Chapter 1

ELECTROMAGNETICS OF METALS

While the optical properties of metals are discussed in most textbooks on condensed matter physics, for convenience this chapter summarizes the most important facts and phenomena that form the basis for a study of surface plasmon polaritons. Starting with a cursory review of Maxwell's equations, we describe the electromagnetic response both of idealized and real metals over a wide frequency range, and introduce the fundamental excitation of the conduction electron sea in bulk metals: *volume plasmons*. The chapter closes with a discussion of the electromagnetic energy density in dispersive media.

1.1 Maxwell's Equations and Electromagnetic Wave Propagation

The interaction of metals with electromagnetic fields can be firmly understood in a classical framework based on Maxwell's equations. Even metallic nanostructures down to sizes on the order of a few nanometres can be described without a need to resort to quantum mechanics, since the high density of free carriers results in minute spacings of the electron energy levels compared to thermal excitations of energy $k_{\mathrm{B}}T$ at room temperature. The optics of metals described in this book thus falls within the realms of the classical theory. However, this does not prevent a rich and often unexpected variety of optical phenomena from occurring, due to the strong dependence of the optical properties on frequency.

As is well known from everyday experience, for frequencies up to the visible part of the spectrum metals are highly reflective and do not allow electromagnetic waves to propagate through them. Metals are thus traditionally employed as cladding layers for the construction of waveguides and resonators for electromagnetic radiation at microwave and far-infrared frequencies. In this

low-frequency regime, the *perfect* or *good conductor approximation* of infinite or fixed finite conductivity is valid for most purposes, since only a negligible fraction of the impinging electromagnetic waves penetrates into the metal. At higher frequencies towards the near-infrared and visible part of the spectrum, field penetration increases significantly, leading to increased dissipation, and prohibiting a simple size scaling of photonic devices that work well at low frequencies to this regime. Finally, at ultraviolet frequencies, metals acquire dielectric character and allow the propagation of electromagnetic waves, albeit with varying degrees of attenuation, depending on the details of the electronic band structure. Alkali metals such as sodium have an almost free-electron-like response and thus exhibit an *ultraviolet transparency*. For noble metals such as gold or silver on the other hand, transitions between electronic bands lead to strong absorption in this regime.

These dispersive properties can be described via a complex dielectric function $\varepsilon(\omega)$, which provides the basis of all phenomena discussed in this text. The underlying physics behind this strong frequency dependence of the optical response is a change in the phase of the induced currents with respect to the driving field for frequencies approaching the reciprocal of the characteristic *electron relaxation time* τ of the metal, as will be discussed in section 1.2.

Before presenting an elementary description of the optical properties of metals, we recall the basic equations governing the electromagnetic response, the *macroscopic Maxwell equations*. The advantage of this *phenomenological approach* is that details of the fundamental interactions between charged particles inside media and electromagnetic fields need not be taken into account, since the rapidly varying microscopic fields are averaged over distances much larger than the underlying microstructure. Specifics about the transition from a microscopic to a macroscopic description of the electromagnetic response of continuous media can be found in most textbooks on electromagnetics such as [Jackson, 1999].

We thus take as a starting point Maxwell's equations of macroscopic electromagnetism in the following form:

$$\nabla \cdot \mathbf{D} = \rho_{\text{ext}} \tag{1.1a}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{1.1b}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{1.1c}$$

$$\nabla \times \mathbf{H} = \mathbf{J}_{\mathbf{ext}} + \frac{\partial \mathbf{D}}{\partial t}. \tag{1.1d}$$

These equations link the four macroscopic fields **D** (the dielectric displacement), **E** (the electric field), **H** (the magnetic field), and **B** (the magnetic induc-

tion or magnetic flux density) with the external charge and current densities ρ_{ext} and $\mathbf{J_{ext}}$. Note that we do not follow the usual procedure of presenting the macroscopic equations via dividing the total charge and current densities ρ_{tot} and $\mathbf{J_{tot}}$ into free and bound sets, which is an arbitrary division [Illinskii and Keldysh, 1994] and can (especially in the case of metallic interfaces) confuse the application of the boundary condition for the dielectric displacement. Instead, we distinguish between external ($\rho_{\text{ext}}, \mathbf{J_{ext}}$) and internal ($\rho, \mathbf{J}$) charge and current densities, so that in total $\rho_{\text{tot}} = \rho_{\text{ext}} + \rho$ and $\mathbf{J_{tot}} = \mathbf{J_{ext}} + \mathbf{J}$. The external set drives the system, while the internal set responds to the external stimuli [Marder, 2000].

The four macroscopic fields are further linked via the polarization $\mathbf{P}$ and magnetization $\mathbf{M}$ by

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P} \tag{1.2a}$$

$$\mathbf{H} = \frac{1}{\mu_0}\mathbf{B} - \mathbf{M}, \tag{1.2b}$$

where ε_0 and μ_0 are the electric permittivity[1] and magnetic permeability[2] of vacuum, respectively. Since we will in this text only treat nonmagnetic media, we need not consider a magnetic response represented by $\mathbf{M}$, but can limit our description to electric polarization effects. $\mathbf{P}$ describes the electric dipole moment per unit volume inside the material, caused by the alignment of microscopic dipoles with the electric field. It is related to the internal charge density via $\nabla \cdot \mathbf{P} = -\rho$. Charge conservation ($\nabla \cdot \mathbf{J} = -\partial\rho/\partial t$) further requires that the internal charge and current densities are linked via

$$\mathbf{J} = \frac{\partial \mathbf{P}}{\partial t}. \tag{1.3}$$

The great advantage of this approach is that the macroscopic electric field includes all polarization effects: In other words, both the external and the induced fields are absorbed into it. This can be shown via inserting (1.2a) into (1.1a), leading to

$$\nabla \cdot \mathbf{E} = \frac{\rho_{\text{tot}}}{\varepsilon_0}. \tag{1.4}$$

In the following, we will limit ourselves to linear, isotropic and nonmagnetic media. One can define the constitutive relations

$$\mathbf{D} = \varepsilon_0 \varepsilon \mathbf{E} \tag{1.5a}$$

$$\mathbf{B} = \mu_0 \mu \mathbf{H}. \tag{1.5b}$$

[1] $\varepsilon_0 \approx 8.854 \times 10^{-12}$ F/m
[2] $\mu_0 \approx 1.257 \times 10^{-6}$ H/m

ε is called the dielectric constant or relative permittivity and $\mu = 1$ the relative permeability of the nonmagnetic medium. The linear relationship (1.5a) between $\mathbf{D}$ and $\mathbf{E}$ is often also implicitly defined using the dielectric susceptibility χ (particularly in quantum mechanical treatments of the optical response [Boyd, 2003]), which describes the linear relationship between $\mathbf{P}$ and $\mathbf{E}$ via

$$\mathbf{P} = \varepsilon_0 \chi \mathbf{E}. \tag{1.6}$$

Inserting (1.2a) and (1.6) into (1.5a) yields $\varepsilon = 1 + \chi$.

The last important constitutive linear relationship we need to mention is that between the internal current density $\mathbf{J}$ and the electric field $\mathbf{E}$, defined via the conductivity σ by

$$\mathbf{J} = \sigma \mathbf{E}. \tag{1.7}$$

We will now show that there is an intimate relationship between ε and σ, and that electromagnetic phenomena with metals can in fact be described using either quantity. Historically, at low frequencies (and in fact in many theoretical considerations) preference is given to the conductivity, while experimentalists usually express observations at optical frequencies in terms of the dielectric constant. However, before embarking on this we have to point out that the statements (1.5a) and (1.7) are only correct for linear media that do not exhibit temporal or spatial dispersion. Since the optical response of metals clearly depends on frequency (possibly also on wave vector), we have to take account of the non-locality in time and space by generalizing the linear relationships to

$$\mathbf{D}(\mathbf{r}, t) = \varepsilon_0 \int dt' \mathbf{dr}' \varepsilon(\mathbf{r} - \mathbf{r}', t - t') \mathbf{E}(\mathbf{r}', t') \tag{1.8a}$$

$$\mathbf{J}(\mathbf{r}, t) = \int dt' \mathbf{dr}' \sigma(\mathbf{r} - \mathbf{r}', t - t') \mathbf{E}(\mathbf{r}', t'). \tag{1.8b}$$

$\varepsilon_0 \varepsilon$ and σ therefore describe the impulse response of the respective linear relationship. Note that we have implicitly assumed that all length scales are significantly larger than the lattice spacing of the material, ensuring homogeneity, i.e. the impulse response functions do not depend on absolute spatial and temporal coordinates, but only their differences. For a local response, the functional form of the impulse response functions is that of a δ-function, and (1.5a) and (1.7) are recovered.

Equations (1.8) simplify significantly by taking the Fourier transform with respect to $\int dt \mathbf{dr} e^{i(\mathbf{K}\cdot\mathbf{r} - \omega t)}$, turning the convolutions into multiplications. We are thus decomposing the fields into individual plane-wave components of wave vector $\mathbf{K}$ and angular frequency ω. This leads to the constitutive rela-

tions in the Fourier domain

$$\mathbf{D}(\mathbf{K}, \omega) = \varepsilon_0 \varepsilon(\mathbf{K}, \omega)\mathbf{E}(\mathbf{K}, \omega) \tag{1.9a}$$

$$\mathbf{J}(\mathbf{K}, \omega) = \sigma(\mathbf{K}, \omega)\mathbf{E}(\mathbf{K}, \omega). \tag{1.9b}$$

Using equations (1.2a), (1.3) and (1.9) and recognizing that in the Fourier domain $\partial/\partial t \rightarrow -i\omega$, we finally arrive at the fundamental relationship between the relative permittivity (from now on called the *dielectric function*) and the conductivity

$$\varepsilon(\mathbf{K}, \omega) = 1 + \frac{i\sigma(\mathbf{K}, \omega)}{\varepsilon_0 \omega}. \tag{1.10}$$

In the interaction of light with metals, the general form of the dielectric response $\varepsilon(\omega, \mathbf{K})$ can be simplified to the limit of a *spatially local* response via $\varepsilon(\mathbf{K} = \mathbf{0}, \omega) = \varepsilon(\omega)$. The simplification is valid as long as the wavelength λ in the material is significantly longer than all characteristic dimensions such as the size of the unit cell or the mean free path of the electrons. This is in general still fulfilled at ultraviolet frequencies[3].

