Imperial College London

JAMIE M. FITZGERALD

ELECTROMAGNETIC FIELD ENHANCEMENT IN CLASSICAL AND QUANTUM PLASMONICS

THESIS SUBMITTED IN PARTIAL FULFILMENT FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

CONDENSED MATTER THEORY GROUP
DEPARTMENT OF PHYSICS
IMPERIAL COLLEGE LONDON

Contents

It is no exaggeration to say that every single person I have gotten to know at Imperial has been a delight to socialise and discuss physics with. It has been a pleasure to be part of such a friendly department. In particular, I would like to thank my supervisor Vincenzo Giannini whose enthusiasm and great knowledge of plasmonics has provided a wonderful atmosphere to undertake a PhD in. I would also like to thank all the authors of the various open-source software that I have used throughout the PhD.

Abstract

Plasmonics has ushered in an era of precise control over light using the hybrid light-matter collective oscillations of electrons in metals, sandwiched in-between the realms of electronics and photonics, i.e. nanometre devices operating at frequencies beyond the terahertz. It offers the promise to mirror the huge technological impact that electronics and photonics have had on all our lives, providing an interface between the two that combines the strengths of each. One of the most defining features of plasmonics are 'hot spots': regions of extreme field enhancement confined to sub-diffraction volumes. Simulations of these highly localised fields are numerically challenging and need sophisticated theory involving a firm understanding of the intricacies of the near-field. In this work, the physics of extreme field enhancement is explored by considering three exemplary systems in classical and quantum plasmonics: the cavity, the nanolens, and the nanorod. For the cavity geometry, the role of new nanophotonic materials, graphene and polar dielectrics, are explored for building a novel platform for molecular sensing in the mid-infrared. Furthermore, the role of strong-coupling between the constituent modes is analysed. For the nanolens, the limits of extreme field enhancement are explored in the quantum regime where nonlocal losses kill strong nanolensing, and in the micron regime where localised surface phonon polaritons excited in polar dielectrics lead to unprecedented field enhancements on the order of 10^4. Far superior to disappointing metallic nanolenses, which are found to offer no better performance than the simpler spherical dimer geometry. Finally, the ultimate small size limit of the nanorod is explored: the single-atom-thick atomic chain. Electronic structure methods are used to identify quantum plasmons and explore the main loss channel at these sizes: plasmon-phonon coupling. This allows *ab initio* calculation of field enhancements, a first in quantum plasmonics.

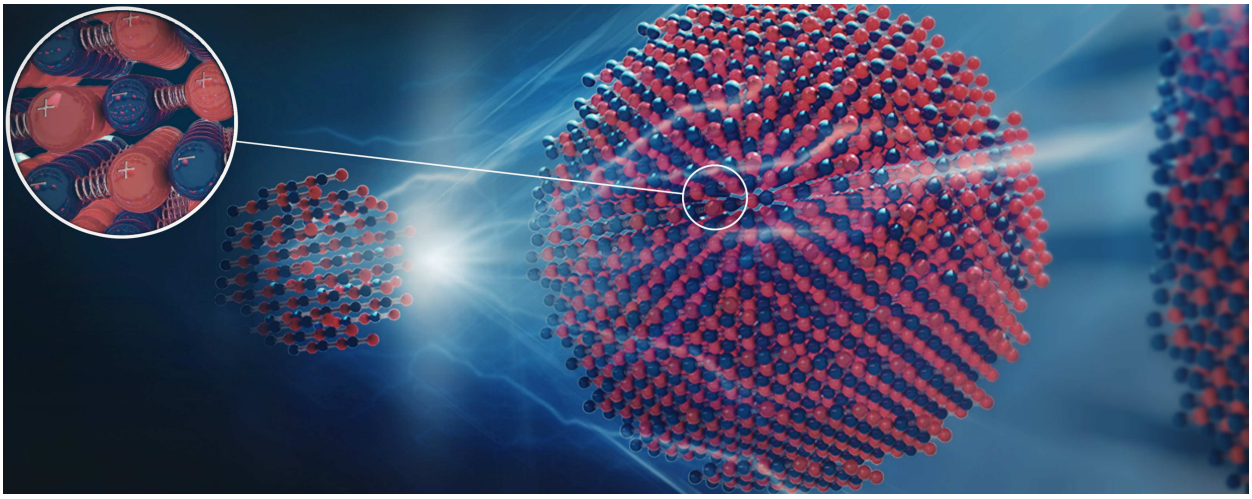


Figure 1: Artistic representation of extreme cascade nanolensing using localised surface phonon polaritons excited in polar dielectrics.

List Of Publications

- Fitzgerald, J. M., Narang, P., Craster, R. V., Maier, S. A., & Giannini, V. (2016). Quantum plasmonics. *Proceedings of the IEEE*, 104(12), 2307-2322.
- Fitzgerald, J. M., & Giannini, V. (2017). Perspective on molecular quantum plasmonic nanoantennas. *J. Opt*, 19(060401), 060401.
- Fitzgerald, J. M., Azadi, S., & Giannini, V. (2017). Quantum plasmonic nanoantennas. *Physical Review B*, 95(23), 235414.
- Fitzgerald, J. M., & Giannini, V. (2018). Battling Retardation and Nonlocality: The Hunt for the Ultimate Plasmonic Cascade Nanolens *ACS Photonics* 5(6), 2459-2467
- Li, K., Fitzgerald, J. M., Xiao, X., Caldwell, J. D., Zhang, C., Maier, S. A., Li, X. & Giannini, V. (2017). Graphene Plasmon Cavities Made with Silicon Carbide. *ACS Omega*, 2(7), 3640-3646. (**Joint first author**)
- Kerber, R. M., Fitzgerald, J. M., Reiter, D. E., Oh, S. S., & Hess, O. (2017). Reading the orbital angular momentum of light using plasmonic nanoantennas. *ACS Photonics*, 4(4), 891-896.
- Lazzarini, C. M., Tadzio, L., Fitzgerald, J. M., Sánchez-Gil, J. A., & Giannini, V. (2017). Linear ultrafast dynamics of plasmon and magnetic resonances in nanoparticles. *Physical Review B*, 96(23), 235407.
- Gisbert-Quilis, P., Masetti, M., Morla-Folch, J., Fitzgerald, J. M., Pazos-Perez, N., Garcia-Rico, E., Giannini, V., Alvarez-Puebla, R.A. & Guerrini, L. (2017). The Structure of Short and Genomic DNA at the Interparticle Junctions of Cationic Nanoparticles. *Advanced Materials Interfaces*, 4(17).
- Kerber, R. M., Fitzgerald, J. M., Oh, S. S., Reiter, D. E., & Hess, O. (2018). Orbital angular momentum dichroism in plasmonic nanoantennas (**submitted**)
- Kerber, R. M., Fitzgerald, J. M., Xiao, X., Oh, S. S., Maier, S. A., Giannini, V., Reiter, D. E. (2018). Interaction of an Archimedean spiral structure with orbital angular momentum light (**submitted**)

List Of Conferences And Talks

- Invited talk at the University of Münster. July 2017.
- Talk at "MRS Spring Meeting", April 2017, Phoenix, Arizona
- Talk at "Photon16", September 2016, Leeds University.
- Poster presented at "Quantum Plasmonics" workshop, August 2016 Imperial College
- Poster presented at "Complex Nanophotonics Science Camp", September 2015, Cumberland Lodge, Windsor.
- Poster presented at "Physics by the lake", Cumberland Lodge, Windsor, August, 2015. EPRSC UK graduate summer school in condensed matter theory.

Introduction

Having spent close to four years on my PhD, and gaining a few pounds of weight for my troubles, it is a pleasure to present this thesis, my *magnum opus*, to the world. It would be amiss not to begin with a few obligatory words in order to give some context, indicate what the reader can expect from reading this work and hopefully catch their fancy. We will be predominately interested in the optical response of metals, or more generally materials that can behave as 'optical metals' such as graphene and polar dielectrics, structured on the micron and nanometre scale. The interaction of electromagnetic (EM) radiation with the boundaries of metallic structures to excite collective oscillations of the electron density is known as plasmonics, which is a subfield of nanophotonics. Plasmonics, via the conversion of EM energy into electronic degrees of freedom, allows sub-diffraction confinement of light, leading to enhanced optical near-fields that can be controlled with nanometre precision. In contrast, traditional photonics is limited by the diffraction limit, leading to bulky devices that are difficult to incorporate with electronic circuits. This is a very relevant problem as future electronic devices will be increasingly limited by fundamental speed and bandwidth limitations; using photons as information carriers is seen as a solution to this problem (Zia et al., 2006; Gramotnev and Bozhevolnyi, 2010). Plasmonics promises to form the backbone of next-generation optoelectronic circuits that merge photonics and electronics in sub-diffraction volumes, see figure 2 for an overview of the different spatial and spectral regimes. By control of nanostructure geometry, made possible by the development of state-of-the-art fabrication methods in the last two decades that allow for routine and precise construction on the nanoscale, the near- and far-field properties of light can be manipulated with exquisite control. This has been complemented by the ever-increasing computational power available to researchers[1], improved algorithms, and a plethora of easy to use commercial and open-source software for advanced numerical modelling. Numerical simulations that a few years ago would have been daunting can now be performed on short timescales, allowing simple comparison of theory and experiment, and the sweeping of large parameter spaces making structure optimisation possible. This has allowed research beyond the spherical geometry; the sphere, for so long, was the treasured plaything of theorists due to an exact analytical

[1] As plasmonics involves sub-wavelength spatial variation it is an inherently challenging multi-scale modelling problem, this should be compared to traditional photonics that has a single defining length scale given by the wavelength.

[2] The famous Mie theory.

description[2], the 'fruit fly' of plasmonics along with its slightly more complicated cousin: the ellipsoid. Now we can explore more challenging geometries, leading to novel ways to control properties of light such as the polarization, phase, magnetic response and field enhancement (FE). As well as structuring matter, it is now becoming increasingly viable to structure the light field. The development of phase-structured beams, such as Bessel beams that carry orbital angular momentum, offer yet another pathway to intriguing new light-matter interactions (Kerber et al., 2017). If I may be a tad hyperbolic: plasmonics has made electromagnetism, as a research topic in its own right, new, exciting and sexy!

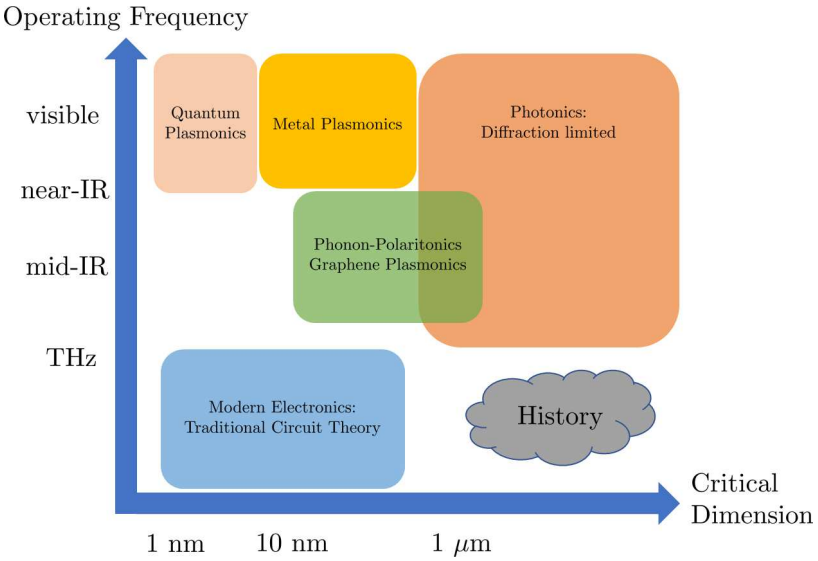


Figure 2: A rough illustration of the operating frequencies and critical dimensions of the topics in this thesis, compared to traditional photonics and electronics. Inspired by figure 8 from (Zia et al., 2006) and expanded.

Plasmonics is an application-driven field with utilisation already in sensing (Anker et al., 2008), such as in pregnancy tests where antibody conjugated gold nanoparticles can be seen by the human eye due to a strong colour change, and localised heating (Baffou and Quidant, 2013), where it has been used for photothermal therapy (Liang et al., 2015). There are further potential applications in lasing (Bergman and Stockman, 2003), controllable and localised chemistry (Cortés et al., 2017), metamaterials/metasurfaces (Boltasseva and Atwater, 2011), information transfer and processing with nanoscale photonic circuits (Ozbay, 2006; Gramotnev and Bozhevolnyi, 2010), coupling to quantum emitters (Pelton, 2015), energy conversion (Atwater and Polman, 2010), and precise photon control in quantum optics (Tame et al., 2013). In particular, thermoplasmonic and hot electron generation based applications are exciting as they actually benefit from the inherent loss in plasmonics, which can plague other applications.

The confinement of light in plasmonics may be so extreme that the quantization of both light and matter may be necessary for an accurate description. Cutting-edge experiments now probe these regimes and push existing theories to the limits (Scholl et al., 2012;

Ciracì et al., 2012; Savage et al., 2012; Raza et al., 2013a; Scholl et al., 2013). This opens up possibilities to explore the nature of many-body collective oscillations as well as develop next-generation quantum-based plasmonic devices that use the particle quality of light and/or the wave quality of matter. Despite the underlying fundamental theories of quantum mechanics and electromagnetism being well understood, extreme light-matter states offer the potential for exciting new physics. Noteworthy work in recent years include the demonstrations of quantum interference between surface plasmon polaritons (SPPs) (Heeres et al., 2013; Di Martino et al., 2014), and strong-coupling between plasmons and single molecules (Chikkaraddy et al., 2016). Unfortunately, these interesting topics must largely be ignored in this thesis and from now on, I reserve the term "quantum plasmonics" to indicate the quantisation of the matter part of plasmonics.

Quantum calculations in plasmonics come in two flavours. One is to continue using the intuitive classical framework and add quantum effects as a correction, this is the nonlocal approach and is commonly implemented within the hydrodynamic model (HM). The second approach relies on performing a complete quantum calculation from first principles[3] and considering the plasmon as constructed from a linear combination of electron-hole (eh) transitions. Time-dependent density-functional theory (TDDFT) is a popular method for evaluating optical properties of quantum systems and allows the description of the correlated dynamics of a many-body system via the solution of an auxiliary system of independent particles in a mean-field potential. Which approach to take, *ab initio* or quasi-quantum, often comes down to necessity, rather than taste. TDDFT is usually only feasible for a few hundred electrons, core electrons can often be considered via a suitable pseudopotential and do not need to be explicitly calculated, which is a huge computational saving. For larger systems[4], full quantum calculations become increasingly infeasible. In this tricky region, which is too large for *ab initio* calculations and too small to be treated classically, quasi-quantum models like the HM come into their own. Interestingly, this encompasses a size regime where both quantum effects and retardation may be important to include.

In this thesis, we will explore the limits of electric field concentration in classical and quantum systems. We will look at the ultimate small size limit of plasmonic components: few atom/molecule systems. They behave in many respects like the analogous classical plasmonic antennas (Giannini et al., 2011), supporting a 'proto-plasmonic' response and efficiently (for their size) converting light from the far-field into a concentrated near-field due to a strong dipole moment, which results from constructive interference between near degenerate molecular states of the same symmetry. This is both of interest to our understanding of fundamental quantum plasmon theory and could help pave the way for the design of efficient molecules for strong light–matter interaction. Plasmonic

[3] I have had some fun debates about whether density-functional theory (DFT) really is an *ab initio* approach.

[4] What is meant by large here is a little fuzzy, it will depend on the material, whether the system has any symmetries and the computational resources available. It will also depend on the type of calculation needed; a one-off calculation of an excitation spectrum to compare to an experimental result is a very different proposition to repeated calculations exploring a particular parameter, as is typical in theoretical studies. I will therefore loosely attach to the meaning of 'large' the range $\sim 1 \rightarrow 5$ nm for some characteristic length of the system.

clusters[5] could act as a useful middle-ground between classical plasmonic nanoparticles, which have a massive oscillator strength but low tunability due to many free carriers, and standard quantum systems, like molecular dyes and quantum dots, which have great tunability but a small oscillator strength.

