Chapter 3

Introduction to plasmons and plasmonics

The aim of this chapter is to give an overview of the physics of plasmons and their applications, so-called plasmonics. The emphasis will be on concepts rather than methods, since the goal is primarily to understand what plasmons are, which types of plasmons can exist under different conditions, and what their relevance to SERS is. Reading this chapter is necessary neither to understand SERS, nor to follow the other chapters of this book. Readers with a more immediate interest in SERS may read only Section 3.2 and then jump directly to Chapter 4.

There are however many reasons, why understanding the physics of plasmons is important to SERS. Firstly, plasmon resonances (in fact a certain type of plasmon resonance) are 'what makes SERS possible', and are mentioned and debated as the origin of the enhancement in almost all SERS discussions. Secondly, *plasmonics* is currently an expanding and very active area, from which SERS can gain further insight and where SERS can play an important role. In fact, this chapter covers many of the 'related plasmonic effects' mentioned in the subtitle of this book.

We discuss first in Section 3.2 the optical properties of noble metals, and in particular of gold and silver, which are most used for SERS. This section is the most directly relevant one to SERS. In Section 3.3, we then give a detailed account of the various types of plasmons, and emphasize their relevance to SERS and other applications[1]. The two most important types of plasmons are then discussed in detail: firstly, Section 3.4 focuses on

---

[1] As explained in the introductory chapter, SERS is a technique that exists at the boundary between physics, chemistry, biology, engineering, etc. As a result, many terms related to the physics of plasmons have 'evolved' into a terminology that is in many cases ambiguous or inaccurate. The detailed classification of different types of plasmons given in Section 3.3 is 'unnecessary' from a SERS standpoint. Having said this, the present book is aiming at a description of the fundamentals of SERS and, therefore, we felt that it was necessary to

the surface plasmon–polaritons at planar interfaces. These are the basis for most applications of plasmonics related to wave propagation and guiding, and some sensing applications. Section 3.5 then discusses the localized surface plasmon–polaritons, which play a major role in all applications using electromagnetic field enhancements, including SERS and surface-enhanced fluorescence (SEF). We then conclude by giving in Section 3.6 a few examples of plasmonic effects and possible applications (other than SERS).

3.1. PLASMONICS AND SERS

Plasmonics is a relatively new term, encompassing all areas of research and technology concerned with the study, fabrication, and applications of plasmon-supporting structures. This is not exactly new research since, in essence, it is simply the study of the optical properties of noble metals, in particular gold and silver. But the recent advances in nano-technologies and nano-science mentioned in the introductory chapter have opened up new possibilities in the design and fabrication of metallic structures with features in the nanometer size scale. This is sometimes referred to as *nano-plasmonics.* As a consequence, the general interest in plasmonics has dramatically increased over the last few years, and has spread to various other areas of physics, chemistry, and biology. This is reflected in the increase in the (normalized) number of plasmon-related publications as illustrated in Fig. 3.1. Furthermore, this is also evident in the number of recent reviews of the topic published in a wide range of scientific publications [11,62,126–133].

Plasmonics and SERS are two areas of research with a strong overlap and can benefit mutually from each other. SERS can be used as a tool for the study of plasmonic substrates and, reciprocally, some plasmonic substrates can turn out to be very good SERS substrates. Indeed, plasmons are at the core of SERS electromagnetic effects and/or enhancements. Most articles, talks, or even informal discussions about SERS are bound to mention plasmons or plasmon resonances at one point or another. Many different expressions appear in this context, among them: plasmon, plasmon resonance, radiative plasmon resonance, surface plasmon, surface plasmon–polariton, localized surface plasmon–polariton, propagating surface plasmon, and in many cases arbitrary combinations of the aforementioned! It is easy sometimes to lose sight of the main concepts and understand what is really meant from a specific standpoint, especially considering that they are not always used consistently or even correctly across the literature. With this in mind, we will attempt in the following to give a brief overview – at an introductory level – of the various concepts related to plasmons.

---

attempt to provide a rigorous classification of plasmons for future reference; an undertaking that would never be pursued in conventional papers in the field. This chapter is in that sense a 'bonus' to the content of the book and a prerequisite if a deeper understanding is sought. It is our hope that it will contribute in the future to a more precise vocabulary and denomination of plasmon-related phenomena.

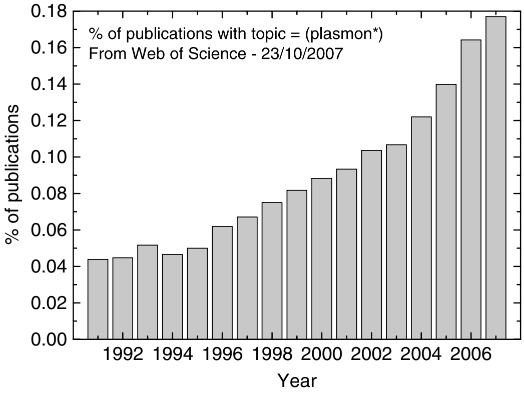


Figure 3.1. Evolution of the fraction of publications related to plasmons in the Web of Science database as of 23 October 2007.

3.2. THE OPTICAL PROPERTIES OF NOBLE METALS

Metals such as gold (Au), silver (Ag), copper (Cu), or aluminum (Al), have long been known to have different optical properties from standard dielectrics. They, for example, reflect light very efficiently in the visible, making them good materials for mirrors of various types (except possibly for their cost, etc.). These particular optical properties, along with many other physical properties (such as heat or electrical conductivity) all have the same physical origin: *the presence of free conduction electrons*. It seems fairly intuitive that these free electrons result in large heat or electrical conductivity[2]. Their connection with the optical response of metals may however appear less obvious at first sight.

The free electrons of a metal move in a background of fixed positive ions (the vibrations of ions, or phonons, are ignored here in a first approximation), which ensures overall neutrality. This forms, by definition, a *plasma* and can be called a free-electron plasma, or solid-state plasma [135]. The study of a solid-state plasma is different from that of a gaseous plasma, since the emphasis is on equilibrium phenomena in the first case, while it is more on instabilities or steady state properties in the latter [135]. The optical response of this free-electron plasma will govern all the optical properties of metals, at least in the visible part of the spectrum where its characteristic resonant energies reside.

To model the optical response of a free-electron plasma, one needs to determine the constitutive equations relating the currents and charges in the plasma to the electromagnetic fields. This is a very difficult undertaking in general

---

[2] The presence of free electrons in metals is linked to their ability to conduct both electricity and heat. However, it must be kept in mind that heat conduction in solids is in general a much more complex problem that does not only relate to the presence of free electrons. In fact, the best known heat conductor is diamond, which is also one of the best insulators.

because of many possible complications, including: the interaction of electrons with the underlying periodic structure of ions, the electron–electron correlations and the fermionic nature of electrons, the interaction of electrons with impurities and phonons, and the possible presence of surfaces. This response can be described with various degrees of refinements [135,136]. We will simply give here a brief account of the simplest model, namely: the Drude model.

3.2.1. The Drude model of the optical response

The Drude model is described in detail in many textbooks [96,136]. It is sufficient in many instances and for our purpose here and provides, in addition, a simple framework to understand electrical conductivity. It actually leads to the same result as a more elaborate approach (the random phase approximation) for the local dielectric function. We review here the main ingredients, in particular those relating to the optical response.

One simple way to introduce the Drude model is by using the Lorentz model for the atomic polarizability, presented in Appendix D . This model describes the optical response of an electron in an atom or molecule, *bound* with a restoring force characterized by a resonant frequency ω_0. The conduction electrons in a metal are not bound and can therefore, in a first approximation, be described by the Lorentz model , *without restoring force* (i.e. $\omega_0 \approx 0$). Moreover, because the free electrons are distributed uniformly and randomly throughout the metal, their contributions to the total optical susceptibility are simply the sum of their individual polarizabilities, without any local field correction. The Drude relative dielectric function of a metal can then be obtained by taking $\omega_0 = 0$ in the Lorentz model (see Appendix D for more details), i.e.:

$$\epsilon(\omega) = 1 - \frac{ne^2}{m\epsilon_0} \frac{1}{\omega^2 + i\gamma_0\omega}, \tag{3.1}$$

where n [m^{-3}] is the number of free electrons per unit volume and m [kg] their mass[3]. The damping term, γ_0 [rad s^{-1}], here corresponds to the collision rate of free electrons with the crystal or impurities (which also leads to the electrical resistivity in this simple model [96,136]). It is usually small compared to ω in the regions of interest here. The optical response of the positive ions in the crystal has so far been ignored. In a first approximation (which is at least correct at long wavelengths), these contribute to a constant background real dielectric function $\epsilon_\infty \geq 1$. This affects the optical response of the crystal and the dynamics of the free electrons. This can easily be incorporated in the

---

[3] Strictly speaking, m is the effective mass to partially account for the effect of the surrounding crystal structure and interactions beyond the free-electron approximation.

Drude model and leads to a slightly modified expression for $\epsilon(\omega)$, namely:

$$\boxed{\epsilon(\omega) = \epsilon_\infty \left(1 - \frac{\omega_p^2}{\omega^2 + \mathrm{i}\gamma_0\omega}\right),} \tag{3.2}$$

where we have defined ω_p [rad s^{-1}] as

$$\omega_p = \sqrt{\frac{ne^2}{m\epsilon_0\epsilon_\infty}}. \tag{3.3}$$

In the absence of an external perturbation, the charge density of a plasma is uniform and zero. It can be shown that ω_p is the natural oscillation frequency of the free-electron-plasma charge density and it is therefore called the *plasma frequency.* One can also define the corresponding wavelength $\lambda_p = 2\pi c/\omega_p$.

Taking the real and imaginary parts of the previous expression, we have:

$$\mathrm{Re}(\epsilon(\omega)) = \epsilon_\infty \left(1 - \frac{\omega_p^2}{\omega^2 + \gamma_0^2}\right), \tag{3.4}$$

and

$$\mathrm{Im}(\epsilon(\omega)) = \frac{\epsilon_\infty\omega_p^2\gamma_0}{\omega(\omega^2 + \gamma_0^2)}. \tag{3.5}$$

Keeping in mind that γ_0 is small compared to ω, we see that for a plasma described by a Drude model, the plasma frequency can be obtained from the condition $\mathrm{Re}(\epsilon(\omega_p)) \approx 0$. We also see that in the region where $\omega < \omega_p$ (wavelength longer than λ_p), we have $\mathrm{Re}(\epsilon(\omega)) < 0$. Moreover, if ω is not too small, the absorption, characterized by $\mathrm{Im}(\epsilon(\omega))$, is also small in this region. It is these two conditions, $\mathrm{Re}(\epsilon(\omega)) < 0$ and small $\mathrm{Im}(\epsilon(\omega))$, that make possible a whole range of interesting optical effects, including plasmon resonances. These conditions are never fulfilled in 'standard' dielectrics where $\mathrm{Re}(\epsilon)$ is typically between 1 and ∼10. For many metals, the plasma frequency is in the UV part of the electromagnetic spectrum, and the region of interest is therefore in the visible (and close UV, or near infrared, depending on the metal).

3.2.2. The optical properties of real metals

Inter-band and intra-band transitions

The Drude model describes in a relatively simple way the optical response of a plasma, and in our case of the free conduction electrons of a metal.

This corresponds to *intra-band optical transitions*, since the excited electrons remain in the same electronic bands (the conduction band). We have already seen that the optical response of the fixed ions can also be included simply when their contributions are a constant background dielectric function ϵ_∞.

In a real metal, additional optical processes are likely to occur and contribute to the optical response, the most common being *inter-band transitions*, i.e. bound electrons optically excited to a higher energy band. Such transitions can for example be described – in a first approximation – as a collection of Lorentz oscillators, or by more complex expressions that take into account the *band-like* nature of these transitions, as illustrated in Appendix E . Their contribution, $\epsilon_b(\omega)$, adds to the free-electron contribution (from the Drude model). The dielectric function then takes the more general form:

$$\epsilon(\omega) = \epsilon_b(\omega) - \epsilon_\infty \frac{\omega_p^2}{\omega^2 + i\gamma_0\omega}, \tag{3.6}$$

where the term $\epsilon_b(\omega)$ represents the *inter-band transitions* that coexist with the free-electron contribution modeled by the second term. In many cases, inter-band transitions occur at energies in the UV, much higher than the plasma frequency. In this case, their contribution to ϵ in the visible is simply constant and real (no absorption off resonance), i.e. $\epsilon_b(\omega) = \epsilon_\infty$ and the expression reduces to the Drude model of the previous section. This is for example approximately the case for silver (see Appendix E).

If, however, there are inter-band transitions close to or below the plasma frequency, then their optical response needs to be added to the free-electron response as above. This is for example the case for gold (see Appendix E). Note that in this case, $\epsilon_b(\omega_p) \neq \epsilon_\infty$, and the condition $\mathrm{Re}(\epsilon(\omega)) = 0$ then no longer yields the free-electron-plasma frequency ω_p, as for the simple Drude model. This means physically that the natural oscillations of the free-electron plasma are 'affected' by the presence of the inter-band transitions.

Plasma frequency for Ag and Au

This description, although over-simplified, is nevertheless successful in explaining semi-quantitatively the optical response of many real metals. For example, at wavelengths longer than all inter-band transitions, Eq. (3.2) should be valid. The ω-dependence should then be described by the Drude term (and vary as ω^{-2}). This has been verified experimentally [137] for Ag and Au, from which the same value of $\sqrt{\epsilon_\infty}\omega_p \approx 1.4 \times 10^{16}$ rad s^{-1} is derived. This is in remarkable agreement with the expression for ω_p in Eq. (3.3) derived from the Drude model. Taking an effective mass equal to the electron mass and a density of conduction electrons of the order of $n \approx 6 \times 10^{28}$ m^{-3} (corresponding to a full d band with one free s electron per atom) give precisely

$\sqrt{\epsilon_\infty}\omega_p \approx 1.4 \times 10^{16}$ rad s^{-1}, or $\hbar\omega_p \approx 9.1/\sqrt{\epsilon_\infty}$ eV, or a corresponding wavelength of $\lambda_p \approx 136\sqrt{\epsilon_\infty}$ nm. The Drude model therefore gives an excellent description of the long-wavelength optical response of Ag and Au. Only the value of ϵ_∞ is missing. This is understandable though, for ϵ_∞ comes from high energy contributions in the deep UV-range, which are not included in this simple treatment of the problem.

The frequency-dependent dielectric functions of Ag and Au are of crucial importance for many plasmonics problems. They are further discussed in Appendix E, where analytical models are also provided. These analytical models will be used extensively in this book for the modeling of plasmonics and SERS effects.

3.2.3. Non-local optical properties

Many optical properties of metals can be understood using a *local* dielectric function, as described above. This implicitly assumes that the dielectric function of the metal is *only* frequency-dependent. There are, however, a few cases where this approximation is unsuccessful in providing an explanation of the experimental observations. A non-local description, where the dielectric function ϵ of the material depends both on frequency (ω) and wave-vector ($\mathbf{k}$) (see Appendix C for more details), is then required. One such example is the anomalous skin effect [136]. Another situation where non-local effects can be important is that of an emitter very close to a metal surface, which is precisely the case of SERS. However, the inclusion of non-local effects in SERS complicates tremendously the electromagnetic problem, which is already not trivial in the local approximation. The easiest solution is to ignore them altogether, which in many cases does not affect the qualitative and even quantitative conclusions. It is however useful to bear in mind that these effects do exist, and can be occasionally the source of either problems or unexplained phenomena. Their influence can be studied separately for very simple cases, such as emitters close to plane or spherical surfaces. A discussion of non-local effects would take us too far from our main purpose here and can be found (to some degree) in the literature [138].

3.2.4. What makes the metal–light interaction so special?

A brief survey of various metals

A comparison of the optical properties (real and imaginary parts of the dielectric function) of various metals (from Ref. [139]) is given in Fig. 3.2. The monotonous decay of Re(ϵ) from small values in the UV to negative values in the visible, and very negative values in the infrared, is common to all metals and is in fact predicted by the Drude model (for typical plasma frequencies in the UV). *This is one of the most important characteristics of metals*, as far as optical properties are concerned, and it is a consequence of

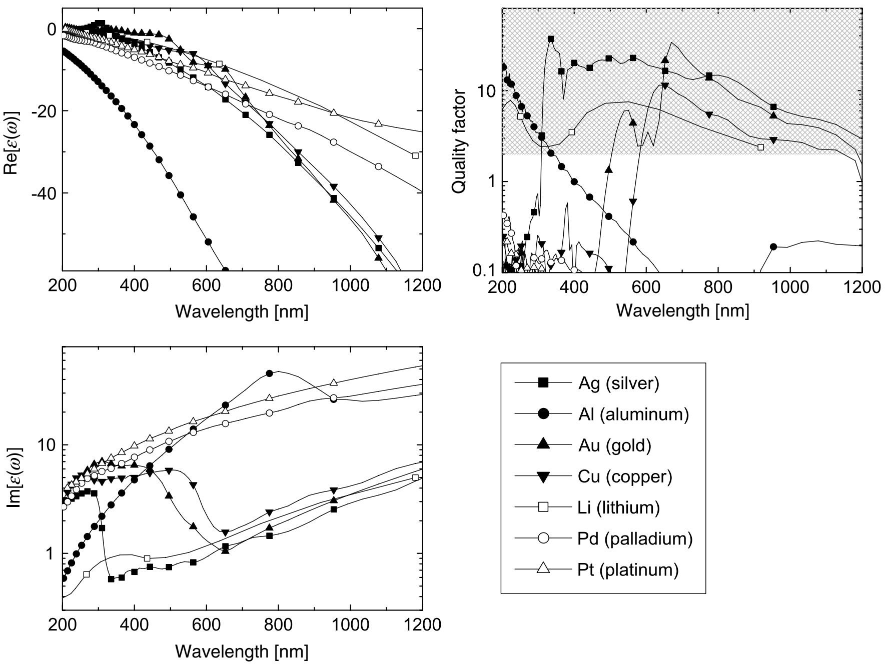


Figure 3.2. Overview of the optical properties of a selection of metals in the (extended) visible range. The real (top) and imaginary (bottom) parts of ϵ are plotted against wavelength on the left (from Ref. [139]). Also shown (top-right) is the predicted approximate quality factor Q of localized surface plasmon resonances for a metal/air nano-particle, as defined in Eq. (3.8). The shaded area is the area of interest to many plasmonics applications.

the optical response of the free (conduction) electrons, as explained simply by the Drude model.

The negativity of the real part of $\epsilon(\omega)$ at visible wavelengths is also the origin of many of the known optical properties of metals, including plasmon-related effects. For example, it implies that the refractive index is smaller than one and even close to zero, while most 'common' materials have a refractive index of ∼1 or more. The reflection coefficient at a dielectric/metal interface is derived simply from the relative refractive indices (see for example Appendix F), and the small refractive index of metals directly leads to a reflection coefficient close to one (almost perfect reflector), which is arguably one of the best known and most 'visible' properties of metals.

Plasmon resonances and plasmonics

There are two important types of plasmons, which will be discussed extensively in the rest of this chapter: localized surface plasmon–polaritons

(LSPP or LSP for short) and propagating surface plasmon–polaritons (PSPP). Here we have to anticipate some of these results to carry on the general discussion of the optical properties of metals. As we shall see, the negative real part of $\epsilon(\omega)$ is also linked to the existence of plasmons and plasmon resonances.