Equation (1.10) reflects a certain arbitrariness in the separation of charges into bound and free sets, which is entirely due to convention. At low frequencies, ε is usually used for the description of the response of bound charges to a driving field, leading to an electric polarization, while σ describes the contribution of free charges to the current flow. At optical frequencies however, the distinction between bound and free charges is blurred. For example, for highly-doped semiconductors, the response of the bound valence electrons could be lumped into a static dielectric constant $\delta\varepsilon$, and the response of the conduction electrons into σ', leading to a dielectric function $\varepsilon(\omega) = \delta\varepsilon + \frac{i\sigma'(\omega)}{\varepsilon_0\omega}$. A simple redefinition $\delta\varepsilon \rightarrow 1$ and $\sigma' \rightarrow \sigma' + \frac{\varepsilon_0\omega}{i}\delta\varepsilon$ will then result in the general form (1.10) [Ashcroft and Mermin, 1976].

In general, $\varepsilon(\omega) = \varepsilon_1(\omega) + i\varepsilon_2(\omega)$ and $\sigma(\omega) = \sigma_1(\omega) + i\sigma_2(\omega)$ are complex-valued functions of angular frequency ω, linked via (1.10). At optical frequencies, ε can be experimentally determined for example via reflectivity studies and the determination of the complex refractive index $\tilde{n}(\omega) = n(\omega) + i\kappa(\omega)$ of the medium, defined as $\tilde{n} = \sqrt{\varepsilon}$. Explicitly, this yields

[3]However, spatial dispersion effects can lead to small corrections for surface plasmons polaritons in metallic nanostructures significantly smaller than the electron mean free path, which can arise for example at the tip of metallic cones (see chapter 7).

$$\varepsilon_1 = n^2 - \kappa^2 \tag{1.11a}$$

$$\varepsilon_2 = 2n\kappa \tag{1.11b}$$

$$n^2 = \frac{\varepsilon_1}{2} + \frac{1}{2}\sqrt{\varepsilon_1^2 + \varepsilon_2^2} \tag{1.11c}$$

$$\kappa = \frac{\varepsilon_2}{2n}. \tag{1.11d}$$

κ is called the extinction coefficient and determines the optical absorption of electromagnetic waves propagating through the medium. It is linked to the absorption coefficient α of Beer's law (describing the exponential attenuation of the intensity of a beam propagating through the medium via $I(x) = I_0 e^{-\alpha x}$) by the relation

$$\alpha(\omega) = \frac{2\kappa(\omega)\omega}{c}. \tag{1.12}$$

Therefore, the imaginary part ε_2 of the dielectric function determines the amount of absorption inside the medium. For $|\varepsilon_1| \gg |\varepsilon_2|$, the real part n of the refractive index, quantifying the lowering of the phase velocity of the propagating waves due to polarization of the material, is mainly determined by ε_1. Examination of (1.10) thus reveals that the real part of σ determines the amount of absorption, while the imaginary part contributes to ε_1 and therefore to the amount of polarization.

We close this section by examining traveling-wave solutions of Maxwell's equations in the absence of external stimuli. Combining the curl equations (1.1c), (1.1d) leads to the *wave equation*

$$\nabla \times \nabla \times \mathbf{E} = -\mu_0 \frac{\partial^2 \mathbf{D}}{\partial t^2} \tag{1.13a}$$

$$\mathbf{K}(\mathbf{K} \cdot \mathbf{E}) - K^2 \mathbf{E} = -\varepsilon(\mathbf{K}, \omega)\frac{\omega^2}{c^2}\mathbf{E}, \tag{1.13b}$$

in the time and Fourier domains, respectively. $c = \frac{1}{\sqrt{\varepsilon_0 \mu_0}}$ is the speed of light in vacuum. Two cases need to be distinguished, depending on the polarization direction of the electric field vector. For transverse waves, $\mathbf{K} \cdot \mathbf{E} = 0$, yielding the generic dispersion relation

$$K^2 = \varepsilon(\mathbf{K}, \omega)\frac{\omega^2}{c^2}. \tag{1.14}$$

For longitudinal waves, (1.13b) implies that

$$\varepsilon(\mathbf{K}, \omega) = 0, \tag{1.15}$$

signifying that longitudinal collective oscillations can only occur at frequencies corresponding to zeros of $\varepsilon(\omega)$. We will return to this point in the discussion of volume plasmons in section 1.3.

1.2 The Dielectric Function of the Free Electron Gas

Over a wide frequency range, the optical properties of metals can be explained by a *plasma model*, where a gas of free electrons of number density n moves against a fixed background of positive ion cores. For alkali metals, this range extends up to the ultraviolet, while for noble metals interband transitions occur at visible frequencies, limiting the validity of this approach. In the plasma model, details of the lattice potential and electron-electron interactions are not taken into account. Instead, one simply assumes that some aspects of the band structure are incorporated into the effective optical mass m of each electron. The electrons oscillate in response to the applied electromagnetic field, and their motion is damped via collisions occurring with a characteristic collision frequency $\gamma = 1/\tau$. τ is known as the relaxation time of the free electron gas, which is typically on the order of 10^{-14} s at room temperature, corresponding to $\gamma = 100$ THz.

One can write a simple equation of motion for an electron of the plasma sea subjected to an external electric field $\mathbf{E}$:

$$m\ddot{\mathbf{x}} + m\gamma\dot{\mathbf{x}} = -e\mathbf{E} \tag{1.16}$$

If we assume a harmonic time dependence $\mathbf{E}(t) = \mathbf{E}_0 \mathrm{e}^{-i\omega t}$ of the driving field, a particular solution of this equation describing the oscillation of the electron is $\mathbf{x}(t) = \mathbf{x}_0 \mathrm{e}^{-i\omega t}$. The complex amplitude $\mathbf{x}_0$ incorporates any phase shifts between driving field and response via

$$\mathbf{x}(t) = \frac{e}{m(\omega^2 + i\gamma\omega)}\mathbf{E}(t). \tag{1.17}$$

The displaced electrons contribute to the macroscopic polarization $\mathbf{P} = -ne\mathbf{x}$, explicitly given by

$$\mathbf{P} = -\frac{ne^2}{m(\omega^2 + i\gamma\omega)}\mathbf{E}. \tag{1.18}$$

Inserting this expression for $\mathbf{P}$ into equation (1.2a) yields

$$\mathbf{D} = \varepsilon_0\left(1 - \frac{\omega_\mathrm{p}^2}{\omega^2 + i\gamma\omega}\right)\mathbf{E}, \tag{1.19}$$

where $\omega_\mathrm{p}^2 = \frac{ne^2}{\varepsilon_0 m}$ is the *plasma frequency* of the free electron gas. Therefore we arrive at the desired result, the dielectric function of the free electron gas:

$$\varepsilon(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + i\gamma\omega}. \tag{1.20}$$

The real and imaginary components of this complex dielectric function $\varepsilon(\omega) = \varepsilon_1(\omega) + i\varepsilon_2(\omega)$ are given by

$$\varepsilon_1(\omega) = 1 - \frac{\omega_p^2\tau^2}{1 + \omega^2\tau^2} \tag{1.21a}$$

$$\varepsilon_2(\omega) = \frac{\omega_p^2\tau}{\omega(1 + \omega^2\tau^2)}, \tag{1.21b}$$

where we have used $\gamma = 1/\tau$. It is insightful to study (1.20) for a variety of different frequency regimes with respect to the collision frequency γ. We will limit ourselves here to frequencies $\omega < \omega_p$, where metals retain their metallic character. For large frequencies close to ω_p, the product $\omega\tau \gg 1$, leading to negligible damping. Here, $\varepsilon(\omega)$ is predominantly real, and

$$\varepsilon(\omega) = 1 - \frac{\omega_p^2}{\omega^2} \tag{1.22}$$

can be taken as the dielectric function of the undamped free electron plasma. Note that the behavior of noble metals in this frequency region is completely altered by interband transitions, leading to an increase in ε_2. The examples of gold and silver will be discussed below and in section 1.4.

We consider next the regime of very low frequencies, where $\omega \ll \tau^{-1}$. Hence, $\varepsilon_2 \gg \varepsilon_1$, and the real and the imaginary part of the complex refractive index are of comparable magnitude with

$$n \approx \kappa = \sqrt{\frac{\varepsilon_2}{2}} = \sqrt{\frac{\tau\omega_p^2}{2\omega}}. \tag{1.23}$$

In this region, metals are mainly absorbing, with an absorption coefficient of

$$\alpha = \left(\frac{2\omega_p^2\tau\omega}{c^2}\right)^{1/2}. \tag{1.24}$$

By introducing the dc-conductivity σ_0, this expression can be recast using $\sigma_0 = \frac{ne^2\tau}{m} = \omega_p^2\tau\varepsilon_0$ to

$$\alpha = \sqrt{2\sigma_0\omega\mu_0}. \tag{1.25}$$

The application of Beer's law of absorption implies that for low frequencies the fields fall off inside the metal as $e^{-z/\delta}$, where δ is the skin depth

$$\delta = \frac{2}{\alpha} = \frac{c}{\kappa\omega} = \sqrt{\frac{2}{\sigma_0\omega\mu_0}}. \tag{1.26}$$

A more rigorous discussion of the low-frequency behavior based on the Boltzmann transport equation [Marder, 2000] shows that this description is indeed valid as long as the mean free path of the electrons $l = v_F\tau \ll \delta$, where v_F is the Fermi velocity. At room temperature, for typical metals $l \approx 10$ nm and $\delta \approx 100$ nm, thus justifying the free-electron model. At low temperatures however, the mean free path can increase by many orders of magnitude, leading to changes in the penetration depth. This phenomenon is known as the anomalous skin effect.

If we use σ instead of ε for the description of the dielectric response of metals, we recognize that in the absorbing regime it is predominantly real, and the free charge velocity responds in phase with the driving field, as can be seen by integrating (1.17). At DC, relaxation effects of free charges are therefore conveniently described via the real DC-conductivity σ_0, whereas the response of bound charges is put into a dielectric constant ε_B, as discussed above in the examination of the interlinked nature between ε and σ.

