

# 1 Plasmons

Plasmons are a direct solution of Maxwell's equations at the boundary between a dielectric and a metal. Despite plasmons existing on length scales from  $\mathcal{O}(100 \text{ nm})$  to  $\mathcal{O}(1 \text{ nm})$ , the high free electron density of metals means that energy levels still retain their characteristic continuous conduction bands and **quantum/quantisation** effects can be ignored. Classical theory is able to accurately describe physical phenomena until the characteristic length scale drops below  $\sim 0.5 \text{ nm}$ . A phenomenological approach using Maxwell's equations [maxwell1865] therefore forms the basis of the mathematical description of plasmons.

## 1.1 Electromagnetic Waves

Maxwell's equations universally describe the classical, dynamical behaviour of electromagnetic (EM) waves and form the foundations of electromagnetism. In their differential form they are given by,

$$\nabla \cdot \mathbf{E} = \frac{\rho_{\text{tot}}}{\varepsilon \varepsilon_0}, \quad (1a)$$

$$\nabla \cdot \mathbf{B} = 0, \quad (1b)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad (1c)$$

$$\nabla \times \mathbf{B} = \mu \mu_0 \left( \mathbf{J} + \varepsilon \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right), \quad (1d)$$

where  $\mathbf{E}$  is the electric field,  $\mathbf{B}$  is the magnetic flux density,  $\varepsilon_0$  is the permittivity of free space,  $\mu_0$  is the permeability of free space,  $\rho_{\text{tot}}$  is the (volume) charge density,  $\mathbf{J}$  is the current density and  $t$  is time. The variables  $\varepsilon$  and  $\mu$  are the relative permittivity and permeability, respectively, and describe the electromagnetic properties of the medium in which the wave exists. This set of partial differential equations describe the microscopic fields within an electromagnetic system. Two new quantities can be introduced to include material dependencies and describe the macroscopic fields. These are the *electric displacement field*  $\mathbf{D}$  and the *magnetic field*  $\mathbf{H}$ , defined as,

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P}, \quad (2a)$$

$$\mathbf{H} = \frac{1}{\mu_0} \mathbf{B} + \mathbf{M}, \quad (2b)$$

where  $\mathbf{P}$  is the polarisation (dipole moment per unit volume) and  $\mathbf{M}$  the magnetisation. The displacement field arises due to polarisation of a material in response to an applied field and is related to the internal charge density by  $\nabla \cdot \mathbf{P} = \rho_{\text{int}}$ . Conservation of charge means that  $\nabla \cdot \mathbf{J} = \partial \rho / \partial t$ , which requires that  $\mathbf{J} = \partial \mathbf{P} / \partial t$  (a result also achievable by differentiating (2a)). The final equation of importance is the relationship between the electric field and the current density, given by,

$$\mathbf{J} = \sigma \mathbf{E}, \quad (3)$$

where  $\sigma$  is the conductivity. These few relations are sufficient to understand the behaviour of electromagnetic waves in media.

By utilising (2a), equations (1a) and (1d) can be redefined to include material dependencies as,

$$\nabla \cdot \mathbf{D} = \rho_{\text{ext}}, \quad (4a)$$

$$\nabla \times \mathbf{H} = \mathbf{J}_{\text{ext}} + \frac{\partial \mathbf{D}}{\partial t}. \quad (4b)$$

The charge and current densities now refer to only the external contributions, related to the internal contributions via  $\rho_{\text{tot}} = \rho_{\text{ext}} + \rho_{\text{int}}$  and  $\mathbf{J}_{\text{tot}} = \mathbf{J}_{\text{ext}} + \mathbf{J}_{\text{int}}$ . Propagation of EM waves within a medium is governed by a wave equation relating both the spatial and temporal changes of a wave. In general, a wave equation describing a wave  $\psi(x)$  along an axis  $x$  in time  $t$  is of the form,

$$\frac{\partial^2 \psi}{\partial x^2} = \frac{1}{v^2} \frac{\partial^2 \psi}{\partial t^2}, \quad (5)$$

where  $v$  is the speed of the wave. Combining (1c) and (1d) leads to the general wave equation for EM waves in the time domain,<sup>1</sup>

$$\nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\varepsilon \varepsilon_0 \mu \mu_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} - \mu \mu_0 \mathbf{J}, \quad (6)$$

$$\nabla(\nabla \cdot \mathbf{D}) - \nabla^2 \mathbf{D} = \frac{\partial^2 \mathbf{D}}{\partial t^2} - \mathbf{J}_{\text{ext}}, \quad (7)$$

describing the propagation of an electromagnetic wave in a given medium. In the absence of both charge and current (7) reduces to,

$$\nabla^2 \mathbf{E} = \varepsilon \varepsilon_0 \mu \mu_0 \frac{\partial^2 \mathbf{E}}{\partial t^2}, \quad (8)$$

describing a wave propagating in both space and time with a velocity  $c$ . Comparison to (5) shows that the speed of light in free space ( $\varepsilon = \mu = 1$ ) is  $c = 1/\sqrt{\varepsilon_0 \mu_0}$ . Furthermore, the refractive index is defined as  $\tilde{n} = \sqrt{\varepsilon \mu}$  such that  $\varepsilon \varepsilon_0 \mu \mu_0 = (\tilde{n}/c)^2 = 1/v^2$ . Waves are therefore slowed down by a factor  $\tilde{n}$  inside a medium.

In general  $\varepsilon$  is a complex quantity,  $\varepsilon = \varepsilon_1 + i\varepsilon_2$ , and depends on the frequency of the EM wave  $\omega$ . Plasmons are a phenomena resulting from this frequency dependence in metallic materials. The relative permittivity is therefore denoted  $\varepsilon(\omega)$  and is referred to as the dielectric function of a material from this point onwards. For this reason equations are simplified by removing any magnetic contributions by assuming  $\mu = 1$ . Since  $\varepsilon(\omega)$  is a complex parameter  $\tilde{n} = \sqrt{\varepsilon(\omega)} = \sqrt{\varepsilon_1 + i\varepsilon_2}$ , the complex refractive index can be expressed as  $\tilde{n} = n + i\kappa$ , where  $n$  is the real part causing refraction and  $\kappa$  is the extinction coefficient determining absorption in the medium. The complex refractive index and the dielectric function are then related via  $\varepsilon_1 = n^2 - \kappa^2$  and  $\varepsilon_2 = 2n\kappa$ .