We will also explore the limit of achievable FEs by considering the well-known plasmonic cascade effect (Li et al., 2003). The incident light is approximately uniform on the scale of a typical plasmonic nanoparticle, but the induced field is highly spatially inhomogeneous with 'hot spots' appearing in the near-field. Using a combination of geometries with advantageous morphology, such as sharp points and narrow gaps, along with an on-resonant excitation frequency, leads to intense regions of field concentration. In fact, it was hot spots that drove a lot of the original interest in plasmonics, in the search for the mechanisms behind the huge observed amplification of Raman signals in surface-enhanced Raman spectroscopy (SERS). It is now known that plasmon resonances are the main explanation for the colossal amplification (Le Ru and Etchegoin, 2008). The idea of squeezing light to achieve a strong FE is almost as old as science itself; a "burning lens" is a convex lens that concentrates the sun's rays down to a small point to achieve localised heating. The celebrated ancient Greek mathematician Archimedes is said to have used burning lenses to incinerate the Roman fleet attacking Syracuse in 212 BC[6], see figure 3. The idea is still used today for solar furnaces in industry. Archimedes' concept was based on classical optics, but a plasmonic analogue exists where the field is concentrated into a near-field focal point, rather than far-field, using a finite chain of self-similar nanospheres. Using a quasi-quantum model, we find that nonlocality considerably degrades the cascade effect in small nanolenses, limiting the achievable FE. A promising strategy for improving FEs is found by moving to longer wavelengths in the mid-infrared (IR) to THz and using polar dielectrics, such as silicon carbide (SiC), that support strong resonances due to localised phonon polaritons rather than plasmons. The larger operational wavelengths, which reduces retardation losses, and high quality factors result in an extremely strong nanolensing, with FEs on the order of 10^4 predicted.

Giving a title to this thesis and choosing its content was difficult as I have been fortunate enough to work on several different topics within the broad subject area of nanophotonics. I settled on the title "Electromagnetic Field Enhancement in Classical and Quantum Plasmonics" as the bulk of my work involved, especially in the second part of my PhD, consideration of FEs in classical and quantum plasmonic system. I have chosen three projects involving systems that range over length scales from angstroms to microns, have operational frequencies from the mid-IR to the visible, and are built from a variety of different materials relevant to modern nanophotonics. All are concerned with the manipulation of the near-field using collective light-matter hybrid modes. I use the

[5] I use the term "cluster" to specifically mean an ensemble of atoms in intermediate size regime between molecules and nanoparticles that can be described by bulk properties. Quantum plasmonics can be considered a subfield of cluster science.

[6] It is debated whether this actually is feasible; the TV series MythBusters attempted to replicate Archimedes' feat, with little success.

word "plasmon" in the title for simplicity, but really this thesis is more general with a particular focus on surface phonon polaritons (SPhPs), which offer advantages (and disadvantages) compared to traditional metal-based plasmonics. Unfortunately, I must ruthlessly leave out some very interesting work I have done on the excitation of dark modes by twisted light, the excitation of twisted plasmons, and using relativistic electrons to heat up metallic nanoparticles. I could argue this increases the overall coherence of the thesis but really it is due to laziness and a looming submission deadline!

As quantum plasmonics is a blend of nanophotonics and electronic structure, the readers of this thesis[7] may come from either of these fields and so I have added some extra details to the introduction and derivations, my hope is that this work is quite self-contained. I have included tricks and observations that I found to be useful on my research journey and have structured it like this for my own sake; I don't want this thesis to be a work just to (hopefully!!) pass a viva, it will be of use to myself for many years to come with a precious collection of derivations and consistent notation. This has the unfortunate consequence that it has become longer than it needs to be but in this digital age it is not such a big problem. The non-essential calculations are relegated to the appendices at the end, so it can be read without the reader being swamped by mathematical details. Finally, I would like to remark that I have had a lot of fun on my journey through this PhD, I hope this thesis can convey at least a small amount of this.

[7] If there are any beyond the examiners!

Figure 3: Painting by Giulio Parigi (1571-1635) showing Archimedes' mirror burning Roman ships. Painted in 1600.

"From a long view of the history of mankind, seen from, say, ten thousand years from now, there can be little doubt that the most significant event of the 19th century will be judged as Maxwell's discovery of the laws of electrodynamics."

- Richard Feynman

1 Classical Plasmonic Theory

1.1 Macroscopic Dielectric Theory

Classical plasmonics is based on macroscopic dielectric theory. In this section, we discuss the assumptions of this approximation and the basic equations, from which we will be able to model plasmons and will find that they dominate the optical response of metals near the plasma frequency.

We begin by stating the *macroscopic Maxwell's equations* in non-magnetic media[1]

$$\begin{aligned} \nabla \cdot \mathbf{D}(\mathbf{r},t) &= \rho_{ext}(\mathbf{r},t) \\ \nabla \cdot \mathbf{B}(\mathbf{r},t) &= 0 \\ \nabla \times \mathbf{E}(\mathbf{r},t) &= -\partial_t \mathbf{B}(\mathbf{r},t) \\ \nabla \times \mathbf{B}(\mathbf{r},t) &= \mu_0 \mathbf{J}_{ext} + \mu_0 \partial_t \mathbf{D}(\mathbf{r},t), \end{aligned} \tag{1.1}$$

where we have introduced three macroscopic fields: the electric field $\mathbf{E}$, the electric displacement $\mathbf{D}$ and the magnetic field $\mathbf{B}$. These fields are linked to an external charge density ρ_{ext} and current density $\mathbf{J}_{ext}$[2]. The macroscopic Maxwell's equations are useful in situations when finding the solution of the (exact) microscopic Maxwell's equations is unfeasible, this encompasses the whole of solid state physics due to the vast number of atoms in a typical material sample. Macroscopic fields are averaged over a region small compared to the system of interest, so that we may take them as continuous fields, but large compared to atomic volumes so that microscopic fluctuations, which are usually of little interest to us, are averaged out (Jackson, 2007).

The macroscopic Maxwell's equations must be complemented by constitutive relations. The displacement and electric field can be linked to the *polarization* of the medium via

$$\mathbf{D}(\mathbf{r},t) = \epsilon_0 \mathbf{E}(\mathbf{r},t) + \mathbf{P}(\mathbf{r},t). \tag{1.2}$$

The polarization is a quantity of fundamental importance to macroscopic electrodynamics as it describes the averaged dipole moment, per unit volume, of a material[3]. As such, it describes the optical response of a material to an external electric field and can be written,

[1] The magnetic field in a material is given by $\mathbf{H} = \frac{\mathbf{B}}{\mu_0} - \mathbf{M}$, where $\mathbf{M}$ is the magnetization. For a non-magnetic material, which we will take to be always the case in this work, we may take $\mathbf{M} = 0$ and equations 1.1 are obtained from the more general macroscopic Maxwell's equations.

[2] Also implicitly contained within the macroscopic Maxwell's equations are the internal charge density ρ_{ind} and current density $\mathbf{J}_{ind}$, which respond to external excitation and lead to the macroscopic fields. The total charge and current density is simply the sum of the external and internal contributions.

[3] The polarization is related to both the internal charge density, $\nabla \cdot \mathbf{P} = -\rho_{ind}$ and current density, $\mathbf{J}_{ind} = \partial_t \mathbf{P}$. Normally one knows the external charge and current distributions, rather than internal, hence why equations 1.1 are written in terms of the electric displacement rather than the polarization.

for linear[4] and isotropic media, in the following conventional linear response formalism

$$\mathbf{P}(\mathbf{r},t) = \epsilon_0 \int d^3r' \int dt'\ \chi(\mathbf{r},\mathbf{r}',t,t')\mathbf{E}(\mathbf{r}',t'), \tag{1.3}$$

which includes nonlocality in *both* space and time. χ is the electric susceptibility and is an example of a *response function*. Linear response theory is ubiquitous in both classical and quantum physics, it allows the expression of out-of-equilibrium properties, induced by a perturbation, in terms of the unperturbed system properties[5]. It is, in almost all cases, permissible to assume that the response function does not depend on absolute coordinates in time[6], so that the two coordinates t and t' may be replaced by the difference $|t-t'|$. Fourier transforming and using the convolution theorem[7] gives a simple local form in frequency-space

$$\mathbf{P}(\mathbf{r},\omega) = \epsilon_0 \int d^3r'\ \chi(\mathbf{r},\mathbf{r}',\omega)\mathbf{E}(\mathbf{r}',\omega), \tag{1.4}$$

i.e. if the system is excited at a frequency ω, it will respond at the same frequency ω. Maxwell's equations scale trivially with frequency but this is not the case for the material response, this frequency dependence is known as *dispersion* and is the reason for the rich and varied form of light-matter interactions over the range of the EM spectrum. This material specific dependence on frequency is explained by the inertial response of the internal charged species to the external field. The most pertinent example for us is the case of metals, where the well-known approximations in the radio and microwave domain breakdown at higher frequencies[8].

With some caution we may repeat the above actions for the spatial coordinates if we assume translational invariance, so that the susceptibility may be written as $\chi(\mathbf{r}-\mathbf{r}',\omega)$, this is tantamount to ignoring the surface of the structure. Making use of the convolution theorem again gives

$$\mathbf{P}(\mathbf{q},\omega) = \epsilon_0\chi(\mathbf{q},\omega)\mathbf{E}(\mathbf{q},\omega), \tag{1.5}$$

a system perturbed at a wavenumber $\mathbf{q}$ will respond at the same wavenumber, this is nothing more than a statement of conservation of momentum; without a surface breaking translational invariance there is no mechanism for momentum transfer. If the wavelength of the incoming light is much larger than the characteristic material dimensions, i.e. the lattice spacing and electron mean free path, then we may approximate the response function as $\delta(\mathbf{r}-\mathbf{r}')\chi(\omega)$; the response of the system at point $\mathbf{r}$ depends only on the field at that same point. This is equivalent to taking the $q \to 0$ limit in momentum-space and gives a simple form for the constitutive equation, used extensively in plasmonics (Maier, 2007)

$$\mathbf{P}(\mathbf{r},\omega) = \epsilon_0\chi(\mathbf{r},\omega)\mathbf{E}(\mathbf{r},\omega). \tag{1.6}$$

[4] Typical applied fields are many orders of magnitude weaker than atomic fields, which are on the order 10^{11} V/m, and so may be taken to be small (Jackson, 2007).

[5] Which are in general known.

[6] This would not be the case if the system depends on its past history, this is known as hysteresis.

[7] The Fourier transform of a convolution equals a simple multiplication in the Fourier domain: $\mathcal{F}\left[\int dt' f(|t-t'|)g(t')\right] = \mathcal{F}[f]\mathcal{F}[g]$.

[8] The failure of conventional microwave/circuit theory and techniques starts in the THz regime.

Typically, this *local* approximation is valid for wavelengths down to the ultraviolet (UV), it is only for nanoparticles smaller than the electron mean free path that we can expect some deviations[9]. It is a valid assumption for the huge majority of nanostructures considered in plasmonics.

[9] A famous deviation is the anomalous skin effect, which occurs for pure conductors at low temperatures when the electron mean free path is greater than the skin depth.

The displacement field can also be linked to the electric field via a linear response equation, which in the Fourier domain reads as

$$\mathbf{D}(\mathbf{q},\omega) = \epsilon_0 \epsilon(\mathbf{q},\omega)\mathbf{E}(,\omega). \tag{1.7}$$

We have introduced another response function called the *dielectric function* ϵ, which is also known as the permittivity[10]. It not only allows us to characterise the response of materials in a simple compact fashion, it also enables us to relate measurable spectroscopic quantities such as scattering, absorption, reflection, and ellipsometry with properties of the internal charge carriers, such as the electronic band structure. The permittivity thus acts as a convenient link between experiment and the microscopic world. It is, in general, a complex quantity and is related to the refractive index via $n(\omega) = \sqrt{\epsilon(\omega)}$, for non-magnetic materials. Writing this out explicitly gives (Fox, 2002; Maier, 2007)

[10] More precisely the relative permittivity, which is the preferred quantity to work with, opposed to the absolute permittivity $\epsilon\epsilon_0$, as it is easier to attach physical meaning to its value, i.e. we measure it relative to the vacuum value of 1.

Another common response function is the conductivity σ, which is related to the dielectric function via $\epsilon(k,\omega) = 1 + \frac{i\sigma(k,\omega)}{\epsilon_0\omega}$ (Maier, 2007). The response function of choice will depend on personal taste, the field of research, and the frequencies of interest. In plasmonics it is the dielectric function that rules supreme in the literature.

$$\begin{aligned}
\Re[\epsilon(\omega)] &= \Re[n(\omega)]^2 - \Im[n(\omega)]^2 \\
\Im[\epsilon(\omega)] &= 2\Re[n(\omega)]\Im[n(\omega)] \\
\Re[n(\omega)] &= \sqrt{\frac{|\epsilon(\omega)| + \Re[\epsilon(\omega)]}{2}} \\
\Im[n(\omega)] &= \sqrt{\frac{|\epsilon(\omega)| - \Re[\epsilon(\omega)]}{2}} = \frac{\Im[\epsilon(\omega)]}{2\Re[n(\omega)]},
\end{aligned} \tag{1.8}$$

where $|\epsilon(\omega)|$ denotes the complex modulus of the permittivity. The imaginary part of the refractive index is known as the extinction coefficient and is related to the absorption coefficient of the famous Beer's law[11]; consequently, the imaginary part of the permittivity determines the absorption of an EM wave propagating inside a material. Materials excited at frequencies with a large $\Im[\epsilon(\omega)]$ are described as 'lossy' and will have a large non-radiative decay channel into heat energy, this is called Joule heating. For materials at frequencies with $\Re[\epsilon(\omega)] >> \Im[\epsilon(\omega)]$, which corresponds to a low-loss dielectric, the refractive index is a real number that describes the reduced phase velocity of an EM wave propagating in the medium.

[11] Beer's law states that a planewave propagating in a material with loss will suffer an intensity attenuation of the form $I(x) = I_0 e^{-\alpha x}$, where the absorption coefficient is given by $\alpha(\omega) = \frac{2\omega\Im[n(\omega)]}{c}$ (Fox, 2002). The factor of two comes from the squaring of the electric field to get the intensity.

Having established some material response functions, it will be useful to link these to a response function specific for a localised set of charges such as an atom or nanoparticle. For the response to light, the polarizability is a suitable function for this purpose. It measures the willingness of a system to have its positive and negative charges displaced relative to one another by an external electric field, i.e. the ease of polarizing the object along a certain direction. In general, the polarizability is a rank two tensor; only

in special cases, such as for bodies with spherical symmetry, can it be reduced to a scalar function. It is defined as follows in Cartesian coordinates

$$\mathbf{p} = \bar{\alpha}\mathbf{E} \rightarrow \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix} = \begin{bmatrix} \alpha_{xx} & \alpha_{xy} & \alpha_{xz} \\ \alpha_{yx} & \alpha_{yy} & \alpha_{yz} \\ \alpha_{zx} & \alpha_{zy} & \alpha_{zz} \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix}, \tag{1.9}$$

where $\mathbf{p}$ is the induced dipole moment and the element α_{ij} describes the polarization response in the jth direction from the ith electric field component.

Next, we discuss propagating wave solutions to Maxwell's equations. Simply by combining Faraday's and Ampere's laws in equations 1.1, we can derive the wave equation, which in the Fourier domain can be written as

$$\mathbf{k}(\mathbf{k} \cdot \mathbf{E}) - k^2\mathbf{E} = -\epsilon(k, \omega)\frac{\omega^2}{c^2}\mathbf{E}. \tag{1.10}$$

It turns out we can learn a lot without solving the wave equation explicitly. There are two types of possible excitations: *longitudinal*, with the polarization vector in the same direction as the propagation vector, and *transverse*, with the two vectors perpendicular[12]. For longitudinal waves, the only way equation 1.10 can hold is if

$$\epsilon(k, \omega) = 0, \tag{1.11}$$

which tells us that longitudinal excitations can only occur for zeros of the dielectric function[13]. The particular wavenumbers and frequencies for which this is satisfied will give the dispersion of the excitation. In this case, the displacement field must be zero and the electric field is a pure depolarization field (Maier, 2007). For transverse waves, the first term in equation 1.10 is zero and the following dispersion relation is obtained

$$k^2 = \epsilon(k, \omega)\frac{\omega^2}{c^2}. \tag{1.12}$$

Note that if the permittivity is negative then the wavevector will become complex, indicating attenuation. As an illustration of this, in figure 1.1 the EM wave propagation in bulk silver and gold over the wavelength range 400 $\rightarrow$ 1800 nm is shown. In this spectral region, the real part of the permittivity for both metals is less than zero and monotonically approaches $-\infty$ for increasing wavelength[14]. This leads to a compression of the field oscillations in space for larger wavelength, as can be nicely seen in figure 1.1(a) for the case of silver. Additionally, the imaginary part of silver's permittivity is (roughly) monotonically increasing, so the attenuation smoothly increases for longer wavelengths. The permittivity is slightly more complicated for gold, see figure 1.1(b). Due to the presence of interband transitions, the imaginary part increases below 600 nm; this is reflected by a large increase in

[12] For a visual representation of the two types of excitations in an electron gas, see figure 1.3.