At higher frequencies ($1 \leq \omega\tau \leq \omega_p\tau$), the complex refractive index is predominantly imaginary (leading to a reflection coefficient $R \approx 1$ [Jackson, 1999]), and σ acquires more and more complex character, blurring the boundary between free and bound charges. In terms of the optical response, $\sigma(\omega)$ enters expressions only in the combination (1.10) [Ashcroft and Mermin, 1976], due to the arbitrariness of the division between free and bound sets discussed above.

Whereas our description up to this point has assumed an ideal free-electron metal, we will now briefly compare the model with an example of a real metal important in the field of plasmonics (an extended discussion can be found in section 1.4). In the free-electron model, $\varepsilon \to 1$ at $\omega \gg \omega_p$. For the noble metals (e.g. Au, Ag, Cu), an extension to this model is needed in the region $\omega > \omega_p$ (where the response is dominated by free s electrons), since the filled d band close to the Fermi surface causes a highly polarized environment. This residual polarization due to the positive background of the ion cores can be described by adding the term $\mathbf{P}_\infty = \varepsilon_0(\varepsilon_\infty - 1)\mathbf{E}$ to (1.2a), where $\mathbf{P}$ now represents solely the polarization (1.18) due to free electrons. This effect is therefore described by a dielectric constant ε_∞ (usually $1 \leq \varepsilon_\infty \leq 10$), and we can write

$$\varepsilon(\omega) = \varepsilon_\infty - \frac{\omega_p^2}{\omega^2 + i\gamma\omega}. \tag{1.27}$$

The validity limits of the free-electron description (1.27) are illustrated for the case of gold in Fig. 1.1. It shows the real and imaginary components ε_1 and ε_2 for a dielectric function of this type, fitted to the experimentally determined dielectric function of gold [Johnson and Christy, 1972]. Clearly, at visible

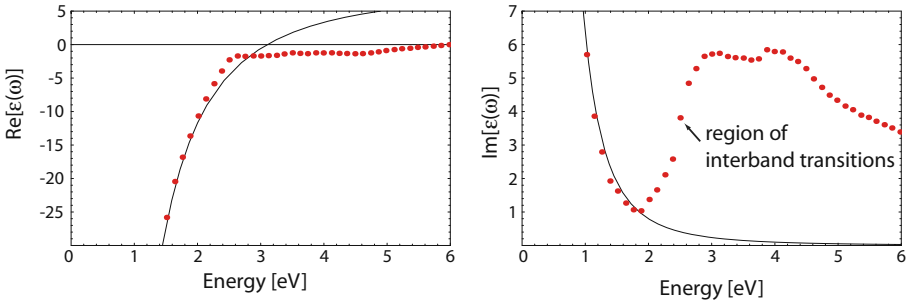

Figure 1.1. Dielectric function $\varepsilon(\omega)$ (1.27) of the free electron gas (solid line) fitted to the literature values of the dielectric data for gold [Johnson and Christy, 1972] (dots). Interband transitions limit the validity of this model at visible and higher frequencies.

frequencies the applicability of the free-electron model breaks down due to the occurrence of interband transitions, leading to an increase in ε_2. This will be discussed in more detail in section 1.4. The components of the complex refractive index corresponding to the fits presented in Fig. 1.1 are shown in Fig. 1.2.

It is instructive to link the dielectric function of the free electron plasma (1.20) to the classical Drude model [Drude, 1900] for the AC conductivity $\sigma(\omega)$ of metals. This can be achieved by recognizing that equation (1.16) can be rewritten as

$$\dot{\mathbf{p}} = -\frac{\mathbf{p}}{\tau} - e\mathbf{E}, \tag{1.28}$$

where $\mathbf{p} = m\dot{\mathbf{x}}$ is the momentum of an individual free electron. Via the same arguments presented above, we arrive at the following expression for the AC conductivity $\sigma = \frac{nep}{m}$,

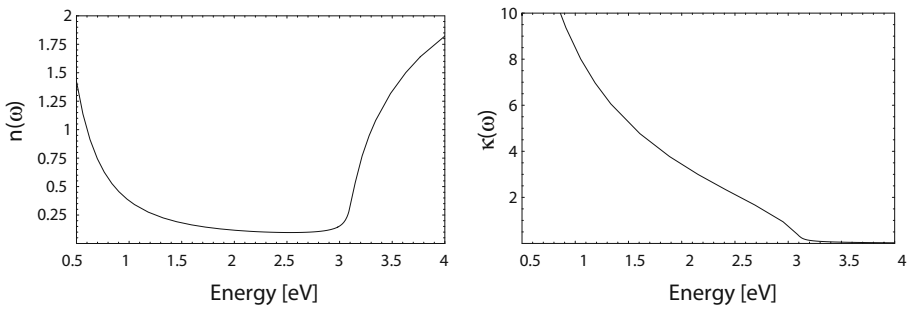

Figure 1.2. Complex refractive index corresponding to the free-electron dielectric function in Fig. 1.1.

$$\sigma(\omega) = \frac{\sigma_0}{1 - i\omega\tau}. \tag{1.29}$$

By comparing equation (1.20) and (1.29), we get

$$\varepsilon(\omega) = 1 + \frac{i\sigma(\omega)}{\varepsilon_0\omega}, \tag{1.30}$$

recovering the previous, general result of equation 1.10. The dielectric function of the free electron gas (1.20) is thus also known as the Drude model of the optical response of metals.

1.3 The Dispersion of the Free Electron Gas and Volume Plasmons

We now turn to a description of the thus-far omitted transparency regime $\omega > \omega_\mathrm{p}$ of the free electron gas model. Using equation (1.22) in (1.14), the dispersion relation of traveling waves evaluates to

$$\omega^2 = \omega_\mathrm{p}^2 + K^2c^2. \tag{1.31}$$

This relation is plotted for a generic free electron metal in Fig. 1.3. As can be seen, for $\omega < \omega_\mathrm{p}$ the propagation of transverse electromagnetic waves is forbidden inside the metal plasma. For $\omega > \omega_\mathrm{p}$ however, the plasma supports transverse waves propagating with a group velocity $v_\mathrm{g} = \mathrm{d}\omega/\mathrm{d}K < c$.

The significance of the plasma frequency ω_p can be further elucidated by recognizing that in the small damping limit, $\varepsilon(\omega_\mathrm{p}) = 0$ (for $\mathbf{K} = 0$). This excitation must therefore correspond to a collective longitudinal mode as shown in the discussion leading to (1.15). In this case, $\mathbf{D} = 0 = \varepsilon_0\mathbf{E} + \mathbf{P}$. We see that

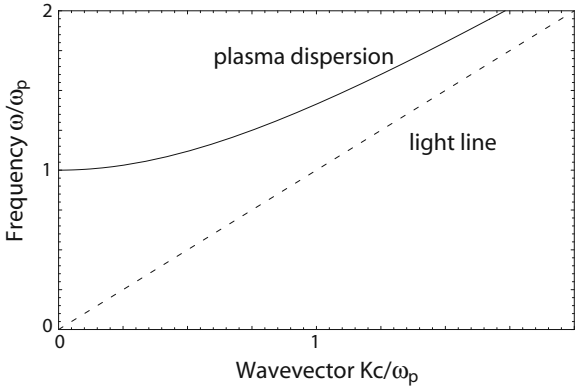

Figure 1.3. The dispersion relation of the free electron gas. Electromagnetic wave propagation is only allowed for $\omega > \omega_\mathrm{p}$.

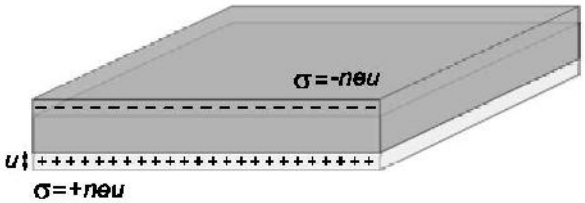

Figure 1.4. Longitudinal collective oscillations of the conduction electrons of a metal: Volume plasmons

at the plasma frequency the electric field is a pure depolarization field, with $\mathbf{E} = \frac{-\mathbf{P}}{\varepsilon_0}$.

The physical significance of the excitation at ω_p can be understood by considering the collective longitudinal oscillation of the conduction electron gas versus the fixed positive background of the ion cores in a plasma slab. Schematically indicated in Fig. 1.4, a collective displacement of the electron cloud by a distance u leads to a surface charge density $\sigma = \pm neu$ at the slab boundaries. This establishes a homogeneous electric field $\mathbf{E} = \frac{neu}{\varepsilon_0}$ inside the slab. Thus, the displaced electrons experience a restoring force, and their movement can be described by the equation of motion $nm\ddot{u} = -ne\mathbf{E}$. Inserting the expression for the electric field, this leads to

$$nm\ddot{u} = -\frac{n^2e^2u}{\varepsilon_0} \tag{1.32a}$$

$$\ddot{u} + \omega_\mathrm{p}^2 u = 0. \tag{1.32b}$$

The plasma frequency ω_p can thus be recognized as the natural frequency of a free oscillation of the electron sea. Note that our derivation has assumed that all electrons move in phase, thus ω_p corresponds to the oscillation frequency in the long-wavelength limit where $\mathbf{K} = 0$. The quanta of these charge oscillations are called plasmons (or *volume* plasmons, to distinguish them from *surface* and *localized* plasmons, which will be discussed in the remainder of this text). Due to the longitudinal nature of the excitation, volume plasmons do not couple to transverse electromagnetic waves, and can only be excited by particle impact. Another consequence of this is that their decay occurs only via energy transfer to single electrons, a process known as Landau damping.

Experimentally, the plasma frequency of metals typically is determined via electron loss spectroscopy experiments, where electrons are passed through thin metallic foils. For most metals, the plasma frequency is in the ultraviolet regime: ω_p is on the order of $5 - 15$ eV, depending on details of the band structure [Kittel, 1996]. As an aside, we want to note that such longitudinal os-

cillations can also be excited in dielectrics, in which case the valence electrons oscillate collectively with respect to the ion cores.