If the material dielectric properties in  $\varepsilon$  are linear then  $\mathbf{D}$  can be expressed in Fourier space as,

$$\mathbf{D}(\mathbf{q}, \omega) = \varepsilon_0 \varepsilon(\mathbf{q}, \omega) \mathbf{E}(\mathbf{q}, \omega). \quad (9)$$

Combining (9) with the differential of (2a), (3) and a harmonic wave solution<sup>2</sup> yields the relationship between a material's conductivity and it's dielectric function,

$$\varepsilon(\mathbf{q}, \omega) = 1 + \frac{i\sigma(\mathbf{q}, \omega)}{\varepsilon_0 \omega}. \quad (10)$$

Depending on which is more convenient, either the dielectric function or the conductivity can be used to describe the optical response of a material. Conductivity is typically used to describe lower frequency phenomena while the dielectric function is used at higher frequencies.

The dispersive properties of a material are found by solving (7) with  $\varepsilon = \varepsilon(\omega)$ , describing the behaviour of a wave propagating through a non-magnetic, dielectric medium. For a monochromatic, harmonic wave with frequency  $\omega$  and wave vector  $\mathbf{q}$  in space  $r$  of the form,

$$\mathbf{E} = \mathbf{E}_0 e^{i(\mathbf{q} \cdot \mathbf{r} - \omega t)}, \quad (11)$$

representing a propagating EM wave, (7) can be expressed in the frequency (Fourier) domain as,<sup>3</sup>

$$\mathbf{q}(\mathbf{q} \cdot \mathbf{E}) - q^2 \mathbf{E} = -\varepsilon(\mathbf{q}, \omega) \frac{\omega^2}{c^2} \mathbf{E}, \quad (12)$$

where  $q = |\mathbf{q}|$  is the magnitude of the wavevector. The variable  $q_0^2 = \omega^2/c^2$  is sometimes used in (12) when all quantities considered are wave vectors. From this equation the propagation behaviour of EM waves in media can be described.

Solutions to (12) depend on the orientation of the wavevector with the field. Transverse wave solutions ( $\mathbf{q} \cdot \mathbf{E} = 0$ ) yield the dispersion relation for light,

$$q = \sqrt{\varepsilon(\mathbf{q}, \omega)} \frac{\omega}{c} = \tilde{n} q_0. \quad (13)$$

<sup>1</sup>Derived using  $\nabla \times \nabla \times \mathbf{E} = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E}$

<sup>2</sup> $\partial/\partial t \rightarrow -i\omega$  and  $\mathbf{J} = \dot{\mathbf{D}} - \varepsilon_0 \dot{\mathbf{E}} = \varepsilon_0 \dot{\mathbf{E}}(\varepsilon - 1) = \sigma \mathbf{E}$

<sup>3</sup>Derived using the identities  $\nabla \times \nabla \times \mathbf{E} = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E}$ ,  $\nabla^2 \mathbf{E} = -q^2 \mathbf{E}$  and  $\partial^2 \mathbf{E}/\partial t^2 = -\omega^2 \mathbf{E}$  where  $\nabla \cdot \mathbf{E} = 0$

Inserting this into (11) gives a general solution for light propagating through a dielectric medium,

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}_0(\mathbf{r}) \exp\left(-\kappa \frac{\omega}{c} \hat{\mathbf{q}} \cdot \mathbf{r}\right) \exp\left(i\omega \left(\frac{n}{c} \hat{\mathbf{q}} \cdot \mathbf{r} - t\right)\right). \quad (14)$$

The real component of the refractive index  $n$  slows the wave whereas the imaginary component corresponds to an exponential decay with characteristic length  $1/\kappa$ , representing loss within a medium. Since the real part of the conductivity is related to the imaginary part of the dielectric function the decay is attributed to energy transfer to move electrons at the surface of the material. Longitudinal wave solutions ( $\mathbf{q} \cdot \mathbf{E} = q|\mathbf{E}|$ ) result in  $\sqrt{\varepsilon(\mathbf{q}, \omega)}\omega/c = 0$ , hence solutions only exist for  $\varepsilon(\mathbf{q}, \omega) = 0$ . Both these conditions are important when describing plasmons. Furthermore, when considering the behaviour of EM waves at an interface between two different media the orientation of the fields with respect to the interface becomes important. Separate solutions exist depending on if a wave is considered transverse magnetic (TM) or transverse electric (TE) (either only  $\mathbf{E}$  or  $\mathbf{H}$  has a component in the direction of propagating, respectively).<sup>4</sup>

Using the framework outlined so far the optical properties of metals can be deduced along with the existence of plasmons. The discussion begins with the Drude model for the optical response of metals [**drude1900**], which is used to first predict the behaviour of plasmons. From there the distinction can be made between plasmons within the volume of a metal and those confined to the surface, which are of most interest in plasmonics.

## 1.2 Bulk Plasmons and the Optical Properties of Metals

Before studying the concept of a surface plasmon it is important to understand the optical properties of metals in general. When light is incident on a metal, free electrons at the surface respond to the field and are displaced in the opposite direction (since  $\mathbf{F} = -e\mathbf{E}$ ). The field of the induced charge distribution cancels the electric field inside the metal. An EM wave impinging on a metal is internally screened, and therefore externally reflected<sup>5</sup>, through the displacement of free electrons inside the metal surface. The reflected wave gives metals their shiny appearance. This behaviour originates from (10), where a high conductivity increases absorption and reduces the field penetration inside the metal. The exponential decay of the wave into the metal, shown in (14), is characterised by the skin depth  $\delta_s = c/2\omega\kappa$ , the point at which the field has decayed by  $1/2e$  of its original value.<sup>6</sup> The small values of  $\delta_s$  exhibited by metals means that they fall within the *perfect conductor* approximation (zero internal field). Light transmission through a metal becomes heavily attenuated once its thickness becomes greater than  $\delta_s$ .

A metal becomes more dielectric once the light oscillates fast enough that the inertia of the **massive** electrons means they cannot respond fast enough, preventing screening and thus transmitting the incident light. Such effects begin to be seen around the visible spectrum of light in noble metals. Fields increasingly penetrate the metal until photon energies in the UV spectrum are reached, at which point most metals become transparent. This is known as the *ultraviolet transparency*. These effects can be seen by simply considering the response of a free electron gas to an applied field.

