springer, 1988.

[18] M. Saillard and D. Maystre. Scattering from metallic and dielectric rough surfaces. *J. Opt. Soc. Am. A*, **7**: 982 – 990, 1990.

[19] D. Sarid and W. Challener. *Modern Introduction to Surface Plasmons: Theory, Mathematica Modelling, and Applications*. Cambridge University Press, 2010.

[20] F.K. Schwering, G.M. Whitman, and A.A. Triolo. New full-wave theory for scattering from rough metal surfaces – the correction current method: the TE-polarization case. *Waves in Random Media*, **14**: 23 – 60, 2004.

[21] B. E. Sernelius. *Surface Modes in Physics*. Wiley-VCH, 2001.

[22] V. M. Shalaev (ed.) *Optical Properties of Nanostructured Random Media*. Topics in Applied Physics 82, Springer 2002.

[23] J. A. Stratton. *Electromagnetic Theory*. IEEE Press, 2007.

[24] J.A. Stratton and L.J. Chu. Diffraction theory of electromagnetic waves. *Phys. Rev.*, **56**, 99–107, 1939.

[25] B. K. Tanner. *Introduction to the Physics of Electrons in Solids*, Cambridge University Press, 1995.

[26] P. Ya. Ufimtsev. *Fundamentals of the Physical Theory of Diffraction*. John Wiley & Sons, 2007.