To understand this in simple terms, let us consider briefly the problem of a small metallic sphere interacting with an electromagnetic field (in a laser beam). The simplest treatment of this problem is to consider a sphere much smaller than the wavelength of the incident beam and solve the problem within the electrostatic approximation (to be treated in Section 5.1.4). It is then equivalent to the electrostatic problem of a sphere, in a medium of relative dielectric constant ϵ_M, and in a uniform external electric field (to be treated in Section 6.2.1). The (complex) electric field inside the sphere can then be shown to be constant and proportional to the incident field E_0 [96]:

$$E_{\mathrm{In}} = \frac{3\epsilon_M}{\epsilon(\omega) + 2\epsilon_M} E_0. \tag{3.7}$$

The important part in this expression is the denominator. If it were close to zero (a condition which can be achieved if $\epsilon(\omega) \approx -2\epsilon_M$) then the fraction would be *very* large (infinite for perfect equality). This is not possible for standard dielectrics, for which ϵ is typically between 1 and ~10. But for metals, this condition can be approximately met if the absorption is small ($\mathrm{Im}(\epsilon(\omega)) \approx 0$) at a wavelength where $\mathrm{Re}(\epsilon(\omega)) \approx -2\epsilon_M$. The optical response (absorption and scattering) at this particular frequency (or wavelength) is then very large, i.e. this is the signature of a *resonance.* This is explicitly illustrated in Fig. 3.3 where the absorption coefficient (proportional to $|E_{\mathrm{In}}|^2$) is shown as a function of wavelength for a small silver sphere in air ($\epsilon_M = 1$). There is a clear resonant response at ≈345 nm, which corresponds to the condition $\mathrm{Re}(\epsilon(\omega)) \approx -2$. It is interesting to note at this stage that while the resonance condition is determined primarily by the real part of $\epsilon(\omega)$, it is actually its imaginary part that limits how large the resonance can be.

This is a simple example of a localized surface plasmon (LSP) resonance. Similar effects will exist for other geometries and configurations, but with different denominators and therefore different resonance conditions. This is an interesting concept that needs to be highlighted: optical resonance conditions in small metallic objects are not purely intrinsic of the material properties but they are strongly linked to the *geometry.* Two objects made with the same metal but with different geometries will have different resonance conditions. These conditions however always correspond to a negative real part of ϵ. Similarly, the existence of propagating surface plasmon–polariton (PSPP) waves at metal/dielectric interfaces also requires a negative real part of ϵ.

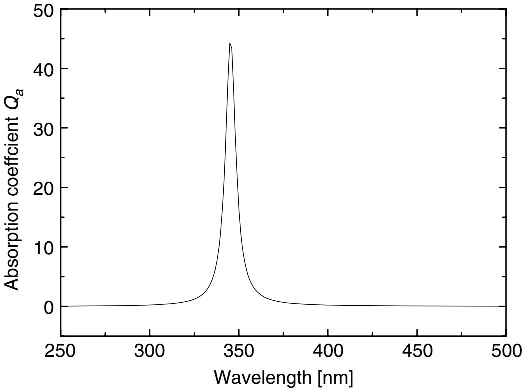


Figure 3.3. Absorption coefficient, $Q^{\mathrm{NP}}_{\mathrm{Abs}}$ (see Section 5.1.3) as a function of wavelength for a silver sphere of radius 25 nm in air in the electrostatic approximation. $Q^{\mathrm{NP}}_{\mathrm{Abs}}$ can be expressed as (see Section 6.2.1) $Q^{\mathrm{NP}}_{\mathrm{Abs}} = 12(\omega a \mathrm{Im}(\epsilon(\omega)))/(c|\epsilon(\omega) + 2|^2)$.

Moreover, both types of plasmons (LSP and PSPP) are strongly affected by optical absorption, i.e. the larger the $\mathrm{Im}(\epsilon)$, the more 'lossy' they are (the meaning of this will be made clearer later). For LSP resonances, this damping can further be characterized by a *quality factor* defined as [140]:

$$Q = \frac{\omega(d\epsilon'/d\omega)}{2(\epsilon''(\omega))^2}, \tag{3.8}$$

where $\epsilon' = \mathrm{Re}(\epsilon)$ and $\epsilon'' = \mathrm{Im}(\epsilon)$. In simple terms, Q is large when $\mathrm{Im}(\epsilon)$ is small and, therefore, characterizes the strength (and width) of the resonance. Q is also plotted in Fig. 3.2 for various metals.

Which metals are good for plasmonics and SERS?

From these qualitative arguments (which will be further justified in the rest of this chapter), one can consider that a metal is good for plasmonics if:

- $\mathrm{Re}(\epsilon)$ is negative in the wavelength range of interest (typically the visible and near infrared). For LSP applications, one in fact needs (as a rule of thumb) $-20 \leq \mathrm{Re}(\epsilon) \leq -1$.
- $\mathrm{Im}(\epsilon)$ is small (or equivalently Q is large) in the range of interest. Typically, Q must be larger than $\sim$2, preferably larger than $\sim$10.

From the examples of Fig. 3.2, this rules out a number of metals, such as aluminum, palladium, platinum, etc. The latter two present too much absorption, while aluminum would only be suitable for applications in the UV.

Note that these metals do show a metallic behavior (including in their optical properties) but are not expected to exhibit any plasmonic properties[4].

Among the other metals studied in Fig. 3.2, silver is the most promising one (and is in fact widely used for SERS and plasmonics). Gold and copper are also suitable, but only at longer wavelengths (typically more than ∼600 nm). At such wavelengths, the optical absorption of gold in fact becomes comparable to that of silver. Finally, lithium also exhibits suitable properties across the whole visible range, as silver, but has not been used much. Lithium reacts easily in water and does not occur freely in nature due to its chemical activity; it is therefore not very easy to be used as a plasmon-supporting material.

To these theoretical considerations, one should add the (very important) practical issues: availability, ease of manipulation, especially for the fabrication of nano-structures, toxicity, durability, cost, etc. Gold is certainly the most promising in these categories, and should therefore be the material of choice for applications beyond ∼600 nm (in the red and near IR). Silver, whose absorption is the smallest especially below ∼600 nm, can be used for large field enhancement applications (such as single-molecule detection). These two metals are by far the most widely used metals in plasmonics applications, including SERS.

Finally, it is interesting to remark that the conditions listed above also appear in other materials, but at different wavelengths. Similar effects are therefore expected, but are usually not considered as being part of the field of plasmonics. This is the case for example of doped semiconductors, whose conduction electrons result in a similar optical response as metals but with a plasma frequency in the far infrared (due to the smaller density of carriers). Another example is the optical response of phonons in ionic crystals, which also leads to negative $\mathrm{Re}(\epsilon)$, but again at much longer wavelengths. In this final case, plasmons do not even play a role, since the optical response is not dictated by free electrons but by phonons; the related effects may be referred to as ‘phononics’.

3.3. WHAT ARE PLASMONS?

In the rest of this chapter, we will focus on plasmons and their relation to SERS and plasmonics.

3.3.1. The plasmon confusion

In the modern SERS literature (and in many other areas), many effects are attributed to plasmons or plasmon resonances, without further details about

---

[4] For example, SERS has been observed on these, but the enhancements are either ‘chemical’ (see Chapter 4) or small electromagnetic enhancements not arising from plasmon resonances.

what plasmons actually are. One can sometimes find a sentence or two about the origin of plasmons; the most common assertion being: ‘plasmon resonances are due to collective charge oscillations of the free electrons of the metal’. However, there is hardly ever a real concern about these ‘electron oscillations’ in typical SERS studies. SERS enhancements are usually explained as an electromagnetic effect, not an electron oscillation effect. There seems to be a missing link between the electron oscillations and the large electromagnetic field enhancements. To add to this, the term plasmon is also used in contexts other than SERS; for example as a way to guide light (in a plasmonic waveguide) or in electron energy loss (EELS) experiments. Finally, adjectives are used in many cases to qualify the plasmons, in an attempt to clarify the situation. Examples of these are: surface plasmon, plasmon–polariton, radiative, non-radiative, propagating, and localized plasmons. Unfortunately, these terms are not always used consistently by authors or across disciplines (chemistry and physics in the first instance), and have even evolved over time to designate something different from what they were 40 years ago.

This general vagueness leads to numerous confusions about what plasmons are, which types of plasmons can be encountered, and what their respective importance is for applications such as SERS. This can be the source of frustration for people new to the field. The aim of the following discussion is therefore three-fold:

- Firstly, we will attempt to define clearly what plasmons are, and describe the different types of plasmons. Because of the confusion discussed above, some of the definitions and descriptions may however be slightly biased and not correspond exactly to the choices of other researchers.
- Secondly, we will highlight some of the most important applications of plasmons, including the main topic of this book which is SERS. We will emphasize in particular the different nature of plasmons used for different applications.
- Finally, one important message we wish to convey is that in most cases of interest to SERS and plasmonics, all plasmon-related effects can be understood as *electromagnetic effects*. The relation to the free electrons of a metal is only secondary (although important from a fundamental point of view). All information on plasmons, and plasmon resonances, is fully contained in the dielectric function and the geometry of a specific problem.

3.3.2. Definition and history

Plasmons

The use of the term plasmon has evolved somewhat over the years, mainly because the types of experiments carried out on metals have changed. It is however useful to look at the original definition of the term, which is still valid although it may now be used in a wider sense and in other frameworks.

The term plasmon was introduced by Pines in 1956 in the introduction of a review article [141] about collective energy losses. In Pines' work we find the following definition:

> The valence electron collective oscillations resemble closely the electronic plasma oscillations observed in gaseous discharges. We introduce the term 'plasmon' to describe the quantum of elementary excitation associated with this high-frequency collective motion.

A plasmon is therefore a quantum quasi-particle representing the elementary excitations, or modes, of the charge density oscillations in a plasma. Note that the study of these oscillations started earlier, even if they were not known or identified as plasmons [141]. We will come back to the notion of elementary excitations or modes of a system in the next section.

Although the term 'plasmon' is sometimes used in a broader context, the formal definition given above is the definition of reference. It draws its origin from quantum mechanics, even though we will see that quantum mechanics is, in fact, not necessary to study plasmons. A useful analogy to understand the meaning of this definition is to recall the formal definition of a photon: it is the quantum particle representing the elementary excitations, or modes, of the free electromagnetic field oscillations. *A plasmon is therefore simply to the plasma charge density what photons are to the electromagnetic field.* Many properties of photons can be studied within a classical framework, using Maxwell's equations. Similarly, many properties of plasmons can be studied within a classical description of the plasma and its interactions. There is, may be, a small difference in the vocabulary between plasmons and photons, but it is only artificial: people typically only use the term 'photon' when dealing with quantum aspects of the electromagnetic fields (such as absorption or emission by an atom). In classical situations, the term electromagnetic wave, or electromagnetic mode, is usually preferred. For reasons that are more historical than scientific, the term 'plasmon' tends to be used in all situations, quantum or classical, instead of equivalent (classical) denominations such as *charge density oscillations* . A more important and fundamental difference is that a photon is a real quantum particle while a plasmon is a *quasi-particle* because it is always *'lossy'* and highly interacting. A charge density oscillation, if not maintained by an external source of energy, will always decay because of various loss mechanisms (collisions, etc.).

Plasmon–polaritons

Another important type of elementary excitations, related to plasmons, is that of an electromagnetic wave propagating in a medium. By medium, we mean here an optically responsive medium (with a relative dielectric function $\epsilon \neq 1$ or relative magnetic permeability $\mu \neq 1$). From classical electromagnetism, it is known that such a situation can be described by Maxwell's equations for media, which introduce the notion of internal polarization, **P**, or magnetization **M** (see Appendix C). These account for the fact that the electromagnetic wave excites the internal degrees of freedom of the medium (typically bound electrons in a dielectric). The energy of such a wave is therefore shared between the electromagnetic field oscillations (of **E** and **B**), and the internal excitations of the medium (typically represented by oscillations of **P** and **M**). The corresponding quantum particle is no longer a photon, but a photon coupled to the internal degrees of freedom of the medium. Such modes are usually called *polaritons*[5]. If the internal excitations of the medium are identified, then the polariton can sometimes be further qualified. For example, the optical response of an ionic crystal in the mid-infrared is dominated by the interaction of light with crystal vibrations (phonons). The electromagnetic waves in such a medium are then called phonon–polaritons (mixed photon–phonon modes). Similarly, the optical response of a metal in the visible and infrared is dominated by the interaction of light with the free-electron plasma. The electromagnetic waves in a metal are then called *plasmon–polaritons* (mixed photon–plasmon modes). As we shall see, SERS and plasmonics are mostly concerned with plasmon–polaritons, rather than 'pure' plasmon modes.

Surface Plasmon–polaritons

Finally, in 1957, shortly after the introduction of the term plasmon, Ritchie [143] predicted the existence of another family of plasma modes in thin films, corresponding to longitudinal charge density waves propagating at a metal/dielectric interface. This prediction was confirmed experimentally in 1959 [144] and these modes (once quantized) were called *surface plasmons* a year later in 1960 [145]. In fact, these surface plasmon modes were originally introduced [143] within the electrostatic approximation (to be treated in Section 5.1.4). If however retardation effects are not negligible, then these charge density waves cannot exist without being associated with a transverse electromagnetic wave (a photon). This then corresponds to a mixed mode where the energy is shared between the charge density wave (plasmon) and the electromagnetic wave (photon), and they should therefore be called *surface plasmon–polaritons*.

---

[5] The exact definition of a polariton remains a matter of choice. Some people reserve this name for media with a strongly resonant optical response (with phonons or plasmons for example), and keep the term photon for 'standard' dielectrics (with a constant relative dielectric function $\epsilon \neq 1$). However, strictly speaking, a photon corresponds only to an electromagnetic wave in vacuum, and it becomes a polariton in any media. See for example Ref. [142] for a detailed discussion on the nature of polaritons.

There is therefore a fundamental difference between plasmon modes and the surface plasmon modes as introduced by Ritchie [143]. Plasmons can exist either by themselves without mixing with a photon, or as a mixed plasmon–photon mode (plasmon–polariton). Surface plasmons however are always strictly speaking surface plasmon–polaritons (mixed modes). The 'pure' surface plasmon modes are only an approximation of a surface plasmon–polariton for which the photon contribution is small or negligible (for example in the electrostatic approximation). We will come back to this distinction later.

3.3.3. The relation between plasmons and the dielectric function

The previous discussion in terms of elementary excitations and their quantization is useful to understand the origin of the term plasmon and the definition of the various types of plasmons. We will indeed come back to it shortly in more detail. However, one should bear in mind that plasmons and plasmon–polaritons are rarely viewed as quasi-particles, and are in fact mostly described as charge density oscillations (for plasmons), or electromagnetic waves in a medium (for plasmon–polaritons).

The body of early work on plasmons in the late 1950s, both theoretical and experimental, was concerned primarily with electron energy loss (EELS) in metals, and not directly with the optical properties of metals themselves. The former relates more to the dynamics of the free-electron plasma by itself rather than to its interaction with an external electromagnetic field. However, this dynamics is partially governed by electromagnetic interactions within the electron gas and with its environment. Charges, static or moving, are sources of electromagnetic field and a charge density wave cannot exist without an associated electromagnetic (or at least electric) wave. Reciprocally, the optical response of the free electrons is determined by their dynamical properties. *The dynamics of the plasma is therefore intricately linked to its optical properties and both can be entirely described using the dielectric function of the metal.*

Depending on the context, it may therefore be useful to emphasize one aspect (charge density and free-electron-gas dynamics) or the other (optical response). In the case of plasmonics, and even more so for SERS, the 'optical response approach' is usually the most relevant. In this sense, *the only thing needed to study SERS and plasmonics effects with a given metal is a knowledge of its optical response described by a relative dielectric function* $\epsilon(\omega)$ (and possibly $\epsilon(\mathbf{k}, \omega)$ if non-local effects are considered). One could then 'forget' about the fact that this optical response is the result of the free-electron dynamics, about the presence or not of charge and/or surface charge oscillations, or collective charge oscillations. Within this 'optical response approach', the plasmons and plasmon–polaritons can then simply be viewed as *electromagnetic modes* of the system under consideration. A detailed general discussion of electromagnetic modes will therefore be given in the rest of

Section 3.3. Its purpose is two-fold: firstly to show that plasmon-related effects are simply a specific class of a wider family of electromagnetic effects. Secondly, to introduce the terminology that is commonly used to qualify plasmon and plasmon–polariton modes. Note that this is somewhat a long theoretical digression in the fundamentals of plasmonics and it is possible to jump directly to the more practical aspects of plasmonics discussed in the rest of this chapter (Section 3.4 and beyond). In fact, this digression may be more digestible (and more useful) as a second read once the rest of the chapter has been understood.

3.3.4. Electromagnetic modes in infinite systems

The concept of elementary excitations or modes in infinite systems

The study of the elementary excitations or modes of a system is common in many areas of physics as part of linear response theory or Fourier analysis. In a nutshell, it consists in finding specific solutions (eigenvectors) of the physical system under study. In infinite systems with translational invariance, these solutions are propagating plane waves, i.e. solutions where all quantities (usually scalars like charge density, or vectors like electric field) have an oscillatory dependence (in space and time) of the form: $\cos(\mathbf{k} \cdot \mathbf{r} - \omega t + \phi)$, or in complex notation $\exp(\mathrm{i}\mathbf{k} \cdot \mathbf{r} - \mathrm{i}\omega t)$ (see Appendix C). Such solutions in general exist only for specific values (eigenvalues) of ω and $\mathbf{k}$. These solutions can usually be described by one or more *dispersion relations* $\omega(\mathbf{k})$. Each of the allowed solution with a given ω and $\mathbf{k}$ is then called an *elementary excitation or mode* of the system, and corresponds to a propagating plane wave. The reason why the modes are indexed here by their $\mathbf{k}$ vector is that the system is assumed to be invariant by translation in all directions.

When such a system is described in quantum mechanics, the elementary excitations are quantized, and they can be viewed either as plane waves, with frequency ω and wave-vector $\mathbf{k}$, or as particles with energy $\hbar\omega$ and momentum $\hbar\mathbf{k}$, linked by the dispersion relation $\omega(\mathbf{k})$. This step is, however, usually not necessary to discuss the classical properties of a system. Despite this, it is common to use the name of the quantum particle to designate the modes or elementary excitations, even when studied within a classical approach. This is in particular the case for plasmons.

Particles and quasi-particles

Many physical systems are affected by damping or losses in one form or another. In this case, a wave cannot propagate unchanged forever. Its amplitude must therefore decay in time and/or space. Such an excitation is called a *quasi-particle* (since it cannot exist ‘forever’ by itself). Note that plasmons are always quasi-particles (except in ideal non-absorbing metals).

This damping translates mathematically by the fact that ω and $\mathbf{k}$, which are related by the dispersion relation, cannot be both real for a quasi-particle. There are two possible points of view in this case:

- The first and most common approach is to consider ω real, which leads to $\mathbf{k}$ being complex. This implies an exponential decrease of the field amplitudes as $\exp(-(\mathrm{Im}(\mathbf{k}) \cdot \mathbf{r}))$. Such waves are called *evanescent waves or modes*, since they only propagate over a limited distance characterized by $1/|\mathrm{Im}(\mathbf{k})|$. The wave-vector for propagation is then $\mathrm{Re}(\mathbf{k})$.
- The second approach is to take $\mathbf{k}$ real, which leads to a complex frequency $\omega = \omega' - \mathrm{i}\omega''$ (with $\omega'' > 0$). This implies that the field amplitudes decay in time as $\exp(-\omega'' t)$ and such modes are then called *virtual modes* (this is the classical denomination, equivalent to a quasi-particle in the quantum point of view). The theory of virtual modes is common in nuclear and particle physics. These modes cannot exist as such (which is why they are called virtual), but they appear as resonances in the response of the system when the (real frequency) is equal to ω'. ω'' then characterizes the width of the resonance, or the lifetime of the virtual excitation ($\tau = 1/(2\omega'')$).