[13] A trivial consequence of this relation is that propagating waves in vacuum are purely transverse, longitudinal fields rely on the presence of charge.

[14] See the next subsection for more details.

absorption in this region, i.e. gold becomes very absorbing. At longer wavelengths, where Drude behaviour dominates, it shows the same overall trend as silver.

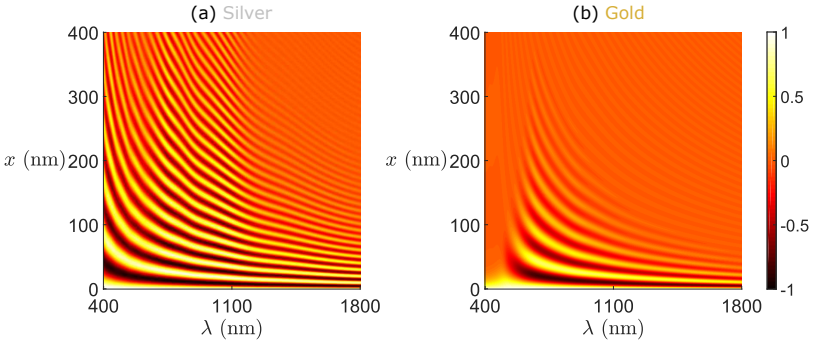


Figure 1.1: The electric field of a propagating wave, $e^{ik_x x}$, with $k_x = \frac{2\pi\sqrt{\epsilon(\omega)}}{\lambda}$, in (a) silver and (b) gold. In both cases the dielectric function from (Johnson and Christy, 1972) is used.

The results in figure 1.1 are our first example of a hybrid light-matter excitation, it is incredible, given the simplicity, that this formulation is capable of accurately describing the propagation of light in the complicated arrangement of charges present in matter. For instance, it is truly remarkable that for a planewave propagating in a transparent linear medium, we need only multiply $\mu_0\epsilon_0$ by a factor $\epsilon\mu$ and solve the usual wave equation, despite all the complicated absorption and emission from the polarized atoms in the material. As noted by (Griffiths, 2008), there is an extraordinary 'conspiracy' and the resulting emitted waves combine with the original field in such a way to create a single wave at the exact same frequency but with a reduced velocity[15].

[15] There is a fantastic paper by (James and Griffiths, 1992) using a perturbation expansion of the field to track how this occurs.

1.1.1 *Optical Properties Of Metals*

Now we will develop the macroscopic theory for the interesting case of metals. The optical properties of metals are dominated by the presence of *free carriers*; in this case, the conduction electrons in the sp bands. The free electrons move against the background of fixed positively charged ion cores, this can be considered a *plasma*. The plasma model has proved to be remarkably successful in explaining the optical response of metals over a wide range of the spectrum; for noble metals, it is valid up to visible frequencies before interband transitions occur. For alkali metals, the range of validity extends to the UV (Maier, 2007). The recipe for deriving any dielectric function is to link the microscopic (either classical or quantum) response to the external perturbation using the macroscopic polarization. To derive the Drude dielectric function, which is the long wavelength ($q \to 0$) response of a free electron gas, we make some rather drastic approximations and assumptions:

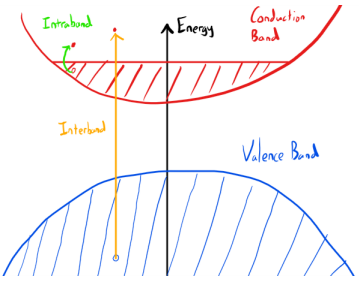


Illustration of an interband (orange) and intraband (green) transition in a metal.

- We describe the microscopic response of the electrons using a classical equation of motion.
- Electrons suffer abrupt collisions with a frequency $\gamma = \frac{1}{\tau}$, where

τ is the relaxation time of the electrons[16].

- The electron-ion interactions are neglected, i.e. band structure is ignored[17].
- The electron-electron interactions are ignored.

The last two approximations mean that an electron, in the absence of an externally applied field and between collisions, will move at a constant velocity. When an external field is applied we use the Newtonian laws of motion to describe the dynamics.

In the interest of generality, let us assume the material contains N sets of electrons[18], which are bound by different potentials with resonant frequencies ω_i and with damping coefficients γ_i. This is a simple model that can represent both intraband and interband transitions. Away from any of the frequencies ω_i, the optical response will be that of a free electron gas. The classical equations of motion are

$$m\ddot{\mathbf{r}}_i(t) + m\gamma_i\dot{\mathbf{r}}_i(t) + m\omega_i^2\mathbf{r}_i(t) = -e\mathfrak{f}_i\mathbf{E}(\mathbf{r},t), \tag{1.13}$$

where m and e are the electron mass and charge respectively, and $\mathfrak{f}_i$ are the (dimensionless) *oscillator strengths*[19] for each transition and quantify the coupling to the external field. We take index 0 to denote the equation of motion for the valence electrons, so $\gamma_0 = \gamma$, $\omega_0 = 0$ and $\mathfrak{f}_0 = 1$. The polarization in the metal is caused by a displacement of all the electrons, so we sum up the different contributions, $\mathbf{P}_i = -en_i\mathbf{r}_i$, where n_i is the concentration of the ith species. The resulting differential equation in time for the polarization[20] is simple to solve by Fourier transforming into the frequency domain, from which the electric displacement, using equation 1.2, can be written as

$$\begin{aligned}\mathbf{D}(\mathbf{r},\omega) &= \epsilon_0\mathbf{E}(\mathbf{r},\omega) + \sum_{i=0}^{N}\mathbf{P}_i(\mathbf{r},\omega)\\ &= \left(\epsilon_0 + \sum_{i=0}^{N}\frac{n_ie^2}{m}\frac{\mathfrak{f}_i}{\omega_i^2-\omega^2-i\omega\gamma_i}\right)\mathbf{E}(\mathbf{r},\omega),\end{aligned} \tag{1.14}$$

giving us our final result: the *Drude-Lorentz* dielectric function

$$\epsilon(\omega) = \underbrace{1 - \frac{\omega_P^2}{\omega^2+i\omega\gamma}}_{Drude} + \underbrace{\sum_{i=1}^{N}\frac{\mathfrak{f}_i\Omega_i^2}{(\omega_i^2-\omega^2)-i\omega\gamma}}_{Lorentz\ Terms}. \tag{1.15}$$

Here I have defined the *plasma frequency*

$$\omega_P = \sqrt{\frac{n_0e^2}{\epsilon_0 m}}, \tag{1.16}$$

as well as the generalised plasma frequencies for the Lorentzian terms, $\Omega_i = \sqrt{\frac{n_ie^2}{\epsilon_0 m}}$. This is a very general result and can be used to model a wide range of different materials. In fact, the transitions need not even be electronic in nature; transitions linked to

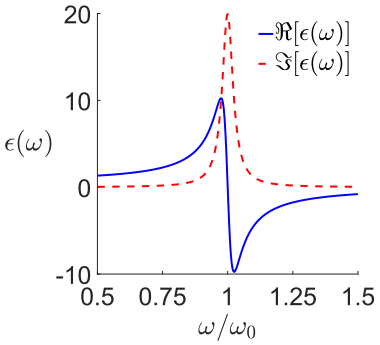
[16] The usual tacit approximation is that the relaxation time is frequency independent. Actually, electron-electron interaction leads to a frequency dependence; for instance, Fermi liquid theory predicts a quadratic dependence on frequency (Yang et al., 2015). At this point, the mechanisms responsible for collisions are not important in formulating the dielectric function and so we may take the collision time to be a phenomenological value. In fact, in his original work Drude assumed the electrons bounced from ion to ion, which is very far from reality (Ashcroft and Mermin, 1976).

[17] Except possibly via an effective optical mass of the electrons (Maier, 2007).

[18] Or more generally sets of charged species, for simplicity we will consider only electrons.

[19] Within classical models, the oscillator strength must be assigned a phenomenological value. In quantum theory, it is possible to calculate them from knowledge of the state transitions and will be proportional to the dipole matrix element, see section 2.3.4.

[20] This is a very convenient form of the response equation for inputting material data into finite difference time domain (FDTD) codes.



Plot of the real and imaginary part of $\epsilon(\omega) = \frac{1}{(\omega_0^2-\omega^2-i\omega\gamma)}$, with $\gamma = \omega_0/20$. There is a strong absorption as the resonant frequency ω_0.

vibrational degrees of freedom in the mid-IR are well described by Lorentz dielectric functions, see section 1.4.2. Commonly in plasmonics one is interested in frequencies well below interband transitions, i.e $\{\omega_i >> \omega\}$. In this case, the Lorentz part of the dielectric function can be taken to be a frequency independent real number allowing us to write

$$\epsilon(\omega) = \epsilon_\infty - \frac{\omega_P^2}{\omega^2 + i\omega\gamma}, \tag{1.17}$$

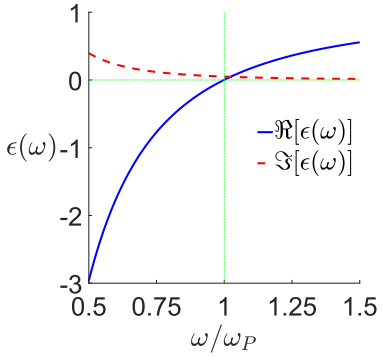
where $\epsilon_\infty = 1 + \sum_{i=1}^{N} \frac{f_i \Omega_i^2}{\omega_i^2}$. This allows us to model the free electron dynamics via a Drude formula and take into account the residual polarization of the ionic core background at high frequencies, ϵ_∞ typically takes a value between 1 to 10 and for noble metals arises from polarization of the d-band electrons (Maier, 2007).

The Drude-Lorentz equation is used heavily in plasmonics (Rakić et al., 1998; Vial et al., 2005); in particular, because of the ease of incorporation into FDTD simulations[21], and for simple analytical models. The parameters are most commonly fitted to experimental data[22] rather than obtained from first principles[23], for methods other than FDTD it is normally simpler to just interpolate experimental data. It should be noted that it often takes a large number of Lorentzians to achieve a satisfactory fit, in which case the physical meaning of the terms becomes somewhat unclear. There are more complicated variations that can achieve a superior fit with fewer terms (Etchegoin et al., 2006).

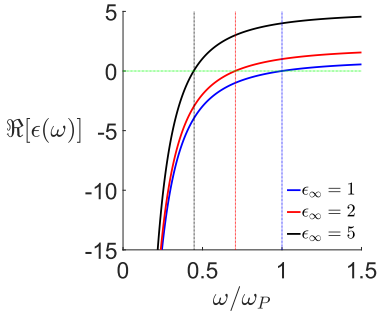
Now, for simplicity, let us consider the extreme limit of a completely free electron gas (so all $\Omega_i = 0$ and $\epsilon_\infty = 1$), which captures the main characteristics of metallic optical properties; we will discover it is the Drude term that gives us plasmonic behaviour. We can observe the following trends with respect to frequency, which reveals a wealth of interesting physics:

- For high frequencies, $\omega >> \gamma$, the permittivity becomes purely real, $\epsilon(\omega) \approx 1 - \frac{\omega_P^2}{\omega^2}$. This is a good approximation near the plasma frequency as it is typically around a hundred times larger than the collision frequency for noble metals. Physically, this means there is no time in an oscillation for the electrons to interact with their environment to exchange energy, hence the damping disappears.

- For $\omega < \omega_P$, the real part of the dielectric function will be negative, revealing that the electrons do not oscillate in phase with the external field as a result of their finite inertia. This leads to a large imaginary part of the refractive index, which means planewaves can only penetrate a finite length into the plasma. This is the reason for the large reflectivity ($R = |\frac{n-1}{n+1}|^2 \sim 100\%$) of metal surfaces. Silver and aluminium are good reflectors in the visible and hence are good materials for making mirrors, gold and copper are more suitable as IR mirrors due to devia-



Plot of the real and imaginary part of $\epsilon(\omega) = 1 - \frac{\omega_P^2}{\omega^2 + i\omega\gamma}$, with $\gamma = \omega_P/20$. The green dotted line indicates the real part of the permittivity crossing 0, which corresponds to the plasma frequency.



Plot of the real part of $\epsilon(\omega) = \epsilon_\infty - \frac{\omega_P^2}{\omega^2 + i\omega\gamma}$, with $\gamma = \omega_P/20$, for different values of ϵ_∞. The real part of the permittivity crosses 0 at $\omega_P/\sqrt{\epsilon_\infty}$.

[21] Which, in contrast to frequency domain methods, cannot use tabulated data. See section 1.3.

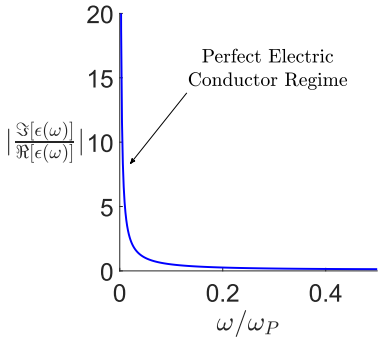
[22] See appendix A for a short discussion on the experimental data used in plasmonics.

[23] Although, for instance, such studies have been performed using DFT for a number of different metals and compare well to experimental data (Werner et al., 2009).

tions from Drude behaviour in the visible. A 'good' optical metal is characterised by $|\Re[\epsilon]| >> |\Im[\epsilon]|$ i.e. a small penetration depth and low ohmic losses.

- For $\omega > \omega_P$, the real part of the dielectric function will be positive and the metal will behave as a dielectric for large enough ω, this is famous UV transparency of metals. This switch from high reflectivity to transparency with decreasing wavelength is illustrated for gold and silver in figure 1.1. At these high frequencies, the finite inertia of the electrons means they cannot keep up with the rapidly varying field and consequently fail to screen it out.

- In the limit of $\omega \sim 0$, the metal completely screens out the external field. This is the realm of the *perfect electric conductor* approximation and is heavily used in radio and microwave engineering, which underpins much of modern technology. In this region $|\Re[\epsilon]| << |\Im[\epsilon]|$. As $\sigma(\omega) \propto \Im[\epsilon(\omega)]$ in this limit, this corresponds to the conductivity going to ∞; the well-known definition of a perfect conductor (Griffiths, 2008).



A plot showing the ratio of the imaginary and real parts of the Drude permittivity at low frequencies, using $\gamma = \omega_P/20$.

In figure 1.2(a) a plot of the Drude dielectric function of silver around the plasma frequency is shown. The corresponding refractive index is presented in figure 1.2(b), which is obtained using equations 1.8. The behaviour of the response functions around the plasma frequency manifests as a very sharp change in reflectivity in this spectral region, as shown in 1.2(c). This is characteristic of a Drude metal; it is experimentally observable and is strong evidence of the validity of the Drude model for noble metals. In figure 1.2(d) the dispersion for a transverse wave propagating in bulk silver is shown, this can be found by subbing the Drude formula (equation 1.17) into equation 1.12

$$\omega = \sqrt{\frac{\omega_P^2 + c^2k^2}{\epsilon_\infty}}. \tag{1.18}$$

This is nothing more than an alternative representation of the results we found in figure 1.1, but now as a dispersion relation rather than in real-space. Below $\omega_P/\sqrt{\epsilon_\infty}$ there is a forbidden frequency gap and propagation of EM waves in the plasma is not possible. We can also see some typical behaviour of hybrid light-matter states; away from the energy crossing point, the mode becomes more light- or matter-like. For low wavenumbers, the mode is matter-like and approaches the bulk plasmon energy. In contrast, for large wavenumbers, the polariton becomes light-like and approaches the dispersion of a free planewave (the red dashed line).