Unlike many other materials whose optical properties are determined by the response of bound electrons (described by the Lorentz oscillator model), the properties of metals are dominated by the response of free electrons delocalised from the positive nuclei background. The Drude model [**drude1900**] is a simple description of the optical properties of metals and describes the motion of a free electron gas in response to an applied field. The equation of motion for a single electron in a time-varying applied field is given by,

$$m\ddot{\mathbf{x}}(t) + m\gamma\dot{\mathbf{x}}(t) = -e\mathbf{E}(t), \quad (15)$$

where  $m$  is its effective optical mass and  $\gamma = 1/\tau$  is the electron collision frequency, the inverse of the relaxation time  $\tau$ . Use of an effective optical mass as opposed to the actual electron mass incorporates band structure effects into the model. The electron collision frequency amounts to an effective coefficient of damping as in a mechanical oscillator. Inserting a harmonic driving field and assuming a similar oscillatory behaviour in the free electron

<sup>4</sup>TE and TM are also known as *s*- and *p*-polarisations, respectively.

<sup>5</sup>Light absorbed by electrons will be re-emitted, or from a different point of view, the prevention of light from entering the metal means the incident field must be reflected.

<sup>6</sup>The skin depth is defined using  $1/2e$  rather than  $1/e$  to consider power instead of field.

displacement ( $\mathbf{x} = \mathbf{x}_0 e^{-i\omega t}$ ) leads to a solution,

$$\mathbf{x}(t) = \frac{e}{m(\omega^2 + i\gamma\omega)} \mathbf{E}(t). \quad (16)$$

There is a resulting polarisation  $\mathbf{P} = -ne\mathbf{x}$  induced in the free electron gas, where  $n$  is the number density of electrons. The resulting displacement field, **acquired** by substituting (16) into (9)<sup>7</sup>, defines the dielectric function of a metal,

$$\varepsilon(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + i\gamma\omega}, \quad (17)$$

where  $\omega_p$  is the plasma frequency of the metal, given by,

$$\omega_p^2 = \frac{ne^2}{\varepsilon_0 m}. \quad (18)$$

The optical properties of a metal can be discerned from the real and imaginary components of  $\varepsilon(\omega)$ , given by,

$$\text{Re}[\varepsilon(\omega)] = \varepsilon_1(\omega) = 1 - \frac{\omega_p^2 \tau^2}{1 + \omega^2 \tau^2}, \quad (19a)$$

$$\text{Im}[\varepsilon(\omega)] = \varepsilon_2(\omega) = \frac{\omega_p^2 \tau}{\omega(1 + \omega^2 \tau^2)}. \quad (19b)$$

Metallic physical characteristics correspond to  $\text{Re}[\varepsilon(\omega)] < 0$ , **i.e. D negative compared to E**, where electrons move to oppose an incident field. The plasma frequency defines the point at which the metal transitions into a dielectric. For  $\omega < \omega_p$ ,  $\text{Re}[\varepsilon(\omega < \omega_p)] < 0$  and a free electron gas remains metallic in character. Once  $\omega > \omega_p$  the free electron gas, **limited by inertia**, cannot respond fast enough to follow the field and becomes dielectric in character. For an ideal free electron gas with negligible damping the expression,

$$\varepsilon(\omega) = 1 - \frac{\omega_p^2}{\omega^2}, \quad (20)$$

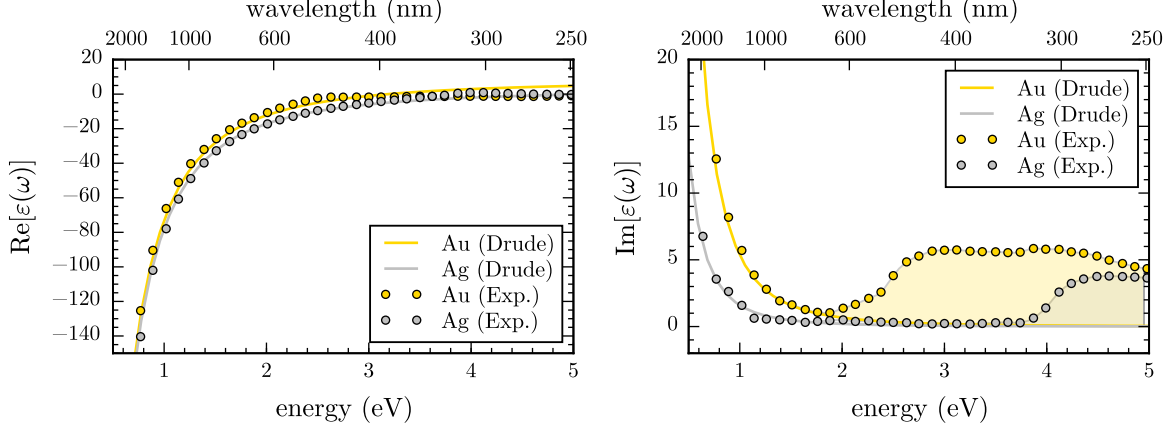
is often used. This is the often the case for large **(optical)** frequencies close to  $\omega_p$  where the imaginary component of  $\varepsilon(\omega)$ , dominated by  $\omega\tau \gg 1$ , becomes negligible. In actually, however, interband transitions in real metals increase  $\text{Im}[\varepsilon(\omega)]$ . The expression for the dielectric function can be modified to account for interband absorption caused by bound electrons by the inclusion of a constant  $\varepsilon_\infty$ . The dielectric function then has the form,

$$\varepsilon(\omega) = \varepsilon_\infty - \frac{\omega_p^2}{\omega^2 + i\gamma\omega}. \quad (21)$$

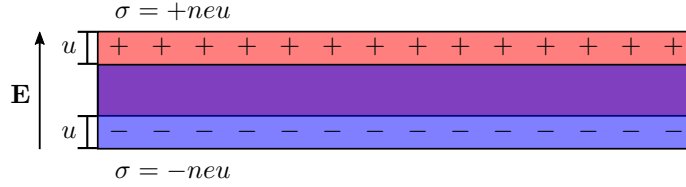
Plotting (21), along with empirical data, illustrates both why noble metals exhibit high quality visible spectrum (400–700 nm, 1.5–3 eV) plasmonics as well as the failings of the Drude model. Noble metals have  $\text{Re}[\varepsilon(\omega)] < 0$  and small  $\text{Im}[\varepsilon(\omega)]$  in the visible region, hence behave very similar to an ideal free electron gas. The Drude model fails at higher energies as interband transitions are not included in the basic model. These transitions increase the absorption ( $\propto \text{Im}[\varepsilon(\omega)]$ ) and are significant for  $\lambda < 500$  nm in Au and  $\lambda < 300$  nm in Ag. A measure of the quality of a plasmonic material can be determined from its quality factor  $Q = |\text{Re}[\varepsilon]|/|\text{Im}[\varepsilon]|$ . The high  $Q$  of noble metals in the visible regions means they can easily respond to an incident field and screen it, behaving plasmonically. Interband transitions lead to a reduction in  $Q$  at higher energies therefore preventing good plasmonics.