Both approaches are simply an attempt to represent *damped electromagnetic modes* or, from a quantum perspective, *quasi-particles.*

These two points of view are equivalent and which one is used depends on the exact physical situation and, in particular, on how the mode is excited:

- If an elementary excitation is maintained over time by an external source (acting in a given region of space, and driven at a given frequency ω), then it is logical to take ω real and $\mathbf{k}$ complex, and view the excitation as a wave decaying in space, i.e. an evanescent wave.
- If an elementary excitation is created at a given time by an external source (which is then switched off), then it is more logical to take ω complex and $\mathbf{k}$ real, and view the excitation as propagating in space but decaying over time, i.e. as a virtual mode. This point of view is also the only possible one for problems where $\mathbf{k}$ is not well defined (no translational invariance).

Finally, in infinite systems without losses and damping, elementary excitation can propagate 'forever' without decay in space and time. They are then called *propagating waves*, or from the quantum perspective, simply *particles.*

Longitudinal and transverse modes

One important concept for plasmon modes is that of longitudinal and transverse modes. A mode is described primarily by its frequency ω and wave-vector $\mathbf{k}$, linked by the dispersion relation $\omega(\mathbf{k})$. If the oscillating quantity is a vector, as is the case for the electric field $\mathbf{E}$, then one can distinguish two situations in isotropic and homogeneous media:

- $\mathbf{E}//\mathbf{k}$ everywhere ($\mathbf{k} \times \mathbf{E} = 0$). This is then called a *longitudinal mode or wave.*
- $\mathbf{E}\perp\mathbf{k}$ everywhere ($\mathbf{k} \cdot \mathbf{E} = 0$), which then corresponds to a *transverse mode or wave.*

The origin of the name longitudinal and transverse is then clear; it refers to the orientation of $\mathbf{E}$ with respect to the direction of propagation $\mathbf{k}$. These definitions can also be extended mathematically to a general vector field $\mathbf{E}$, without the need for a wave-vector $\mathbf{k}$ (and therefore also valid in the absence of translational invariance):

- $\nabla \times \mathbf{E} = 0$ for a longitudinal field, and
- $\nabla \cdot \mathbf{E} = 0$ for a transverse field.

For a propagating mode, with a dependence on complex notation of the type $\exp(i\mathbf{k}\cdot\mathbf{r})$, this simply reduces to the previous simple definition in terms of $\mathbf{k}$. Finally, there is a theorem from vector analysis stating that any vector field can be decomposed (uniquely) into the sum of a transverse and a longitudinal field.

Electromagnetic modes in infinite (3D) vacuum – photons

One simple example of an infinite physical system is the electromagnetic field in vacuum. The electromagnetic modes are then derived easily from Maxwell's equations (see Section F.1). In particular, the equations $\nabla \cdot \mathbf{E} = 0$ and $\nabla \cdot \mathbf{B} = 0$ imply that $\mathbf{E}$ and $\mathbf{B}$ are transverse fields. The modes are then *transverse propagating plane waves* characterized by a wave-vector $\mathbf{k}$ and frequency ω, related by the dispersion relation: $\omega = c|\mathbf{k}|$. Each pair of real ω and $\mathbf{k}$ satisfying the dispersion relation corresponds to a *propagating electromagnetic mode.* After quantization, these modes can also be viewed as particles and are then called photons, but most of their properties can also be described classically using Maxwell's equations.

Electromagnetic modes in an infinite (3D) medium – polaritons

A similar situation occurs for transverse electromagnetic waves in a medium[6] with (local) relative dielectric function $\epsilon(\omega)$. The dispersion relation

---

[6] The media are always assumed to be non-magnetic (with relative magnetic permeability $\mu = 1$), unless otherwise stated.

is then modified to give (Eq. (F.2) in Section F.1):

$$\epsilon(\omega)\omega^2 = c^2\mathbf{k}\cdot\mathbf{k}. \tag{3.9}$$

Such an electromagnetic wave creates in the medium an internal polarization wave: $\mathbf{P} = \epsilon_0(\epsilon(\omega) - 1)\mathbf{E}$. These modes are then called polaritons because they couple transverse electromagnetic excitations (photons) with an internal polarization **P**, which originates physically from internal excitations of the medium, such as excited bound electrons, phonons or plasmons. In non-absorbing dielectrics, such as glass, the polaritons are quantum particles similar in many ways to photons. For metals, in the region where the optical response is dominated by the free-electron plasma, these modes are usually called *bulk plasmon–polaritons* and are quasi-particles. Note that for these modes, **E** and **P** are transverse, and the internal charge density is therefore $\rho_{\text{int}} = 0$ everywhere. *There are no macroscopic charge density oscillations.* The denomination plasmon–polariton can be misleading in this respect since there is no net charge density wave, but only a polarization wave.

Longitudinal electric wave in an infinite (3D) medium

Finally, in an infinite medium, there is another family of electromagnetic modes that do not exist in vacuum. Maxwell's equations (C.27) – (C.27) state that the electric displacement $\mathbf{D} = \epsilon_0\epsilon\mathbf{E}$ is transverse: $\nabla\cdot\mathbf{D} = 0$. This condition was previously fulfilled by assuming that the electric field **E** was also transverse, but an alternative possibility is that $\epsilon(\omega) = 0$. Such a condition can be fulfilled in some media (and in particular in metals) at one or more specific frequencies ω. If this is the case, then $\mathbf{D} = 0$, and Eq. (C.27) implies that **H** is a longitudinal field. Because **H** is also a transverse field from Eq. (C.27), it must therefore be zero: $\mathbf{H} = 0$. The remaining equation (C.27) then implies that **E** is a longitudinal field. Such a solution therefore corresponds to a *longitudinal electric wave* (with zero magnetic field). This wave is associated with an internal polarization wave $\mathbf{P} = -\epsilon_0\mathbf{E}$ (since $\mathbf{D} = 0$). Moreover, because **P** is longitudinal, it also corresponds to an internal charge density wave $\rho_{\text{int}} = -\nabla\cdot\mathbf{P} \neq 0$. These modes therefore correspond to a *real charge density wave*, with an associated longitudinal electric wave. Because of the longitudinal nature of these modes (and of the structure of Maxwell's equations), these modes *cannot couple to light* or photons, which are transverse electromagnetic excitations. For metals, these modes are the 'pure' plasmon modes, as defined previously, and are usually called *bulk plasmons.* In the local approximation, they exist only at specific frequencies for which $\epsilon(\omega) = 0$, and **k** can take any values (small enough for the local approximation to remain valid). For a more detailed non-local treatment, the dispersion relation of these modes is given by $\epsilon(\omega,\mathbf{k}) = 0$. Bulk plasmons are longitudinal modes and therefore do not couple to light.

Electromagnetic modes in an infinite (3D) metal

We have introduced in the previous two subsections the two types of electromagnetic modes that exist in an infinite 3D medium. Here we discuss briefly further their properties for the simple case of an ideal metal whose optical properties are described by a simple Drude model. $\epsilon(\omega)$ is then given by Eq. (3.6), where we neglect the damping term ($\gamma_0 \approx 0$), i.e.:

$$\epsilon(\omega) = \epsilon_\infty \left(1 - \frac{\omega_p^2}{\omega^2}\right). \tag{3.10}$$

- The first types of modes are the *bulk plasmon–polaritons.* Using the above expression for $\epsilon(\omega)$, their dispersion relation, Eq. (3.9), can then be simplified as:

  $$\omega^2 = \omega_p^2 + \frac{c^2}{\epsilon_\infty} k^2. \tag{3.11}$$

  These are transverse modes and can essentially be viewed as light or photons, whose properties are modified by the interaction with the electrons of the metal. For a real metal, these modes are damped (i.e. they are quasi-particles) and the corresponding waves are evanescent. Note that it has been pointed out [146,147] that these modes arise from the interaction of the photons with single-electron excitations and have therefore no direct relation to collective excitations as often assumed.

- The second types of modes are the *bulk plasmons*, which are longitudinal modes corresponding to internal charge density oscillations and an associated electric wave (no magnetic field). These modes *do not couple to photons* and are therefore mostly irrelevant to SERS and many plasmonics effects. The condition $\epsilon(\omega) = 0$ reduces simply in the Drude case to $\omega = \omega_p$. These modes are therefore simply collective oscillations of the plasma charge density, and occur only at the plasma frequency ω_p. This is not a surprise since they are the natural modes of oscillation of the free-electron plasma, in the absence of interaction with light or any other electromagnetic sources[7].

  The dispersion relation for these modes could be determined more accurately using a non-local dielectric function. For example, including

---

[7] Note that for a real metal, the bulk plasmon frequency (determined by $\epsilon(\omega) = 0$) may differ from the plasma frequency ω_p because of the presence of inter-band transitions. This is for example the case for gold.

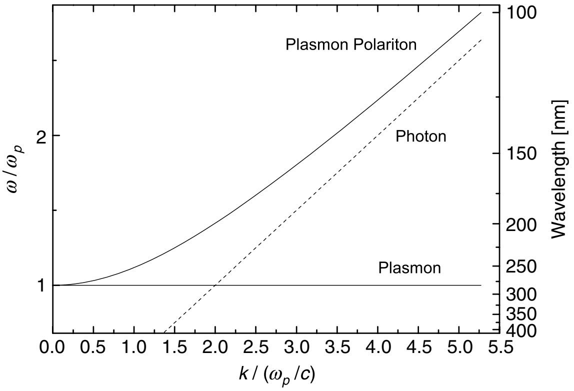


Figure 3.4. Dispersion relations of electromagnetic modes in an infinite 3D metal (silver here), showing the bulk plasmon–polaritons and the bulk plasmons. The dashed line is the dispersion of hypothetic photons non-interacting with the free-electron plasma.

hydrodynamic interactions in the electron plasma, the relation $\epsilon(\omega, \mathbf{k}) = 0$ leads to [146]:

$$\omega^2 = \omega_p^2 + \beta^2 k^2, \tag{3.12}$$

where β is a constant that depends on the chosen hydrodynamic model. In a first approximation, one can assume $\beta \approx \sqrt{0.6} v_F$, where v_F is the Fermi velocity of the metal [146]. β is therefore much smaller than c, which means that the local approximation $\omega \approx \omega_p$ is very good in the optical range.

The dispersion relations of the bulk plasmon–polaritons and bulk plasmons are illustrated in Fig. 3.4 for $\hbar\omega_p \approx 4.4$ eV and $\epsilon_\infty = 4$ (these are the values for silver). One sees that bulk plasmon–polariton modes occur only for energies *larger* than the plasmon frequency. For most metals, these modes are therefore in the UV, or even deep UV. Bulk plasmon–polaritons (like bulk plasmon) are therefore in general not relevant for optical applications, including SERS. They are discussed here as an introduction to the surface modes (of interest to SERS and plasmonics), on which we now focus.

3.3.5. Electromagnetic modes of a system of material bodies

Electromagnetic modes

We have so far discussed the concept of modes or elementary excitation focusing only on infinite systems with translational invariance. To understand the various types of plasmons and plasmon–polaritons, it is now useful to consider in more detail the general problem of the electromagnetic modes of a system of material bodies, i.e. in the presence of interfaces. These considerations can then be applied to metallic structures (and therefore to plasmons and plasmon–polaritons), but the same modes can also exist in other kinds of systems.

The system considered here is restricted to one or more entities of different materials described by a local dielectric function and separated by ideal boundaries. Any such electromagnetic problem can therefore be formally studied by solving Maxwell's equations with the appropriate boundary conditions at the various interfaces (see Appendix C). The modes, or eigenvectors, of the system consist of specific solutions from which any general solution can be inferred. These are no longer necessarily plane waves because there is no translational invariance. Here we consider only electromagnetic fields oscillating at a frequency ω, and the modes then correspond to specific values of ω (continuous or discrete). All physical quantities can be described by a complex amplitude, where a $\exp(-\mathrm{i}\omega t)$ dependence is assumed.

2D and 1D systems with translational invariance

In systems where the translational invariance only exists in 2 dimensions (layered structure), or even 1D (for example for a cylinder) then the modes can again be indexed by a **k** vector parallel to the directions of translational invariance. This leads to a dispersion relation for the modes of the type $\omega(k_x)$, corresponding to propagating modes along (Ox), which can as before be fully propagating, evanescent, or virtual modes.

Studying such systems with a perfect translational invariance may appear as a purely academic exercise, since infinite plane or cylinders never exist in real life. However, it is sufficient that the invariance holds over a typical length scale of the problem, which in many cases is simply the wavelength λ of the electromagnetic radiation, for the approximation to be valid and meaningful. In many experimental situations, all interfaces can indeed be approximated by plane surfaces at length scales of the order of $\sim\lambda$.

Systems without translational invariance

Only when the objects exhibit features on length scales of the order of the wavelength does the translational invariance really fail. In such systems without translational invariance, and in particular for particles of dimensions comparable with λ, the description in terms of a **k** vector becomes irrelevant and inadequate. The electromagnetic modes of the system then correspond to discrete values of ω. If $\omega = \omega' - \mathrm{i}\omega''$ is complex, they again correspond to *virtual modes*. As discussed previously, these are not real modes of the system but resonances with a width ω''. In problems where the oscillation of the fields is not imposed externally, these virtual modes can also be viewed as elementary excitation with a lifetime $\tau = 1/(2\omega'')$.

3.3.6. Classification of electromagnetic modes

We now discuss in more detail a possible classification of some common electromagnetic modes in systems of material bodies as defined before. Here we follow Ref. [148], which gives a detailed description of this classification.

Purely longitudinal modes

Similarly to the case of infinite media, one class of solution regroups those for which $\epsilon(\omega) = 0$ inside the material bodies. They have similar characteristics to those obtained in infinite media, i.e. they are longitudinal electric waves with $\mathbf{H} = 0$ (no magnetic field) and $\nabla \times \mathbf{E} = 0$. Since $\nabla \cdot \mathbf{E} \neq 0$, they are also associated with a bulk internal charge density (ρ_{int}) wave (bulk plasmon in metals). In addition, the boundary conditions at interfaces imply the presence of surface charge density (ρ_s) waves. The condition $\epsilon(\omega) = 0$ is the same as for infinite media. These modes are therefore simply *modified bulk modes*, which are confined inside the material bodies by the interfaces. As before, because these modes are longitudinal, they do not couple to light.

Incident wave modes

We now focus on the situations where $\epsilon(\omega) \neq 0$, and therefore $\nabla \cdot \mathbf{E} = 0$ in the materials. The condition $\nabla \cdot \mathbf{E} = 0$ is, however, not satisfied at the interfaces where the boundary conditions may imply the presence of surface charges. These solutions, although transverse inside the materials, are in fact a *mixture of longitudinal and transverse waves.*

Sufficiently far from all interfaces, 3D electromagnetic modes can exist (plane waves described by $\mathbf{k}$ and ω). A typical experiment will involve sending an *incident wave* of amplitude E_{Inc} toward the interfaces under study, which in general leads to an *outgoing (or scattered) wave* of amplitude E_{Sca}. Due to the linearity of Maxwell's equations, the scattered wave amplitude is proportional to the incident wave amplitude: $E_{\text{Sca}} = fE_{\text{Inc}}$, where f is a factor, possibly complex, depending on the geometry, the optical properties of materials, ω, and $\mathbf{k}$. f characterizes the *optical response* of the system. Note that we have neglected for simplicity here the vectorial nature of the field (i.e. the field polarization). In reality, f should be a matrix and is related to the amplitude scattering matrix [149].

Some electromagnetic modes only exist in the presence of such an incident wave and are therefore called *incident wave modes.* These modes would still exist (as a simple incident wave) if the interface was removed. They can therefore be viewed as a modification of the incident wave due to the interface. The standard reflection or refraction of a plane wave at a dielectric/dielectric interface is an example of an incident wave mode.

Bound modes or surface modes

There can also be modes that exist even in the absence of an incident wave (with only a scattered wave), and they are called *bound modes*. The bound modes owe their existence to the interface and, in contrast to incident wave modes, have no equivalent in infinite media if the interface is removed. They are therefore also called *surface modes*. Note that this denomination of incident wave vs bound mode is similar to the classification of the solutions in quantum mechanics.

The conditions $E_{\mathrm{Sca}} \neq 0$ and $E_{\mathrm{Inc}} = 0$ imply that the optical response, f, is infinite for bound modes. This appears to violate energy conservation, but in reality, the condition $f = \infty$ only occurs for complex values of ω or k. This means that bound modes are damped , and should therefore be viewed either as evanescent waves (ω real, k complex) or virtual modes (k real, ω complex). Because of this damping, it is necessary in practice (but not in theory) to have an incident wave to excite and maintain these modes.

If both ω and k are real, then f presents a sharp peak, instead of a real infinity, for values approaching those of the damped bound modes. This results in strong optical resonances, instead of unrealistic infinite optical response. For example, it can be shown that for a virtual bound mode with imaginary frequency $\omega = \omega' - \mathrm{i}\omega''$, the resonant response occurs for real frequencies equal to ω' and with a half-width of ω''. The smaller the ω'', the sharper the resonance.

Surface plasmon–polaritons at a metal/dielectric interface (discussed in Sections 3.4 and 3.5) are examples of bound or surface modes. In most cases, the *plasmon resonances* mentioned in the context of SERS or plasmonics are resonances associated with these surface plasmon–polariton modes. The resonant optical response can manifest itself differently depending on the type of surface plasmon–polariton creating it, i.e. large optical absorption, large far-field scattering, or large local field. For this reason, it can be useful to further classify the various types of surface modes.

3.3.7. Other properties of electromagnetic modes

Radiative and non-radiative modes

Bound or surface modes and incident wave modes can be associated with an outgoing wave or scattered wave. If this outgoing wave is propagating, it will correspond to a standard 3D mode (photon) when far from the interface, and such a bound mode is then called *radiative*. However, if the outgoing wave is evanescent, then the field decays exponentially away from the surface and this is called a *non-radiative mode*.

For a *bound radiative mode*, the scattered wave radiates energy in the far field but there is no incident wave to provide this energy. Therefore, these modes cannot be real and are always *virtual modes*. If an incident wave excites such a mode at resonance, the strong optical responses will translate in a large far-field radiation, i.e. there is a resonance in scattered intensity. Note that the damping of the mode is in general due not only to radiation, but also to optical absorption in the materials. The resonance in scattering can therefore be accompanied by a resonance in absorption. Radiative modes can be called more or less absorptive depending on the proportion of radiated to absorbed power.

The limiting case of a very absorptive mode is, in fact, a non-radiative mode. When exciting a bound non-radiative mode at resonance, the incident energy is transferred to the mode, but is not re-emitted in the far field because it is non-radiative. There is a resonance in optical absorption, not in scattering.

Localized modes

We have seen that a **k** vector could be defined in the directions of translational invariance. In each of these directions, the mode can either be propagating if **k** is real, or evanescent if not. If the mode is not propagating (evanescent) in all of these directions, then we can call it a *localized mode*. One particular case is that of systems without translational invariance, for which **k** is not relevant. All surface modes are then *localized modes*. This is for example the case of particles small or comparable to the wavelength (i.e. most nano-particles), where all surface modes are necessarily localized modes.

Excitation of electromagnetic modes

The description of a system in terms of electromagnetic modes may appear at first sight to be purely academic in nature. However, once the physical nature of the modes is understood, it is actually a powerful framework to understand the response of a complex system or the interactions between different subsystems. The reason is that a complex system can be divided into subsystems. Furthermore, when the modes of the individual subsystems are known, the response of the whole system can be studied as an interaction or coupling between subsystem modes. For weak coupling, the modes of the subsystems are essentially unchanged, while for stronger coupling the interaction can lead to new types of modes. For two modes to couple, they need to have the same frequency ω (which corresponds to energy conservation), and in the case of translational invariance the same wave-vector $\mathbf{k}$ (for momentum conservation) is also required. We can give some simple examples in terms of the types of modes discussed previously:

- An incident wave mode is by construction automatically coupled to an incident photon with the same ω and $\mathbf{k}$.
- Similarly, a radiative mode is coupled to an outgoing photon with appropriate ω and $\mathbf{k}$.