I have always considered it extraordinary that the Drude-Lorentz formula can work so well. Here I offer some thoughts on why it is so successful: the key idea is that while I have presented the Drude-Lorentz formula as a crude physical model, we can alternately derive it as a mathematical consequence of some general constraints

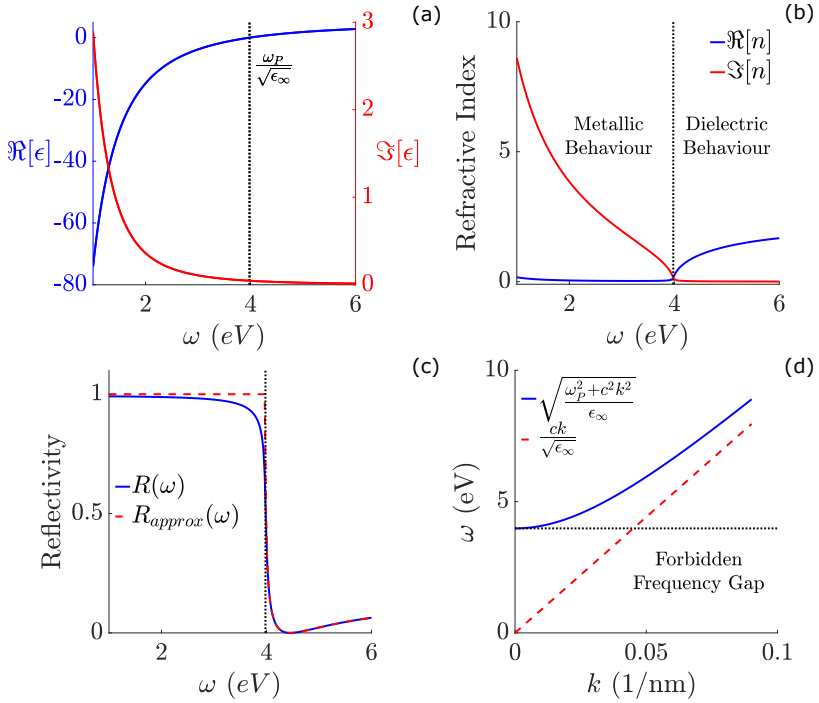


Figure 1.2: (a) Drude dielectric function for silver with parameters $\omega_P = 8.9$ eV, $\tau = 17$ fs and $\epsilon_\infty = 5$. (b) The corresponding refractive index calculated using equations 1.8. (c) The corresponding reflectivity for a silver surface. (d) The dispersion relation for a polariton in bulk silver. The Drude fit is taken from (Yang et al., 2015).

on response functions for them to be physically reasonable. The principle of causality is sacrosanct in physics, any good and honest response function must obey it, we can state it as:

THE CAUSE ALWAYS PRECEDES THE EFFECT!

This temporal constraint must have consequences in the frequency domain. We can write the following causal relation between the electric displacement and the electric field (Jackson, 2007)

$$\mathbf{D}(\mathbf{r},t) = \epsilon_0 \mathbf{E}(\mathbf{r},t) - \epsilon_0 \int_0^\infty \chi(\tau)\mathbf{E}(\mathbf{r},t-\tau)d\tau, \tag{1.19}$$

which is linked to the permittivity via $\epsilon(\omega)/\epsilon_0 = 1 + \chi(\omega) = 1 - \int_0^\infty \chi(\tau)e^{i\omega\tau}d\tau$. τ denotes the difference in time between the system response and the perturbation that caused it, hence causality demands that the response function is zero for $\tau < 0$. This is enforced via the integration limits in the above equations. Expressing the frequency as a complex variable means that the complex exponential can be written in the upper half plane as $e^{i\Re[\omega]\tau - \Im[\omega]\tau}$, the decaying exponential term, due to $\Im[\omega] > 0$, means that the integral must converge for $\tau > 0$ and hence there cannot be any poles in the upper half plane[24]; causality demands that the susceptibility, $\chi(\omega)$, and hence the permittivity, $\epsilon(\omega)$, is an analytic function[25] in the upper half-plane. This conditions directly leads to the beautiful Kramers-Kronig relations, which link the real and imaginary parts

[24] And that $\chi^*(\omega) = \chi(-\omega^*)$.

[25] A function is analytic, at a given point in the complex plane, if it is single-valued and differentiable in the local region.

of the permittivity (Arfken, 1985)

$$\begin{aligned}\Re[\epsilon(\omega)] &= 1 + \frac{2}{\pi}\mathcal{P}\left[\int_0^\infty \frac{\omega'\Im[\epsilon(\omega')]}{\omega'^2-\omega^2}d\omega'\right] \\ \Im[\epsilon(\omega)] &= -\frac{2\omega}{\pi}\mathcal{P}\left[\int_0^\infty \frac{\Re[\epsilon(\omega')]-1}{\omega'^2-\omega^2}d\omega'\right],\end{aligned} \tag{1.20}$$

where $\mathcal{P}$ denotes the principal part of the function. From these considerations it is possible to derive two insightful limits:

- A narrow absorption line at $\omega = \omega_0$ is described by $\Im[\epsilon] \propto \delta(\omega - \omega_0)$ and must have a real part described by $\Re[\epsilon] - \epsilon_{\text{inf}} \propto \frac{1}{\omega_0^2-\omega^2}$.
- The response function $\chi(\tau)$ can be expanded for short time differences, i.e. around $\tau = 0$. This links to high frequency behaviour and gives $\Re[\epsilon - 1] \propto \frac{1}{\omega^2}$. Thus, we can always split off a high frequency part of a response function of the form $\epsilon(\omega) = 1 - \frac{\omega_p^2}{\omega^2}$. In fact, this can be seen as the general definition of the plasma frequency[26].

[26] This is why one sometimes will see discussions about the plasma frequency of non-conductive materials.

Thus, as long as each absorption contribution of a material is defined in a narrow range, which we know to be often the case, we may write any response function as a collection of Lorentz-Drude terms as we did in equation 1.15. Furthermore, the form of the integrand in the Kramers-Kronig relations means that, while technically all frequencies contribute, it is permissible to take only a small frequency range of the imaginary part to determine the real part of the permittivity; this is what allows us to neatly divide the optical response of a material into different physical contributions.

1.2 *The Plasmon Family*

HAVING ESTABLISHED THE MACROSCOPIC DIELECTRIC THEORY FOR METALS, WE NOW LOOK FOR PARTICULAR SOLUTIONS TO MAXWELL'S EQUATIONS NEAR THE PLASMA FREQUENCY WHICH CORRESPOND TO COLLECTIVE OSCILLATIONS OF THE ELECTRON GAS.

The analysis of the Drude function in the last section revealed that the optical response of metals is hugely frequency dependent, with the natural scale determined by the plasma frequency; plasmonics is the study of metals near the plasma frequency. More generally we may state:

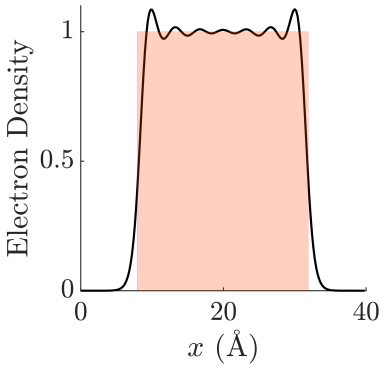
PLASMONICS IS THE STUDY OF THE OPTICAL RESPONSE OF MATERIALS WITH FREE CARRIERS IN FREQUENCY REGIMES WHERE $\Re[\epsilon] < 0$ AND $|\Re[\epsilon]| > |\Im[\epsilon]|$.

We begin with a brief discussion of the *volume plasmon* (also known as a bulk plasmon). On the whole, we are more interested in surface plasmons (SPs), but an understanding of volume plasmons will be of use and, in fact, for small nanoparticles, it is possible to excite both surface and volume plasmons with light. We will then discuss both propagating and localised SPs, which are supported on metal extended surfaces and nanoparticles respectively.

The classical theory of plasmonics relies on two cornerstone assumptions:

- The first is sharp boundary conditions, this corresponds to approximating the electron density as a step function. Electron spill-out is on the order of a few Å so in the vast amount of cases, it is safe to ignore it. Only for very small and nearly touching nanoparticles will it be important to include.
- The second key assumption is that the permittivity is a local response function, this physically corresponds to treating the electrons as point-like particles and means that the surface screening charge is precisely localised to an infinitesimal layer on interfaces. In reality, electrons have a finite spatial extent due to their wavelike nature, this wavelength will be on the order of the Fermi wavelength. This means in nonlocal response theories that the screening charge is smeared out over a length on this scale. Again this can be ignored as long as the systems characteristic sizes are large enough.

As a rough rule of thumb, quantum effects can be ignored for particles sizes above $\sim 10 \to 20$ nm and gap sizes above $\sim 2 \to 5$ nm. In a classical picture, plasmons, despite the fancy name, are simply solutions to Maxwell's equations for particular boundary conditions. Using the assumptions listed above, we can write



Electron density of a sodium slab (of thickness 24 Å and $r_s = 3.93$, where r_s is the Wigner-Seitz cell radius), calculated using DFT within the local density approximation (LDA) (Perdew and Wang, 1992). The coloured box indicates the jellium background and the electron density is normalised relative to this density. Note the electron spill-out at the slab edges, it is on the order of a few Å.

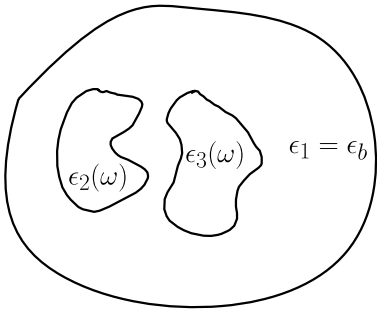


Illustration of a typical plasmonic calculational domain, within each closed region the dielectric function is taken to be spatially homogeneous. The Helmholtz equation is solved in each of the regions 1, 2 and 3 and the solutions must match at the boundaries according to the standard EM boundary conditions discussed in appendix B.

the domain of interest as a piecewise construction of different *homogeneous* dielectric regions, $\{\epsilon_i(\omega)\}$. In each region, we solve the Helmholtz equation

$$\nabla^2\mathbf{E}(\mathbf{r},\omega) - k^2\epsilon_i(\omega)\mathbf{E}(\mathbf{r},\omega) = 0, \tag{1.21}$$

which is a simplification of the wave equation in real-space if the spatial variation of the permittivity can be neglected.

1.2.1 *The Volume Plasmon*

A bulk plasma oscillation is a longitudinal excitation, which consists of the coherent motion of an electron gas; all the electrons move with the same wavenumber and frequency. The energy quanta are called volume plasmons. We have already seen that the plasma frequency of an electron gas is a special frequency that is a natural scale to define the optical response, so it is no surprise to find it corresponds to the frequency of the volume plasmon in the long wavelength (and low damping) limit. This can be seen by subbing in the Drude function (with $\gamma = 0$ and for simplicity we set $\epsilon_\infty = 1$) into the condition for a longitudinal excitation, $\epsilon(\omega) = 0$, which is satisfied for one frequency point at $\omega = \omega_P$[27]. Away from $k = 0$, the local approximation breaks down and the dispersion deviates from a flat line in ω-k space, this can be analysed using the Lindhard dielectric function (Giuliani and Vignale, 2005). Typical values for the plasma frequency in metals is around 10 eV. This energy is large, relative to the eh excitation energy, and is crucial for the stability of the plasmon[28]. It is simply a result of the large number of electrons that make up the plasmon (Pines, 1964). It is also large compared to other collective excitations such as phonons, which are typically on the order of ~ 10 meV, as well as the thermal energy, k_BT, at room temperature, meaning we can take the zero-temperature limit. The early pioneering work on collective modes in degenerate electron gases was done by Bohm and Pines in the 1950s. Using the random phase approximation (RPA), they found, depending on the wavenumber relative to a cutoff wavenumber k_c, a quantum plasma may behave collectively for long wavelengths or as a system of free particles at shorter wavelengths. The Coulomb interaction means each electron experiences an interaction with a large number of other electrons, resulting in a perturbation of its position and momentum. These developments were crucial for advancements in many-body physics, where the long-range nature of the electron-electron interaction led to apparent divergences. The plasmon, along with the related concept of screening, allowed the division of the many-body problem into two (approximately) separate simpler problems of collective excitations[29] and individual electrons interacting via a screened Coulomb potential. This means, via a suitable canonical transformation of the many-body Hamiltonian, that the divergences can be dealt with in an elegant fashion

[27] That the low k energy is finite is a non-trivial consequence of the long-range Coulomb interaction. In contrast, systems such as Boson liquids and superfluid ^{3}He (which is a neutral Fermi liquid) have a zero energy mode at $k = 0$. This type of mode is known as a "Goldstone boson" (Anderson, 1984).

[28] At low wavenumbers the high energy of the plasmon means it lays outside of the eh continuum and so decay into eh pairs is forbidden by momentum conservation. Above a critical wavenumber, the plasmon dispersion enters the eh continuum and the volume plasmon ceases to be a well-defined mode (Giuliani and Vignale, 2005).

[29] Which can be modelled using a simple harmonic oscillator for a collective coordinate.

(Pines, 1964).

In the small k limit, we can build an intuitive physical picture of the volume plasmon as it must be uniform in real-space. We consider the electron cloud immersed in a smeared out positive background[30] and rigidly displaced by a distance x in one direction. This will result in a layer of uniform charge on each surface[31] of opposite sign. The polarization of a uniform charge layer, of thickness x, is $P = \pm nex$, which is associated with an electric field that acts as a restoring force on the electron cloud. This behaviour is governed by the following differential equation in time

$$\frac{d^2\mathbf{x}(t)}{dt^2} = -\frac{e}{m}\mathbf{E}(t) = -\omega_P^2\mathbf{x}(t), \tag{1.22}$$

which is nothing more than simple harmonic motion and has the solution of oscillatory motion with frequency ω_P. Thus, the volume plasmon can be considered as a 'sloshing' of the electron gas in real-space at the plasma frequency.

As the volume plasmon is a longitudinal excitation it cannot be excited by standard planewave illumination; instead, it must be excited by an EM field with a longitudinal component. To give some insight, an illustration of longitudinal and transverse density perturbations is presented in figure 1.3. An ideal tool for detection is electron energy loss spectroscopy (EELS), as the electric near-field of the travelling electron has a longitudinal component. Typically these electrons will be fast and have energies on the order of a hundred KeV[32]. The energy loss can be measured in either reflection or transmission experiments. For metals where interband transitions are negligible around ω_P, reflection experiments will show distinct peaks corresponding to the bulk plasmons and lower energy SP modes. Transmission experiments have the benefit of allowing information on the momentum transfer to be deduced by the angular distribution of the inelastically scattered electrons, this allows the plasmon dispersion to be mapped out (Nozières and Pines, 1999). For metals in which the effect of the background ionic lattice is expected to be small, the agreement between theory and experiment is good; the free electron value for the plasma frequency typically agrees within 10% with the experimental value (see table 4.2 in (Nozières and Pines, 1999)), and the RPA qualitatively predicts the plasmon dispersion correctly[33].

In extremely small nanoparticles ($\lesssim 10$ nm) it becomes possible to excite volume plasmons using light, this is a consequence of the large curvature of the surface, which is proportional to $1/R$, giving the necessary momentum transfer to break the usual selection rules. Therefore, we can expect the spectroscopy of ultra-small nanoparticles to be much richer than that of larger particles, which are dominated by the dipolar surface mode. Unfortunately dealing with small nanoparticles is difficult, they are tough to fabricate and the scattering cross section has a strong dependence on size (for a sphere it scales with R^6), meaning to pick out a small object

[30] To mimic the ionic background and keep the model electrically neutral, this is the famous jellium approximation

[31] As we are considering a bulk system the surface is to be taken to infinity.

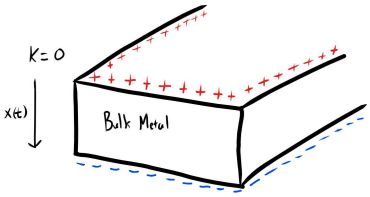
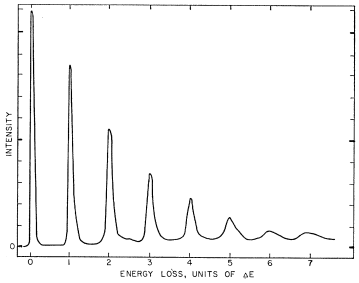


Illustration of a volume plasmon oscillation in the long wavelength limit.

[32] The kinetic energy of an electron, including relativistic effects, is the total energy minus the mass-energy, $E_{ke} = mc^2\left(\left(1-\frac{v^2}{c^2}\right)^{1/2}-1\right)$. For instance, an electron with energy 80 keV will have a velocity of 0.5c.

[33] Although there are interesting exceptions; for instance, the heavier alkali metals surprisingly do not behave free electron like (Baltz, 1997).