The significance of the plasma frequency  $\omega_p$  in  $\varepsilon(\omega)$  is that it not only describes a metal-to-dielectric transition but also dictates the frequency of the collective longitudinal mode of oscillation. By substituting (20) into the dispersion relations for transverse and longitudinal waves it is clear that transverse waves are only supported if  $\omega > \omega_p$  with dispersion  $\omega^2 = \omega_p^2 + q^2 c^2$ . However, a collective longitudinal oscillation is allowed at  $\omega = \omega_p$  since  $\varepsilon(\omega) = 0$  in the absence of damping. In this case the free electron gas is displaced from the ionic core background

<sup>7</sup>Derived from  $\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P} = \varepsilon_0 \mathbf{E} - [ne^2/m(\omega^2 + i\gamma\omega)]\mathbf{E} = \varepsilon_0 \varepsilon(\omega) \mathbf{E}$



**Figure 1: Plot of the dielectric function, given by the Drude model, for Au and Ag compared with empirical data.**  $\varepsilon(\omega)$  is calculated using (21). The plasma frequency is calculated using (18). The parameters of the curves are  $n = 5.90 \times 10^{28} \text{ m}^{-3}$ ,  $m = 9.11 \times 10^{-31} \text{ kg}$ ,  $\gamma = 1/\tau = 1/1 \times 10^{-14} \text{ s}$  and  $\varepsilon_\infty = 8$  for Au and  $n = 5.86 \times 10^{28} \text{ m}^{-3}$ ,  $m = 9.11 \times 10^{-31} \text{ kg}$ ,  $\gamma = 1/\tau = 1/3 \times 10^{-15} \text{ s}$  and  $\varepsilon_\infty = 3$  for Ag. Empirical data (Johnson and Christy, 1972 [johnson1972optical]) is shown for comparison to illustrate the importance of interband transitions. Deviations from the Drude model due to interband transitions are shaded.



**Figure 2: Charge displacement of a free electron gas under an applied field.** The optical electric field displaces the electrons leaving behind the positive cores. The slab becomes polarised with opposing surface charge densities  $\sigma$ . The charge oscillations resonate when the field frequency is  $\omega_p$ .

a distance  $u$  due to the applied field to form surface charge densities  $\sigma = \pm neu$ . The resulting depolarisation field<sup>8</sup> is  $\mathbf{E} = neu/\varepsilon_0$  and the motion of the free electrons are defined by,

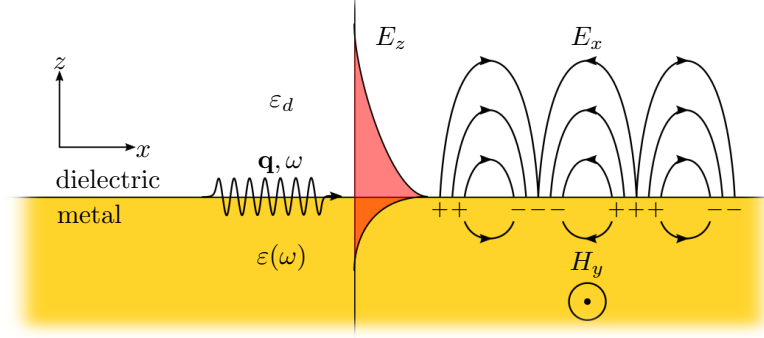
$$nm\ddot{u} = -ne\mathbf{E} = -\frac{n^2e^2u}{\varepsilon_0}. \quad (22)$$

Simplifying this relation leads to,

$$\ddot{u} + \omega_p^2 u = 0, \quad (23)$$

hence  $\omega_p$  is considered the natural frequency of the system and the electrons resonate when driving at  $\omega = \omega_p$ . This is known as the bulk or *volume plasmon*. Since this is a longitudinal oscillation, however, light cannot couple with a volume plasmon. Experimentally these are measured using electron energy loss spectroscopy (EELS), whereby colliding electrons impart energy and excite a volume plasmon [egerton2011electron]. Observation of optical plasmonic phenomena therefore means that light-matter plasmonics is the result of a different kind of plasmon.

<sup>8</sup>Gauss' law  $\int \mathbf{E} \cdot d\mathbf{A} = Q/\varepsilon_0 = \sigma A/\varepsilon_0$ , hence  $\mathbf{E} = \sigma/\varepsilon_0$ . Alternatively  $D = 0 = \varepsilon_0 \mathbf{E} + \mathbf{P}$  therefore  $\mathbf{E} = -\mathbf{P}/\varepsilon_0 = -neu/\varepsilon_0$ .



**Figure 3: Diagram of a surface plasmon polariton (SPP).** TM surface electron density waves (surface plasmons) couple with an evanescent wave originating from an EM wave to form a SPP. The SPP remains confined to the interface but can propagate across the surface.

### 1.3 Surface Plasmons, Surface Plasmon Polaritons and Localised Surface Plasmon Polaritons

Surface plasmons (SPs), unlike bulk plasmons, are tightly confined to the surface of the metal. As stated previously, light at an interface is not necessarily restricted by the diffraction limit. The maximum magnitude of the wavevector is set by  $q_0 = 2\pi/\lambda$  with individual components restricted by  $q = \sqrt{q_x^2 + q_y^2 + q_z^2}$ . Consider (13) in the form,

$$q_x^2 = \tilde{n}^2 q_0^2 - q_y^2 - q_z^2. \quad (24)$$

If a wave propagates freely in all three dimensions then it remains diffraction limited and the propagation constant  $q_x < \tilde{n}q_0$ . However if one or more of its wavevector components become imaginary ( $q_{y,z}^2 < 0$ ) then it becomes possible that  $q_x > \tilde{n}q_0$ . This behaviour can occur at an interface, where surface waves take on evanescent character in the  $z$ -direction whilst propagating in the  $xy$  plane.<sup>9</sup> By coupling light into surface waves the diffraction limit can be beaten. Waves of frequency  $\omega$  can acquire wavelengths many times smaller than their excitation wavelength. The SP is one such case of this phenomenon and occurs at metal-dielectric interfaces. Unlike in the bulk of a metal, electrons displaced by an applied field at the surface of a metal do feel a restoring force with the positive nuclei background. Transverse fields impinging on the metal surface at an angle are then able to manipulate the electron motion. SPs can therefore be manipulated by light as well as by the longitudinal waves needed to excite bulk plasmons. They also have the ability to form polariton quasiparticles under strong coupling with photons, hence they are optically excitable under the right conditions.<sup>10</sup> This optical excitation is known as the surface plasmon polariton (SPP) and, as a result of it being optically accessible, is one of the most commonly studied plasmonic phenomena.