- An incident photon can also couple to a bound mode provided ω (and if relevant $\mathbf{k}$) conservation is fulfilled. For virtual modes ($\omega = \omega' - \mathrm{i}\omega''$ complex), the ω conservation applies to real parts ($\omega = \omega'$) and is broadened with a width ω''. Similarly, for evanescent modes ($\mathrm{Im}(k) > 0$), the $\mathbf{k}$ conservation is broadened.

There are specific configurations where one electromagnetic mode may be excited, for example at a given incident angle, incident polarization, or wavelength. One then expects a resonant optical response (for example in reflectivity or absorption) when the parameters match the excitation condition of such a mode. When the electromagnetic modes are plasmons, or plasmon–polaritons, this is called a *plasmon resonance.* The nature and characteristics of such a resonance depends on the nature of the electromagnetic mode giving rise to it. The term *plasmon resonance* can, therefore, have different meanings depending on the context. It can, for example, correspond to enhanced energy loss at a specific electron beam energy, or to a decreased reflectivity at a specific incident angle on a surface, or to an increased field intensity at the surface (which can give rise to a SERS signal) at a specific wavelength, etc. In the context of SERS and plasmonics, plasmon resonances *refer in most cases to bound modes called surface plasmon–polaritons*, discussed in detail in Sections 3.4 and 3.5.

3.3.8. Summary and discussion

The somewhat formal discussion given in the last few subsections should hopefully become clearer when discussing specific examples in the following sections.

The various types of plasmon excitations are tentatively summarized in Fig. 3.5 along with their main properties. The *'pure' plasmon and surface plasmon modes*, as originally defined and studied by Pines [141], Ritchie [143], and others are not directly relevant to SERS and plasmonics because they do not interact with light[8]. This is because they correspond to *longitudinal* excitation of the electric field (electric wave) and cannot therefore couple to photons. These plasmons are also in fact those associated with true collective charge density oscillations.

The electromagnetic modes that interact with light are the *plasmon–polaritons*, which mix photons with internal excitations of the metal. *Bulk plasmon–polaritons* are those modes that exist in an infinite metal and can be viewed simply as photons propagating in a metal. These modes are again not directly relevant to SERS and plasmonics.

---

[8] They however play an important role in many other properties of metals and metal surfaces [129], in particular for other types of spectroscopies like Electron Energy Loss Spectroscopy (EELS).

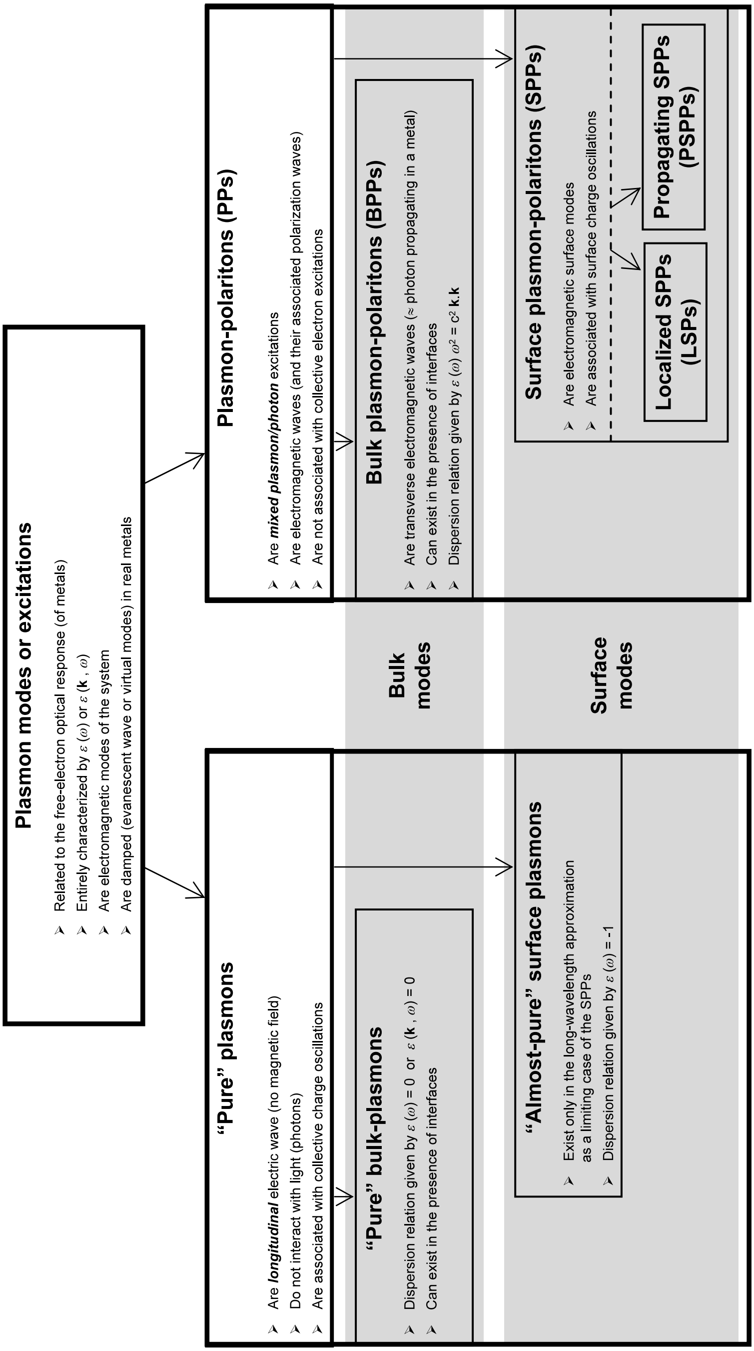


Figure 3.5. Schematic representation of the various plasmon types and their main properties.

The presence of an interface (typically metal/dielectric) gives rise to a new family of electromagnetic modes, which happens to be much more interesting for SERS and plasmonics. These additional modes are bound modes or *surface modes*, as described in the previous section in a more general context. In the case of metals, they are called *surface plasmon–polaritons* (SPPs) .

These electromagnetic surface modes are transverse inside and outside the metal (because $\epsilon \neq 0$). Note however that they do also have a longitudinal component because of the discontinuities at the interface, and are in this sense also partly longitudinal surface plasmon waves. A non-local description of ϵ, where the interface is no longer treated as a singularity, is necessary to identify clearly this mixed longitudinal–transverse nature, see for example Ref. [146]. The 'pure' longitudinal surface plasmons introduced by Ritchie [143] are simply a limiting case where the transverse contribution is negligible (in the electrostatic approximation). The strict decoupling between longitudinal ('pure' plasmons) and transverse (plasmon–polaritons) modes, which was natural for bulk modes in infinite media, is no longer possible for surface modes. Instead of introducing another terminology, these modes are simply called *surface plasmon–polaritons.*

Finally, let us note that there are typically three reasons (two good ones, and the last one partly flawed) that are put forward to justify the denomination of *surface* plasmon–polariton or *surface* mode:

- Firstly, they would not exist without the interface.
- Secondly, the characteristics of these modes depend not only on the optical properties of the metal, but also on that of the dielectric forming the interface.
- Finally, in many cases, these surface modes are localized at the interface (for example the electric field decays exponentially from it). This is however not really a good criterion, since as we will see later, for small metallic objects the electric field of such a surface mode can be almost uniform inside the object and extend infinitely outside (i.e. it is a radiative mode).

After this long digression, these SPP modes are finally the ones that are relevant to SERS and plasmonics. Whenever plasmons or surface plasmons as such are mentioned in a SERS context, it is arguably an abuse of language (although a very common one), and it is strictly speaking referring to surface plasmon–polaritons. As for general electromagnetic modes, these can come in several flavors: propagating, localized, radiating, non-radiating, bound, virtual, or evanescent. They can give rise to various resonance effects, which may be used for various applications. All these different cases will be the subject of the following sections, which focus on a more practical description of SPP modes in metals.

3.4. SURFACE PLASMON–POLARITONS ON PLANAR INTERFACES

The previous discussion was aimed at introducing plasmons from an historical point of view, and emphasized the intricate relationship between optical properties and the free-electron-plasma dynamics. This helped in understanding the origin of the term plasmon and its relation with 'charge oscillations', as often loosely stated. We also highlighted the fact that surface plasmon–polaritons (SPPs) can simply be viewed as *electromagnetic surface modes* of the system under consideration. The study of SPPs then simply reduces to an electromagnetic problem, where the actual role of the electrons can be ignored and is simply contained in the dielectric function describing the metal. This more pragmatic approach is particularly suited to study SPPs in various geometries. One important case is that of a planar metal/dielectric interface, firstly because it is reasonably easy to solve, secondly because many interfaces can be considered as planar over a characteristic length scale of the problem (typically the wavelength λ), and finally because it supports the types of SPPs (propagating SPPs) most useful for several plasmonics applications.

3.4.1. Electromagnetic modes for a planar dielectric/metal interface

Description of the electromagnetic problem

To illustrate this, we will restrict ourselves to a local relative dielectric function $\epsilon(\omega)$ and consider first the textbook example of a plane metal/dielectric interface [150]. The metal (region 2), described by $\epsilon(\omega)$ occupies the half-plane $z > 0$, and a non-absorbing dielectric (region 1), with relative dielectric function $\epsilon_M \geq 1$, real and constant, forms the outside medium in $z < 0$, as shown schematically in Fig. 3.6. Bulk electromagnetic modes exist far away from the interface both in the dielectric (photon modes) and in the metal (bulk plasmon and bulk plasmon–polaritons). Bulk plasmon modes for which $\epsilon(\omega) = 0$ may also exist in the presence of the interface, but we ignore these longitudinal modes here since they do not interact with light. Because of the translational invariance along directions in the plane $z = 0$, the electromagnetic modes should be characterized by their frequency ω, and tangential wave-vector, k_x, in the plane (we assume $k_y = 0$ without loss of generality). Various technical aspects of this problem are treated in Appendix F, which may therefore be read in conjunction with this. As discussed before, the presence of the surface can introduce several types of modes:

- Incident wave modes are those where an incident wave (and possibly a scattered/reflected wave) is present. From the standpoint of a metal/dielectric interface, these modes correspond simply to the classic problem of reflection/refraction of a plane wave at the interface. This problem is treated in many electromagnetic textbooks [96,151], and the most important features are summarized in Appendix F . The refracted/transmitted wave in the metal is always evanescent because of optical absorption in the metal ($\epsilon(\omega)$ complex or real negative). If the incident wave is propagating, and if there is a reflected wave, then it is also propagating, and the mode can be called radiative, and more or less absorptive depending on the reflection coefficient. These incident wave modes exist for all ω and k_x compatible with the dispersion relation in the dielectric, i.e. $k_x^2 + k_z^2 = \epsilon_M \omega^2/c^2$. Note that for TM polarization[9] they are associated with a propagating surface charge density (ρ_s) wave at the interface, but they are not surface modes (or bound modes) as defined previously.

- A particular case of incident wave modes is that for which there is no reflected wave (only a refracted wave). This happens only (see Appendix F) for a specific angle of incidence (called the Brewster angle) . The corresponding Brewster modes are an example of non-radiative incident wave modes.
- The other family of modes, the bound or surface modes, is the one of interest here. They correspond to solutions where no incident wave is present, only the scattered wave. These are derived and discussed in detail in Section F.2.5 of Appendix F and we now discuss in more detail their properties for a metal/dielectric interface.

Surface modes of a metal/dielectric interface

It is shown in Appendix F that there are no surface modes with TE (s-) polarization; only with TM (p-) polarization and we therefore focus on this latter case only. Such a surface mode consists of a scattered wave (no incident wave) in the dielectric with wave-vector $\mathbf{k}'_1 = k_x\mathbf{e}_x + k'_{1z}\mathbf{e}_z$ and a transmitted wave in the metal with wave-vector $\mathbf{k}_2 = k_x\mathbf{e}_x + k_{2z}\mathbf{e}_z$, as shown in Fig. 3.6; see Appendix F for more details. The electromagnetic fields of these modes are given by (see Appendix F):

$$\begin{cases} \mathbf{H}_2 = (H_y\mathbf{e}_y)\exp(\mathrm{i}k_x x + \mathrm{i}k_{2z}z) \\ \mathbf{E}_2 = \dfrac{H_y}{\omega\epsilon_0\epsilon_2}\,(k_{2z}\mathbf{e}_x - k_x\mathbf{e}_z)\exp(\mathrm{i}k_x x + \mathrm{i}k_{2z}z), \end{cases} \tag{3.13}$$

and similar expressions for $\mathbf{H}'_1$ and $\mathbf{E}'_1$ (the fields have been expressed in terms of a single amplitude: $H_y = H'_{1y} = H_{2y}$).

---

[9] TM and TE polarizations are defined in Section F.2.1 of Appendix F . For TM polarization, the magnetic field is perpendicular to the plane of incidence, meaning along y here. An example is shown in Fig. 3.6.

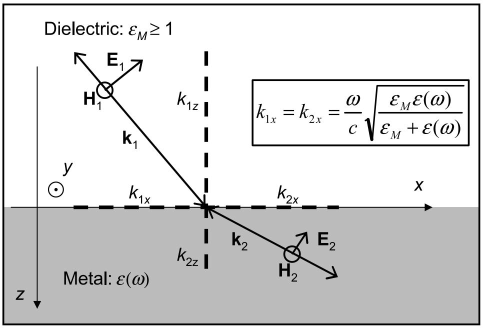


Figure 3.6. Schematic illustration of the electromagnetic problem of two TM waves, on either side of a metal (region 2)/dielectric (region 1) interface. The dispersion relations in each medium, together with the boundary conditions lead to the dispersion relation for SPPs (framed equation). The arrows representing the directions of the k vectors are shown for both directions to show that no assumption is made on the nature of the waves (for example incident or scattered wave in Region 1).

Moreover, the condition for the existence of such a surface mode is (see Appendix F):

$$k_x^2 = k_0^2 \frac{\epsilon \epsilon_M}{\epsilon + \epsilon_M}. \tag{3.14}$$

There are two possible solutions for k_x (of opposite sign). Since they are physically equivalent, we will only consider, by convention, the one given by[10]:

$$\boxed{k_x = \frac{\omega}{c}\sqrt{\frac{\epsilon(\omega)\epsilon_M}{\epsilon(\omega) + \epsilon_M}}.} \tag{3.15}$$

Since $\mathrm{Im}(\epsilon) \geq 0$, we have $\mathrm{Re}(k_x) \geq 0$ and $\mathrm{Im}(k_x) \geq 0$.

One can moreover deduce the corresponding expressions for k'_{1z} and k_{2z}. This is not as trivial as may seem, mostly because of sign issues, a problem often swept under the carpet. This is discussed extensively in Section F.2.5.

This set of expressions entirely defines the TM electromagnetic surface modes for each frequency ω. Each mode is composed of two electromagnetic waves, one on either side of the interface, with $\mathbf{k}$ vectors given by the above

---

[10] In all these expressions, there are in principle two choices for the square root of a complex number. We use the 'standard' complex square root convention, i.e. the one with a positive real part (or if it is zero, the one with the positive imaginary part). More explicitly: for $-\pi < \phi \leq \pi$, $\sqrt{r \exp(i\phi)} = \sqrt{r}\exp(i\phi/2)$. We then have $\mathrm{Re}(\sqrt{z}) \geq 0$, and $\mathrm{Im}(\sqrt{z})$ has the same sign as $\mathrm{Im}(z)$.

relations. The dispersion relation for the surface modes is Eq. (3.15), which links the frequency ω to its tangential wave-vector k_x.

Finally we note that these modes are associated with a surface charge density ρ_{Surf} created by the discontinuity of the normal component of the electric field. They, therefore, correspond to a surface charge density wave along the interface with wave-vector $k_x\mathbf{e}_x$, propagating, pseudo-propagating, or evanescent depending on k_x (see the next subsection). Note however that this is not a defining characteristic of the surface modes since a similar surface charge density wave exists for most TM incident wave modes (for example for reflection at the metal/dielectric interface).

Classification of surface modes

The exact nature of each mode will depend on whether the wave-vector components k_x, k_{1z}, and k_{2z}, are real or complex. A mode will correspond to a *propagating* wave along a given direction if its wave-vector along this direction is real. Contrarily, if it is imaginary, it is *evanescent* (with an exponential decay of the intensity). If this decay occurs over long distances compared to the wavelength, the wave can be called *pseudo-propagating.* In our example here, if k'_{1z} is real, then we have a propagating scattered wave in medium 1. Contrarily, if k'_{1z} has a non-zero imaginary part, then we have an evanescent wave along the z direction. The field amplitudes then decay as $\exp(-|\text{Im}(k'_{1z})z|)$ and are therefore negligible in the far field ($z \to -\infty$). This corresponds to a *non-radiative* mode, or trapped surface wave. A detailed discussion of the various cases is given in Section F.2.5. Here we only highlight the main results, which are summarized schematically in Fig. 3.7.

It can be useful in this context to consider first the ideal case where the metal (medium 2) is non-absorbing, i.e. its dielectric function is real. Such a metal can in principle support infinitely propagating waves (with k_x and k_{2z} real). In reality, this cannot happen since a real metal always presents a small amount of absorption and all waves are strictly speaking evanescent. However, if $|\text{Im}(k)| \ll |\text{Re}(k)|$ in one direction, then the wave will propagate over long distances (many spatial wavelengths), without substantial decay in amplitude (it is a pseudo-propagating wave).