A typical transmission EELS spectrum for an aluminium slab. It clearly shows the quantisation of volume plasmons at integer multiplies of $\hbar\omega_P$. Reprinted figure with permission from (Marton et al., 1962), copyright 1962 by the American Physical Society.

from its background can be challenging (Maier, 2007). The 'cluster science approach'[34] to this problem is to create an ensemble of nanoparticles either embedded in a suitable matrix (such as glass (Kreibig, 1974)) or in a gas phase via high-pressure vapour flow from which the desired size can be selected via a mass spectrometer[35]. A nice recent example of such methods is given by Xia *et al*, who detected the presence of volume plasmons in Na_{20} and Na_{92} clusters via the photoabsorption spectra (Xia et al., 2009). Another example is provided by the measured photoionization cross section of C_{60} (Buckminsterfullerene) ions, which show a volume plasmon near 40 eV (Scully et al., 2005), although this interpretation has been challenged (Korol and Solov'Yov, 2007). Of course, to theoretically describe such tiny systems we will need to employ the machinery of quantum mechanics, so we defer further discussion of clusters to chapter 2.

[34] By this I mean the sort of typical approaches used in cluster science from the 1960s onwards (De Heer, 1993).

[35] This approach allows the study of clusters all the way down to the molecular level (Robbins et al., 1967; De Heer, 1993). Unfortunately, there is a size distribution that can 'wash-out' quantum effects, plus it is hard to control environmental effects which can be difficult to incorporate in theoretical models. It is partly for these reasons that there was so much excitement about more recent quantum plasmonic experiments analysing *individual* nanoparticles (Scholl et al., 2012), see section 2.1.

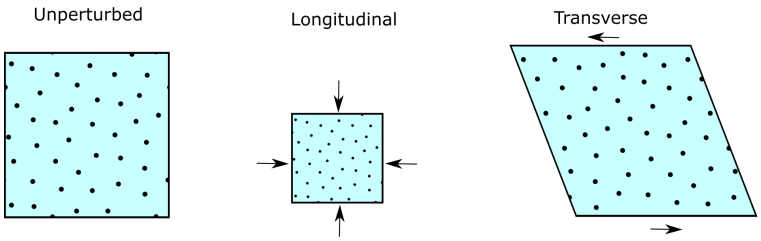


Figure 1.3: Illustration of the two types of density excitations possible in an electron gas, the density is indicated by the black dots. The middle figure shows a compression of the gas, which leads to a closer spacing of the black dots and hence a higher density but does not add any rotation to the gas; mathematically this can be described by the pair of equations $\nabla \cdot \mathbf{J}(\mathbf{r},t) = -\partial_t n(\mathbf{r},t) \neq 0$ and $\nabla \times \mathbf{J}(\mathbf{r},t) = 0$. In contrast, the right hand figure shows a shear stress to the gas which imparts a rotation $\nabla \times \mathbf{J}(\mathbf{r},t) \neq 0$, but there is no change in density, $\nabla \cdot \mathbf{J}(\mathbf{r},t) = 0$.

1.2.2 *Surface Plasmon Polaritons*

Soon after the discovery of the volume plasmon, Ritchie predicted that the interface between metal-dielectric half-spaces support lower energy collective modes (Ritchie, 1957). This was experimentally confirmed[36] using the reflection of fast electrons by Powell and Swan for Aluminium (Powell and Swan, 1959a) and Magnesium (Powell and Swan, 1959b), the new type of mode was given the name: surface plasmon (SP) (Stern and Ferrell, 1960). The present 'boom' of interest in plasmonics was ignited by an experimental work in 1998 by Ebbesen *et al* who observed that, under certain conditions, light can have a transmission efficiency larger than unity through subwavelength holes milled in metal films (Ebbesen et al., 1998). This was explained via the coupling of light to SPs, and a subsequent theoretical work in 2001 explored this in detail (Martin-Moreno et al., 2001).

In figure 1.4 an illustration of a surface plasmon polariton (SPP) on a flat surface is shown, we will take the surface to be in the xy plane. The presence of a surface breaks translational invariance and allows the plasmon wavevector component normal to the surface to be imaginary, this corresponds to a mode bound to the surface[37]. This means each component of the electric and magnetic field must

[36] Actually, the first experimental evidence of SPs seems to have been by Wood in 1902 who reported the anomalous behaviour of diffracted light from metallic diffraction gratings (Wood, 1902). At least some of these results can be explained by light coupling to SPs (Barnes et al., 2003).

[37] No power (or at least very little) is radiated to the far-field perpendicular to the surface, the mode is near-field in nature

have a factor $e^{-\operatorname{sgn}(z)k_{z,i}z}$, where $\operatorname{sgn}(z)$ is the sign function and is -1 for $z < 0$ and $+1$ for $z > 0$. The index i denotes the half-space, the interface is located at $z = 0$. Parallel to the surface the wave is propagating and can transport energy, the EM components must have a planewave form, $\propto e^{ik_x x}$, we take the propagation direction to be along the x direction for simplicity. This type of surface wave can only be supported by transverse magnetic (TM) modes[38], as there clearly needs to be a z component for the electric field. Crucially, the SP has a transverse character, this means it can couple to light, with a suitable structuring of the surface to overcome the momentum mismatch, and the resulting non-radiative and propagating hybrid mode is called an SPP.

[38] The magnetic field points purely in the y direction.

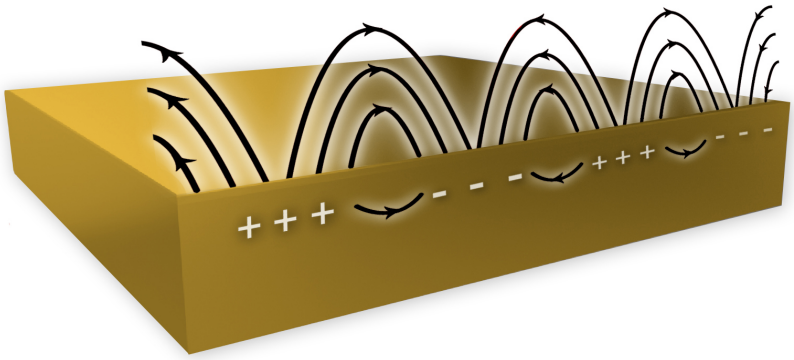


Figure 1.4: An illustration of an SPP on a metal-dielectric surface.

By considering a field of the form hypothesised above and subbing into Maxwell's equations, application of the standard macroscopic boundary conditions[39] leads to the following condition for a bound TM mode at an interface[40]

$$\frac{k_{z,2}(\omega)}{k_{z,1}(\omega)} = -\frac{\epsilon_2(\omega)}{\epsilon_1(\omega)}, \tag{1.23}$$

[39] A brief overview of the classical EM boundary conditions at a sharp interface is presented in appendix B.

[40] The derivation is presented in appendix C.

which reveals that the real part of the permittivities must have opposite signs. The SPP condition is fulfilled for metal-dielectric surfaces in the visible and near-IR. The evanescent decay in the normal direction into medium i occurs over a length scale $\frac{1}{\Re[k_{z,i}]}$. The wavevector components can be related to one another via the wave equation

$$k_{z,i}^2(\omega) = k_x^2(\omega) - \frac{\omega^2}{c^2}\epsilon_i(\omega), \tag{1.24}$$

which reveals that a bound surface mode is a consequence of a momentum component along the propagation direction *greater* than the free space momentum. This expression combined with equation 1.23 gives the SPP dispersion relation

$$k_x(\omega) = \frac{\omega}{c}\sqrt{\frac{\epsilon_1(\omega)\epsilon_2(\omega)}{\epsilon_1(\omega) + \epsilon_2(\omega)}}. \tag{1.25}$$

It is valid for both real and complex ϵ_i. In general, the solution to

[41] i.e. an explicit equation for ω.

equation 1.25 will not be expressible in closed form[41], but for a simple Drude dielectric function with no damping we need only solve a simple quadratic equation to get (Pitarke et al., 2006)

$$\omega_{\pm}^2 = c^2k_x^2 + \omega_{SP}^2 \pm \sqrt{c^4k_x^4 + \omega_{SP}^4}, \tag{1.26}$$

where ω_{SP} is the SP frequency. In figure 1.5 the SPP dispersion for a silver-air interface is shown, from which we can distinguish a number of different regions and limits:

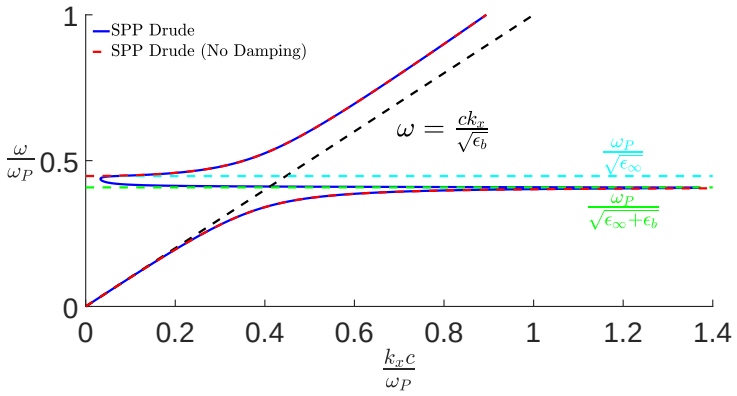


Figure 1.5: The SPP dispersion for a silver-air interface modelled with the Drude function (same parameters as used for figure 1.2).

- For small wavevectors ($k << \frac{\omega_P}{c}$), which for typical plasmonic metals corresponds to mid-IR and lower, the SPP dispersion is close to the light line and behaves as grazing-incidence light (Maier, 2007). In this region the SPP is light-like.
- For large wavevectors ($k >> \frac{\omega_P}{c}$), the frequency approaches a constant value known as the SP frequency, $\omega_{SP} = \frac{\omega_P}{\sqrt{\epsilon_\infty+\epsilon_b}}$. The group velocity approaches zero and the wavenumber goes to infinity in the limit of no damping[42]. In this region the SPP has an electrostatic character and is plasmon-like, i.e. it is a surface charge oscillation.
- Between these two limits the SPP exhibits both light-like and plasmon-like behaviour, it is a hybrid light-matter state.
- In the limit of no damping, there is an energy gap between the SP energy and the volume plasmon energy, $\frac{\omega_P}{\sqrt{\epsilon_\infty}}$, that corresponds to a purely imaginary k_x. The inclusion of damping leads to a back bending of the dispersion curve and the SPP is quasibound in this region (Maier, 2007). There is now a maximum possible wavevector, which sets a limit on the amount of field confinement and the plasmon wavelength.
- For frequencies above $\frac{\omega_P}{\sqrt{\epsilon_\infty}}$ in the upper branch (called a Brewster mode) radiation into the metal can occur. This is no longer a surface wave as the z component of the wavevector is no longer purely imaginary (Novotny and Hecht, 2012).

[42] This is a flawed prediction of the classical model. For large wavenumbers, the plasmon will 'sense' deviations away from classical macroscopic theory. Even with no damping, nonlocality would impose a maximum wavevector on the order of the Fermi wavevector.

For regions of the dispersion to the right of the light line, k_x is larger than the corresponding wavevector of the equal frequency

freely propagating light. Unless the lines cross at a particular frequency, it is not possible to couple in or out light due to the momentum mismatch[43].

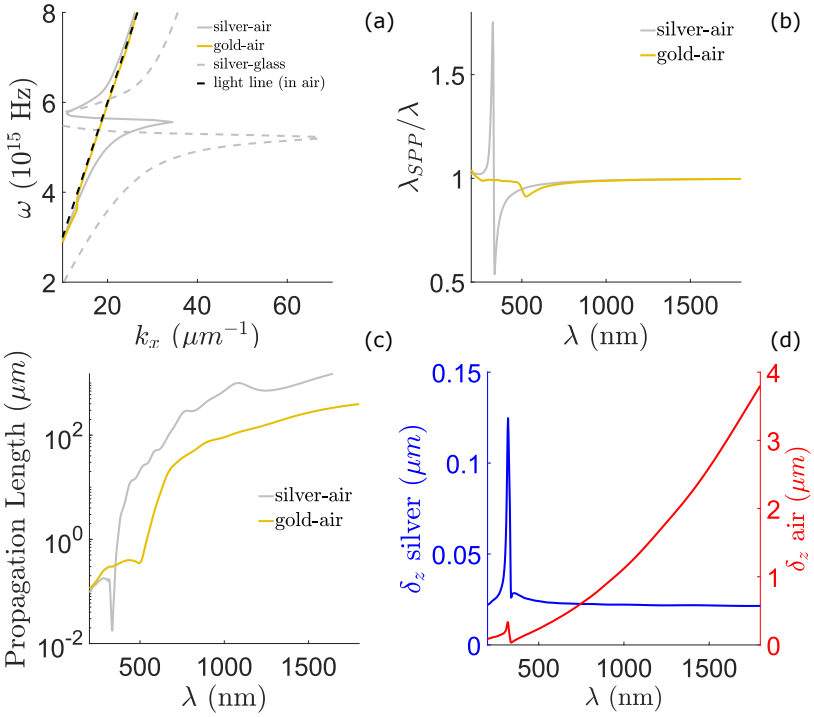


Figure 1.6: (a) SPP dispersion for silver-air, gold-air and silver-glass ($n_b = 1.5$) surfaces, using experimental data for the dielectric functions (Johnson and Christy, 1972). (b) The SPP wavelength, normalised to the incident wavelength, for a gold and silver surface in air. (c) The corresponding propagation length of the SPP. (d) The penetration depth of the SPP into the metal and dielectric half-spaces, for a silver-air surface.

To gain some more physical insight into SPPs, we can plot some relevant physical quantities. In figure 1.6(a) the SPP dispersion for silver-air, gold-air and silver-glass ($n_b = 1.5$) interfaces are shown. Silver shows a much more pronounced SP resonance, compared to gold, and is well recognised as the superior plasmonic metal near the plasma frequency due to lower losses[44]. The role of the dielectric half-space is also illustrated, using a dielectric with a larger permittivity than air leads to a redshift of the SP resonance ($\frac{\omega_p}{\sqrt{1+\epsilon_b}} < \frac{\omega_p}{\sqrt{2}}$) as well as an increased maximum wavevector, which is a consequence of the wavelength being squeezed already in the dielectric relative to the free wavelength in air. In figure 1.6(b) the SPP wavelength, $\lambda_{SPP} = 2\pi/k_x$, normalised to the incident wavelength, is shown as a function of incident wavelength for gold-air and silver-air surfaces. For all wavelengths longer than the SP resonance wavelength, the SPP wavelength is always smaller than the incident wavelength, although only by a small amount at longer wavelengths away from the resonance. This reflects the bound nature of the surface polariton and breaks down above the SP frequency. As expected, λ_{SPP} in silver can be many times smaller than for gold near the SP resonance. The propagation length of the SPP, which is defined as $\delta_{SPP} = \frac{1}{2\Im[k_x]}$ (Maier, 2007) and measures how far the SPP travels before its intensity falls by an amount e^{-1}, is shown in figure 1.6(c). Away from the SP resonance, SPPs in both silver and gold exist over many wavelengths ($\delta_{SPP} >>$

[43] Theoretically, the easiest way to overcome this is to use the evanescent wave created at the interface of two dielectrics just above the metal surface, this will tilt the light line. This is known as the Otto configuration but it is experimentally inconvenient to control the small gaps. More suitable alternatives include the Kretschmann configuration where a thin metal film is deposited on a glass prism or grating couplers. Moving away from planewave excitation, dipole-like sources close to the surface can be used to access near-field components to match the SPP wavevector and couple locally. Suitable near-field probes include metal nanoparticles, subwavelength apertures (such as used in scanning near-field optical microscopy) and quantum emitters, see (Maier, 2007) and (Novotny and Hecht, 2012) for more details. Fast electrons can also be used for exciting and probing SPPs (Chu et al., 2008).

[44] Gold becomes more competitive as longer wavelengths away from interband transitions and is often preferred to silver in experiments due to a more desirable surface chemistry; silver is susceptible to oxidation effects that limit its use in non-vacuum conditions.

λ_{SPP}). This reflects their less bound state at these wavelengths, a smaller amount of the EM energy is in the lossy metal. This point is further emphasised by 1.6(d), which shows the penetration depth, $\delta_{z,i} = \frac{1}{\Re[k_{z,i}]}$, into the metal and dielectric half-spaces, for a silver-air interface. As the wavelength is increased the penetration depth into the dielectric increases. In contrast, the penetration depth into the silver becomes approximately independent of wavelength; it shows only a slow decrease towards a constant value given by the skin depth of the metal (Jackson, 2007). Together these results illustrate the typical trade-off seen in plasmonics: **the more bound a surface mode is, the greater the losses suffered**.