In addition to the in-phase oscillation at ω_p, there exists a whole class of longitudinal oscillations at higher frequencies with finite wavevectors, for which (1.15) is fulfilled. The derivation of the dispersion relation of volume plasmons is beyond the scope of this treatment and can be found in many textbooks on condensed matter physics [Marder, 2000, Kittel, 1996]. Up to quadratic order in $\mathbf{K}$,

$$\omega^2 = \omega_\mathrm{p}^2 + \frac{6E_\mathrm{F}K^2}{5m}, \tag{1.33}$$

where E_F is the Fermi energy. Practically, the dispersion can be measured using inelastic scattering experiments such as electron energy loss spectroscopy (EELS).

1.4 Real Metals and Interband Transitions

We have already on several occasions stated that the dielectric function (1.20) of the Drude model adequately describes the optical response of metals only for photon energies below the threshold of transitions between electronic bands. For some of the noble metals, interband effects already start to occur for energies in excess of 1 eV (corresponding to a wavelength $\lambda \approx 1\ \mu$m). As examples, Figs. 1.1 and 1.5 show the real and the imaginary parts $\varepsilon_1(\omega)$, $\varepsilon_2(\omega)$ of the dielectric function for gold and silver [Johnson and Christy, 1972] and Drude model fits to the data. Clearly, this model is not adequate for describing either ε_1 or ε_2 at high frequencies, and in the case of gold, its validity breaks down already at the boundary between the near-infrared and the visible.

We limit this comparison between the Drude model and the dielectric response of real metals to the cases of gold and silver, the most important metals for plasmonic studies in the visible and near-infrared. Above their respective

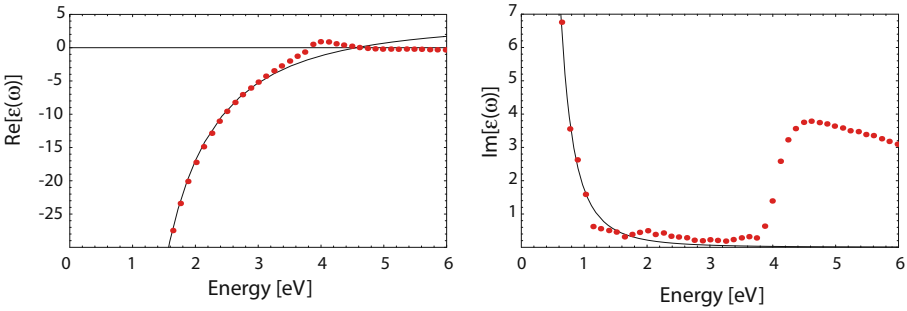

Figure 1.5. The real and imaginary part of $\varepsilon(\omega)$ for silver determined by Johnson and Christy [Johnson and Christy, 1972] (dots) and a Drude model fit to the data.

band edge thresholds, photons are very efficient in inducing interband transitions, where electrons from the filled band below the Fermi surface are excited to higher bands. Theoretically, these can be described using the same approach used for direct band transitions in semiconductors [Ashcroft and Mermin, 1976, Marder, 2000], and we will not embark on a more detailed discussion. The main consequence of these processes concerning surface plasmon polaritons is an increased damping and competition between the two excitations at visible frequencies.

For practical purposes, a big advantage of the Drude model is that it can easily be incorporated into time-domain based numerical solvers for Maxwell's equations, such as the finite-difference time-domain (FDTD) scheme [Kashiwa and Fukai, 1990], via the direct calculation of the induced currents using (1.16). Its inadequacy in describing the optical properties of gold and silver at visible frequencies can be overcome by replacing (1.16) by

$$m\ddot{\mathbf{x}} + m\gamma\dot{\mathbf{x}} + m\omega_0^2\mathbf{x} = -e\mathbf{E}. \tag{1.34}$$

Interband transitions are thus described using the classical picture of a bound electron with resonance frequency ω_0, and (1.34) can then be used to calculate the resulting polarization. We note that a number of equations of this form might have to be solved (each resulting in a separate contribution to the total polarization) in order to model $\varepsilon(\omega)$ for noble metals accurately. Each of these equations leads to a Lorentz-oscillator term of the form $\frac{A_i}{\omega_i^2-\omega^2-i\gamma_i\omega}$ added to the free-electron result (1.20) [Vial et al., 2005].

1.5 The Energy of the Electromagnetic Field in Metals

We finish this chapter by taking a brief look at the energy of the electromagnetic field inside metals, or more generally inside dispersive media. Since the amount of field localization is often quantified in terms of the electromagnetic energy distribution, a careful consideration of the effects of dispersion is necessary. For a linear medium with no dispersion or losses (i.e. (1.5) holds), the total energy density of the electromagnetic field can be written as [Jackson, 1999]

$$u = \frac{1}{2}(\mathbf{E}\cdot\mathbf{D} + \mathbf{B}\cdot\mathbf{H}). \tag{1.35}$$

This expression enters together with the Poynting vector of energy flow $\mathbf{S} = \mathbf{E}\times\mathbf{H}$ into the conservation law

$$\frac{\partial u}{\partial t} + \nabla\cdot\mathbf{S} = -\mathbf{J}\cdot\mathbf{E}, \tag{1.36}$$

relating changes in electromagnetic energy density to energy flow and absorption inside the material.

In the following, we will concentrate on the contribution u_{E} of the electric field $\mathbf{E}$ to the total electromagnetic energy density. In metals, ε is complex and frequency-dependent due to dispersion, and (1.35) does not apply. For a field consisting of monochromatic components, Landau and Lifshitz have shown that the conservation law (1.36) can be held up if u_{E} is replaced by an *effective* electric energy density u_{eff}, defined as

$$u_{\mathrm{eff}} = \frac{1}{2}\mathrm{Re}\left[\frac{d(\omega\varepsilon)}{d\omega}\right]_{\omega_0} \langle \mathbf{E}(\mathbf{r},t)\cdot\mathbf{E}(\mathbf{r},t)\rangle\,, \tag{1.37}$$

where $\langle \mathbf{E}(\mathbf{r},t)\cdot\mathbf{E}(\mathbf{r},t)\rangle$ signifies field-averaging over one optical cycle, and ω_0 is the frequency of interest. This expression is valid if $\mathbf{E}$ is only appreciable in a narrow frequency range around ω_0, and the fields are slowly-varying compared to a timescale $1/\omega_0$. Furthermore, it is assumed that $|\varepsilon_2| \ll |\varepsilon_1|$, so that absorption is small. We note that additional care must be taken with the correct calculation of absorption on the right side of (1.36), where $\mathbf{J}\cdot\mathbf{E}$ should be replaced by $\omega_0\mathrm{Im}\left[\varepsilon(\omega_0)\right]\langle \mathbf{E}(\mathbf{r},t)\cdot\mathbf{E}(\mathbf{r},t)\rangle$ if the dielectric response of the metal is completely described via $\varepsilon(\omega)$ [Jackson, 1999], in line with the discussion surrounding (1.10).

The requirement of low absorption limits (1.37) to visible and near-infrared frequencies, but not to lower frequencies or the regime of interband effects where $|\varepsilon_2| > |\varepsilon_1|$. However, the electric field energy can also be determined by taking the electric polarization explicitly into account, in the form described by (1.16) [Loudon, 1970, Ruppin, 2002]. The obtained expression for the electric field energy of a material described by a free-electron-type dielectric function $\varepsilon = \varepsilon_1 + i\varepsilon_2$ of the form (1.20) is

$$u_{\mathrm{eff}} = \frac{\varepsilon_0}{4}\left(\varepsilon_1 + \frac{2\omega\varepsilon_2}{\gamma}\right)|\mathbf{E}|^2\,, \tag{1.38}$$

where an additional factor $1/2$ is included due to an implicit assumption of harmonic time dependence of the oscillating fields. For negligible ε_2, it can be shown that (1.38) reduces as expected to (1.37) for time-harmonic fields. We will use (1.38) in chapter 2 when discussing the amount of energy localization in fields localized at metallic surfaces.

Chapter 2

SURFACE PLASMON POLARITONS AT METAL / INSULATOR INTERFACES

Surface plasmon polaritons are electromagnetic excitations propagating at the interface between a dielectric and a conductor, evanescently confined in the perpendicular direction. These electromagnetic surface waves arise via the coupling of the electromagnetic fields to oscillations of the conductor's electron plasma. Taking the wave equation as a starting point, this chapter describes the fundamentals of surface plasmon polaritons both at single, flat interfaces and in metal/dielectric multilayer structures. The surface excitations are characterized in terms of their dispersion and spatial profile, together with a detailed discussion of the quantification of field confinement. Applications of surface plasmon polaritons in waveguiding will be deferred to chapter 7.

2.1 The Wave Equation

In order to investigate the physical properties of surface plasmon polaritons (SPPs), we have to apply Maxwell's equations (1.1) to the flat interface between a conductor and a dielectric. To present this discussion most clearly, it is advantageous to cast the equations first in a general form applicable to the guiding of electromagnetic waves, the *wave equation*.

As we have seen in chapter 1, in the absence of external charge and current densities, the curl equations (1.1c, 1.1d) can be combined to yield

$$\nabla \times \nabla \times \mathbf{E} = -\mu_0 \frac{\partial^2 \mathbf{D}}{\partial t^2}. \tag{2.1}$$

Using the identities $\nabla \times \nabla \times \mathbf{E} \equiv \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E}$ as well as $\nabla \cdot (\varepsilon \mathbf{E}) \equiv \mathbf{E} \cdot \nabla \varepsilon + \varepsilon \nabla \cdot \mathbf{E}$, and remembering that due to the absence of external stimuli $\nabla \cdot \mathbf{D} = 0$, (2.1) can be rewritten as

$$\nabla\left(-\frac{1}{\varepsilon}\mathbf{E}\cdot\nabla\varepsilon\right)-\nabla^2\mathbf{E}=-\mu_0\varepsilon_0\varepsilon\frac{\partial^2\mathbf{E}}{\partial t^2}. \tag{2.2}$$

For negligible variation of the dielectric profile $\varepsilon = \varepsilon(\mathbf{r})$ over distances on the order of one optical wavelength, (2.2) simplifies to the central equation of electromagnetic wave theory,

$$\nabla^2\mathbf{E}-\frac{\varepsilon}{c^2}\frac{\partial^2\mathbf{E}}{\partial t^2}=0. \tag{2.3}$$

Practically, this equation has to be solved separately in regions of constant ε, and the obtained solutions have to been matched using appropriate boundary conditions. To cast (2.3) in a form suitable for the description of confined propagating waves, we proceed in two steps. First, we assume in all generality a harmonic time dependence $\mathbf{E}(\mathbf{r},t)=\mathbf{E}(\mathbf{r})\mathrm{e}^{-i\omega t}$ of the electric field. Inserted into (2.3), this yields

$$\nabla^2\mathbf{E}+k_0^2\varepsilon\mathbf{E}=0, \tag{2.4}$$

where $k_0=\frac{\omega}{c}$ is the wave vector of the propagating wave in vacuum. Equation (2.4) is known as the *Helmholtz equation*.