We can then distinguish three cases for an ideal non-absorbing metal, and extend this classification to a real (absorbing) metal. These situations are represented schematically in Fig. 3.7 and discussed below:

- $\epsilon(\omega) > 0$ (Fig. 3.7(a)) corresponds to an ideal metal at high frequencies, which then behaves as a standard dielectric with no absorption. All wave-vector components, k_x, k'_{1z}, and k_{2z} are then real. k'_{1z} and k_{2z} must also both be positive. These solutions are discussed further in Section F.2.5. They are called *Brewster's modes* and are not strictly speaking surface modes for an ideal metal since k'_{1z} then represents

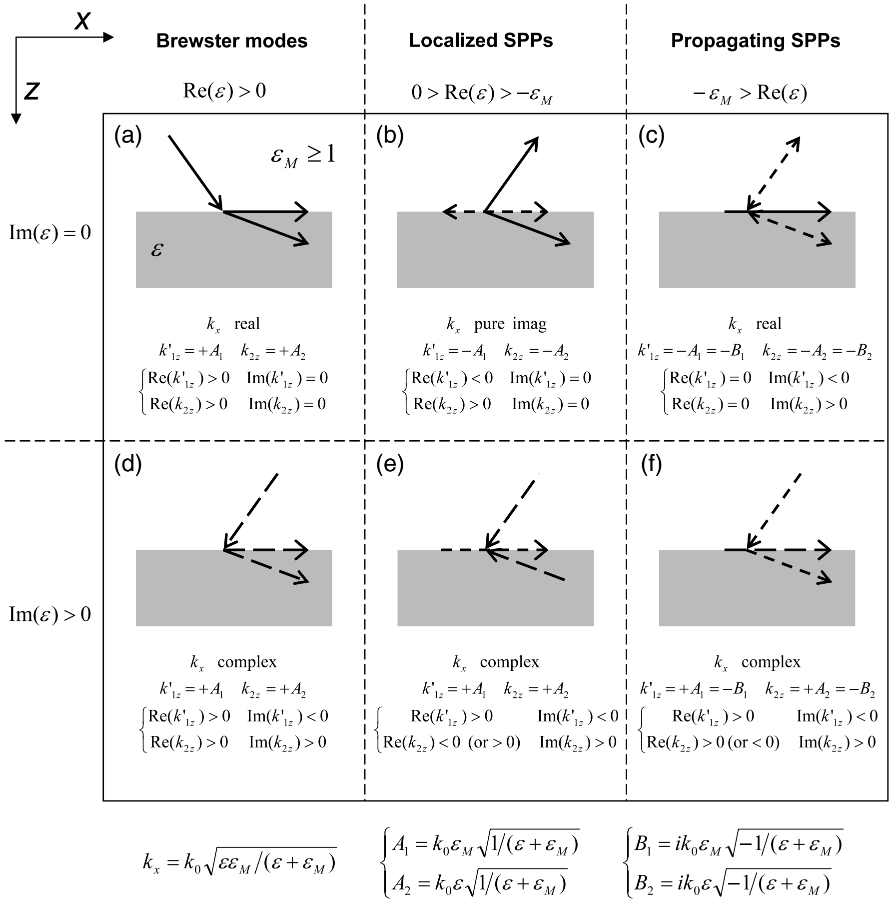


Figure 3.7. Schematic illustrating various types of electromagnetic modes existing at a planar metal/dielectric interface. Solid (dashed) lines indicate propagating (evanescent) waves. For all evanescent waves, the field amplitudes decay exponentially when moving away from the interface. The direction of propagation (given by $\mathrm{Re}(k_z)$) is indicated by arrows. Note that for SPPs, this may depend on the exact value of $\mathrm{Im}(\epsilon)$ (see Section F.2.5). Long-dashed lines are used for pseudo-propagating waves (only when $\mathrm{Im}(\epsilon)$ remains small). The Brewster modes for $\mathrm{Im}(\epsilon) = 0$ are strictly speaking incident wave modes, but the other situations correspond to surface modes. The appropriate expressions for k'_{1z} and k_{2z} are given for each case, along with the nature (real or imaginary) of the wave-vector components. See Section F.2.5 for more details.

an incident wave. They do become surface modes however, for a non-ideal metal with $\mathrm{Im}(\epsilon) > 0$ (Fig. 3.7(d)), since the wave in region 1 is then evanescent. In this sense Brewster's modes may be viewed as surface modes (then this becomes, ultimately, only a question of vocabulary). Note that these modes are associated with a propagating surface charge wave created by the discontinuity of the electric field. They are a mixture of propagating photons and surface charge waves, and are therefore surface polaritons. They, however, are usually not considered as surface plasmon–polaritons, a denomination reserved for the case where $\mathrm{Re}(\epsilon) < 0$. Their existence is pointed out here to show that surface charge density waves are not a defining characteristic of SPPs.

- $-\epsilon_M < \epsilon(\omega) < 0$ corresponds to the case of an ideal metal at frequencies just below the plasma frequency (Fig. 3.7(b)). We then have $k'_{1z} < 0$ and $k_{2z} > 0$ real of opposite sign, and $k_x = \mathrm{i}\kappa_x$ is pure imaginary. The corresponding surface wave does not propagate at all. These modes therefore correspond to localized modes, and can be called *localized SPPs*[11]. The energy of the evanescent surface wave is dissipated into the two waves propagating away from the surface. These modes are therefore *radiative*. They can in principle be excited by a wave with $k_x \approx 0$, i.e. at normal incidence, but the resonance condition is largely broadened by the strong evanescent nature of the modes. For a real metal with $\mathrm{Im}(\epsilon) > 0$ (Fig. 3.7(e)), both waves in region 1 and 2 become evanescent (pseudo-propagating if the absorption is small), and the corresponding modes are then non-radiative. If the absorption is large, the theory then predicts that these surface modes may become pseudo-propagating but this situation has not been studied in detail.

- Arguably, the most interesting case for plasmonics is when $\epsilon(\omega) < -\epsilon_M$ (Fig. 3.7(c)). We then have for an ideal metal k_x real and therefore a truly *propagating surface wave*. Moreover, k'_{1z} and k_{2z} are both pure imaginary, and therefore correspond to *evanescent waves perpendicular to the surface*. These are *non-radiative surface modes* that are fully trapped at the surface, and that propagate along the interface. There is again a propagating surface charge density wave at the surface. A non-local treatment [146] can show that this corresponds to a longitudinal surface plasmon wave. These modes are the *propagating surface plasmon–polariton* modes of a metal/dielectric planar interface. The energy is trapped at the surface and shared between photon and surface plasmon oscillations. For a real metal (Fig. 3.7(f)), k_x is no longer real, and the surface wave is then a *pseudo-propagating* wave (if the absorption is not too large).

In summary, surface modes exist for (almost) all values of ϵ (and therefore ω). All of them are surface polaritons (surface charge oscillations coupled to electromagnetic fields). When $\mathrm{Re}(\epsilon) < 0$, they are called surface

---

[11] The term SPP is sometimes reserved for *propagating* or *pseudo-propagating* surface waves and what we call here localized SPPs would then be excluded. This is again only a matter of vocabulary.

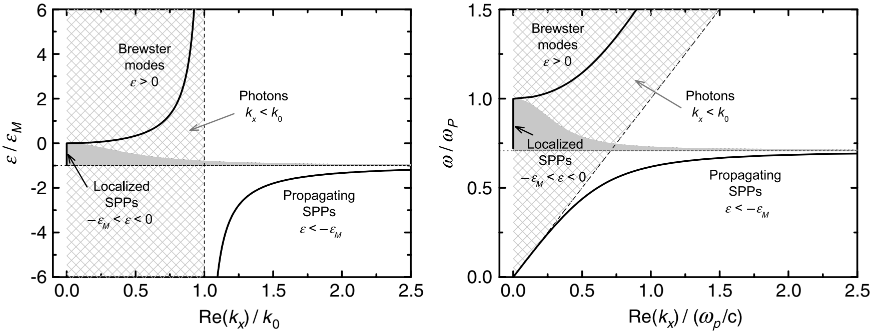


Figure 3.8. Two possible representations of the dispersion relation $\omega(k_x)$ for the surface modes of an ideal metal/dielectric interface ($\mathrm{Im}(\epsilon) = 0$, i.e. no absorption). On the left, ϵ/ϵ_M is plotted against $\mathrm{Re}(k_x)/k_0$. The advantage here is that this plot applies to any (ideal) metal. However, the dispersion relation can only be visualized indirectly since it requires the additional dependence of $\epsilon(\omega)$ (and k_0 also depends on ω). On the right, we show a more direct plot of the dispersion relation for an interface of air ($\epsilon_M = 1$) and an ideal metal following the Drude model (Eq. (3.2)); with $\epsilon_\infty = 1$, a plasma frequency ω_p and no absorption. The plot is shown in terms of adimensional quantities: ω/ω_p and $\mathrm{Re}(k_x)/(\omega_p/c)$. In both plots, the area covered by the dispersion relations of incident photons (with any possible incident angle) is shown as a hatched area. It corresponds to $k_x < k_0\sqrt{\epsilon_M}$ (since $k_x = k_0\sqrt{\epsilon_M}\sin\theta$). The condition $\epsilon(\omega) = -\epsilon_M$, which delimits the region of propagating SPPs is also shown as a straight dashed line. Finally, the broadening of the modes is indicated by a gray shaded area corresponding to the region between $\mathrm{Re}(k_x) - \mathrm{Im}(k_x)/2$ and $\mathrm{Re}(k_x) + \mathrm{Im}(k_x)/2$. Only localized SPP modes are broadened for an ideal metal.

plasmon–polaritons (SPPs) and may either be localized ($-\epsilon_M < \mathrm{Re}(\epsilon) < 0$, if absorption is low) or (pseudo-)propagating ($\mathrm{Re}(\epsilon) < -\epsilon_M$, if absorption is low). If $\mathrm{Re}(\epsilon) > 0$, they correspond to Brewster's modes, but do not play a role in the context of plasmonics. These considerations are summarized in Fig. 3.7 and further discussed in Section F.2.5. In the context of planar metal/dielectric interfaces, the propagating (or pseudo-propagating) SPPs (PSPPs), occurring for $\mathrm{Re}(\epsilon) < -\epsilon_M$, are the most important ones for most plasmonics applications. The rest of Section 3.4 is primarily aimed at discussing some of their properties.

3.4.2. Properties of the SPP modes at planar metal/dielectric interfaces

Dispersion relations for ideal metals

The dispersion relations for the surface modes, which relate the frequency ω, with the wave-vector for propagation k_x are given by Eq. (3.15), provided that the frequency dependence of $\epsilon(\omega)$ is known. The dispersion relations of surface modes for an ideal metal/dielectric interface are shown in Fig. 3.8

where $\mathrm{Re}(k_x)/k_0$ is plotted against ϵ/ϵ_M (which in practice depends on ω). Also shown is the broadening of the mode, characterized here by adding $\pm(1/2)\mathrm{Im}(k_x)/k_0$ to $\mathrm{Re}(k_x)/k_0$. Such plots are more commonly shown as ω as a function of k_x, but for this, one needs to relate ϵ to ω. This is, for example, illustrated in Fig. 3.8 for an ideal Drude metal with plasma frequency ω_p.

The three types of modes that have just been discussed appear clearly in these plots:

- The Brewster modes for $\epsilon > 0$, with no broadening in the ideal case. They appear above the plasma frequency.
- The localized SPPs for $-1 < \epsilon/\epsilon_M < 0$, with $\mathrm{Re}(k_x) = 0$ and a large broadening.
- The propagating SPPs for $\epsilon/\epsilon_M < -1$, with no broadening. The frequency of the SPPs modes approaches for large k_x the limiting frequency $\omega_{\mathrm{SP}} < \omega_p$ corresponding to the condition $\epsilon(\omega_{\mathrm{SP}}) = -\epsilon_M$. Large k_x correspond to the electrostatic (long spatial wavelength) approximation, and these modes therefore resemble the 'pure surface plasmons' introduced by Ritchie [143] and discussed previously. For a perfect Drude model with $\epsilon_\infty = 1$, $\omega_{\mathrm{SP}} = \omega_p/\sqrt{2}$, which is often referred to as the surface plasmon frequency. For a real metal, ω_{SP} can be different due to ϵ_∞ or inter-band transitions. At lower frequencies (longer wavelength), the SPP mode wave-vector approaches that of a photon with grazing angle of incidence and the propagating SPPs are more 'photon-like'. These SPP modes are those most useful for a number of applications in plasmonics.

For a photon incident from the dielectric (region 1) at an angle of incidence θ (angle with respect to the normal of the interface), the tangential component of k is $k_x = k_0\sqrt{\epsilon_M}\sin\theta$. Such a photon would appear on these plots as a straight line of equation $\omega = ck_x/\sin\theta/\sqrt{\epsilon_M}$. The extreme case of a photon with grazing incidence ($\theta = \pi/2$), i.e. $\omega = ck_x/\sqrt{\epsilon_M}$, is shown as a dashed line on the plots in Fig. 3.8. Every point on the graph with $\omega \geq ck_x/\sqrt{\epsilon_M}$ also corresponds to a photon with a given incident angle θ. This region is hatched on the graph and corresponds to modes that can couple (with ω and k_x conservation) to an incident photon from the dielectric. Modes in the other region cannot be directly excited with photons, because of energy and momentum conservation. This is the case of propagating SPP modes and the next section will be entirely dedicated to the important question of how to circumvent this problem and couple photons to PSPPs.

Dispersion relations for real metals

In real metals, ϵ has inevitably a small imaginary part across the visible range. In some cases, the absorption can even be quite large at some

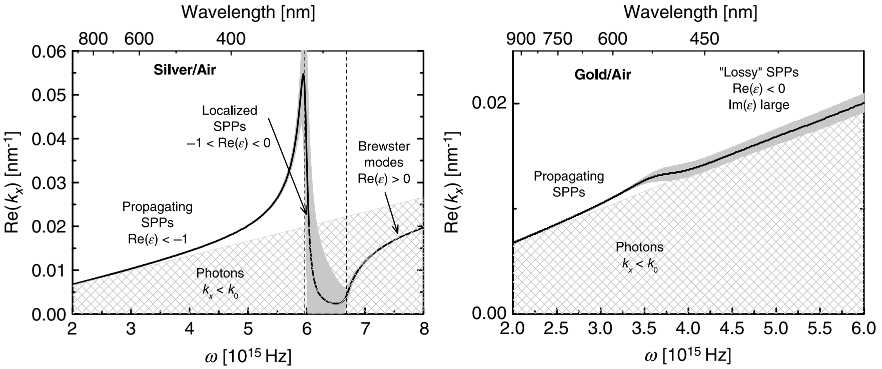


Figure 3.9. Dispersion relations $\omega(k_x)$ for the surface modes of a silver/air (left) or gold/air (right) interface. The legend is the same as that for Fig. 3.8, but note that the axes have been inverted (purely for technical reasons: it is easier to produce the plots).

wavelengths, for example in the case of gold due to inter-band transitions. The dispersion relation of SPP modes can still be obtained from Eq. (3.15) and is plotted in Fig. 3.9 for the model examples of silver and gold in air.

For silver, first, the plot is very similar to the ideal metal case, owing to the relatively small $\mathrm{Im}(\epsilon)$. The Brewster and propagating SPP modes now present a broadening, but it remains very small. One interesting difference is that the dispersion relation no longer diverges for $\epsilon = -\epsilon_M$, and it is actually back-bending in the region of the localized SPPs, i.e. $\mathrm{Re}(k_x)$ decreases with ω, a situation that would never occur for ideal metals. This anomalous dispersion has been observed in experiments [152,153] and bears some similarity with the concept of negative refraction.

For gold, the dispersion plot is quite different from that of an ideal metal, mostly because of the large absorption at wavelengths smaller than 600 nm (due to inter-band transitions). The propagating SPPs still exist for $\lambda > 600$ nm, with similar properties as already discussed. For $\lambda < 600$ nm (for which ϵ lies approximately between -3 and 0), the large absorption significantly affects and broadens the dispersion relation. The previous distinction between propagating and localized SPPs loses its strict meaning. The corresponding modes present a strong damping and are therefore referred to as ‘lossy’ SPPs, and they are quite localized in nature.

Propagation lengths for propagating SPPs

We now focus more specifically on the properties of propagating SPPs (PSPPs). When absorption is neglected, the surface wave associated with PSPP modes propagates forever along the x direction. Once losses are included ($\mathrm{Im}(\epsilon) > 0$), however, the surface wave is damped by absorption in the

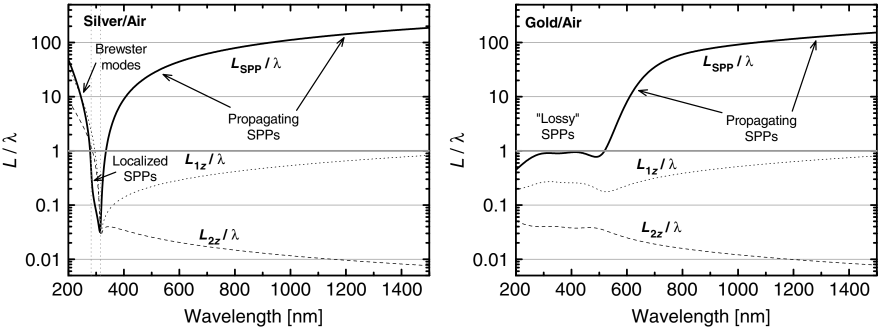


Figure 3.10. Propagation/decay lengths (normalized to the wavelength λ) for surface modes for a silver/air (left) and a gold/air (right) planar interface. L_{SPP} is the propagation length of the surface wave (along x). L_{1z} and L_{2z} are the decay lengths (along z) away from the interface in the dielectric and in the metal, respectively.

metal and the field intensity then decays as $\exp(-2\text{Im}(k_x)x)$. This defines a *propagation length* for the PSPP wave as:

$$\boxed{L_{\text{SPP}} = \frac{1}{2\text{Im}(k_x)}.} \tag{3.16}$$

Using Eq. (3.15), this can be expressed as a function of ϵ in a number of forms, none of which are particularly simple. Writing $\epsilon = \epsilon' + i\epsilon''$, a much simpler (and arguably more useful) expression can be obtained [11] in the case where ϵ'' remains small compared to ϵ' and $\epsilon' < -\epsilon_M$:

$$L_{\text{SPP}} \approx \frac{\lambda}{2\pi}\left(\frac{\epsilon'(\omega)+\epsilon_M}{\epsilon'(\omega)\epsilon_M}\right)^{\frac{3}{2}}\frac{\epsilon'(\omega)^2}{\epsilon''(\omega)}. \tag{3.17}$$

This propagation length can be much larger than the wavelength if ϵ'' is sufficiently small. It is also larger when $|\epsilon'(\omega)|$ is large, i.e. at longer wavelengths for metals.

To illustrate this, the propagation length (normalized to the wavelength) is plotted in Fig. 3.10 for silver and gold interfaces with air (using the exact expressions in Eq. (3.16)). The definition of L_{SPP} is extended to the region of localized SPP and Brewster modes, although when $L_{\text{SPP}} < \lambda$, it should be viewed as a decay length rather than a propagation length. We first note that the results are similar for silver and gold in the long-wavelength region $\lambda > 600$ nm, where the inter-band transitions no longer play a role. L_{SPP} for the PSPPs is then in the range ∼10–100 μm for both silver and gold and can be as large as 0.3 mm in the near infrared ($\lambda \approx 1.5$ μm).

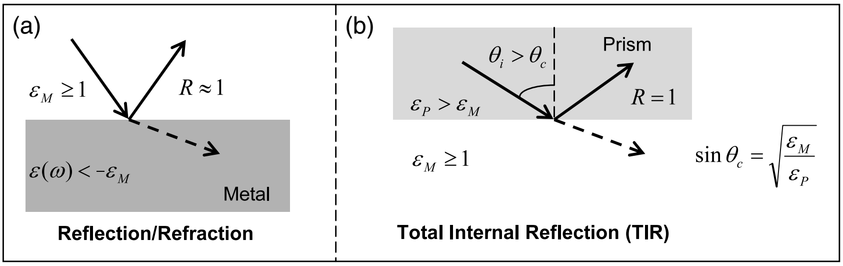


Figure 3.11. (a) Schematic illustrating the standard problem of reflection/refraction at a dielectric/metal interface. No PSPP modes can be excited here because of k_x conservation, and most of the energy is reflected ($R \approx 1$). (b) Schematic illustrating the phenomenon of *total internal reflection* (TIR) for a wave impinging from a high-refractive index dielectric (such as a prism, here with ϵ_P) onto a low-refractive index dielectric (such as air or water, here with ϵ_M). TIR occurs only for incident angles θ_i larger than the critical angle θ_c. In this case, k_x conservation cannot be met for a propagating transmitted wave. The transmitted wave is therefore evanescent and all power is reflected ($R = 1$ exactly for non-absorbing dielectrics).

It is interesting to compare this propagation length along the interface to the confinement of the fields along the z direction. The fields intensities decay as $\exp(-2|\mathrm{Im}(k'_{1z})z|)$ in the dielectric and as $\exp(-2|\mathrm{Im}(k_{2z})z|)$ in the metal. The corresponding decay lengths $L_{1z} = 1/(2|\mathrm{Im}(k'_{1z})|)$ and $L_{2z} = 1/(2|\mathrm{Im}(k_{2z})|)$ are also shown in Fig. 3.10. For PSPPs, it is clear that this decay length or penetration depth in the metal, L_{2z}, is very small ($\approx$10–15 nm). The confinement on the dielectric side is also quite good with a decay length smaller than one wavelength and more than two orders of magnitude shorter than the propagation length along the surface. These PSPP modes are therefore truly trapped electromagnetic surface waves. Provided light can be coupled in and out of them, they can therefore be used as high confinement light wave-guides, called *plasmonic wave-guides.*

3.4.3. Coupling of PSPP modes with light

In the previous description of propagating SPP modes, we have already hinted at the important issue of coupling them to light, a desirable step for many applications, which we now discuss in detail.