1.2.3 *Localised Surface Plasmons Part 1: Quasistatic Approximation*

The main bulk of this thesis will be devoted to the topic of localised surface plasmons (LSPs). A LSP is simply a confined SPP in a small nanoparticle[45], the mode does not propagate and hence has a flat dispersion at the LSP resonance, which allows for direct excitation by light. What this means is that the surface can give the necessary momentum transfer to couple external light to SPs. Localised plasmon excitations may also be associated with the finite dimensions of infinite size nanostructures. For example, this is the case for an infinitely long cylinder which supports an LSP confined in the direction perpendicular to the infinite axis[46]. By controlling the geometry of the nanostructure, method of excitation and its dielectric environment, the near- and far-field properties can be manipulated. A simple picture of an LSP is shown in figure 1.7, an impinging EM wave will periodically shift the electron cloud, with the ionic background acting as a restoring force. This is nothing more than a harmonic oscillator, with the external light acting as the driving force, and will have a natural frequency given by the LSP resonance. Furthermore, the induced charge leads to an induced dipole moment and the LSP can couple strongly to light.

A substantial understanding can be achieved by considering the case of a small ($R << \lambda$) nanosphere in the *quasi-static* limit, also known as the electrostatic limit. In this case, the electric field can be taken to have a spatially homogeneous magnitude and the magnetic field can be ignored; the problem reduces to an electrostatic one[47]. More precisely, exactly when this approximation is valid will depend on the material[48]

$$R << \frac{\sqrt{2}c}{\omega\sqrt{\Im[\epsilon(\omega)]}}. \tag{1.27}$$

For a metal like aluminium, which has a large plasma frequency, the electrostatic approximation for plasmonic applications will start to break down for $R \gtrsim 5$ nm. For gold and silver at visible frequencies, it is valid for $R \lesssim 20$ nm. As we shall see in later chapters, at small sizes the macroscopic dielectric theory breaks

[45] We can consider an SPP propagating over the surface of a sphere. At certain excitation frequencies, the SPP path length will correspond to an integer number of SPP wavelengths and constructive interference will occur, setting up a standing wave. These frequencies correspond to the LSP resonances (Fang and Tian, 2014).

[46] Metal wires also support a propagating SPP mode along the infinite axis (Novotny and Hecht, 2012).

[47] All time dependence is contained in a trivial factor of $e^{-i\omega t}$.

[48] This expression can be found in (Bohren and Huffman, 2008), see appendix D for details of the simple derivation for this expression.

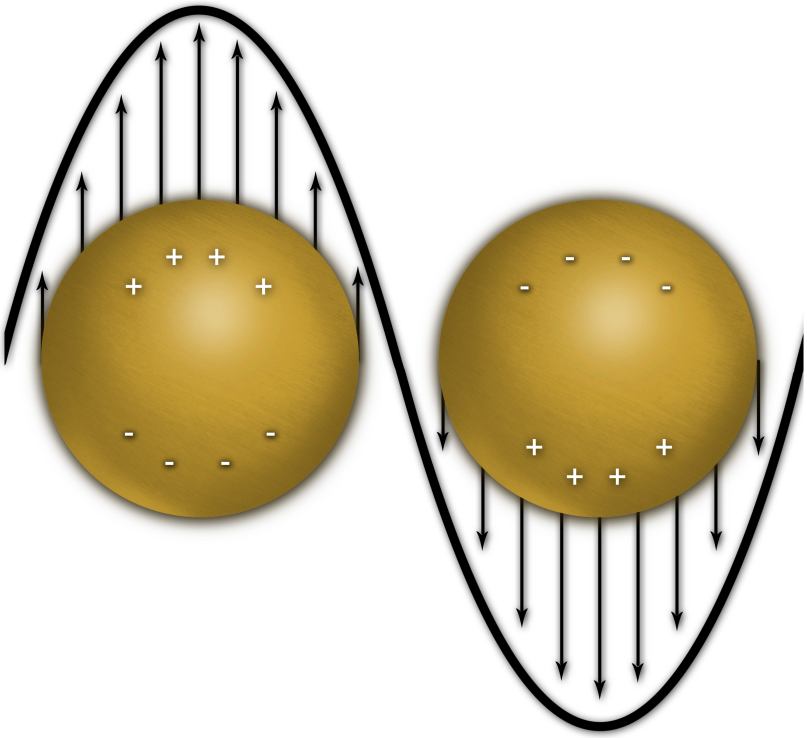


Figure 1.7: An illustration of an LSP in a metal nanosphere at two points in time during the charge oscillation.

down as quantum size effects become important. Therefore, the electrostatic approximation is only valid over a very small window of particle sizes. Fortunately, this is enough to give us a valuable insight into plasmonic behaviour in nanoparticles, much of the current understanding and intuition in modern plasmonics is built upon it. The solution of the Helmholtz equation (equation 1.21) in each dielectric region reduces[49] to the solution of the Laplace equation

$$\nabla^2 \phi_i = 0, \tag{1.28}$$

where ϕ_i is the scalar potential in region i and completely determines the electric field, $\mathbf{E}_i = -\nabla\phi_i$. In the case of a single metal sphere there are two regions: inside the sphere, which has a dispersive dielectric function $\epsilon(\omega)$, and outside, which has a real dielectric constant ϵ_b[50]. The derivation of the LSP resonances for a sphere is a textbook calculation based on an expansion of the potential in spherical harmonics, the details can be found in appendix E. The condition for an LSP within the electrostatic approximation is

$$-\frac{\epsilon(\omega)}{\epsilon_b} = \frac{l+1}{l}. \tag{1.29}$$

There is an infinite family of modes characterised by the angular momentum number $l = 1, 2...\infty$, these correspond to the dipole, quadrupole, octupole... modes. For an ideal Drude metal nanosphere in vacuum with no damping ($\epsilon_\infty = 1,\ \gamma = 0$), we get

[49] The electrostatic approximation corresponds to taking the limit $k \to 0$.

[50] This corresponds to the most common case of a dielectric background. Common examples are air, water and glass which can be taken to have frequency independent permittivities and are effectively lossless in the frequency ranges we are interested in. Note that no dispersive material can be completely loss-free for all frequencies due to the requirements of the Kramers-Kronig relations (Jackson, 2007).

the condition

$$\omega_{LSP} = \sqrt{\frac{l}{2l+1}}\omega_P. \tag{1.30}$$

This result reveals the LSP energies of the sphere, but it does not tell us which modes will be excited by an external field; this will be determined by the symmetry of the perturbation[51]. For planewave excitation it is simple to see that only the dipole mode can be excited, $l = 1$. Using the above relation then gives us the Frölich condition for an LSP

$$\omega_{LSP} = \frac{\omega_P}{\sqrt{3}}. \tag{1.31}$$

Note that for $l \to \infty$ we obtain the SP energy of an infinite plane in vacuum, $\omega_{SP} = \frac{\omega_P}{\sqrt{2}}$. Calculating the electric field for planewave excitation gives[52]

$$\begin{aligned} \mathbf{E}_{in}(\mathbf{r},\omega) &= \mathbf{E}_{ext}\frac{3\epsilon_b}{\epsilon(\omega)+2\epsilon_b} \\ \mathbf{E}_{out}(\mathbf{r},\omega) &= \mathbf{E}_{ext} + \frac{3\hat{\mathbf{n}}(\hat{\mathbf{n}}\cdot\mathbf{p}(\omega)) - \mathbf{p}(\omega)}{4\pi\epsilon_0\epsilon_b r^3}, \end{aligned} \tag{1.32}$$

where $\hat{\mathbf{n}}$ is the unit normal to the surface. This reveals some interesting insights:

- The internal field $\mathbf{E}_{in}$ is spatially uniform and is a consequence of the nanosphere being smaller than the metal's skin depth. This is also true for any small ellipsoidal nanoparticle under uniform excitation, but for any other geometry the field inside will be spatially dependent; this is known as the *Eshelby conjecture* (Moroz, 2009).
- The external field is the exciting field plus a dipole field, corresponding to that of a dipole moment $\mathbf{p}(\omega)$ located at the origin[53], where $p(\omega) = \alpha(\omega)|\mathbf{E}_{ext}|$. Both the external and internal fields have a radial component, they are therefore near-field in nature. This is a consequence of the magnetic field being small so that the electric field is irrotational and resembles an electrostatic field, $\nabla \times \mathbf{E} = 0$.
- The field drops off rapidly with $1/r^3$, meaning that the field is well localised around the nanosphere, which is much smaller than the external field wavelength. The light has been squeezed well past the diffraction limit. This is a consequence of coupling of EM energy to free carriers, this physics is captured by the dielectric function used in the nanoparticle's polarizability, $\alpha(\omega) = 4\pi\epsilon_0\epsilon_b R^3\left(\frac{\epsilon(\omega)-\epsilon_b}{\epsilon(\omega)+2\epsilon_b}\right)$. Note that this is the only place that the radius of the particle enters and that the LSP energies are surprisingly size independent.
- Both the internal and external electric field experience a resonance at the same frequency that satisfies $\epsilon(\omega) = -2\epsilon_b$. At the LSP resonance, the electric field can be enhanced signifi-

[51] This will be a common theme in this thesis, we will find that linear response DFT is similar in the respect that one calculates all the modes of the system and then selects only the ones that can be excited by the external field, from which the relevant quantities of interest can be calculated.

[52] See appendix E for the derivation.

[53] In electrostatic theory the potential outside a spherically symmetrical charge distribution, $\rho(r)$, is the same as if the total charge, Q, was concentrated at the centre of the sphere, $\int d^3\mathbf{r}\frac{\rho(r')}{|\mathbf{r}-\mathbf{r}'|} = \frac{Q}{r}$.

cantly; the value of the FE is dictated by the quality factor of the nanosphere, which within the quasistatic approximation is solely determined by material loss.

The quality factor is a crucial quantity in the physics of FEs and is determined by the complex permittivity of the material, it is independent of the nanostructure geometry and dielectric environment within the quasistatic approximation[54] (Wang and Shen, 2006). It is a dimensionless parameter that characterises the damping of a resonance, which is equivalent to measuring the resonance linewidth relative to the resonant frequency. It can be calculated for plasmonic nanoparticles using the same formalism as used for dielectric cavities. It is defined as 2π times the ratio of the time-averaged energy in the resonator to the power loss per cycle (Jackson, 2007)

$$Q = \omega_0 \frac{\text{Stored Energy}}{\text{Power Loss}} = \frac{\omega}{2\Im[\epsilon(\omega)]} \frac{d\Re[\epsilon(\omega)]}{d\omega}. \tag{1.33}$$

It gives a measure of how long a photon spends in a cavity, or equivalently how many reflections back-and-forth from the cavity walls that it makes. The coupling of EM energy to free carriers must come with an increase in damping, as energy is now lost at the rate of the scattering of electrons: **higher loss is the inevitable price paid for the increased field confinement and enhancement in plasmonics**, this is also what we found in the case of SPPs in the last section. To give some context: a typical plasmonic resonator will have $Q \sim 10$, while a typical microwave cavity will be a few hundred or thousand (Jackson, 2007) although state-of-the-art dielectric cavities used to explore strong-coupling can reach much higher, see chapter 3.

[54] Although the geometry will change the frequency of the LSP resonance and hence the frequency at which we are interested in the quality factor will change with geometry. For instance, metals have superior quality factors in the IR so it can be advantageous for some applications to choose a geometry that has an LSP resonance in this frequency region.

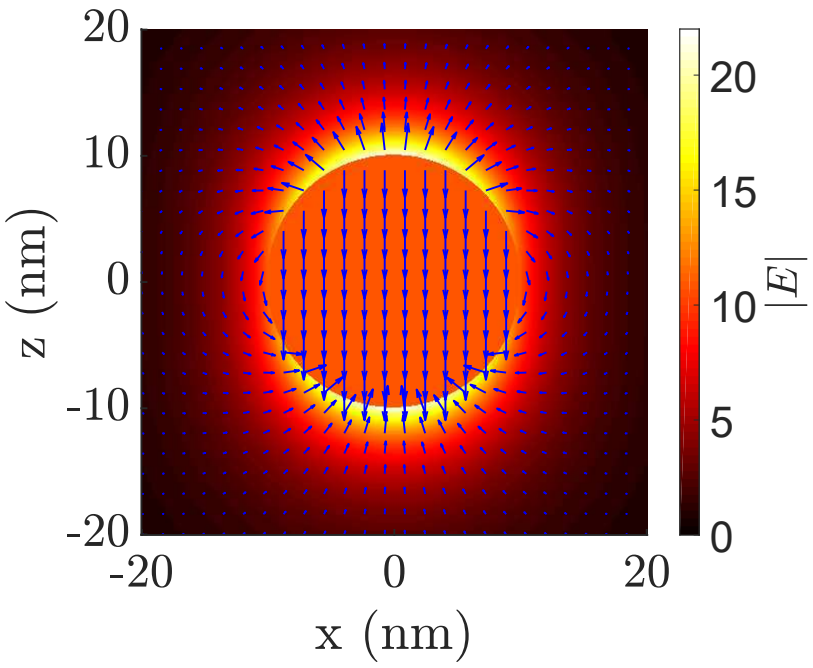


Figure 1.8: A vector plot of the electric field, superimposed on the electric field magnitude, for the dipole LSP resonance of a 10 nm radius silver nanoparticle. The dielectric function used is from (Johnson and Christy, 1972).

To further appreciate how a plasmon couples to the EM field, we can construct a simple model for the LSP based on energy considerations. Using the Poynting vector, it is simple to construct the following energy balance equation in the time domain for a region[55] enclosed by the surface $\mathcal{S}$

$$\oint \mathbf{E} \times \mathbf{H}^* d\boldsymbol{\mathcal{S}} = -\int dV \left[\mu_0 \mathbf{H} \cdot \partial_t \mathbf{H}^* + \epsilon_0 \mathbf{E} \cdot \partial_t \mathbf{E}^* + \mathbf{E} \cdot \partial_t \mathbf{P}^*\right] \quad (1.34)$$

[55] We take the imaginary enclosing box to be large so that any EM energy escaping can be taken to be far-field radiation.

As a simple model for the polarization, we consider a forced and damped oscillator[56] with natural frequency ω_{LSP}

$$\partial_t^2 \mathbf{P} + \Gamma \partial_t \mathbf{P} + \omega_{LSP}^2 \mathbf{P} = \epsilon_0 \mathbf{E}. \quad (1.35)$$

[56] It is a similar idea to the model used earlier for the long wavelength limit of the volume plasmon.

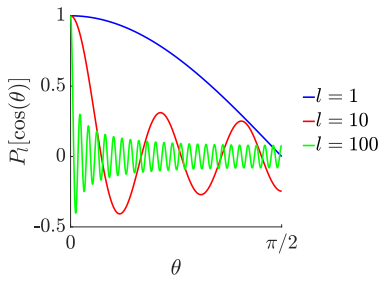
Combining these two equations it can be shown that the following decay equation holds for the energy

$$\overbrace{\overbrace{-\oint \mathbf{E} \times \mathbf{H}^* d\boldsymbol{\mathcal{S}}}^{\text{Radiative Loss}} - \overbrace{\int dV \frac{\Gamma}{\epsilon_0} |\partial_t \mathbf{P}|^2}^{\text{Material Loss}}}^{\text{Total Loss}} = \underbrace{\partial_t \int dV \left[\underbrace{\mu |\mathbf{H}|^2 + \epsilon_0 |\mathbf{E}|^2}_{\text{EM Field Energy}} + \underbrace{\frac{1}{2\epsilon_0}\left(|\partial_t \mathbf{P}|^2 + \omega_{LSP}^2 \mathbf{P}^2\right)}_{\text{Energy In The Plasmon}}\right]}_{\text{Total Energy}}. \quad (1.36)$$

The polarization field, oscillating in time at the LSP frequency, can store energy from the field *and* leads to an extra decay channel via Ohmic losses. Crucially, the more energy is given to the dynamics of the polarization field the more the loss grows, as both are proportional to $|\partial_t \mathbf{P}|^2$.

Let us look at the plasmonic field (equations 1.32) in more detail. In figure 1.8 a vector plot of the electric field, superimposed on the magnitude, is shown for the dipole LSP resonance of a silver nanoparticle with a radius of 10 nm. The excitation frequency is chosen to be slightly below the resonance, above the resonance the field will flip in direction. We see a strong FE in the z direction (the same direction as the external field polarization) and on the order ~ 20, which is a typical value for a noble metal at resonance[57]. The field shows a typical dipolar character and will couple strongly to light; it is a bright mode. The field is localised strongly to the nanosphere surface showing a strong field concentration below the diffraction limit.