Next, we have to define the propagation geometry. We assume for simplicity a one-dimensional problem, i.e. ε depends only on one spatial coordinate. Specifically, the waves propagate along the x-direction of a cartesian coordinate system, and show no spatial variation in the perpendicular, in-plane y-direction (see Fig. 2.1); therefore $\varepsilon = \varepsilon(z)$. Applied to electromagnetic surface problems, the plane $z = 0$ coincides with the interface sustaining the

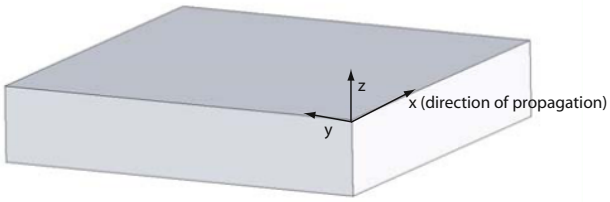

Figure 2.1. Definition of a planar waveguide geometry. The waves propagate along the x-direction in a cartesian coordinate system.

propagating waves,which can now be described as $\mathbf{E}(x, y, z) = \mathbf{E}(z)e^{i\beta x}$. The complex parameter $\beta = k_x$ is called the *propagation constant* of the traveling waves and corresponds to the component of the wave vector in the direction of propagation. Inserting this expression into (2.4) yields the desired form of the wave equation

$$\frac{\partial^2 \mathbf{E}(z)}{\partial z^2} + \left(k_0^2 \varepsilon - \beta^2\right) \mathbf{E} = 0. \tag{2.5}$$

Naturally, a similar equation exists for the magnetic field $\mathbf{H}$.

Equation (2.5) is the starting point for the general analysis of guided electromagnetic modes in waveguides, and an extended discussion of its properties and applications can be found in [Yariv, 1997] and similar treatments of photonics and optoelectronics. In order to use the wave equation for determining the spatial field profile and dispersion of propagating waves, we now need to find explicit expressions for the different field components of $\mathbf{E}$ and $\mathbf{H}$. This can be achieved in a straightforward way using the curl equations (1.1c, 1.1d).

For harmonic time dependence $\left(\frac{\partial}{\partial t} = -i\omega\right)$, we arrive at the following set of coupled equations

$$\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} = i\omega\mu_0 H_x \tag{2.6a}$$

$$\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} = i\omega\mu_0 H_y \tag{2.6b}$$

$$\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = i\omega\mu_0 H_z \tag{2.6c}$$

$$\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} = -i\omega\varepsilon_0\varepsilon E_x \tag{2.6d}$$

$$\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} = -i\omega\varepsilon_0\varepsilon E_y \tag{2.6e}$$

$$\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} = -i\omega\varepsilon_0\varepsilon E_z. \tag{2.6f}$$

For propagation along the x-direction $\left(\frac{\partial}{\partial x} = i\beta\right)$ and homogeneity in the y-direction $\left(\frac{\partial}{\partial y} = 0\right)$, this system of equation simplifies to

$$\frac{\partial E_y}{\partial z} = -i\omega\mu_0 H_x \tag{2.7a}$$

$$\frac{\partial E_x}{\partial z} - i\beta E_z = i\omega\mu_0 H_y \tag{2.7b}$$

$$i\beta E_y = i\omega\mu_0 H_z \tag{2.7c}$$

$$\frac{\partial H_y}{\partial z} = i\omega\varepsilon_0\varepsilon E_x \tag{2.7d}$$

$$\frac{\partial H_x}{\partial z} - i\beta H_z = -i\omega\varepsilon_0\varepsilon E_y \tag{2.7e}$$

$$i\beta H_y = -i\omega\varepsilon_0\varepsilon E_z. \tag{2.7f}$$

It can easily be shown that this system allows two sets of self-consistent solutions with different polarization properties of the propagating waves. The first set are the transverse magnetic (TM or p) modes, where only the field components E_x, E_z and H_y are nonzero, and the second set the transverse electric (TE or s) modes, with only H_x, H_z and E_y being nonzero.

For TM modes, the system of governing equations (2.7) reduces to

$$E_x = -i\frac{1}{\omega\varepsilon_0\varepsilon}\frac{\partial H_y}{\partial z} \tag{2.8a}$$

$$E_z = -\frac{\beta}{\omega\varepsilon_0\varepsilon}H_y, \tag{2.8b}$$

and the wave equation for TM modes is

$$\frac{\partial^2 H_y}{\partial z^2} + \left(k_0^2\varepsilon - \beta^2\right) H_y = 0. \tag{2.8c}$$

For TE modes the analogous set is

$$H_x = i\frac{1}{\omega\mu_0}\frac{\partial E_y}{\partial z} \tag{2.9a}$$

$$H_z = \frac{\beta}{\omega\mu_0}E_y, \tag{2.9b}$$

with the TE wave equation

$$\frac{\partial^2 E_y}{\partial z^2} + \left(k_0^2\varepsilon - \beta^2\right) E_y = 0. \tag{2.9c}$$

With these equations at our disposal, we are now in a position to embark on the description of surface plasmon polaritons.

2.2 Surface Plasmon Polaritons at a Single Interface

The most simple geometry sustaining SPPs is that of a single, flat interface (Fig. 2.2) between a dielectric, non-absorbing half space ($z > 0$) with positive real dielectric constant ε_2 and an adjacent conducting half space ($z < 0$) described via a dielectric function $\varepsilon_1(\omega)$. The requirement of metallic character implies that $\mathrm{Re}\,[\varepsilon_1] < 0$. As shown in chapter 1, for metals this condition is fulfilled at frequencies below the bulk plasmon frequency ω_p. We want to look for propagating wave solutions confined to the interface, i.e. with evanescent decay in the perpendicular z-direction.

Let us first look at TM solutions. Using the equation set (2.8) in both half spaces yields

$$H_y(z) = A_2 \mathrm{e}^{i\beta x} \mathrm{e}^{-k_2 z} \tag{2.10a}$$

$$E_x(z) = i A_2 \frac{1}{\omega \varepsilon_0 \varepsilon_2} k_2 e^{i\beta x} \mathrm{e}^{-k_2 z} \tag{2.10b}$$

$$E_z(z) = -A_1 \frac{\beta}{\omega \varepsilon_0 \varepsilon_2} e^{i\beta x} \mathrm{e}^{-k_2 z} \tag{2.10c}$$

for $z > 0$ and

$$H_y(z) = A_1 \mathrm{e}^{i\beta x} \mathrm{e}^{k_1 z} \tag{2.11a}$$

$$E_x(z) = -i A_1 \frac{1}{\omega \varepsilon_0 \varepsilon_1} k_1 e^{i\beta x} \mathrm{e}^{k_1 z} \tag{2.11b}$$

$$E_z(z) = -A_1 \frac{\beta}{\omega \varepsilon_0 \varepsilon_1} e^{i\beta x} \mathrm{e}^{k_1 z} \tag{2.11c}$$

for $z < 0$. $k_i \equiv k_{z,i} (i = 1, 2)$ is the component of the wave vector perpendicular to the interface in the two media. Its reciprocal value, $\hat{z} = 1/|k_z|$, defines the evanescent decay length of the fields perpendicular to the interface,

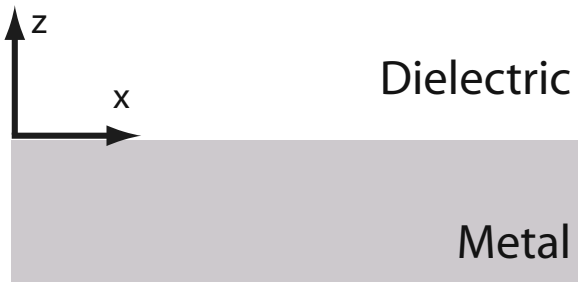

Figure 2.2. Geometry for SPP propagation at a single interface between a metal and a dielectric.

which quantifies the confinement of the wave. Continuity of H_y and $\varepsilon_i E_z$ at the interface requires that $A_1 = A_2$ and

$$\frac{k_2}{k_1} = -\frac{\varepsilon_2}{\varepsilon_1}. \tag{2.12}$$

Note that with our convention of the signs in the exponents in (2.10,2.11), confinement to the surface demands $\mathrm{Re}\,[\varepsilon_1] < 0$ if $\varepsilon_2 > 0$ - the surface waves exist only at interfaces between materials with opposite signs of the real part of their dielectric permittivities, i.e. between a conductor and an insulator. The expression for H_y further has to fulfill the wave equation (2.8c), yielding

$$k_1^2 = \beta^2 - k_0^2 \varepsilon_1 \tag{2.13a}$$

$$k_2^2 = \beta^2 - k_0^2 \varepsilon_2. \tag{2.13b}$$

Combining this and (2.12) we arrive at the central result of this section, the dispersion relation of SPPs propagating at the interface between the two half spaces

$$\beta = k_0 \sqrt{\frac{\varepsilon_1 \varepsilon_2}{\varepsilon_1 + \varepsilon_2}}. \tag{2.14}$$

This expression is valid for both real and complex ε_1, i.e. for conductors without and with attenuation.