Let us first note that PSPP modes are TM (or p-polarized) electromagnetic waves. Because the nature of polarization is conserved at planar interfaces (see Appendix F), *only TM waves can excite PSPP modes.* In what follows, we therefore implicitly assume that all incident waves are TM polarized.

Moreover, one of the peculiarities of PSPP modes is that their wave-vector k_x (or $\mathrm{Re}(k_x)$) is larger than the wave-vector of a photon in the dielectric (which is equal to $k_0\sqrt{\epsilon_M}$). This is evident from the dispersion relation in Eq. (3.15), recalling that $\epsilon(\omega) < -\epsilon_M$ for PSPP modes, or more visually

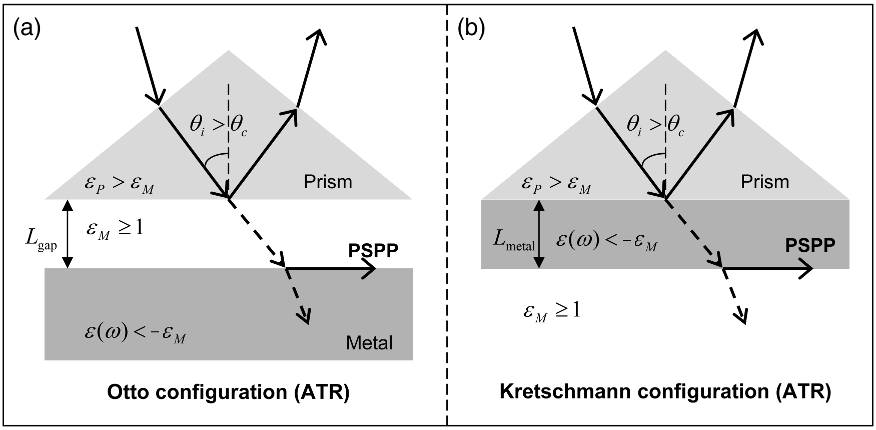


Figure 3.12. Schematic illustrating the two common configurations based on attenuated total reflection (ATR) for exciting PSPP modes with photons: (a) Otto configuration, (b) Prism coated with a thin metal film (Kretschmann configuration). In both situations, the PSPP is excited thanks to the larger momentum of the photon in the prism, which is transferred by the evanescent field, using total internal reflection either through a dielectric spacer (a) or directly through the metal (b).

from the plots of the dispersion relation in Figs 3.8 and 3.9. Let us also recall that $\hbar k$ represents the momentum of the particle or mode (photon or PSPP). The momentum along the x direction for a photon in the dielectric (which is the projection of its total momentum) is therefore always smaller than that of the PSPP modes. Note that the mismatch is usually small in the wave-guiding region (at longer wavelengths for longer propagation, see Fig. 3.10). For example, in the red for silver, we have $k_x \approx 1.03\,k_0$ (in air). However, because momentum must be conserved here (due to the translational invariance), it means that *a photon cannot directly excite the PSPP modes.* In fact, in the situation of an incident photon (electromagnetic wave) impinging on the surface, it is an incident wave mode corresponding to usual reflection/refraction that will be excited as illustrated in Fig. 3.11(a). Moreover for metals in the visible, the reflection coefficient (R^p, see Appendix F) is usually close to 1, i.e. almost no energy is transferred to the metal or to a surface wave.

Several techniques have been devised to impart the missing momentum and excite the PSPP modes with a photon, and we will briefly describe here some of them.

Coupling by total internal reflection

The first technique relies on the phenomenon of total internal reflection (also discussed in Section F.3.5) at a dielectric interface, which is illustrated in

Fig. 3.11(b). This occurs when an incident electromagnetic wave in a medium with refractive index n_1 impinges with a large angle of incidence θ onto a planar interface with a second dielectric of smaller refractive index $n_2 < n_1$. The classical theory of optics (through Snell's law) tells us that there is a critical incident angle defined by $\sin\theta_c = n_2/n_1$, beyond which there is no transmitted wave propagating in medium 2. A more detailed electromagnetic treatment shows that the wave in medium 2 is in this case evanescent (see Section F.3.5). Note that the momentum along the x direction must again be conserved. In medium 1, it is $\hbar k_x = n_1 \hbar k_0 \sin\theta$, and is therefore larger for $\theta \geq \theta_c$ than that allowed for a propagating photon in medium 2 (which is always smaller than $n_2 \hbar k_0$). This is precisely why the transmitted wave must be evanescent.

This technique therefore allows one to create an evanescent excitation in medium 2 with a parallel momentum $\hbar k_x$ larger than that normally allowed for a propagating photon. Such an excitation could therefore be used to excite the PSPP modes at a (dielectric M)/metal interface. This can be realized as a (dielectric P)/(dielectric M)/metal interface with $n_P > n_M$ (P stands for prism, see also Fig. 3.12(a)). The parallel momentum $\hbar k_x = n_P \hbar k_0 \sin\theta$ in the prism could be sufficiently large to excite PSPP modes of the (dielectric M)/metal interface. However, because the excitation in dielectric M is evanescent in the z direction after going through the P/M interface (because of TIR), the field decays exponentially and the (dielectric M)/metal interface needs to be close enough to the P/M interface to be excited efficiently by this evanescent field. Note however that if the two surfaces are too close to each other (typically smaller than the decay length L_{1z} of the PSPPs modes), then the PSPPs of the metal/(dielectric M) interface are strongly modified and different electromagnetic modes of the (dielectric P)/(dielectric M)/metal system arise, with different properties. There is therefore an optimum separation, of the order of L_{1z} to obtain maximum coupling without affecting the nature of the PSPP modes, and it is typically of the order of $\sim$1 μm.

The simplest way to realize this experimentally is to use a prism on top of a metal surface with a small air gap, as depicted schematically in Fig. 3.12(a). This was first proposed and demonstrated by Otto in 1968 [154] and is now called the *Otto configuration*. This work showed for the first time that PSPP modes could be easily and efficiently coupled to light and triggered a renewed interest in PSPPs and, in some respect, marked the beginning of what is now known as 'plasmonics'.

Thin metal films

One of the problems of the Otto configuration is that it is not straightforward to create an air gap or dielectric spacing with controlled thickness of the order of a micron. It was soon after realized [155,156] that

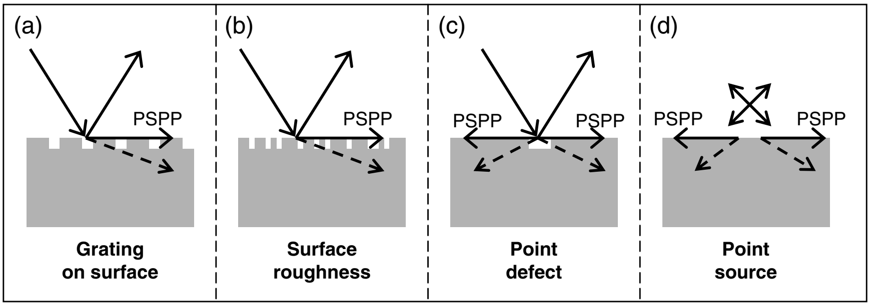


Figure 3.13. Schematic illustrating other techniques of exciting PSPP modes with photons, all based on breaking the translational invariance (and therefore the momentum conservation) by the presence of either a grating (a), surface roughness (b), a point defect (c), or a point source (d).

the metal itself could be used to deliver the photon (through an evanescent field again) using a configuration of the type (dielectric P)/metal/(dielectric M) with $n_P > n_M$ again. This configuration, sometimes called *Kretschmann configuration*, is depicted in Fig. 3.12(b). It only requires to deposit a thin metal film on the prism surface (dielectric P), which is reasonably easy experimentally. The film thickness is much smaller, typically 50 nm, than the dielectric gap in the Otto configuration, due to the stronger decay of the evanescent field in the metal, but must remain larger than the decay length L_{2z} for the metal/(dielectric M) PSPPs to avoid strong coupling between the two interfaces. The parallel momentum $\hbar k_x = n_P \hbar k_0 \sin\theta$ in the prism is then transferred by the evanescent field through the metal and can excite (on the other side) PSPP modes of the metal/(dielectric M) interface.

Both Otto and Kretschmann configurations are based on the same principle of using an evanescent wave to excite the PSPP modes, and will both be referred to as the *ATR configuration, after attenuated total reflection.*

Surface gratings and surface roughness

Another somewhat different approach to coupling PSPPs to light is to relax the conservation of momentum restriction, i.e. break the translational invariance. This can for example be achieved by engraving a periodic structure on the surface along the x direction, i.e. by forming a *surface grating* as illustrated in Fig. 3.13(a). As before, the perturbation to the surface needs to be small enough not to alter substantially the nature of the SPP modes. This is possible for example for a grating whose depth remains small compared to its spatial period, λ_g, and to L_{1z}. Such a periodicity does not fully break translational invariance, but modifies the law of conservation of k_x, allowing addition or subtraction of integer multiples of $k_g = 2\pi/\lambda_g$, and therefore making the coupling of light to PSPPs possible for some specific frequencies and/or angles. The coupling condition remains very sharp due to the discrete nature of the modified conservation law.

Such a conservation law can be removed completely in the case of a *random surface roughness* as illustrated in Fig. 3.13(b). Such a surface behaves like a combination of gratings with any arbitrary λ_g or k_g; k_x conservation is then always possible. Photons then always couple to PSPP modes, but this coupling is not as strong and does not show any sharp resonant effects, which can be a problem for some applications.

Point defects and point sources

A third approach to break the translational invariance is to introduce a spatially localized point defect, such as a bump or hole in the surface (Fig. 3.13(c)). This defect can also be created experimentally by a metal-coated tip of a scanning probe microscope placed in close proximity to the metallic surface. Such a defect breaks locally the law of momentum conservation and enables coupling of light to PSPP modes that will then propagate away from the point. It therefore acts as a *point source for PSPP modes*.

A similar situation happens when a light source (such as a dipolar emitter) is sufficiently close to the metal surface (Fig. 3.13(d)). The translational symmetry of the problem, and k_x conservation, is again broken locally, which enables the light emitted from the source to excite PSPP modes at the dielectric/metal interface. This again acts as a point source for PSPP modes. This can be realized experimentally for example using the optical probe of a scanning near-field optical microscope as the localized source [157].

Strictly speaking the PSPP modes created in this way are slightly different (although similar in many aspects) from those studied so far because they have a symmetry of revolution, instead of translational invariance. A separate mathematical treatment in cylindrical coordinates would be needed to describe them accurately.

A final remark about coupling light to PSPPs

Note that in many of the cases discussed above, the geometry of the problem is changed, for example, by addition of one or more interfaces, or by modification of the surface itself. In this respect, the electromagnetic modes of the system are also changed and are strictly speaking no longer the PSPP modes of a single planar dielectric/metal interface. However, if these modes are weakly coupled to other parts of the system, they then retain essentially their nature. The main change is that they are no longer strictly non-radiative modes because they couple (at least weakly) to radiation. This coupling can affect the propagation length of these modes, since there are now radiative losses in addition to the intrinsic decay through absorption in the metal. If these additional losses are too large, then the modes may lose their propagating nature (when the propagation length is of the order of the wavelength) and should then be considered as localized SPP modes.

3.4.4. PSPP resonances at planar interfaces

Origin of PSPP resonances

One of the important aspects of coupling light with PSPPs is that the coupling condition (conservation of momentum and energy) sets very stringent requirements on the parameters for optimal coupling. When these requirements are met, then the energy of the incident light is efficiently transferred to the PSPP modes, and this can have a dramatic effect on the optical response of the system. This is usually referred to as a *surface plasmon resonance* (SPR). Note that the denomination SPR may sometimes be used in a much more general context, i.e. for resonances arising from coupling to any SPP modes (propagating or localized SPPs for any types of geometries). We reserve it here for resonances arising from coupling to propagating SPP modes at planar interfaces.

In this particular case of PSPPs at planar interfaces, the simultaneous conservation of both momentum and energy (frequency) together with the very small broadening of the modes makes any resonances arising from them particularly sharp. Moreover, the main characteristic of these modes is that they are *non-radiative.* This means that the energy that is transferred to them cannot be re-radiated. It propagates at the surface and unless another coupling mechanism is used to collect it, it will simply dissipate by optical absorption in the metal (typically due to resistive losses related to the imaginary part of the dielectric function). The SPRs for PSPPs at planar interface therefore usually appear as *resonances in absorption*, or equivalently as a *resonant decrease in reflectivity.*

The main parameters that can be varied in a typical experiment are the wavelength λ, the angle of incidence θ_i, the dielectric constant of the outside medium ϵ_M, or the thicknesses of the dielectric layers in a multi-layer configuration (like the ATR configurations). SPRs can appear as functions of any of these parameters depending on the setup. Moreover, in regions of interest (for example visible or near infrared for Ag), the broadening of the PSPP modes, characterized by $\mathrm{Im}(k_x)$ (see Eq. (3.15)), is typically very small. This means that these resonances can be extremely sharp and therefore *very sensitive to the value of the parameters* under study. This makes such SPRs very attractive in applications such as sensors.

The resonance condition

The most common types of optical resonances for PSPPs at planar interfaces are found in the measurement of the reflectivity (i.e. of the reflection coefficient R^p for TM waves, see Appendix F). For a typical metallic plane (with ϵ negative), the reflection coefficient is normally close to 1 (see Appendix F) and does not vary much with λ or θ_i. This is why metals are known as good reflectors. Light does not couple to PSPP modes in this case, and is mostly reflected as illustrated in Fig. 3.11(a).

We now consider a reflectivity experiment in an ATR setup (Otto or Kretschmann configurations, as in Fig. 3.12(a–b)) and look at the resonance condition for coupling to PSPP modes. k_x in the prism is given by $\sqrt{\epsilon_P} k_0 \sin(\theta_i)$, and this is transferred through the dielectric gap or the metal by an evanescent field. The PSPPs excited in the ATR configuration are those of the interface of the metal with a dielectric of lower refractive index than the prism ($\epsilon_M < \epsilon_P$).

The resonance condition resulting from k_x conservation is then obtained from the PSPP dispersion relation for such an interface, given in Eq. (3.15), and takes the form:

$$\sqrt{\epsilon_P} \sin\theta_i = \mathrm{Re}\left(\sqrt{\frac{\epsilon(\omega)\epsilon_M}{\epsilon(\omega)+\epsilon_M}}\right). \tag{3.18}$$

If the absorption is small, i.e. $\epsilon = \epsilon' + \mathrm{i}\epsilon''$ with ϵ'' small compared to $|\epsilon'|$, then this simplifies to:

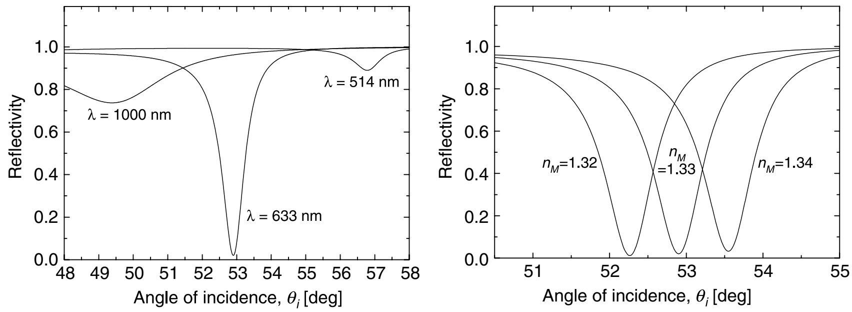
$$\epsilon_P \sin^2\theta_i \approx \frac{\epsilon'(\omega)\epsilon_M}{\epsilon'(\omega)+\epsilon_M}. \tag{3.19}$$

Moreover, for long wavelengths, where $\epsilon'(\omega)$ is negative and large, the resonance condition can further be approximated by:

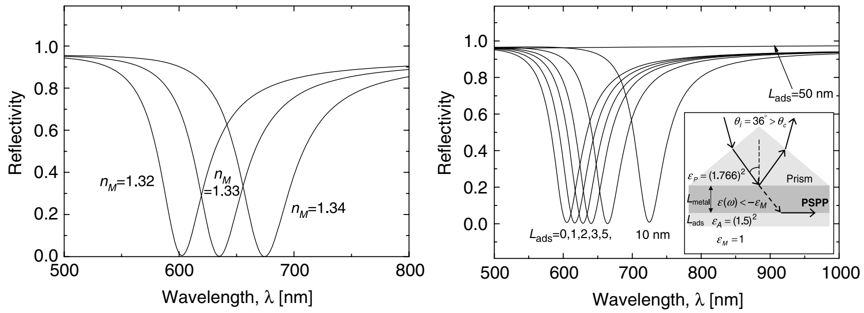
$$\sin\theta_i \approx \sqrt{\frac{\epsilon_M}{\epsilon_P}} = \frac{n_M}{n_P}, \tag{3.20}$$

i.e. θ_i at resonance is close to (but still above) the critical angle for total internal reflection at the prism/dielectric interface.

Because the broadening of k_x for the PSPP modes is very small, the resonance condition is extremely sharp, i.e. it is very sensitive to the parameters: angle of incidence, wavelength, and refractive index of the dielectric medium. If this condition is met, then one expects the incident light to couple and transfer its energy to the PSPP modes, resulting in a drop in reflectivity. This drop in reflectivity can be dramatic (and easily detectable), when the configuration is optimized for maximum coupling of the incident light to the PSPP modes. In the Otto configuration, this optimization is achieved by adjusting the size of the dielectric gap L_{Gap}. If L_{Gap} is too large, then the coupling will be poor, even at resonance, because of the evanescent



(a) Otto setup: L_{Gap}=500 nm, n_M=1.33 (b) Otto setup: L_{Gap}=500 nm, λ = 633 nm



(c) Kretschmann setup: θ_i=53, L_{Ag}=50 nm (d) Kretschmann setup: θ_i=36, L_{Ag}=50 nm, n_M=1

Figure 3.14. Examples of SPR experiments involving the excitation of propagating SPP modes at a planar silver/dielectric interface. The ATR configuration is used for coupling light to PSPPs, either in the Otto (a and b) or Kretschmann (c and d) setup (see Fig. 3.12). Angle-modulation (AM) is illustrated in (a) and (b) and wavelength-modulation (WM) in (c) and (d). (b) and (c) illustrate the possibility of refractive-index sensing using SPRs. In (d) an adsorbed layer on the metal changes the effective refractive index of the dielectric medium, thereby resulting in a shift of the SPR condition. Such a configuration can be used for adsorbate detection and/or thickness analysis.

nature of the exciting wave in the gap. If L_{Gap} is too small, then the proximity of the two interfaces may modify the nature of the SPP modes and broaden the resonance. In between the two limits, there is an optimum coupling condition for a given set of experimental variables.

Examples of surface plasmon resonances with PSPPs

The reflectivity in the ATR setup (Otto or Kretschmann configurations) can in fact be calculated by solving Maxwell's equations for the corresponding three-layer system, the details of which are given in Appendix F along with a possible Matlab implementation. Such calculations are shown as an example in Fig. 3.14 for a sapphire prism ($n_P = 1.766$, $\epsilon_P = n_P^2$), using silver as the metal.

In a typical setup, this reflectivity will be measured as a function of angle of incidence, sometimes called angle-modulation (AM), as in Fig. 3.14(a–b), or wavelength, called wavelength-modulation (WM), as in Fig. 3.14(c–d). The position of the resonance can then be used to deduce other parameters, such as the refractive index of the dielectric material, n_M. The angle-dependent reflectivity (R^p) is for example shown for the Otto configuration (see Fig. 3.12(a)) in Fig. 3.14(a). A clear sharp resonance, where the reflectivity drops from 1 to virtually 0 is observed for excitation at 633 nm. This combination of parameters therefore offers optimum coupling to the silver/water PSPP modes at 633 nm. Weaker resonances are also observed (at a different angle of incidence) for 514 nm and 1000 nm excitation, for which the parameters are not optimized. The refractive-index sensitivity of the angle-dependent resonance is illustrated in Fig. 3.14(b) for different values of n_M. The shifts in the resonance are clearly resolved and such a setup could therefore be used to measure the refractive index n_M with high accuracy. A similar example of refractive-index sensing is shown in Fig. 3.14(c) for wavelength-dependent reflectivity in the Kretschmann configuration.