#### 1.3.1 Surface Plasmon Polaritons

A SPP is a propagating TM wave confined to the surface of a metal and the bound state between a photon and a SP. While confined in two dimensions to the planar boundary between a metal and a dielectric, the SPP can either propagate or become stationary as a result of interference **with itself or other plasmons**. The latter stationary form of the SPP is similar to the localised surface plasmons described later.

The SPP itself is described through its dispersion. As a TM wave propagating in the  $x$ -direction along a metal/dielectric interface its field profile is given by  $\mathbf{E}(\mathbf{x}) = \mathbf{E}(z)e^{i\beta x}$  where  $\beta = q_x$  is the propagation constant. The magnetic field in this configuration is then  $\mathbf{H}(\mathbf{x}) = \mathbf{H}(y)e^{i\beta x}$ . The behaviour of such a wave in (12) is

<sup>9</sup>Evanescent meaning imaginary  $q_z$  therefore exponentially decaying amplitude in the  $z$ -direction.

<sup>10</sup>Polaritons are the name given to quanta or quasiparticles of light-matter interactions. Strong coupling describes the point at which a quasiparticle is no longer distinguishable between its two constituent components.

described by,

$$\frac{\partial^2 \mathbf{E}(z)}{\partial z^2} + (q_0^2 \varepsilon - \beta^2) \mathbf{E} = 0, \quad (25a)$$

$$\frac{\partial^2 \mathbf{H}(y)}{\partial y^2} + (q_0^2 \varepsilon - \beta^2) \mathbf{H} = 0, \quad (25b)$$

which can be solved under the appropriate boundary conditions to yield the properties of a SPP. Assuming a TM wave ( $\partial E_y / \partial z = 0$ ,  $\partial H_z / \partial z = 0$ ), propagating only in the  $x$ -direction with symmetry in  $y$ -direction, results in outcomes from (1c) and (1d) of,

$$E_x = -i \frac{1}{\omega \varepsilon_0 \varepsilon} \frac{\partial H_y}{\partial z}, \quad (26a)$$

$$E_z = -\frac{\beta}{\omega \varepsilon_0 \varepsilon} H_y. \quad (26b)$$

In this instance the TM wave equation, and requirement for evanescent decay in the  $z$ -direction, mean that  $H_y = A_d e^{i\beta x} e^{q_d z}$  for  $z > 0$  (in the dielectric) and  $H_y = A_m e^{i\beta x} e^{q_m z}$  for  $z < 0$  (in the metal). The components of the electric field can therefore be expressed as,

$$E_x(z) = i A_d \frac{1}{\omega \varepsilon_0 \varepsilon_d} q_d e^{i\beta x} e^{-q_d z}, \quad (27a)$$

$$E_z(z) = -A_d \frac{\beta}{\omega \varepsilon_0 \varepsilon_d} e^{i\beta x} e^{-q_d z}, \quad (27b)$$

inside the dielectric and,

$$E_x(z) = -i A_m \frac{1}{\omega \varepsilon_0 \varepsilon_m} q_m e^{i\beta x} e^{q_m z}, \quad (28a)$$

$$E_z(z) = -A_m \frac{\beta}{\omega \varepsilon_0 \varepsilon_m} e^{i\beta x} e^{q_m z}, \quad (28b)$$

inside the metal, where  $\varepsilon_1 = \varepsilon_d$  and  $\varepsilon_m = \varepsilon(\omega)$  in the previously used notation. Continuity across the boundary ( $z = 0$ ) dictates that  $A_d = A_m$  and yields the relation,

$$\frac{q_2}{q_1} = -\frac{\varepsilon_m}{\varepsilon_d}, \quad (29)$$

hence the ratio between wavevectors inside and outside of the metal depends on the relative change in dielectric constant across the boundary. It is this continuity relation that allows for the existence of SPPs in the TM configuration, unlike for TE waves for which this result is not possible. To determine the SPP propagation constant  $\beta$  a further relation is needed to fix the wave vectors in a given medium with respect to the dielectric constant. Using the expression for  $H_y$  in the wave equation (12) yields a set of relations,

$$q_m^2 = \beta^2 - q_0^2 \varepsilon_m, \quad (30a)$$

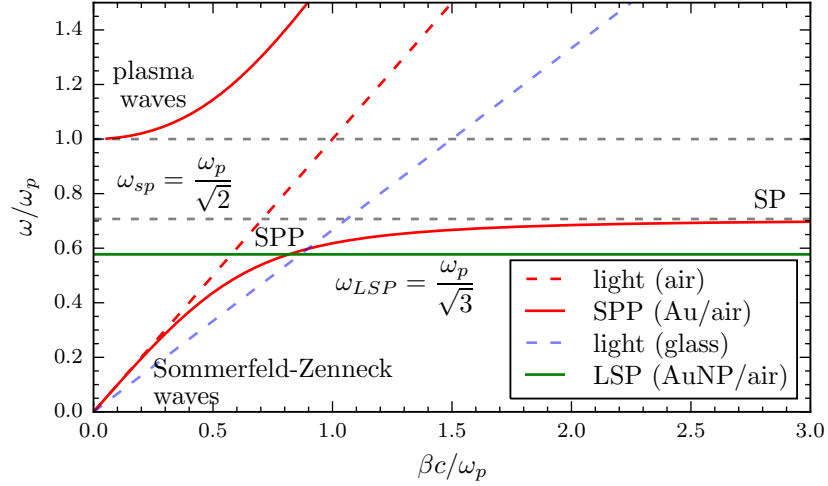
$$q_d^2 = \beta^2 - q_0^2 \varepsilon_d, \quad (30b)$$

which, when combined with (29), fully describe the fields around the interface. The propagation constant is then given by,

$$\beta = \frac{\omega}{c} \sqrt{\frac{\varepsilon_d \varepsilon(\omega)}{\varepsilon_d + \varepsilon(\omega)}}. \quad (31)$$

This is the dispersion relation for a SPP and describes many of their important properties.