Before discussing the properties of the dispersion relation (2.14) in more detail, we now briefly analyze the possibility of TE surface modes. Using (2.9), the respective expressions for the field components are

$$E_y(z) = A_2 \mathrm{e}^{i\beta x} \mathrm{e}^{-k_2 z} \tag{2.15a}$$

$$H_x(z) = -i A_2 \frac{1}{\omega \mu_0} k_2 e^{i\beta x} \mathrm{e}^{-k_2 z} \tag{2.15b}$$

$$H_z(z) = A_2 \frac{\beta}{\omega \mu_0} e^{i\beta x} \mathrm{e}^{-k_2 z} \tag{2.15c}$$

for $z > 0$ and

$$E_y(z) = A_1 \mathrm{e}^{i\beta x} \mathrm{e}^{k_1 z} \tag{2.16a}$$

$$H_x(z) = i A_1 \frac{1}{\omega \mu_0} k_1 e^{i\beta x} \mathrm{e}^{k_1 z} \tag{2.16b}$$

$$H_z(z) = A_1 \frac{\beta}{\omega \mu_0} e^{i\beta x} \mathrm{e}^{k_1 z} \tag{2.16c}$$

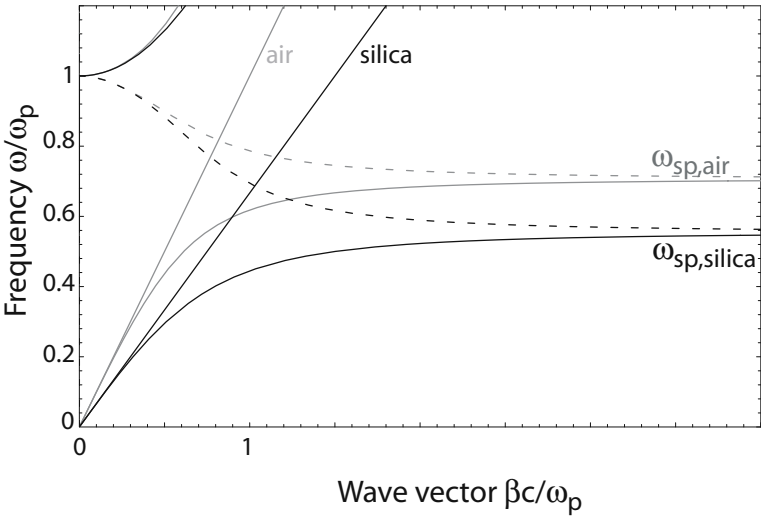

Figure 2.3. Dispersion relation of SPPs at the interface between a Drude metal with negligible collision frequency and air (gray curves) and silica (black curves).

for $z < 0$. Continuity of E_y and H_x at the interface leads to the condition

$$A_1 (k_1 + k_2) = 0. \tag{2.17}$$

Since confinement to the surface requires $\mathrm{Re}\,[k_1] > 0$ and $\mathrm{Re}\,[k_2] > 0$, this condition is only fulfilled if $A_1 = 0$, so that also $A_2 = A_1 = 0$. Thus, no surface modes exist for TE polarization. *Surface plasmon polaritons only exist for TM polarization.*

We now want to examine the properties of SPPs by taking a closer look at their dispersion relation. Fig. 2.3 shows plots of (2.14) for a metal with negligible damping described by the real Drude dielectric function (1.22) for an air ($\varepsilon_2 = 1$) and a fused silica ($\varepsilon_2 = 2.25$) interface. In this plot, the frequency ω is normalized to the plasma frequency ω_p, and both the real (continuous curves) and the imaginary part (broken curves) of the wave vector β are shown. Due to their bound nature, the SPP excitations correspond to the part of the dispersion curves lying to the right of the respective light lines of air and silica. Thus, special phase-matching techniques such as grating or prism coupling are required for their excitation via three-dimensional beams, which will be discussed in chapter 3. Radiation into the metal occurs in the transparency regime $\omega > \omega_\mathrm{p}$ as mentioned in chapter 1. Between the regime of the bound and radiative modes, a frequency gap region with purely imaginary β prohibiting propagation exists.

For small wave vectors corresponding to low (mid-infrared or lower) frequencies, the SPP propagation constant is close to k_0 at the light line, and the

waves extend over many wavelengths into the dielectric space. In this regime, SPPs therefore acquire the nature of a grazing-incidence light field, and are also known as *Sommerfeld-Zenneck waves* [Goubau, 1950].

In the opposite regime of large wave vectors, the frequency of the SPPs approaches the characteristic *surface plasmon frequency*

$$\omega_{\mathrm{sp}} = \frac{\omega_{\mathrm{p}}}{\sqrt{1+\varepsilon_2}}, \tag{2.18}$$

as can be shown by inserting the free-electron dielectric function (1.20) into (2.14). In the limit of negligible damping of the conduction electron oscillation (implying $\mathrm{Im}\,[\varepsilon_1(\omega)] = 0$), the wave vector β goes to infinity as the frequency approaches ω_{sp}, and the group velocity $v_{\mathrm{g}} \to 0$. The mode thus acquires electrostatic character, and is known as the *surface plasmon*. It can indeed be obtained via a straightforward solution of the Laplace equation $\nabla^2\phi = 0$ for the single interface geometry of Fig. 2.2, where ϕ is the electric potential. A solution that is wavelike in the x-direction and exponentially decaying in the z-direction is given by

$$\phi(z) = A_2 \mathrm{e}^{i\beta x}\mathrm{e}^{-k_2 z} \tag{2.19}$$

for $z > 0$ and

$$\phi(z) = A_1 \mathrm{e}^{i\beta x}\mathrm{e}^{k_1 z} \tag{2.20}$$

for $z < 0$. $\nabla^2\phi = 0$ requires that $k_1 = k_2 = \beta$: the exponential decay lengths $\left|\hat{z}\right| = 1/k_z$ into the dielectric and into the metal are equal. Continuity of ϕ and $\varepsilon\partial\phi/\partial z$ ensure continuity of the tangential field components and the normal components of the dielectric displacement and require that $A_1 = A_2$ and additionally

$$\varepsilon_1(\omega) + \varepsilon_2 = 0. \tag{2.21}$$

For a metal described by a dielectric function of the form (1.22), this condition is fulfilled at ω_{sp}. Comparison of (2.21) and (2.14) show that the surface plasmon is indeed the limiting form of a SPP as $\beta \to \infty$.

The above discussions of Fig. 2.3 have assumed an ideal conductor with $\mathrm{Im}\,[\varepsilon_1] = 0$. Excitations of the conduction electrons of real metals however suffer both from free-electron and interband damping. Therefore, $\varepsilon_1(\omega)$ is complex, and with it also the SPP propagation constant β. The traveling SPPs are damped with an energy attenuation length (also called propagation length) $L = (2\mathrm{Im}\left[\beta\right])^{-1}$, typically between 10 and 100 μm in the visible regime, depending upon the metal/dielectric configuration in question.

Fig. 2.4 shows as an example the dispersion relation of SPPs propagating at a silver/air and silver/silica interface, with the dielectric function $\varepsilon_1(\omega)$ of silver

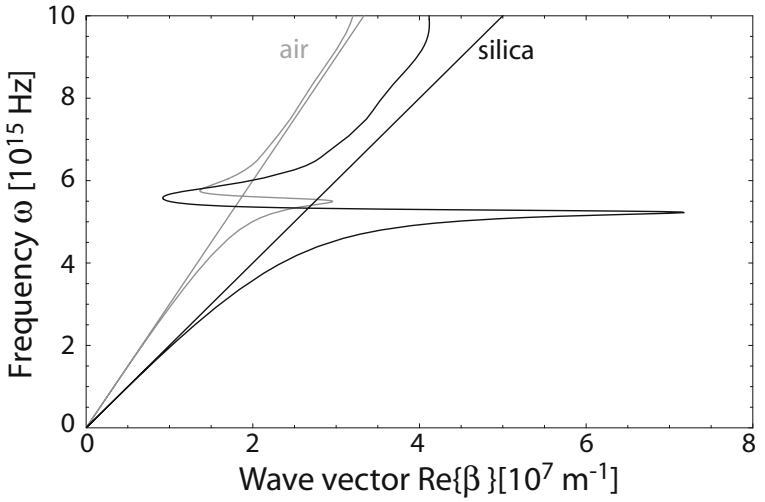

Figure 2.4. Dispersion relation of SPPs at a silver/air (gray curve) and silver/silica (black curve) interface. Due to the damping, the wave vector of the bound SPPs approaches a finite limit at the surface plasmon frequency.

taken from the data obtained by Johnson and Christy [Johnson and Christy, 1972]. Compared with the dispersion relation of completely undamped SPPs depicted in Fig. 2.3, it can be seen that the bound SPPs approach now a maximum, *finite* wave vector at the the surface plasmon frequency ω_{sp} of the system. This limitation puts a lower bound both on the wavelength $\lambda_{\mathrm{sp}} = 2\pi/\mathrm{Re}\left[\beta\right]$ of the surface plasmon and also on the amount of mode confinement perpendicular to the interface, since the SPP fields in the dielectric fall off as $e^{-|k_z||z|}$ with $k_z = \sqrt{\beta^2 - \varepsilon_2\left(\frac{\omega}{c}\right)^2}$. Also, the *quasibound*, leaky part of the dispersion relation between ω_{sp} and ω_{p} is now allowed, in contrast to the case of an ideal conductor, where $\mathrm{Re}\left[\beta\right] = 0$ in this regime (Fig. 2.3).

We finish this section by providing an example of the propagation length L and the energy confinement (quantified by $\hat{z}$) in the dielectric. As is evident from the dispersion relation, both show a strong dependence on frequency. SPPs at frequencies close to ω_{sp} exhibit large field confinement to the interface and a subsequent small propagation distance due to increased damping. Using the theoretical treatment outlined above, we see that SPPs at a silver/air interface at $\lambda_0 = 450$ nm for example have $L \approx 16\ \mu$m and $\hat{z} \approx 180$ nm. At $\lambda_0 \approx 1.5\ \mu$m however, $L \approx 1080\ \mu$m and $\hat{z} \approx 2.6\ \mu$m. The better the confinement, the lower the propagation length. This characteristic trade-off between localization and loss is typical for plasmonics. We note that field-confinement below the diffraction limit of half the wavelength in the dielectric can be achieved close to ω_{sp}. In the metal itself, the fields fall off over distances

on the order of 20 nm over a wide frequency range spanning from the visible to the infrared.