Finally, it is worth pointing out that such a high- refractive-index sensitivity can in fact be used to measure the thickness of an adsorbed layer (of known refractive index n_A), down to only one monolayer in optimized conditions. The adsorbed layer is embedded in the dielectric medium with refractive index n_M and modifies slightly the properties of the metal/(dielectric M) PSPP modes, thereby affecting the resonance condition. A simple way of studying this effect is to add a fourth layer (the adsorbate layer of thickness L_{Ads} and refractive index n_A) to the EM problem and calculate the reflectivity[12]. This is illustrated in Fig. 3.14(d) for WM reflectivity in the Kretschmann configuration. The shift in resonance is clearly observable as a function of adsorbate layer thickness and can therefore be used as a sensor for the presence of adsorbed molecules. In practice, one usually wants to detect one specific type of adsorbed molecule, and it is therefore necessary to functionalize the metallic surface, so that only the desired molecules bind to it. The surface functionalization is in fact a necessary and important step in many sensing applications of plasmonics (including SERS).

3.4.5. Local field enhancements and SPPs at planar interfaces

We have so far concentrated on far-field properties of the SPPs at planar interfaces, i.e. how they couple to incident light, and what their effect is on scattering properties like the reflectivity. For some applications, the near-field properties are also important, since they rely on large local field enhancements at the surface (this is in particular the case of SERS).

---

[12] This approach has limitations, especially for the smallest thicknesses, but nevertheless illustrates the point.

Due to the strong confinement of the field at the surface (which decays exponentially from it), one could intuitively expect that the field intensity close to the surface is large. In a sense, the electromagnetic energy is confined at the surface and should result in a large local field.

In order to quantify this assertion more accurately, we need to be able to compare the local field at the surface to an exciting field. To do so, one could choose one of the setups previously discussed for coupling an incident wave with amplitude E_{Inc} to the SPP modes, calculate the local field at the surface E_{Loc}, and deduce the local field intensity enhancement factor (LFIEF) $M_{\mathrm{Loc}} = |E_{\mathrm{Loc}}/E_{\mathrm{Inc}}|^2$. We will do that soon, but shall first start with a more general and physical approach.

Local fields for SPP modes

Let us first analyze the local field at a planar metal/dielectric interface for a SPP mode (propagating or localized). The field expressions are given in Eq. (3.13), from which we deduce the electric fields $\mathbf{E}_1$ (in the dielectric) and $\mathbf{E}_2$ (in the metal) at the surface ($z = 0$). They can be decomposed as tangential and perpendicular components, and using the expressions of k_x, k_{1z}, and k_{2z}, we have:

$$\frac{|\mathbf{E}_1^{\perp}(0)|^2}{|\mathbf{E}_1^{\|}(0)|^2} = \frac{|k_x|^2}{|k_{1z}|^2} = \frac{|\epsilon|}{\epsilon_M}, \tag{3.21}$$

and

$$\frac{|\mathbf{E}_2^{\perp}(0)|^2}{|\mathbf{E}_2^{\|}(0)|^2} = \frac{|k_x|^2}{|k_{2z}|^2} = \frac{\epsilon_M}{|\epsilon|}. \tag{3.22}$$

Moreover, we deduce from these and from the fact that $\mathbf{E}_2^{\|}(0) = \mathbf{E}_1^{\|}(0)$ (imposed by the boundary condition at the interface) that:

$$\frac{|\mathbf{E}_1(0)|^2}{|\mathbf{E}_2(0)|^2} = \frac{|\mathbf{E}_1^{\perp}(0)|^2 + |\mathbf{E}_1^{\|}(0)|^2}{|\mathbf{E}_2^{\perp}(0)|^2 + |\mathbf{E}_2^{\|}(0)|^2} = \frac{|\epsilon|}{\epsilon_M}. \tag{3.23}$$

These equalities are useful to bear in mind when considering the local fields for SPP modes at planar interfaces. In particular, there are many common situations where $|\epsilon| \gg \epsilon_M$, for example for PSPP modes at long wavelengths, where $\mathrm{Re}(\epsilon)$ is negative and large, or when losses are important ($\mathrm{Im}(\epsilon)$ large). In these cases, we deduce that:

- The surface electric field inside the metal, $\mathbf{E}_2$, is mostly *parallel* to the interface.

- The surface electric field outside the metal (in the dielectric), $\mathbf{E}_1$, is mostly *perpendicular* to the interface.
- The intensity of the field outside is much larger than that inside the metal. This is the most important effect and can be loosely viewed as an expulsion of the field from the metal and a concomitant concentration of energy at the surface just outside (within a distance spanned by the evanescent field).

We will now attempt to place this final statement on a more rigorous footing, using simple energy considerations.

Some general energy considerations

We consider an experiment where an incident wave in a dielectric (ϵ_{Inc}) is coupled (using any of the methods described before) to *propagating SPP* modes at a dielectric(ϵ_M)/metal(ϵ) interface. The coupling efficiency, i.e. power injected into the PSPP mode over incident power, denoted by η, can in principle be close to 1 at resonance (under optimal coupling conditions). Let us analyze the energy balance in a volume covering a small surface area S on the interface. Because the PSPP modes are non-radiative, no energy is lost to the far field in the dielectric or the metal. Moreover, because of invariance by translation, there is no net flow of energy laterally. The only source of energy loss for the PSPP modes is therefore due to optical absorption in the metal. In the volume delimited by the surface S, the power absorbed is:

$$P_{\text{Abs}} = \int_0^\infty \frac{1}{2}\omega\epsilon_0\epsilon''|\mathbf{E}_2(z)|^2 S dz = S\frac{\epsilon_0 c}{2}\epsilon'' k_0 L_{2z}|\mathbf{E}_2(0)|^2, \tag{3.24}$$

where $k_0 = \omega/c$ is the free-space wave-vector, $L_{2z} = 1/(2\text{Im}(k_{2z}))$ is the penetration depth in the metal defined previously, and $\mathbf{E}_2(0)$ is the electric field at the interface, just inside the metal, (see Eq. (3.13)).

This absorbed power must, by energy conservation, be balanced by the power coupled into the SPP mode, i.e. $P_{\text{Abs}} = \eta P_{\text{Inc}}$. For an incident wave incoming from a dielectric (ϵ_{Inc}) with incident angle θ_{Inc}, the incident power on surface S is:

$$P_{\text{Inc}} = S\frac{\epsilon_0 c}{2}\sqrt{\epsilon_{\text{Inc}}}|E_{\text{Inc}}|^2\cos\theta_{\text{Inc}}, \tag{3.25}$$

where $|E_{\text{Inc}}|$ is the magnitude of the incident electric field. Expanding $P_{\text{Abs}} = \eta P_{\text{Inc}}$, we therefore deduce:

$$\frac{|\mathbf{E}_2(0)|^2}{|E_{\text{Inc}}|^2} = \frac{\eta\sqrt{\epsilon_{\text{Inc}}}\cos\theta_{\text{Inc}}}{\epsilon'' k_0 L_{2z}}. \tag{3.26}$$

Because ϵ'' and $k_0 L_{2z}$ are in typical situations both small, this simple argument naturally predicts a field enhancement at the surface (with respect to the incident field E_{Inc}. But here $\mathbf{E}_2(0)$ is the surface field inside the metal. We are more interested in the surface field $\mathbf{E}_1(0)$ just outside, in the dielectric. These two are related by Eq. (3.23) derived above, which is then the source of an additional enhancement by a factor $|\epsilon|/\epsilon_M$. For PSPPs for example, $\epsilon' < -\epsilon_M$ and usually $|\epsilon'| \gg \epsilon_M$ and this factor is the source of a further field enhancement of the surface field outside compared to the surface field inside. Using the exact expression for $k_0 L_{2z}$, we can finally write the local field intensity enhancement factor (LFIEF) as:

$$M_{\text{Loc}} = \frac{|\mathbf{E}_1(0)|^2}{|E_{\text{Inc}}|^2} = \frac{2|\epsilon|}{\epsilon_M \epsilon''} \text{Im}\left(\epsilon \sqrt{\frac{1}{\epsilon + \epsilon_M}}\right) \eta \sqrt{\epsilon_{\text{Inc}}} \cos\theta_{\text{Inc}}. \tag{3.27}$$

In cases where $|\epsilon'| \gg \epsilon_M$, $\mathbf{E}_1$ is almost perpendicular to the interface, and the LFIEF can be approximated by:

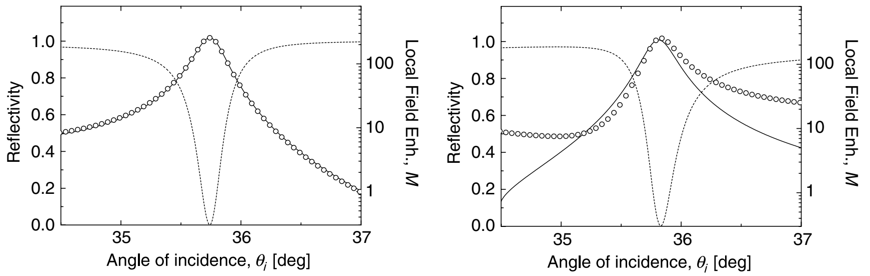
$$M_{\text{Loc}} \approx M_{\text{Loc}}^{\perp} \approx \frac{2|\epsilon'|^{3/2}}{\epsilon_M \epsilon''} \eta \sqrt{\epsilon_{\text{Inc}}} \cos\theta_{\text{Inc}}. \tag{3.28}$$

For a silver/air interface at 633 nm, excited from air, this leads to values of the order of $M_{\text{Loc}} \approx 180$ (from Eq. (3.27)) at normal incidence $\theta_{\text{Inc}} = 0$ for optimum coupling $\eta \approx 1$. Note that this is the maximum LFIEF that can be expected in such a situation. This increases to $M_{\text{Loc}} \approx 200$ at 800 nm. For the same interface in the ATR configuration with a sapphire prism ($\epsilon_P = (1.766)^2$), the incident wave is coming from the prism, i.e. $\epsilon_{\text{Inc}} = \epsilon_P$, and the incident angle must be chosen as $\theta_{\text{Inc}} \approx 36°$ for optimum coupling ($\eta \approx 1$). We then obtain $M \approx 255$ at 633 nm from Eq. (3.27). Note however that the LFIEF is in this case calculated with respect to $E_{\text{Inc}} = E_P$, i.e. the electric field amplitude in the prism. In practice, this needs to be related to the true incident wave that excited the one in the prism from outside, which will be discussed briefly later.

The advantage of the previous argument is that it is very general and gives an upper limit on the field enhancement for SPPs on planar interfaces. In practice however, it can sometimes be difficult to estimate the coupling efficiency η to SPP modes. One must then model the electromagnetic problem to predict the field enhancement. We give here a couple of examples.

Example in the Otto configuration

We first consider the Otto configuration with a sapphire prism ($\epsilon_P = n_P^2 = (1.766)^2$), and an air ($\epsilon_M = 1$)/silver ($\epsilon$) interface. Here the air gap is optimized for PSPP coupling at 633 nm: $L_{\text{gap}} = 850$ nm. We calculate the



(a) Otto setup: L_{Gap}=850 nm, n_M=1.0, λ=633 nm (b) Kretsch. setup: L_{Ag}=50 nm, n_M=1.0, λ=633 nm

Figure 3.15. Calculated reflection coefficient, R, (dashed line) and local field intensity enhancement factor, M_{Loc}, in the dielectric at the metal/dielectric interface (solid line) when exciting PSPP modes in either the Otto (a) or Kretschmann (b) configurations. The symbols show the predicted LFIEF using Eq. (3.27) and $\eta = 1 - R$.

LFIEF at the metal surface (on the air side), using the tools of Appendix F . This is shown as a function of incident angle θ_i along with the reflection coefficient in Fig. 3.15(a). It is clear that the LFIEF, M_{Loc}, exhibits a sharp resonance, like the reflectivity, R. At the reflectivity minimum, we have $R \approx 0$ and $\theta_i \approx 35.7°$. Assuming all the incident power is coupled into the SPP modes, ($\eta \approx 1$), we can estimate from the simple energy conservation argument (Eq. (3.27)) that $M_{\mathrm{Loc}} \approx 256$, which is precisely what is predicted by the calculation.

Here we can expand further the energy arguments of the previous section. If we assume that even outside resonance the power that is not reflected is coupled into the SPP modes, then we have $\eta = 1 - R$, and the LFIEF can therefore be estimated from Eq. (3.27). This estimation, also shown in Fig. 3.15(a), fits the calculated M_{Loc} extremely well.

Example in the Kretschmann configuration

Finally, to illustrate the limits of this approach, we now study a similar experiment in the Kretschmann configuration. The calculated LFIEF and reflectivity are shown in Fig. 3.15(b) along with the prediction using Eq. (3.27) and $\eta = 1 - R$. The agreement is quite good at resonance, but not that good for other angles. The reason for this discrepancy can be understood, at least qualitatively. In the Otto configuration, the reflection/refraction at the first interface (prism/air) can only excite non-dissipative modes (with no optical absorption), i.e. all power that is not reflected is transmitted to the air/metal interface. In the Kretschmann configuration, however, the excitation must go through the thin metal layer before reaching the metal/air interface supporting the PSPP modes. Dissipation may occur in this layer by optical absorption in the metal. In fact, optical modes other than PSPPs are excited in this configuration. Because these additional modes are dissipative, $\eta = 1 - R$ is only an upper estimate of the coupling to the SPP modes (part of this energy may couple to the other modes and dissipate in the form of heat). Moreover, these modes may contribute to the local electric field at the metal surface, and therefore increase or decrease the LFIEF. This is the reason why, at resonance where SPP coupling dominates, our simple estimate based on Eq. (3.27) is correct, but away from resonance, the contribution of other types of modes becomes important and our simple energy argument is no longer useful.

These two examples illustrate both the power of simple energy arguments in connecting local field enhancements and resonant mode coupling, and its limitations in more complex systems. We will come back to these arguments when discussing localized SPPs in Section 3.5.

These examples also further confirm that: (i) PSPP modes are associated with relatively large electric field at the interface, (ii) This can result in a large local field enhancement *provided the incoming power is efficiently coupled into the PSPP modes.*

The air/prism interface

Finally, we make here a brief digression to discuss the air/prism interface in relation to the energy conservation argument and local field enhancements. In many practical situations, the ATR configuration is used to couple to the PSPP modes, and we must therefore have a prism with $\epsilon_{\text{Inc}} = \epsilon_P > \epsilon_M$. The angle of incidence θ_{Inc} is then not free to vary since it must be chosen appropriately to couple efficiently to the PSPP modes (otherwise η would be very small). In fact, k_x conservation requires:

$$\sqrt{\epsilon_P}\sin\theta_i = \text{Re}\left(\sqrt{\frac{\epsilon\epsilon_M}{\epsilon + \epsilon_M}}\right). \tag{3.29}$$

The expression for the LFIEF can then again be simplified if $|\epsilon'| \gg \epsilon_M, \epsilon''$:

$$M_{\text{Loc}} \approx M_{\text{Loc}}^{\perp} \approx \frac{|\mathbf{E}_1(0)|^2}{|E_P|^2} = \frac{2|\epsilon'|^{3/2}}{\epsilon_M\epsilon''}\eta\sqrt{\epsilon_P - \epsilon_M}. \tag{3.30}$$

Moreover, in practice, the real source of the incident wave is rarely embedded in the high-dielectric constant material (i.e. the prism), but typically comes from air (with $\epsilon_{\text{Source}} = 1$) and is injected into the high-dielectric constant material (ϵ_P) using a prism configuration. To study the true field enhancement in this situation, one must also study the coupling of the light from air into the prism itself, i.e. relate the field amplitude $|E_{\text{source}}|$ of the incident light, to the field amplitude $|E_P|$ refracted in the prism, which then serves as excitation for the PSPP modes. This should be assessed on a case-by-case basis, and should be calculated independently of the rest of the problem because the coupling to the prism breaks the translational invariance and cannot therefore be included as an additional layer. The coupling at the air/prism interface is a simple reflection/refraction problem between two dielectrics and can be modeled using the tools of Appendix F .

We only provide here a model example (in relation to the two previous examples discussed above) for ATR coupling to the silver/air interface at 633 nm. We consider a triangular sapphire prism ($n_P = 1.766$), with (base) angle of 67°, and an incident beam *perpendicular* to the metal/dielectric interface (the base of the prism), i.e. $\theta_{\text{Source}} = 0$. The angle of incidence of this beam with respect to the prism side surface is 67°, and from Snell's law the refracted angle (again with respect to the prism side) is 31.4°, which results in an incident angle for the bottom prism/dielectric interface of $\theta_P = 35.6°$, i.e. the angle required for PSPP coupling at 633 nm (we in fact chose the base angle of the prism for this to be the case).

Assuming optimal coupling from the prism to the PSPP modes, there are two ways of viewing the energy conservation argument with regard to coupling from the air to the prism:

- One may consider the wave from outside the prism as the incident wave. In this case $\epsilon_{\text{Inc}} = \epsilon_{\text{Source}} = 1$ (air) and $\theta_{\text{Inc}} = \theta_{\text{Source}} = 0$. However, the reflection/refraction at the air/prism interface on the side of the prism changes the incident flux of energy (along z) by a factor $(1/n_P)|t^p|^2 \cos\theta_P \approx 0.368$ (this is derived from the expressions of Section F.3.3). This results in a decrease in the eventual coupling to the PSPP modes to $\eta \approx 36.8\%$. Equation (3.27) then gives $M_{\text{Loc}} \approx 66$ for the LFIEF.
- One may consider alternatively the wave in the prism as the incident wave, as was done in the discussion of the examples above. In this case $\epsilon_{\text{Inc}} = \epsilon_P$ and $\theta_{\text{Inc}} = \theta_P$. Equation (3.27) can then be used to estimate the LFIEF and we obtain $M_P \approx 256$, but here this is with respect to the electric field inside the prism E_P, not E_{Source}. These are related (see Section F.3.3) by $|E_P|^2/|E_{\text{Source}}|^2 \approx 0.256$ from which we deduce again the true LFIEF $M_{\text{Loc}} \approx 66$.

Both views are, fortunately, consistent, but they highlight the necessity to take special care when handling this energy conservation argument, and in the definitions of the LFIEF.

3.4.6. SPP modes on planar interfaces: A brief summary

In this section we have discussed mostly the propagating SPP modes at a planar metal/dielectric interface. These modes are important for two reasons:

- They can propagate over relatively long distances on the surface, whilst remaining strongly confined at the surface. This opens up the possibility to use these modes to manipulate and guide light and design plasmonic wave-guides and even plasmonic chips.

- These modes exhibit an extremely narrow broadening and their coupling with light is therefore very sensitive to external parameters. This leads to very sharp surface plasmon resonances, which can be exploited for example as sensors.

Although a planar interface may appear as a very particular case of the possible geometries that may be encountered in real experiments, its reach is in fact much wider: as discussed earlier, small perturbations to the surface (like a shallow grating, roughness, or point defects) do not affect significantly the nature of the PSPP modes. Moreover, there are many situations where a metallic surface can be approximated by a plane over distances of the order of the propagation length of the SPPs (say around 10 μm). In all these cases, the PSPP modes will play an important role. We will now focus on situations where this is no longer the case.