The SPP dispersion described by (31) is shown in Figure 4 along with the dispersion of light for both air and glass mediums. From the dispersion curve it is clear that SPPs cannot couple with light within the same medium as their dispersion curves do not cross. However, light from within a higher refractive index medium such as



**Figure 4: Plasmon dispersion relations for the SPP and LSP.** The dashed lines indicate the dispersion of light in both glass and air (vacuum) along with the surface plasmon frequency. SPPs can be described as photon-like or plasmon-like depending on their point of excitation. SPPs excited with large  $q$  and  $\omega \approx \omega_{SPP}$  are considered plasmon-like while SPPs with low  $q$  are considered more photon-like.

glass can generate evanescent waves and excite SPPs on a nearby metal/air interface. This method of coupling photons with surface plasmons, depending on the specific prism arrangement, is known as the Kretschmann (prism-metal-dielectric) or Otto (prism-dielectric-metal) configuration [otto1968, kretschmann1971]. Since a diffraction grating may also impart momentum onto a photon ( $q_x \rightarrow q_x + n\pi$ ) a metallic grating can launch SPPs along a planar metal-dielectric interface. This phenomenon was first observed in 1902 by Wood, dubbed as Wood's anomaly [wood1902], and only explained via surface waves many years later [fano1941].

Closer inspection of the curve highlights one of the major features of a plasmon. While SPPs retain the frequency of the excitation field, their wavelength is considerably smaller than the diffraction-limited wavelength of light. Depending on where on the curve the SPP lies it can be considered to be either more photon-like or more plasmon-like. For small  $\beta \approx q_0$  the SPP is similar to light grazing the interface whereas SPPs with large  $q$  become more plasmon-like and their frequency saturates at the surface plasmon frequency,

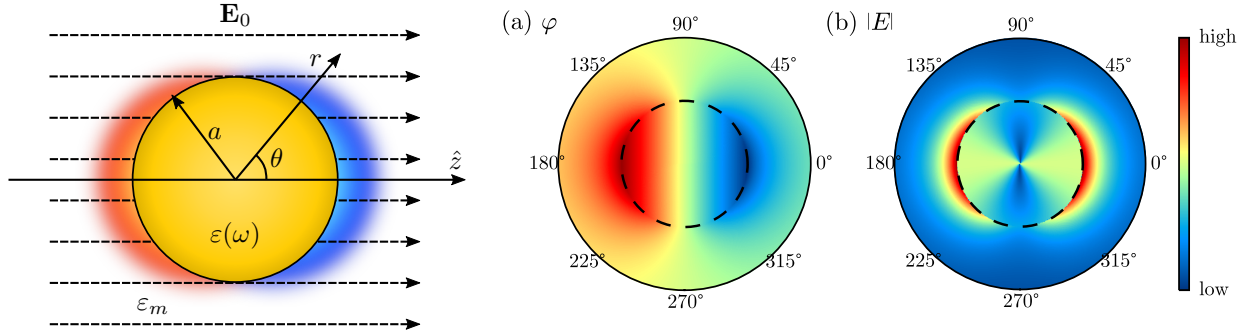
$$\omega_{sp} = \frac{\omega_p}{\sqrt{1 + \epsilon_d}}. \quad (32)$$

At this point the SPP can be considered electrostatic and becomes a SP. To some extent, SPs confined to a finite, continuous, non-planar surface, defining a nanoparticle (NP), can be considered to be the basis for a localised surface plasmon (LSP) or localised surface plasmon polariton (LSPP).

### 1.3.2 Localised Surface Plasmons

LSPs are collective oscillations of conduction electrons confined within a fixed sub-wavelength spatial extent, usually on the surface of a metallic nanoparticle (MNP). Free electrons are displaced from the nuclei in response to an applied field and form a surface charge distribution, polarising the particle. Coulomb interaction between the poles of the surface charge distribution results in a restoring force within the particle. This gives rise to a natural frequency of oscillation, leading to a surface plasmon resonance (SPR) when driven harmonically at the correct frequency. The particle geometry sets which multipolar surface charge distributions are supported while its material properties and the dielectric properties of the surrounding medium set the restoring force. Each different multipolar charge distribution is therefore considered to be a unique LSP mode, identifiable by its SPR [murray2007]. This bears some similarity with the slab of free electron gas supporting longitudinal volume plasmons, shown in Figure 2, except that the sub-wavelength geometry modifies the restoring force and allows





**Figure 5: A spherical metallic particle in an applied electric field.** The sphere is assumed to be in the quasistatic regime ( $a \ll \lambda$ ). The aura around the particle indicates the phase of the free electron oscillations in the plasmon. Calculations of (a) the potential and (b) the magnitude of the electric field for a spherical nanoparticle on resonance ( $\varepsilon(\omega) = -2\varepsilon_d$ ).

transverse modes of oscillation. The plasma frequency in LSPs is then rescaled by a factor depending on the geometry of each LSP mode.

The simplest form of a LSP is the dipole resonance of a spherical MNP. Assuming the sphere radius,  $a \ll \lambda$ , the wavelength of light, the particle is considered to be in the quasistatic regime and electrostatics, rather than electrodynamics, is applicable to solve the problem. The electrostatic potential,  $\varphi$ , of the system is described by the Laplace equation,  $\nabla^2 \varphi = 0$ , with a general solution in a spherical geometry of the form,

$$\varphi_{l,m_l}(r, \theta, \phi) = \sum_{l=0}^{\infty} \sum_{m_l=-l}^l [A_l r^l + B_l r^{-(l+1)}] P_l^m(\cos \theta) e^{im_l \phi}, \quad (33)$$

where  $l$  is the degree of spherical harmonic and  $m_l$  its projection,  $A_l$  and  $B_l$  are constants and  $P_l^m(\cos \theta)$  are associated Legendre polynomials. For a sphere of radius  $a$  and dielectric function  $\varepsilon(\omega)$  in a dielectric medium described by  $\varepsilon_d$  the solution is fixed by the boundary conditions  $\varphi_{out} \rightarrow -E_0 z$  as  $r \rightarrow \infty$  and  $\nabla \varphi_{in}(r)|_{r=a} = \nabla \varphi_{out}(r)|_{r=a}$ . This reduces to a solution [Jackson1999classical],

$$\varphi = \begin{cases} -\frac{3\varepsilon_d}{\varepsilon(\omega) + 2\varepsilon_d} \mathbf{E}_0 \cdot \mathbf{r} & r \leq a \text{ (inside)}, \\ \left( -1 + \frac{\varepsilon(\omega) - \varepsilon_d}{\varepsilon(\omega) + 2\varepsilon_d} \frac{a^3}{r^3} \right) \mathbf{E}_0 \cdot \mathbf{r} & r > a \text{ (outside)}. \end{cases} \quad (34)$$