2.3 Multilayer Systems

We now turn our attention to SPPs in multilayers consisting of alternating conducting and dielectric thin films. In such a system, each single interface can sustain bound SPPs. When the separation between adjacent interfaces is comparable to or smaller than the decay length $\hat{z}$ of the interface mode, interactions between SPPs give rise to coupled modes. In order to elucidate the general properties of coupled SPPs, we will focus on two specific three-layer systems of the geometry depicted in Fig. 2.5: Firstly, a thin metallic layer (I) sandwiched between two (infinitely) thick dielectric claddings (II, III), an insulator/metal/insulator (IMI) heterostructure, and secondly a thin dielectric core layer (I) sandwiched between two metallic claddings (II, III), a metal/insulator/metal (MIM) heterostructure.

Since we are here only interested in the lowest-order bound modes, we start with a general description of TM modes that are non-oscillatory in the z-direction normal to the interfaces using (2.8). For $z > a$, the field components are

$$H_y = A\mathrm{e}^{i\beta x}\mathrm{e}^{-k_3 z} \tag{2.22a}$$

$$E_x = iA\frac{1}{\omega\varepsilon_0\varepsilon_3}k_3\mathrm{e}^{i\beta x}\mathrm{e}^{-k_3 z} \tag{2.22b}$$

$$E_z = -A\frac{\beta}{\omega\varepsilon_0\varepsilon_3}\mathrm{e}^{i\beta x}\mathrm{e}^{-k_3 z}, \tag{2.22c}$$

while for $z < -a$ we get

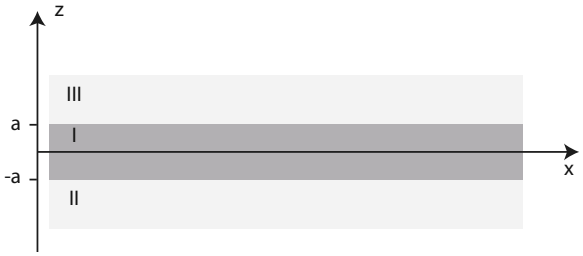

Figure 2.5. Geometry of a three-layer system consisting of a thin layer I sandwiched between two infinite half spaces II and III.

$$H_y = B\mathrm{e}^{i\beta x}\mathrm{e}^{k_2 z} \tag{2.23a}$$

$$E_x = -iB\frac{1}{\omega\varepsilon_0\varepsilon_2}k_2\mathrm{e}^{i\beta x}\mathrm{e}^{k_2 z} \tag{2.23b}$$

$$E_z = -B\frac{\beta}{\omega\varepsilon_0\varepsilon_2}\mathrm{e}^{i\beta x}\mathrm{e}^{k_2 z}. \tag{2.23c}$$

Thus, we demand that the fields decay exponentially in the claddings (II) and (III). Note that for simplicity as before we denote the component of the wave vector perpendicular to the interfaces simply as $k_i \equiv k_{z,i}$.

In the core region $-a < z < a$, the modes localized at the bottom and top interface couple, yielding

$$H_y = C\mathrm{e}^{i\beta x}\mathrm{e}^{k_1 z} + D\mathrm{e}^{i\beta x}\mathrm{e}^{-k_1 z} \tag{2.24a}$$

$$E_x = -iC\frac{1}{\omega\varepsilon_0\varepsilon_1}k_1\mathrm{e}^{i\beta x}\mathrm{e}^{k_1 z} + iD\frac{1}{\omega\varepsilon_0\varepsilon_1}k_1\mathrm{e}^{i\beta x}\mathrm{e}^{-k_1 z} \tag{2.24b}$$

$$E_z = C\frac{\beta}{\omega\varepsilon_0\varepsilon_1}\mathrm{e}^{i\beta x}\mathrm{e}^{k_1 z} + D\frac{\beta}{\omega\varepsilon_0\varepsilon_1}\mathrm{e}^{i\beta x}\mathrm{e}^{-k_1 z}. \tag{2.24c}$$

The requirement of continutity of H_y and E_x leads to

$$A\mathrm{e}^{-k_3 a} = C\mathrm{e}^{k_1 a} + D\mathrm{e}^{-k_1 a} \tag{2.25a}$$

$$\frac{A}{\varepsilon_3}k_3\mathrm{e}^{-k_3 a} = -\frac{C}{\varepsilon_1}k_1\mathrm{e}^{k_1 a} + \frac{D}{\varepsilon_1}k_1\mathrm{e}^{-k_1 a} \tag{2.25b}$$

at $z = a$ and

$$B\mathrm{e}^{-k_2 a} = C\mathrm{e}^{-k_1 a} + D\mathrm{e}^{k_1 a} \tag{2.26a}$$

$$-\frac{B}{\varepsilon_2}k_2\mathrm{e}^{-k_2 a} = -\frac{C}{\varepsilon_1}k_1\mathrm{e}^{-k_1 a} + \frac{D}{\varepsilon_1}k_1\mathrm{e}^{k_1 a} \tag{2.26b}$$

at $z = -a$, a linear system of four coupled equations. H_y further has to fulfill the wave equation (2.8c) in the three distinct regions, via

$$k_i^2 = \beta^2 - k_0^2\varepsilon_i \tag{2.27}$$

for $i = 1, 2, 3$. Solving this system of linear equations results in an implicit expression for the dispersion relation linking β and ω via

$$\mathrm{e}^{-4k_1 a} = \frac{k_1/\varepsilon_1 + k_2/\varepsilon_2}{k_1/\varepsilon_1 - k_2/\varepsilon_2}\frac{k_1/\varepsilon_1 + k_3/\varepsilon_3}{k_1/\varepsilon_1 - k_3/\varepsilon_3}. \tag{2.28}$$

We note that for infinite thickness ($a \to \infty$), (2.28) reduces to (2.12), the equation of two uncoupled SPP at the respective interfaces.

We will from this point onwards consider the interesting special case where the sub- and the superstrates (II) and (III) are equal in terms of their dielectric response, i.e. $\varepsilon_2 = \varepsilon_3$ and thus $k_2 = k_3$. In this case, the dispersion relation (2.28) can be split into a pair of equations, namely

$$\tanh k_1 a = -\frac{k_2 \varepsilon_1}{k_1 \varepsilon_2} \tag{2.29a}$$

$$\tanh k_1 a = -\frac{k_1 \varepsilon_2}{k_2 \varepsilon_1}. \tag{2.29b}$$

It can be shown that equation (2.29a) describes modes of odd vector parity ($E_x(z)$ is odd, $H_y(z)$ and $E_z(z)$ are even functions), while (2.29b) describes modes of even vector parity ($E_x(z)$ is even function, $H_y(z)$ and $E_z(z)$ are odd).

The dispersion relations (2.29a, 2.29b) can now be applied to IMI and MIM structures to investigate the properties of the coupled SPP modes in these two systems. We first start with the IMI geometry - a thin metallic film of thickness $2a$ sandwiched between two insulating layers. In this case $\varepsilon_1 = \varepsilon_1(\omega)$ represents the dielectric function of the metal, and ε_2 the positive, real dielectric constant of the insulating sub- and superstrates. As an example, Fig. 2.6 shows the dispersion relations of the odd and even modes (2.29a, 2.29b) for an air/silver/air geometry for two different thicknesses of the silver thin film. For

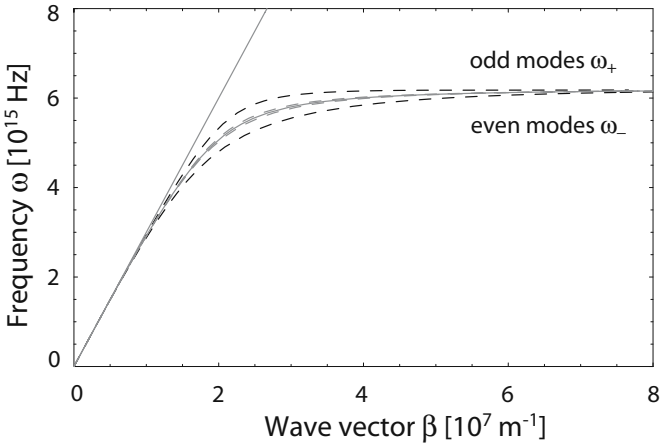

Figure 2.6. Dispersion relation of the coupled odd and even modes for an air/silver/air multilayer with a metal core of thickness 100 nm (dashed gray curves) and 50 nm (dashed black curves). Also shown is the dispersion of a single silver/air interface (gray curve). Silver is modeled as a Drude metal with negligible damping.

simplicity, here the dielectric function of silver is approximated via a Drude model with negligible damping ($\varepsilon(\omega)$ real and of the form (1.22)), so that $\mathrm{Im}\left[\beta\right] = 0$.

As can be seen, the odd modes have frequencies ω_+ higher than the respective frequencies for a single interface SPP, and the even modes lower frequencies ω_-. For large wave vectors β (which are only achievable if $\mathrm{Im}\left[\varepsilon(\omega)\right] = 0$), the limiting frequencies are

$$\omega_+ = \frac{\omega_\mathrm{p}}{\sqrt{1+\varepsilon_2}}\sqrt{1+\frac{2\varepsilon_2 \mathrm{e}^{-2\beta a}}{1+\varepsilon_2}} \tag{2.30a}$$

$$\omega_- = \frac{\omega_\mathrm{p}}{\sqrt{1+\varepsilon_2}}\sqrt{1-\frac{2\varepsilon_2 \mathrm{e}^{-2\beta a}}{1+\varepsilon_2}}. \tag{2.30b}$$

Odd modes have the interesting property that upon decreasing metal film thickness, the confinement of the coupled SPP to the metal film decreases as the mode evolves into a plane wave supported by the homogeneous dielectric environment. For real, absorptive metals described via a complex $\varepsilon(\omega)$, this implies a drastically increased SPP propagation length [Sarid, 1981]. These *long-ranging* SPPs will be further discussed in chapter 7. The even modes exhibit the opposite behavior - their confinement to the metal increases with decreasing metal film thickness, resulting in a reduction in propagation length.