[57] As a rough rule of thumb, the FE will be on the order of the quality factor.



The Legendre functions for different l. Observe the increased squeezing towards $\theta = 0$ for larger l values.

The sphere can also support higher order plasmon modes, as revealed by equation 1.29; the expression for the potential and electric field of mode l is given in appendix E. It is interesting to compare these electric field plots, shown in figure 1.9 for $l = 2$ and $l = 3$, with the dipole mode shown in 1.8. The first thing to notice is that the electric field is increasingly squeezed into the z axis for larger l. Mathematically, this is a consequence of the behaviour of the

Legendre polynomials which become progressively squeezed into $\theta = 0$ as l grows. The field also becomes increasingly concentrated to the surface, with the field dropping off as r^{-l-2}. Also apparent is that for $l > 1$, the internal field is not a constant and shows an increasingly large minimum in the centre of the sphere.

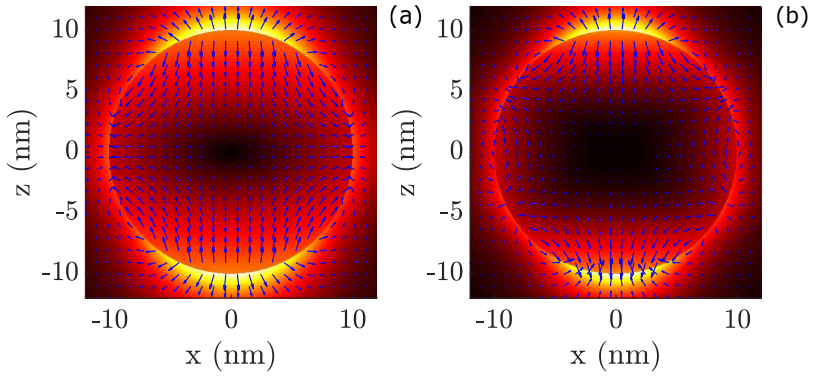


Figure 1.9: The electric field vector plot, superimposed on the electric field magnitude, for the (a) quadrupole and (b) octupole resonances of a 10 nm radius silver nanoparticle. The dielectric function used is from (Johnson and Christy, 1972).

Further physical insight can be obtained by looking at the surface charge[58]. In figure 1.10 we illustrate the surface charge density for the dipole, quadrupole, octupole and $l = 10$ plasmons. Areas of high induced surface charge correspond to points where the electric field lines terminate. For higher l there is an increasing number of surface charge oscillations around the sphere's circumference and an increasing localisation for small angles around the z axis.

[58] See appendix E for a derivation.

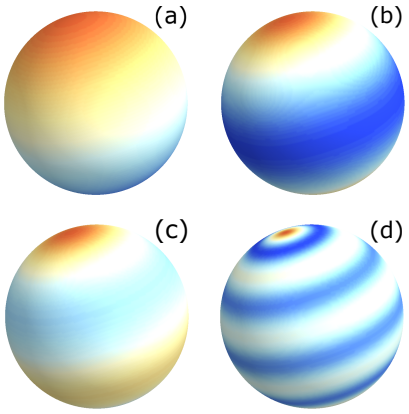


Figure 1.10: The induced surface charge the (a) dipole, (b) quadrupole, (c) octupole and (d) $l = 10$ resonances of a 10 nm radius silver nanoparticle.

The strong localisation of the electric field at the nanosphere surface must inevitably lead to strong field gradients. This is of potential interest in surface-enhanced spectroscopies for exciting dipole-forbidden transitions in atoms and molecules, which could lead to new techniques for chemical fingerprinting (Alabastri et al., 2016; Rivera et al., 2016). This is particularly interesting for molecules that do not possess a dipole moment due to their spherical symmetry, two relevant examples for modern times are the greenhouse gasses: carbon dioxide (CO_2) and methane (CH_4). Using the famous dipole-approximation Hamiltonian, which works incredibly well for atoms and molecules under planewave illumi-

nation, the field variation is neglected and so the gradient, which is proportional to the wavevector, is approximately zero. This is valid if the atom/molecule is much smaller than the wavelength of the impinging EM wave and leads to the familiar dipole selection rules for the angular momentum quantum numbers: $\Delta l = \pm 1$ and $\Delta m = 0$. An atom/molecule near the surface of a metal nanosphere will experience a much larger gradient that allows 'forbidden' transitions not accessible with planewave illumination. The gradient outside the nanosphere and in the radial direction can be written as[59]

[59] See appendix E for the derivation.

$$\nabla \mathbf{E}|_{r,r} \propto \cos(\theta) \frac{R^{2l+1}}{r^{l+3}}. \tag{1.37}$$

In the most common case that the dipole LSP of the nanosphere is excited, we find that $\nabla \mathbf{E}|_{r,r} \propto \frac{1}{R}$, thus if $R << \lambda$ (which must hold for the preceding formulas to be correct), the gradient can be strongly enhanced in small nanoparticles, see figure 1.11 for a map of the radial electric field gradient enhancement for the dipole LSP of a 10 nm radius silver sphere. Interestingly, this result is a peculiarity of the dipole mode. The radial gradient term at the surface ($r = R$) scales as R^{l-2}, so for $l = 2$ the gradient is independent of particle size and for $l > 2$ is enhanced for larger spheres[60].

[60] Remembering, of course, that at larger radii the quasistatic approximation will break down making this analysis invalid.

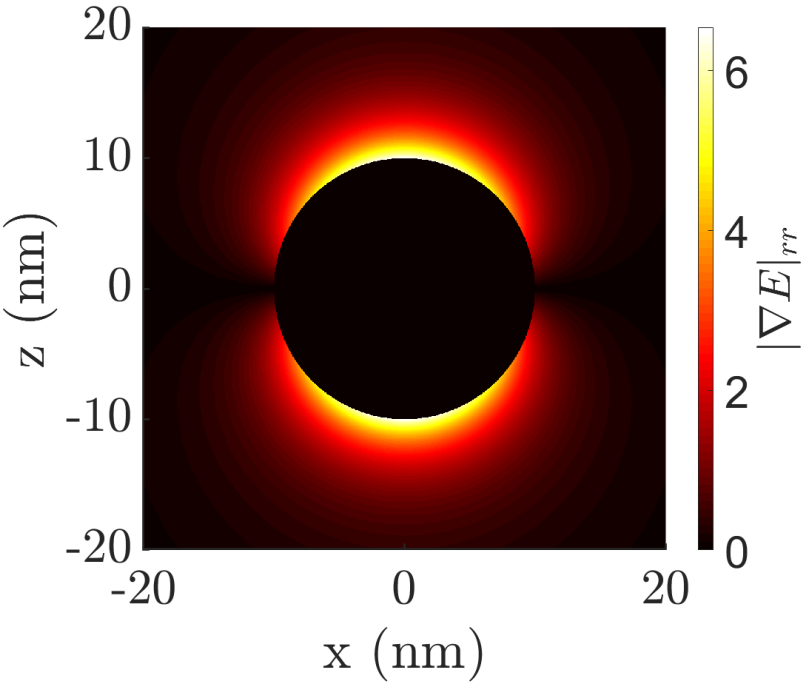


Figure 1.11: The radial electric field gradient for the dipole LSP resonance of a 10 nm radius silver nanosphere. The dielectric function used is from (Johnson and Christy, 1972).

Analytical solutions for geometries other than the sphere are rare; they are only possible in systems which have suitable symmetry. The ellipsoid is a common system to look at as it is the simplest case of an anisotropic nanoparticle and can be treated analytically in the quasistatic approximation. By varying the ra-

tio of the long and short axes, it can support strong and tunable plasmon resonances. For an ellipsoid with semiaxes R_1, R_2, R_3, the case of $R_1 > R_2, R_2 = R_3$ is known as a *prolate* spheroid[61] and is often used as a simple model for nanorods. The other type of spheroid $R_1 = R_2, R_1 > R_3$, are known as *oblate* spheroids and are suitable models for nanodisks. Elongated nanoparticles break spherical symmetry and lift the LSP degeneracy; they can support two plasmon modes called the longitudinal mode (L-mode) and transverse mode (T-mode), which are associated with the oscillation of electrons along the long and short axis respectively. As such, the modes will be selectively excited dependent on the polarization of the external light. Compared to the plasmon resonance of a sphere of equal volume, the L-mode will have a lower energy (redshifted) and the T-mode will have a higher energy (blueshifted), this can be understood as an increase or reduction of the restoring force from the surface charge due to elongation. The solution for an ellipsoid in the electrostatic approximation is well-known (Bohren and Huffman, 2008), the polarizability, along direction i of an ellipsoid with permittivity $\epsilon(\omega)$ in a background medium ϵ_b is

$$\alpha_i(\omega) = 4\pi R_1 R_2 R_3 \frac{\epsilon(\omega) - \epsilon_b}{3\epsilon_b + 3L_i(\epsilon(\omega) - \epsilon_b)}, \tag{1.38}$$

where $i = \{1, 2, 3\}$ and L_i is the depolarization factor given by

$$L_i = \frac{R_1 R_2 R_3}{2} \int_0^\infty \frac{1}{f(x)\left(R_i^2 + x\right)} dx, \tag{1.39}$$

where $f(x) = \sqrt{(x + R_1^2)(x + R_2^2)(x + R_3^2)}$. The three depolarization factors are not independent, $L_1 + L_2 + L_3 = 1$. For a prolate spheroid the integral can be written in analytical form in terms of the eccentricity $\mathfrak{e}^2 = 1 - \frac{R_2^2}{R_1^2}$

$$L_1 = \frac{1 - \mathfrak{e}^2}{\mathfrak{e}^2}\left(\frac{1}{2\mathfrak{e}} \ln\left[\frac{1 + \mathfrak{e}}{1 - \mathfrak{e}}\right] - 1)\right). \tag{1.40}$$

In figure 1.12 a plot for the L-mode resonance (blue line) is shown, obtained by the solution to $\epsilon_b + L_i(\epsilon(\omega) - \epsilon_b) = 0$, i.e. the resonance condition of equation 1.38. The long semiaxis R_1 of a gold prolate ellipsoid is varied while keeping the volume fixed and equal to a 50 nm sphere[62]. The results are compared to the extracted resonance wavelength condition of the absorption cross section obtained using the boundary element method (BEM) (see section 1.3.2) within the electrostatic approximation[63], which shows excellent agreement with the analytical result, and the full retarded solution, which shows a redshift compared to the electrostatic result. All three results shows a redshift of the L-mode for increasing elongation.

For structures that cannot be treated analytically we can fit the

[61] A spheroid is an ellipsoid in the particular case that two axes have equal value.

[62] Which corresponds to the case $R_1 = R_2 = R_3 = 50$ nm.

[63] See section 1.2.4 for a definition of the optical cross section. A field-map of the L-mode and T-mode, calculated with the BEM, is presented in figure 1.15.

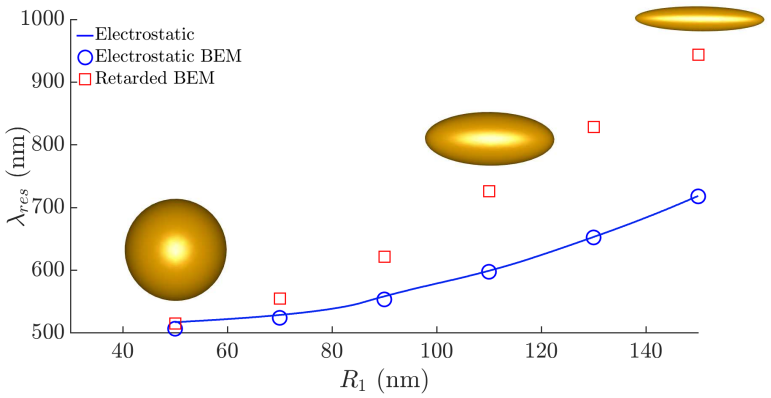


Figure 1.12: Sweep of the long axis R_1 for a gold prolate spheroid in vacuum and of fixed volume (equivalent to a 50 nm radius sphere). Dielectric data used for gold from (Johnson and Christy, 1972).

polarizability using the general form (Davis and Gómez, 2017)

$$\alpha(\omega) = V\epsilon_0\epsilon_b \frac{2\xi(\epsilon(\omega)-\epsilon_b)}{\epsilon_b(\xi+1)+\epsilon(\omega)(\xi-1)}, \tag{1.41}$$

where ξ is the mode eigenvalue, it can be calculated by solving a self-consistent equation for the induced surface charge[64] or can be treated as a fitting parameter for numerical results. The eigenvalue is always greater than 1 and by setting the denominator of the above equation to zero, we find the following resonance condition.

$$\Re[\epsilon(\omega)] = \epsilon_b \frac{1+\xi}{1-\xi} \tag{1.42}$$

For a sphere $\xi = 3$ and for an ellipsoid can be written in terms of the depolarization factor $\xi_i = \frac{1}{1-2L_i}$. This fitting approach can be very useful for extending our deductions from the sphere and ellipsoid to more complicated geometries as well as including retardation effects. Retardation, at least for smaller structures, may only lead to simple shifts of the LSP resonance; the mode can still be readily understood within the electrostatic framework. In these situations equation 1.41 can be of great use.

A quick remark is needed to elucidate how is it possible for such a simple model as the Drude dielectric function to predict such rich plasmonic behaviour: why should a classical system of free particles have the ability to form a coherent collective oscillation? In the derivation of the Drude function, the dynamics of a single free electron were solved for and I made the tacit assumption that we can describe the dynamics of the whole material by multiplying by the number of electrons N, in effect we forced all the electrons to react in phase (coherently) to the perturbation. Coherence is a crucial aspect as an incoherent combination of eh pairs will tend to cancel out the total density change (Allen, 1996). This cancelling out will become perfect in the thermodynamic limit as the amplitude of an incoherent cosines is smaller than a coherent sum by a factor $1/\sqrt{N_{eh}}$, where N_{eh} is the number of eh pairs. We will come across this point again later for small nanoclusters,

[64] The BEM is suitable for this.

where this cancellation will not occur as N_{eh} is a finite number. The Drude electron gas is therefore a rather peculiar gas and not at all what we have in mind (or at least not what I have in my mind) when thinking of a classical gas, which is characterised by particles undergoing random and frequent collisions with one another and therefore having a range of different motions. In fact it is this sort of behaviour that leads to ordinary sound waves where frequent collisions bring about local equilibrium in response to perturbation[65]. In contrast, collisions disrupt the restoring force for a plasma oscillation which require $\omega\tau >> 1$ (Pines, 1964). Of course a more realistic picture is to take the electron gas to be a Fermi gas, but even here we run into problems as a simple back-of-the-envelope style calculation shows that the coulomb energy is significantly larger than the electron kinetic energy; so it does not seem reasonable to neglect electron-electron interactions, or even treat them as a perturbation. The solution to this dilemma was delivered by Landau's Fermi liquid theory which tells us that, as long as we are near the Fermi surface, we can still use the non-interacting picture if we switch to the basis of *quasiparticles*, which are the elementary excitations of the system. So remarkably, by assuming non-interacting particles and an apparently innocent multiplication by N, we are able to capture a lot of the essential physics of the electron gas and describe the optical response of metals. It is only when we are dealing with electron transitions away from the Fermi surface such as d-shell transitions that the Drude approach breaks down.

[65] This is the hydrodynamic limit that we discuss in section 2.2.

The coherence supported by an electron gas may be so large that the coherence of single photon-matter hybrid states can be maintained, something rather remarkable given the huge degrees of freedom contained by a macroscopic collection of electrons. This is often taken for granted, as was astutely noticed by Barnes (Barnes, 2002), in response to it being shown by Altewischer *et al* that photon entanglement survives conversion into SPPs (Altewischer et al., 2002), that we are not surprised by entanglement surviving reflection from a metallic mirror despite the same physics being at play.