For a metal sphere the potential describes an induced dipolar surface charge distribution, as plotted in Figure 5a. The description is simplified by defining the dipole moment as,

$$\mathbf{p} = \varepsilon_0 \varepsilon_d \alpha \mathbf{E}_0 = 4\pi \varepsilon_0 \varepsilon_d a^3 \frac{\varepsilon(\omega) - \varepsilon_d}{\varepsilon(\omega) + 2\varepsilon_d} \mathbf{E}_0, \quad (35)$$

where the polarisability,  $\alpha$ , incorporates the frequency dependent behaviour and is defined as,

$$\alpha(\omega) = 4\pi a^2 \frac{\varepsilon(\omega) - \varepsilon_d}{\varepsilon(\omega) + 2\varepsilon_d}. \quad (36)$$

The outside potential is then expressed as,

$$\varphi_{out} = -\mathbf{E}_0 \cdot \mathbf{r} + \frac{\mathbf{p} \cdot \mathbf{r}}{4\pi \varepsilon_0 \varepsilon_d r^3}. \quad (37)$$

The potential of the system is simply the potential of a dipole superimposed onto the incident field. The electric

field inside and outside of the sphere is calculated using  $\mathbf{E} = -\nabla\varphi$ , resulting in,

$$\mathbf{E} = \begin{cases} \frac{3\varepsilon_d}{\varepsilon(\omega) + 2\varepsilon_d} \mathbf{E}_0 & r \leq a \text{ (inside),} \\ \mathbf{E}_0 + \frac{3\mathbf{n}(\mathbf{n} \cdot \mathbf{p}) - \mathbf{p}}{4\pi\varepsilon_0\varepsilon_d} \frac{1}{r^3} & r > a \text{ (outside),} \end{cases} \quad (38)$$

where  $\mathbf{n} = \mathbf{r}/r$  is the unit vector. As with the potential, the electric field is the superposition of the incident field with the field emitted from a point dipole, shown in Figure 5b.

It is assumed from its sub-wavelength size that electrons in the sphere respond instantaneously to the incident field and phase retardation is ignored (quasistatic approximation). Under this assumption the electrostatic result is simply multiplied with a harmonic time dependence to describe electrodynamic behaviour. Therefore, when illuminated by a plane wave  $\mathbf{E}_0 e^{i\omega t}$  the induced dipole moment oscillates like  $\mathbf{p} e^{i\omega t}$ . In this sense the coherent oscillation of free electrons in the sphere is considered to be an oscillating point dipole source given by  $\mathbf{p} e^{i\omega t} = \varepsilon_0 \varepsilon_d \alpha \mathbf{E}_0 e^{i\omega t}$ . The physical behaviour of the electrons in the sphere, described using the dielectric function  $\varepsilon(\omega)$ , is then simply incorporated into  $\alpha(\omega)$ , at which point its frequency dependence becomes important.

For a good metal  $\text{Re}[\varepsilon(\omega)] < 0$  and the denominator in (36) undergoes resonance whereby  $\alpha \rightarrow -\infty$  when  $\varepsilon(\omega) + 2\varepsilon_d \rightarrow 0$ . Resonance therefore occurs at the Fröhlich condition when,

$$\text{Re}[\varepsilon(\omega)] = -2\varepsilon_d. \quad (39)$$

Resonant excitation of the induced oscillating dipole moment corresponds to excitation of a collective oscillation of conduction electrons on the surface of the sphere - the dipolar LSP. It's strength in realistic metals is restricted by damping of the electron motion leading to a Lorentzian-shaped resonance band - the dipolar SPR. As seen in (38), the field from the induced dipole moment of the plasmon is superimposed onto the incident field leading to a resonant near-field enhancement, both inside and outside the surface of the sphere. This is one of the fundamental properties of the plasmon, and one that is most exploited in sensing and sensor developments.

Under the Drude model, with  $\varepsilon(\omega)$  given by (21), the Fröhlich condition is satisfied when,

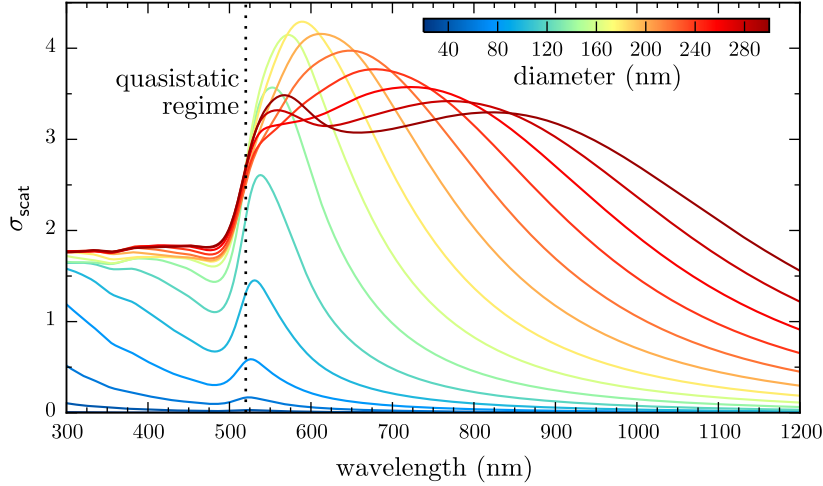
$$\omega_{lsp} = \frac{\omega_p}{\sqrt{1 + 2\varepsilon_d}}, \quad (40)$$

which evaluates to  $\omega = \omega_p/\sqrt{3}$  for a MNP in vacuum. As can be seen in Figure 4, the flat dispersion of a LSP mode means it crosses the light line at a single point. Light of the correct frequency therefore readily couples with LSPs without the need for SPP momentum matching mechanisms. In general, the optical spectrum of a MNP can contain a number of multipolar plasmon modes **other than the dipole mode** for which the resonant frequencies are given by,

$$\omega_l = \omega_p \sqrt{\frac{l}{\varepsilon_d(l+1) + 1}}, \quad (41)$$

where the degree of spherical harmonics  $l$  denote the charge distribution and denoted mode order ( $l = 1$  for dipole,  $l = 2$  for quadrupole, etc.). However, these modes only exist outside of the quasistatic regime in larger MNPs or more complex geometries. For plasmonically-active noble metals, such as Au and Ag, the fundamental  $l = 1$  mode occurs in the visible spectrum ( $\lambda = 520$  nm for Au and  $\lambda = 360$  nm for Ag), leading to them often being the plasmonic metal of choice. Additionally, since the polarisability changes with NP geometry, changes from a spherical shape lead to tuning of the SPRs across the visible spectrum. This geometrical dependence is well known and has been measured experimentally on a number of occasions [mock2002, kuwata2003].