Moving on to MIM geometries, we now set $\varepsilon_2 = \varepsilon_2(\omega)$ as the dielectric function of the metal and ε_1 as the dielectric constant of the insulating core in equations (2.29a, 2.29b). From an energy confinement point of view, the most interesting mode is the fundamental odd mode of the system, which does not exhibit a cut-off for vanishing core layer thickness [Prade et al., 1991]. Fig. 2.7 shows the dispersion relation of this mode for a silver/air/silver heterostructure. This time, the dielectric function $\varepsilon(\omega)$ was taken as a complex fit to the dielectric data of silver obtained by Johnson and Christy [Johnson and Christy, 1972]. Thus β does not go to infinity as the surface plasmon frequency is approached, but folds back and eventually crosses the light line, as for SPPs propagating at single interfaces.

It is apparent that large propagation constants β can be achieved even for excitation well below ω_sp, provided that the width of the dielectric core is chosen sufficiently small. The ability to access such large wave vectors and thus small penetration lengths $\hat{z}$ into the metallic layers by adjusting the geometry indicates that localization effects that for a single interface can only be sustained at excitations near ω_sp, can for such MIM structures also be attained for excitation out in the the infrared. An analysis of various other MIM structures, for example concentric shells, has given similar results [Takahara et al., 1997]. Geometries amendable to easy fabrication such as triangular metal V-grooves

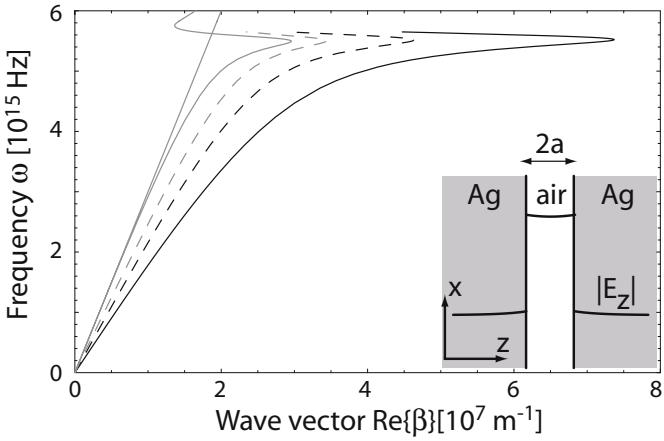

Figure 2.7. Dispersion relation of the fundamental coupled SPP modes of a silver/air/silver multilayer geometry for an air core of size 100 nm (broken gray curve), 50 nm (broken black curve), and 25 nm (continuous black curve). Also shown is the dispersion of a SPP at a single silver/air interface (gray curve) and the air light line (gray line).

on a flat metal surface have already shown great promise for applications in waveguiding, which will be presented in chapter 7.

We have limited our discussion of coupled SPPs in three-layer structures to the fundamental bound modes of the system, with a view on applications in waveguiding and confinement of electromagnetic energy. It is important to note that the family of modes supported by this geometry is much richer than described in this treatment. For example, for IMI structures, we have omitted a discussion of leaky modes, and MIM layers can also exhibit oscillatory modes for sufficient thickness of the dielectric core. Additionally, the coupling between SPPs at the two core/cladding interfaces changes significantly when the dielectric constants of the sub- and superstrates are different, so that $\varepsilon_2 \neq \varepsilon_3$, prohibiting phase-matching between the modes located at the two interfaces. A detailed treatment of these cases can be found in [Economou, 1969, Burke and Stegeman, 1986, Prade et al., 1991].

2.4 Energy Confinement and the Effective Mode Length

In chapter 5 we will see that using localized surface plasmons in metal nanoparticles, electromagnetic energy can be confined or squeezed into volumes smaller than the diffraction limit $(\lambda_0/2n)^3$, where $n = \sqrt{\varepsilon}$ is the refractive index of the surrounding medium. This high confinement leads to a concomitant field enhancement and is of prime importance in plasmonics, enabling a great variety of applications in optical sensing, as will be discussed in chapter 9. In the essentially one-dimensional cases of single interfaces and

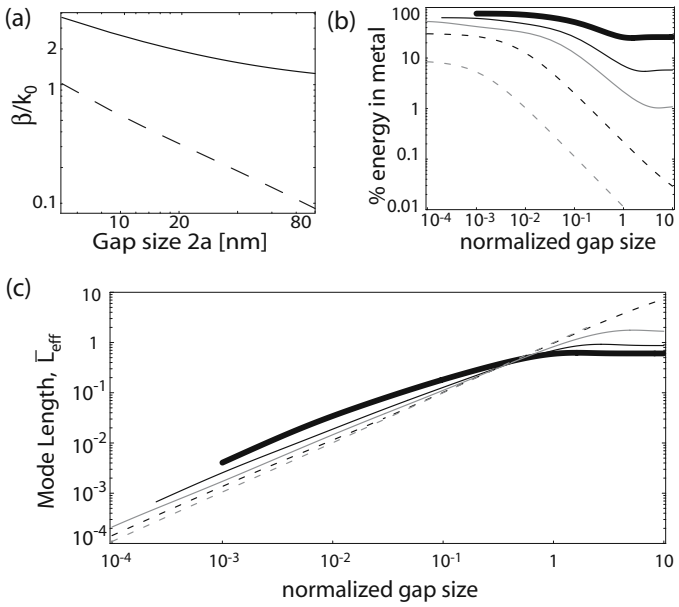

Figure 2.8. Energy confinement in a gold/air/gold MIM structure. (a) Real (solid curve) and imaginary (dashed curve) part of the normalized propagation constant β versus gap size at $\lambda_0 = 850$nm. (b) Fraction of electric field energy residing inside the metallic half spaces as a function of normalized gap size for excitation at $\lambda_0 = 600$ nm (thick curve), 850nm (black curve), 1.5 μm (gray curve), 10 μm (broken black curve), and 100 μm (broken gray curve). (c) Effective mode length L_{eff} normalized to free-space wavelength λ_0 as a function of gap size. Adapted from [Maier, 2006b].

multilayer structures presented above that support propagating SPPs, energy localization below the diffraction limit perpendicular to the interface(s) is also possible. We have already hinted at this phenomenon when stating that the field decay length $\hat{z}$ in the dielectric layers can be significantly smaller than λ_0/n.

However, care must be taken when quantifying energy confinement, since a sub-wavelength field decay length $\hat{z}$ on the dielectric side of the interface implies that a significant amount of the total electric field energy of the SPP mode resides inside the metal. This energy must be taken into account using (1.38) when calculating the spatial distribution of the electric energy density, since for the quantification of the strength of interactions between light and matter (e.g. a molecule placed into the field), the field strength per unit energy (i.e., single photon) is of importance.

Taking a gold/air/gold MIM heterostructure as an example, Fig. 2.8(a) shows the evolution of both the real and imaginary parts of the propagation constant β of the fundamental SPP mode with varying gap size for excitation at a free space wavelength of $\lambda_0 = 850$ nm, calculated using Drude fits to the dielectric

function of gold [Johnson and Christy, 1972, Ordal et al., 1983]. Both parts increase with decreasing gap size, since the mode is becoming more electron-plasma in character, suggesting that the electromagnetic energy is residing increasingly in the metal half-spaces. A plot of the fractional amount of the electric field energy inside the metal regions is shown in Fig. 2.8(b) for excitation at wavelengths $\lambda_0 = 600$ nm, 850 nm, 1.5 μm, 10 μm, and 100 μm (= 3 THz). For a gap of 20 nm for example, at $\lambda_0 = 850$ nm this fraction already reaches 40%. Note that the gap size is normalized to the respective free space wavelength. It is apparent that along with the increased localization of the field to the gold/air interface, either via small gap sizes or excitation closer to ω_{sp}, comes a shift of the energy into the metal regions.

In order to get a better handle on the consequences of increasing fractions of the total energy of the mode entering the metallic cladding upon decreasing size of the dielectric gap, we can define in analogy to the effective mode *volume* V_{eff} used to quantify the strength of light-matter interactions in cavity quantum electrodynamics [Andreani et al., 1999] an effective mode *length* L_{eff}, with

$$L_{eff}(z_0)u_{eff}(z_0) = \int u_{eff}(z)dz. \tag{2.31}$$

$u_{eff}(z_0)$ represents the electric field energy density at a position z_0 of interest within the air core (e.g. the location of an emitter). In this one-dimensional picture, the effective mode length is therefore given as the ratio of the total energy of the SPP mode divided by the energy density (energy per unit length) at the position of interest, which is often taken as the position of highest field. In a quantized picture for normalized total energy, the inverse of the effective mode length thus quantifies the field strength per single SPP excitation. More details can be found in [Maier, 2006b].

A determination of the effective mode length of MIM structures allows an examination how the electric field strength per SPP excitation in the air gap scales as a function of the gap size. Fig. 2.8(c) shows the variation of L_{eff} (normalized to the free-space wavelength λ_0) with normalized gap size. z_0 is taken to be at the air side of the air/gold boundary, where the electric field strength is maximum. The mode lengths drop well below $\lambda_0/2$, demonstrating that plasmonic metal structures can indeed sustain *effective* as well as *physical* mode lengths below the diffraction limit of light. The trend in L_{eff} with gap size tends to scale with the physical extent of the air gap. For large normalized gap sizes and low frequencies, this is due to the delocalized nature of the surface plasmon, leading to smaller mode lengths for excitation closer to the surface plasmon frequency ω_{sp} for the same normalized gap size.

As the gap size is reduced to a point where the dispersion curve of the SPP mode turns over (see Fig. 2.7) and energy begins to enter the metallic half spaces, the continued reduction in mode length is due to an increase in field

localization to the metal-air interface. In this regime, excitations with lower frequencies show smaller mode lengths for the same normalized gap size than excitations closer to the plasmon resonance, due to the fact that more energy resides inside the metal for the latter. We note that for very small gaps with $2a < 2$ nm, the effects of local fields due to unscreened surface electrons become important [Larkin et al., 2004], leading to a further decrease in L_{eff}. This cannot be captured using the dielectric function approach.

To summarize, we see that despite the penetration of a significant amount of energy of a SPP mode into the conducting medium (for excitation near ω_{sp} or in small gap structures), the associated large propagation constants β ensure that the effective extent of the mode perpendicular to the interface(s) drops well below the diffraction limit.