3.5. LOCALIZED SURFACE PLASMON–POLARITONS

3.5.1. Introduction to localized SPPs

The planar approximation is no longer true for small metallic objects, and in particular for nano-particles, where the size becomes comparable or smaller than the wavelength. The nature of the electromagnetic modes of the system is then completely modified. In particular, the description in terms of k vector (k_x for a plane) becomes irrelevant, since the translational invariance is lost. The electromagnetic modes then exist for discrete values of ω (instead of having continuous modes described by the dispersion relation $\omega(k_x)$). These modes are then called *localized surface plasmon–polaritons* (LSPs)[13].

In fact, this is not a property of metals or plasmon–polaritons only. The same happens for photons when the environment exhibits features of the order of the wavelength. Photons correspond to free-space modes of the electromagnetic field (plane waves with well defined ω and $\mathbf{k}$). When boundaries have features much larger than the wavelength, one can apply the ‘standard’ description giving rise to reflection and refraction at interfaces (Snell’s law). All boundaries are approximated by locally planar interfaces, and this ‘ray optics’ approach is perfectly legitimate. However, when the dimensions of the system become comparable to the wavelength, say in a cavity or a wave-guide, this approach fails. The concept of photon is replaced by that of electromagnetic modes of the cavity (characterized by discrete

---

[13] In the acronym LSP, the P should be understood as standing for plasmon–polariton, not plasmon. The use (abuse, in fact) of the denomination localized surface plasmon instead of localized SPP is however very common, so common that we have adopted it in this book.

values of ω, or $\omega(k)$ where k is irrelevant). These modes are highly localized inside the cavity. They are sometimes called cavity polaritons to emphasize their mixed nature of a photon with its optical environment, and are the photon analogs of localized SPPs.

Note that the localized SPPs meet the two criteria introduced previously to characterize SPPs (see Section 3.3.8): they would not exist without the presence of the interfaces, and their properties depend on the optical properties of the outside medium. The third criterion (having a field localized at the interface) can lead to misleading interpretations. Finally, LSP modes, and their effects on SERS and other optical properties, will be extensively discussed in Chapter 6 on practical examples, and we therefore restrict ourselves here to general considerations.

3.5.2. LSP on planar structures

A particular case of localized SPP has already been encountered in the discussion of the surface modes of a planar interface. These modes arise when $-\epsilon_M < \mathrm{Re}(\epsilon) < 0$ and are characterized by a large broadening ($\mathrm{Im}(k_x)$), which essentially means that the description in terms of k_x becomes irrelevant. These particular types of localized SPPs have rarely appeared in applications and we will therefore focus on the more important LSPs arising in metallic nano-particles.

3.5.3. LSP modes of a metallic sphere

To understand better the nature of LSP, it is useful to consider the canonical example of a metallic sphere. A full analytical treatment is then possible using Mie theory [149,158]; its results will be discussed in Chapter 6 while the technical details are given in Appendix H . We only discuss qualitatively these results in terms of the electromagnetic modes of the sphere. This discussion is therefore rather abstract.

Thanks to the analytical solution from Mie theory, it is possible to define and study all the electromagnetic modes of the sphere. Many of these modes are irrelevant to SERS or plasmonics but they provide a nice illustration of the general discussion about electromagnetic modes given in Section 3.3. A similar study was carried out for an ionic crystal sphere in Ref. [159], where more details can be found. The only difference here is the form of the dielectric function $\epsilon(\omega)$ and we therefore adapt the discussion to the case of a metal. The k vector is irrelevant here and the modes correspond to a discrete set of frequencies ω. Because of the spherical symmetry, it is convenient to index these modes with an integer $l \geq 1$ corresponding to the total angular momentum.

The longitudinal modes, as is the case for bulk metals must satisfy $\mathrm{Re}(\epsilon(\omega_{\mathrm{BP}})) = 0$ and therefore all occur at the same frequency ω_{BP} as bulk plasmons. For a Drude model, this is simply the plasma frequency. These are 'pure' spherical plasmon modes and behave exactly like bulk plasmon modes. They correspond to longitudinal spherical electric waves (no magnetic field) inside the metal, associated with collective charge oscillations. The field outside the sphere is zero everywhere. These modes are not relevant to the optical properties because of their longitudinal character.

All the other electromagnetic modes are transverse, inside and outside the sphere, but may have a longitudinal character at the interface because of the boundary conditions (and may therefore be associated with surface charge oscillations). Incident wave modes exist for any ω and simply correspond to the solution of the EM problem (using Mie theory). Here we focus on the surface modes (bound modes), which are solutions for which there are no incident waves (eigen-solutions of the linear problem). For a given l, the frequencies of the surface modes are solutions of a complex equation given in Section H.3.5. Because of intrinsic absorption in the metal, all the surface modes must be virtual modes (with $\omega = \omega' - \mathrm{i}\omega''$ complex). The finite lifetime $1/(2\omega'')$, or broadening $2\omega''$ accounts for the absorptive losses in the metal and possible radiative nature of the mode. Coupling to these modes will occur when the incident frequency (real) matches ω', with a broadening of the order of $2\omega''$. This will be discussed further in Chapter 6 and also partly in Appendix H .

The most important surface mode in most cases is the lowest frequency one, which corresponds to $l = 1$ and whose scattered field is that of an electric dipole and it is, therefore, a radiative mode. It should be called the *dipolar localized surface plasmon–polariton* mode of the sphere, but is often loosely called (at best) 'localized surface plasmon' or simply 'surface plasmon', and even sometimes 'plasmon'. Other surface modes exist for $l > 1$, at increasing frequencies, with quadrupolar, octupolar, etc. nature, but are usually less relevant. They are also radiative modes, but with a larger absorptive character than the dipolar modes (i.e. the proportion of losses through absorption, as opposed to radiation, is larger).

The frequency of the dipolar LSP mode of the sphere depends on several parameters:

- Obviously, the metal (through its frequency-dependent optical properties characterized by $\epsilon(\omega)$).
- The environment, through its dielectric constant ϵ_M.
- The size of the sphere (i.e. its radius a).

For the smallest spheres (typically $a < 10$ nm), the dipolar LSP mode frequency can be obtained approximately from:

$$\mathrm{Re}(\epsilon(\omega_{\mathrm{LSP}})) = -2\epsilon_M, \tag{3.31}$$

and therefore lies at a wavelength longer than the longitudinal plasmon–polaritons ($\mathrm{Re}(\epsilon(\omega_{\mathrm{BP}})) = 0$) and the localized SPPs of the plane interface ($-\epsilon_M < \mathrm{Re}(\epsilon(\omega_{\mathrm{SP}})) < 0$). For the smallest spheres, for which the electrostatic approximation applies (see Section 5.1.4), the LSP mode is almost a "pure surface plasmon" in nature, with only a small "photon component" [160]. As the size increases, the dipolar LSP frequency red-shifts (to longer wavelengths) and this is associated with a larger broadening (ω''). These properties will be reviewed in more detail in Chapter 6.

3.5.4. LSP modes of nano-particles

The description of the LSP modes of non-spherical nano-particles in terms of complex frequencies (eigenvalues) is an extremely difficult problem because of the lack of an analytical solution to the EM problem. The standard approach is then to study the EM problem of excitation of the nano-particle by incident waves. The LSP modes of the nano-particle then appear as *resonances* in the optical response and their nature can be inferred from the field solution at their resonance frequency (if there is not too much overlap with other resonances).

The LSP modes of non-spherical nano-particles have the same qualitative features as that of the sphere, but their frequency depends in addition on the geometry (shape) of the particle. This shape may in addition introduce anisotropies, i.e. their coupling to an external field becomes polarization-dependent. This will also be further discussed in Chapter 6.

3.5.5. LSP resonances

The LSP modes of a nano-particle can be excited by an incident wave with the appropriate polarization and frequency. Efficient coupling to LSP modes will then result in a resonant optical response at the LSP frequency. As opposed to PSPP on planar interfaces, LSPs are radiative modes (with an absorptive component because of optical absorption in the metal). The resonant response therefore, not only appears in absorption (which is analogous to the reflectivity experiments for a plane interface), but also in scattering (or similarly extinction) measurements. These resonances, sometimes called LSP resonances (LSPR) to differentiate them from SPR (based on PSPPs), are sensitive to the environment and, like SPR, can be used for applications in refractive-index and chemical sensing.

The LSP resonances also manifest themselves, as for PSPP modes, as large local field enhancements inside the metal, and more importantly on the surface outside. This effect is the basis for most surface-enhanced spectroscopies, including SERS.

It is interesting to highlight the main differences between SPR and LSPR:

- The SPR condition requires conservation of both k_x and ω. This is more difficult to fulfill than only ω conservation for LSPR. In particular, k_x

conservation typically requires a more complex setup, such as the ATR configuration.

- SPRs offer more liberty in the implementation, either in terms of angle-modulation or wavelength-modulation, whereas only wavelength-modulation can be used for LSPRs.
- SPRs are typically much sharper resonances compared to LSPRs. This can be an advantage or a disadvantage depending on the application. It should for example in principle result in a larger sensitivity but only on a more limited range of parameters. For SERS, resonances must be broad enough to encompass both the exciting laser and the Stokes frequencies, and SPRs are typically too sharp to fulfill that condition.
- The active surface for SPRs is a single planar interface, while for LSPRs it is the nano-particle surface (which can therefore be spread in a 3D volume, for example by dispersing the particles in water).
- There are more degrees of freedom to tailor or engineer the LSPRs (shape, size, etc.) as opposed to the SPRs, which may open more possibilities, but also more problems (such as poly-dispersity).

In summary, the use of SPR vs LSPR will depend on the exact application. LSPRs are more versatile (easier to implement) but the resonances are not as well defined as for SPRs.

3.5.6. Local field enhancements and LSP

The local field enhancements arising from LSP excitation in nano-particles (NPs) and more complex structures will be discussed in detail in Chapter 6. Their link to SERS and surface-enhanced fluorescence is moreover the subject of Chapter 4.

Therefore we will only discuss here the local field intensity enhancement factor (LFIEF) in the context of the simple energy conservation argument discussed earlier for PSPPs. There are several differences that make this approach more difficult for LSP modes. Firstly, the full EM solutions is known in the case of PSPPs, therefore enabling the calculation of, for example, the optical absorption accurately. Moreover, the invariance by translation of PSPPs means that the LFIEF is the same everywhere on the surface. This is no longer the case for NPs, where the LFIEF is expected to be non-uniform on the surface. In addition, LSPs are usually radiative modes, which therefore introduces an additional mechanism of energy loss, to be included in the energy balance. Finally, and perhaps most importantly, it is more difficult to define a coupling efficiency η for a LSP mode of a nano-particle excited by an incident beam. The reason is that the extent of the beam is usually much larger than the NP (because of the diffraction limit) and one must therefore reason in terms of incident power density (and cross-section).

These additional features make it difficult to extend simply the energy argument to the case of LSPs [133], but one can nevertheless rescue the qualitative conclusions. We had concluded, for example, that large local field enhancements can be expected when (i) the incoming power is efficiently coupled into one or more electromagnetic modes with the following additional characteristics: (ii) these modes are confined to a small volume (smaller than the metallic object), and (iii) they are not too lossy, i.e. there is little dissipation by optical absorption or loss by radiation. Moreover, (iv) at metal surfaces, the local field enhancement outside the metal is even larger when $|\epsilon| \gg \epsilon_M$.

For example, for the PSPP modes discussed in Section 3.4, the coupling of incoming power can be extremely good (condition (i)). There is a small degree of confinement, at least in one dimension (condition (ii)). The losses are relatively small, since these modes are non-radiative and dissipation is small for good metals like silver or gold (condition (iii)). Finally, the condition of a large local field enhancement outside the metal is also met for long-wavelength PSPP modes (condition (iv)).

These considerations can be qualitatively applied to LSP also. The optimum coupling condition (i) implies that the largest LFIEF will be obtained at resonance with the LSP mode, i.e. at the LSP frequency. Condition (ii) may be ambiguous for LSP modes since the fields inside the NP are not necessarily confined at the surface. However, the fact that it is at least confined to the NP dimensions (which are small) contributes to a larger LFIEF. Condition (iii) implies that the LFIEFs will be larger for low-loss metals like silver in the visible, or gold beyond 600 nm, and for smaller objects (for which radiation losses are smaller). Finally, condition (iv) of a large local field enhancement outside the metal should also result in additional enhancements for LSP modes resonant at longer wavelengths. These qualitative conclusions will in fact be confirmed in Chapter 6 in the study of specific examples. These examples will also highlight additional important considerations regarding the LFIEF arising from coupling to LSP modes: (i) the LFIEF is typically larger at tips, corners, or edges, as opposed to flat surfaces (the so-called *lightning rod effect*), (ii) the LFIEF can also be magnified by LSP mode interactions, i.e. at small gaps between nano-particles. These conditions, together with the factors influencing the resonance frequency of the LSP modes, can be used as a guide to 'engineer' and devise structures with large local field enhancements.

3.5.7. Interaction of SPPs – gap SPPs

Gap SPPs and local field enhancements

Finally, we cannot conclude this chapter on plasmons without mentioning an additional type of plasmonic effect that has become increasingly important in recent research, and which will be dubbed *gap SPPs*. We have just hinted at its importance by mentioning that the LFIEF can be greatly enhanced at the gap between two metallic objects. This effect arises when two metallic objects (typically nano-particles) are brought very close to each other. The LSP modes of each object then interact with each other and form (for the closest distances) hybridized modes – in a similar fashion to atomic orbitals, for example. In fact, most of the work to date on topics like *single-molecule SERS* is based on the use and exploitation of gap SPPs; they provide some of the largest known local field enhancements at surfaces. This will be further justified in Chapter 6.

Let us note that the pair of objects could in fact be considered as a single entity with its own electromagnetic modes, and therefore its own LSP modes. This problem could be studied independently of the study of the LSP modes of each individual objects. It may however be simpler (and more intuitive) to try and deduce the properties of the LSP modes of the pair from those of the components. They then arise as a result of the interaction (coupling) of two modes, and many general results can then be invoked from a general mode-interaction theory.

A simple analog of gap SPPs

An analogy with the case of atomic orbitals (of widespread use in chemistry) can be invoked at this stage.

Let us exemplify the concept of interaction theory with the standard example of the hydrogen molecule. If we start from two isolated hydrogen atoms and we push them together until their electrons start interacting, it is possible to build a *symmetric* (bonding) and *anti-symmetric* (anti-bonding) wave-function by taking linear combinations of the atomic *1s* states in both atoms. Through the interaction between the two atoms, the bonding wave-function has an energy lower than the original degenerate energies, while the anti-bonding state is pushed upwards in energy. The bonding state accumulates electronic charge density in the middle of the molecule while the anti-bonding state does the opposite. By accumulating electronic density in the middle of the molecule, the bonding state achieves a better screening of the Coulomb repulsion of the nuclei, thus resulting in a lower overall energy. This is the basic phenomenology of the general theory of chemical bonding.

Similar concepts apply qualitatively to LSP resonances [134,160,207]. Let us take two metallic nano-particles far away from each other. Under these conditions the two NPs couple independently to an external electromagnetic excitation, for example through their respective dipolar LSP resonance. As the NPs approach each other, their responses start to interact to define a coupled object, and its associated coupled resonances. Even in the simplest minded approximation of a dipolar picture, it is not difficult to imagine that we can take linear combinations of the responses that add them *in phase* or *out of phase* to define the equivalent of the bonding and anti-bonding electronic interactions mentioned above. The 'bonding' resonance is the most important in this framework, for it concentrates its electric field distribution at the gap in between the particles. The resulting resonance is red-shifted. The picture of coupled LSP resonances is a lot more complicated than simple orbital interaction theory of electrons, because of the vectorial nature of electromagnetic fields, retardation effects, and the presence of higher order resonances (quadrupolar, etc.), which play a significant role in particular at short distances. Nevertheless, the most red-shifted resonance, resulting from the coupling between the dipolar LSP resonances of each particle, is usually easy to identify.

Taken in the appropriate context, the concept of interaction theory can be a useful shorthand for the discussion of qualitative effects. More detailed results usually require a solution of the EM problem, as discussed in Section 6.4.

3.6. BRIEF SURVEY OF PLASMONICS APPLICATIONS

Finally, we conclude this chapter by discussing briefly the main possible applications of SPP modes, i.e. of plasmonics [133]. They can be separated into three groups:

- Applications based on surface plasmon resonances or localized surface plasmon resonances, such as chemical sensors.
- Application based on surface wave propagation and guiding.
- Applications based on local field enhancements.

Let us now consider these groups separately.

3.6.1. Applications of surface plasmon resonances

PSPP-based resonances

The sharpness and large sensitivity to parameters of resonances arising from PSPPs on planar interfaces make then well suited for applications as sensors, usually called surface plasmon resonance (SPR) sensors.

The simplest and most direct application is probably to use this sensitivity to measure the optical constants of metals. This is obviously limited to the wavelength range where PSPPs can be excited, but can nevertheless be a valuable approach to complement more conventional measurements, such as ellipsometry. This is particularly important since other optical measurements may be affected by PSPP excitations, and therefore be inaccurate in this region. Examples of such measurements are given in Ref. [150], and references therein.

Many of the other applications of PSPPs as sensors are based on the high-refractive-index sensitivity illustrated and discussed in Section 3.4.4. This means that PSPP-based sensors can (in principle) detect very small changes in the local environment at the metal surface, in particular, the adsorption of molecules. Typical implementations are based on a Kretschmann ATR configuration with a prism and may use angle and/or wavelength-modulation. The main obstacle in the practical implementation is the lack of *specificity*, i.e. one cannot distinguish which type of molecule adsorbs on the metal. As a remedy to this problem, a complementary aspect of SPR sensor research is the study of the *surface functionalization* of metallic surfaces (the application of which extends well beyond the field of plasmonics). Using appropriate chemistry, it is possible to coat the metallic surface with molecules that will bind only to one type of chemical group or even one type of molecule. Examples of such specific binding are antibody–antigen, ligand–receptor, or nucleic acid binding. Implementations based on SPR sensors with surface functionalization are therefore numerous and only limited by the functionalization step. We will not review specific applications here, and instead refer the reader to (for example) Ref. [127] and references therein, where an extensive description of this particular type of application of plasmon resonances is provided.

LSP-based resonances

More recently, there has been an increased interest in using LSP resonances – rather than PSPP resonances – for some sensing applications. The respective merits of these two approaches have been discussed in Section 3.5.5 and we shall not come back to it here. More details on applications of LSP-based resonances can be found in recent reviews, for example in Ref. [126].

3.6.2. SPP propagation and SPP optics

Another active area of research in plasmonics is concerned with the use of PSPPs as optical devices, i.e. to propagate, guide, and manipulate light. The driving force behind this effort is the hope that, thanks to the confined nature of the PSPP modes, large miniaturization of optical devices may be possible, even beyond the diffraction limit, which is the fundamental limit of 'conventional' optics. A related subject is also the use of SPPs for the design of negative refractive index materials, another very active field of research. These aspects of plasmonics, although interesting by themselves, are not directly related to SERS, and will therefore not be discussed further. Recent reviews of pure plasmonics topics are Refs. [11,130,131,133] which constitute a good starting point for further information on these specific aspects.

3.6.3. Local field enhancements

The most relevant types of plasmonics applications to us are those based on the large local field enhancements arising from coupling to SPP modes (primarily localized SPPs). SERS belongs to this group, along with many related techniques, including most surface-enhanced spectroscopies and, in particular, surface-enhanced fluorescence (SEF). Based on similar principles there are numerous applications where plasmonic structures are used to engineer and enhance the optical properties of light emitters; for example to enhance the directionality of emission or the quantum yield of solid-state emitters and absorbers such as semiconductor quantum dots, quantum wells, light-emitting diodes and solar cells, see for example [11,130,131] for further details.

The fundamental principles of these techniques and their relation to local field enhancements and SPPs will be discussed extensively in the rest of the book.