As with any oscillating charge distribution, energy is both absorbed and emitted from the plasmon. The oscillating charge has the ability to scatter the incident planar field into spherical waves. The excitation of a resonant dipolar oscillation means that the scattering is also resonantly enhanced. The balance between the absorbance and the scattering from a MNP is dictated by its size. The absorbance and scattering cross sections



**Figure 6: Mie scattering efficiencies/cross-sections for AuNPs of increasing diameter.** The 520 nm resonance position of the dipolar LSP mode of a AuNP in the quasistatic approximation is indicated by the dotted line. The resonance stays at 520 nm until  $d > 80$  nm then redshifts. The emergence of higher order modes following a similar behaviour is seen once  $d > 100$  nm.

are given by [bohren2008absorption],

$$\sigma_{\text{scat}} = \frac{q^4}{6\pi} |\alpha|^2 = \frac{8\pi}{3} q^4 a^6 \left| \frac{\varepsilon(\omega) - \varepsilon_d}{\varepsilon(\omega) + 2\varepsilon_d} \right|^2, \quad (42a)$$

$$\sigma_{\text{abs}} = q \text{Im}[\alpha] = 4\pi q a^3 \text{Im} \left[ \frac{\varepsilon(\omega) - \varepsilon_d}{\varepsilon(\omega) + 2\varepsilon_d} \right]. \quad (42b)$$

The extinction cross section, commonly used in spectroscopy, can be calculated using  $\sigma_{\text{ext}} = \sigma_{\text{scat}} + \sigma_{\text{abs}}$ . Since they depend on  $\alpha$  the size of the cross section, i.e. the spatial extent over which light can interact with the MNP, is increased on resonance. This is why MNPs appear strongly coloured and much larger in optical microscopy than they actually are. The absorbance and scattering cross sections scale as  $a^3$  and  $a^6$ , respectively, hence larger particles scatter more than they absorb whereas absorption dominates in smaller particles.

From the cross-sections the relationship between the LSP near-field and the far-field is made clearer. On resonance with the dipolar LSP, both cross sections are enhanced by the polarisability resonance from  $\mathcal{O}(a \sim 50 \text{ nm})$  to  $\mathcal{O}(500 \text{ nm})$ . The increased size of the cross-section is comparable to the wavelength of light meaning LSPs efficiently couple with photons in the far-field.<sup>11</sup> The LSP mediates energy transfer between the near-field and the far-field and acts to match the electromagnetic modes of nanoscale absorbers/emitters, such as phonons (Raman) and radiative energy levels (quantum emitters, fluorescence), with those of a diffraction-limited photonic mode via an oscillating charge density [berweger2012]. A plasmonic NP is often therefore described as an *optical antenna* in a similar manner to a device that converts between radio waves and an electrical current is named a radio antenna. LSP modes which readily couple with the far-field are then sometimes referred to as *antenna modes* and become important when designing resonant structures for specific sensing applications.

Whilst the quasistatic approach is useful to first demonstrate the enhancing capabilities of a MNP, the description breaks down once the size of the particle becomes more comparable to the excitation wavelength. Retardation effects between the field and the electrons mean that phase differences between the charge oscillations and the incident field become important. At this point Mie theory (electrodynamics) is required to describe the spectral response of spherical MNPs [mie1908]. Mie theory provides a more general description of the optical response of spherical MNPs. Using this approach the spectra of a MNP can be decomposed into superimposed

<sup>11</sup>Consider that  $\sigma_{\text{scat}}$  for a AuNP is enhanced 100 $\times$  on resonance, meaning it's area cross-section is  $\sqrt{100/\pi} = 6\times$  wider than it's radius, hence why a 50 nm AuNP looks like a 300 nm green sphere when imaged.

multipoles, alluding to the existence of higher order LSP modes in larger MNPs. The spectral response of spherical AuNPs of varying sizes is shown in Figure 6, demonstrating the redshift and broadening of lower order modes with increasing particle size and the excitation of higher order modes.

### 1.3.3 Geometrical Influences on Localised Surface Plasmon Resonances

As previously stated, the LSP is a geometrical resonance and exists primarily due to the interplay between the driving field inducing oscillations in the mobile electrons and the resulting restoring force acting between the positive ionic core background and the accumulation of conduction electrons at the surface. While the surface charge oscillation directly correlates to the driving field the restoring force, incorporated into the polarisability, is geometry dependent. The SPRs of a single particle are therefore very geometry dependent [krenn2000, mock2002, kuwata2003]. The larger the separation between opposing surface charge poles in the antenna, the weaker the restoring force. As seen with Mie theory, a larger particle with a smaller restoring force has a lower energy resonance. The same can be said if the particle geometry is elongated (a nanoellipsoid or nanorod). Within the quasistatic regime, the polarisability of an ellipsoidal MNP is adjusted by inserting a geometrical correction, leading to the redefinition,

$$\alpha_i(\omega) = 4\pi a_1 a_2 a_3 \frac{\varepsilon(\omega) - \varepsilon_d}{3\varepsilon_d + 3L_i(\varepsilon(\omega) - \varepsilon_d)}, \quad (43)$$

where  $i$  is the index of each anisotropic axis with a geometrical factor,

$$L_i = \frac{a_1 a_2 a_3}{2} \int_0^\infty \frac{dq}{(a_i^2 + q)\sqrt{(q + a_1^2)(q + a_2^2)(q + a_3^2)}}. \quad (44)$$

The resonance condition along each axis is then changed to,

$$\text{Re}[\varepsilon(\omega)] = -\frac{1 - L_i}{L_i} \varepsilon_d. \quad (45)$$

By decreasing the geometrical factor the resonance condition decreases from  $-2\varepsilon_d$ , corresponding to a decreased resonant frequency and a redshifted SPR. The longer the particle is, the larger the redshift until the particle is large enough to no longer be considered plasmonic. Continual separation of the poles inevitably weakens the restoring force to the point that each lowest-order antenna SPR no longer exists.

By utilising the particle material, geometry and polarisation anisotropy to tune the plasmon resonance, the resonant wavelength band can be shifted across the entire UV–NIR spectrum to tailor individual applications. The limitation to this technique is the relatively small field enhancement that a single particle can provide. An alternative approach to exploiting LSPs is therefore to couple the fields of many plasmons together. Through coupling the confined fields in the gaps between MNPs can be enhanced by many more orders of magnitudes.