1.2.4 *Localised Surface Plasmons Part 2: Including Retardation*

Here we will delineate the role of retardation on plasmonic systems. Within the electrostatic approximation, plasmonics boils down to solving Laplace's equation. For larger structures the induced EM radiation will be emitted in one section of the structure, propagate and pick up a phase, before then interacting with another part of the structure. Another way of seeing this, is that the oscillating plasmon field produces a oscillating magnetic field, which in turn act on the free carriers (a displacement current) and shifts the plasmon resonance[66]. In the electrostatic approximation this magnetic field is ignored. Furthermore, considering the LSP as an oscillating

[66] The alert reader may be a tad puzzled at this moment, is this not very similar to our earlier discussions on nonlocality? In fact a more careful analysis of the response formalism allows the clear separation of two types of space-time correlation: the 'memory' carried by the mediums EM response, which is responsible for the nonlocality in space-time, and

dipole it must radiate into the far-field[67]. This opens up energy decay channels via the emission of photons, distinct from the intrinsic material loss. To first-order in the wavenumber the scattering from a small nanoparticle is zero, we have to work to order k^3 to include it. From these considerations it is clear that the inclusion of retardation will not be trivial. In fact, it is only for spheres and infinite cylinders that tractable analytical solutions exist. Numerical calculation is also now more complicated. We will again use the sphere as an example system, this will allow direct comparison to our findings from the last section.

Before we proceed it will be advantageous to discuss a convenient partitioning of the EM field into three distinct regions, this is best illustrated with the field of an electric dipole. We have already seen that in the electrostatic limit, the field outside a sphere is completely equivalent to that of the near-field of a dipole at the sphere centre; here we will also look at the *intermediate* and *far-field* terms. We will also consider the different physical quantities related to each zone; we have seen in the near-field that the induced field is of interest while in the far-field we will be more interested in integrated quantities such as the cross sections, i.e. what is measured in typical optics experiments. We will borrow the elegant[68] dyadic Green's function formalism (Novotny and Hecht, 2012), as it can be easily generalised to the case of more that one dipole (Draine, 1988). The Green's function $\overline{\mathbf{G}}$ is a tensor relating the electric field at a point $\mathbf{r}$ generated by a radiating electronic dipole at location $\mathbf{r}'$, mathematically we can write this as[69]

$$\mathbf{E} = \underbrace{\frac{e^{ikR}}{4\pi\epsilon_0 R}\left(\left[\bar{\mathbf{I}} - \frac{\mathbf{RR}}{R^3}\right] + \frac{i}{kR}\left[\bar{\mathbf{I}} - 3\frac{\mathbf{RR}}{R^3}\right] + \frac{1}{k^2R^2}\left[-\bar{\mathbf{I}} + 3\frac{\mathbf{RR}}{R^3}\right]\right)}_{\overline{\mathbf{G}}}\mathbf{p}, \tag{1.43}$$

where $\mathbf{RR}$ denotes the outer product of $\mathbf{R}$ with itself. The Green's function has terms on the order $(kR)^{-1}$, $(kR)^{-2}$ and $(kR)^{-3}$, where $R = |\mathbf{r} - \mathbf{r}'|$ denotes the distance separating the source and observation. Distances from the source can thus be measured relative to the wavelength and we can distinguish three distinct regions:

- The near-field is the zone where $R \ll \lambda$. If a nanoparticle, or collection of nanoparticles, are contained within a distance on this order then the discussions of section 1.2.3 hold. The evanescent nature of the near-field is associated with virtual photons, they are responsible for phenomena such as the van der Waals and Casimir forces and have applications in near-field microscopy (Novotny and Hecht, 2012).
- The intermediate-field is defined as $R \approx \lambda$ and the terms $(kR)^{-2}$ dominate in the Green's function. Here the field has both longitudinal and transverse components and is interesting for design of plasmonic components as the field can pick up significant phase changes over a nanostructure, this can be utilized for

[67] This is known as radiation reaction and is due to a recoil force from the dipoles field acting back on itself. The force is given by the Abraham-Lorentz formula and is found to depend on the time derivative of the acceleration: $\mathbf{F}_{rad} = \frac{\mu_0 e^2}{6\pi c}\partial_t^3 x$ (Griffiths, 2008). It is not often one sees a third derivative in time!

[68] Although we will sadly not have reason to see its full power and versatility in this thesis.

[69] In a general system one would sum up all the point current sources, this is the volume-integral approach (Novotny and Hecht, 2012). For a single dipole we have only one contribution.

nanoantenna design.

- The far-field is defined for $R \gg \lambda$. Here the transverse component of the field is dominating and can be identified with a propagating EM wave, i.e. it is EM energy that has escaped from the near-field zone of the system and hence we call it the radiation term. This term is associated with real photons.

It is the energy radiated into the far-field that is most often collected in optical experiments and analysed to give information on the system[70]. If this rate of energy is less than that which would be present without the nanoparticle, i.e. just the incident power, then the presence of the particle has resulted in *extinction*. Extinction is measured via the extinction cross section[71]

$$\sigma_{ext}(\omega) = \frac{k}{\epsilon_0 \epsilon_b} \Im[\alpha(\omega)] \tag{1.44}$$

Note that this expression is often given as the absorption cross section (e.g. in (Maier, 2007; Novotny and Hecht, 2012)), this is only in the extreme quasistatic regime where scattering can be neglected and $\sigma_{ext} \approx \sigma_{abs}$. Cross sections have the units of area; in the limit of geometrical optics, σ_{ext} is simply the area of the shadow the particles casts on the detector. More generally, it is an interference term between incident and scattered light and may be larger than the geometrical area of the particle. For a collection of nanoparticles, if they are non-interacting then one can simply take a sum of the individual cross sections, if they are interacting then one can use an effective polarizability taking into account interactions. For example, a coupled-dipole approach has been utilised to explore polarization conversion (Wiecha et al., 2017) and optical activity (Auguié et al., 2011), as well as providing the theoretical basis of the discrete-dipole approximation (Draine, 1988). Extinction is a sum of two processes: absorption, which is the conversion of EM energy into other forms (usually heat), and scattering into the far-field by the particle[72], this is a form of the optical theorem (Bohren and Huffman, 2008). Both these quantities have their own associated cross section,

$$\sigma_{scat}(\omega) = \frac{k^4}{6\pi\epsilon_0^2\epsilon_b^2} |\alpha(\omega)|^2, \tag{1.45}$$

and the absorption cross section can be calculated from $\sigma_{ext} - \sigma_{scat}$. The various cross sections are important as they are one of the main windows into the microscopic world and can give us information on size, shape, dielectric environment, composition and number of scatterers. Especially relevant to us is the behaviour of the cross sections at or near the LSP resonance. It is clear that when the polarizability has a resonance, so will *both* the absorption and scattering cross sections. The LSP resonance corresponds to a huge enhancement in the ability of nanoparticles to absorb and scatter light. At the resonance the Poynting vector lines strongly converge

[70] It is possible to measure the near-field using near-field scanning microscopy and can even be mapped using SERS and fluorescence (Cang et al., 2011; Weber and Willets, 2012)

[71] The different cross sections can be calculated by finding the power flow though an imaginary sphere, with a large radius ($R >> \lambda$), around the scatterer(s) by calculating the surface integral of the Poynting vector. See (Bohren and Huffman, 2008) for details.

[72] This nice distinction relies on the surrounding medium being non-absorbing.

Would it really be a thesis on plasmonics without mention of the Lycurgus cup? Printed with permission from the British Museum.

towards the particle, enhancing the effective cross section 'seen' by the light. The cross section spectral position can be manipulated by changing the nanoparticles properties and dielectric environment, this has obvious implication for sensing applications and is also the physics behind some of the most visually stunning demonstrations of plasmonics, such as the famous Lycurgus cup, where a combination of different sized gold/silver alloy nanoparticles allows different colour response for reflection (green) and transmission (red).

Next we turn to Mie theory. The full derivation is long and tedious so I banish it to appendix F and provide only some pertinent results here. The physical quantities of interest are written in terms of expansion coefficients that can be calculated from infinite series of spherical Bessel functions. In figure 1.13 the extinction cross section spectrum of a silver nanosphere is plotted against radii, calculated within the electrostatic approximation and full Mie theory. The electrostatic predicts no size dependence for the spectral position of the LSP (indicated by the white dashed line in the figure) and the extinction efficiency[73] grows with increasing R. In contrast, Mie theory predicts a redshift and broadening of the LSP peak due to retardation effects, this means there is a maximum extinction efficiency. Note that the maximum extinction cross section is 15 times larger than the geometrical area! This maximum efficiency is for a radius between 20 to 30 nm. The figure also illustrates that the quasistatic approximation breaks down very quickly, already at 20 nm there is significant deviations.

[73] i.e. extinction cross section divided by the geometrical area of the nanoparticle.

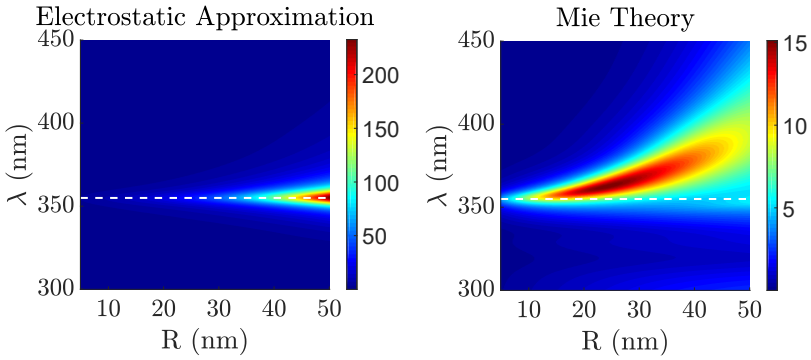


Figure 1.13: Plot of the extinction cross section, normalised to the geometrical area, for a sweep over silver nanosphere radii. The LSP resonance, $\epsilon(\omega) = -2$, calculated within the electrostatic approximation, is shown by the dashed white line. Dielectric data used for silver from (Johnson and Christy, 1972).

It is interesting to see if this behaviour is mimicked in the near-field, we cannot *a priori* assume so. A suitable quantity to look at the is near-field scattering efficiency introduced by Messinger *et al*, which measures the efficiency of a nanoparticle to convert incident light into near-field intensity (Messinger et al., 1981) and is linked to the FE. This is analogous to the scattering cross section, which measures the efficiency of a nanoparticle to convert incident light into scattered radiation. The cross section spectrum is shown in figure 1.14, calculated for a silver nanosphere over a range of radii. As expected, the overall trend of the near-field efficiency is similar to the extinction cross section; the largest FE is found close to the

LSP resonance (dashed white line) and for smaller nanospheres. As retardation becomes more important at larger radii, there is a redshift and broadening of the near-field peak.

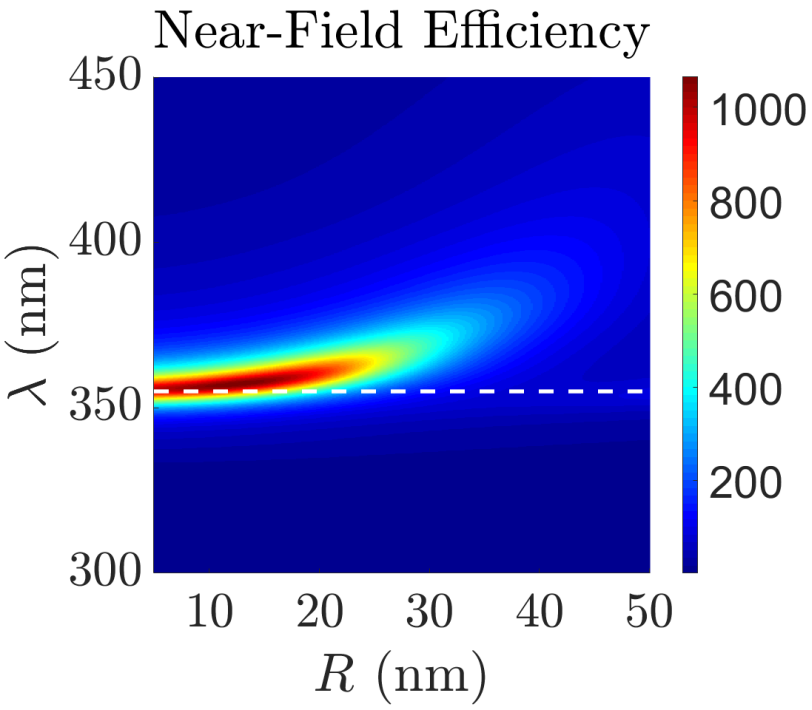


Figure 1.14: Plot of the near-field scattering efficiency, for a sweep over a silver nanosphere radius. The LSP resonance, $\epsilon(\omega) = -2$, calculated within the electrostatic approximation, is shown by the dashed white line. Dielectric data used for silver from (Johnson and Christy, 1972).

1.3 Numerical Techniques In Plasmonics

In many ways plasmonics is tough on the theorist, analytical solutions are rare and when they do exist are often unwieldy, as we have seen for the case of Mie theory. Luckily the field of research into numerical solutions of Maxwell's equations is mature and there exists a large number of computational approaches. In this section we explore different numerical methods that pertain to the work covered in this thesis.

Plasmonics research relies on a wealth of different numerical techniques. There is no one-size-fits-all method, with the choice depending on the problem at hand. Here we look at the three techniques used heavily in this work: finite difference time domain (FDTD), boundary element method (BEM) and the generalised multiparticle Mie (GMM) theory[74]. Having a number of different approaches is a blessing as it allows a nice check of results. This should be compared to electronic structure calculations, where one often does not have this luxury, commonly DFT is the only reasonable method to model a quantum system. In plasmonics, agreement between the different methods is generally good, with FDTD being the hardest to converge. As a rough rule of thumb, agreement of better than $\sim 5-10\%$ between FDTD and another method such as BEM should make one content.

[74] More correctly maybe, this should be called a semi-analytical method as the solution can be written down explicitly, but as this can only be evaluated using a computer and achieving convergence involves some challenges, I place it in this category.

1.3.1 Finite Difference Methods

Finite difference methods are a brute force method to solving Maxwell's equations. Its strength, it is very general method and can be used in a huge number of situations, is also its weakness, often more specialised methods will be much faster and simpler. The method relies on approximating differential equations with finite difference equations, in the limit that spatial and time step size goes to zero, this approach formally becomes exact[75]. The convergence of this will depend on the exact finite difference technique, the geometry of the structure under investigation as well as the dielectric environment[76]. Finite difference methods come in two flavours: *time-domain* and *frequency-domain*. In both case the spatial domain is divided into a grid of nodes, at each node the field will be related to its neighbours. In frequency domain methods this is calculated for each frequency, while in time-domain methods, the equations are propagated in time until the field energy has decayed away from the simulation domain, and the solution Fourier transformed to get spectral information.

Because the whole 3D domain must be discretized, the computational time scales with the spatial step δ_x as $\sim \frac{1}{\delta_x^4}$ for 3D simulations[77]. To make matters worst, the computational domain

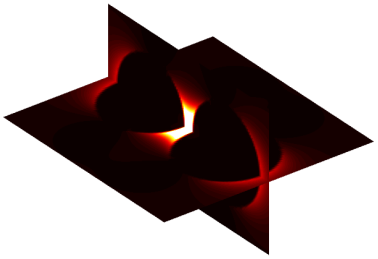
[75] In practise if the mesh size is reduced too far, roundoff errors will start to accumulate due to the increased number of arithmetic operations.

[76] For instance convergence is very hard to achieve for small real refractive indexes, unfortunately this corresponds to low loss systems which are interesting in plasmonics, such as silver. As I have discovered myself, through hours of suffering, ultra-high quality factor materials like SiC in the reststrahlen band can be next to impossible to converge with the FDTD methods!

[77] It goes to the 4th inverse power as the time step also must be proportional to the spatial step via the Courant number $c\frac{\delta_t}{\delta_x}$, which must be less than or equal to 1.

must often be quite large to prevent spurious interactions with the boundary, typically the structure must be at least $\sim \lambda/2$ away from the boundary[78]. It is thus a computationally expensive method; fortunately advanced codes, such as the frequently used commercial package Lumerical (Lumerical Inc.), offer non-uniform meshes as well as sub-cell techniques to take into account curved surfaces, which can lighten the computational burden somewhat. A strength of the FDTD method is its generality, there is no constraint on structure geometry and by changing the boundary conditions various different types of systems can be studied: open domain (absorbing boundary conditions), cavity (perfect electric/magnetic conductor boundary conditions) and periodic. This includes both individual photonic/plasmonic elements as well as metamaterials and photonic crystals. It can also be simply extended to include more complicated material response such an nonlinearity and nonlocality.

[78] This can be rectified partially by using intelligent absorbing boundary conditions.



The electric field intensity enhancement between two silver dimers of diameter 22 nm and gap 2.5 nm excited at wavelength of 532 nm. Calculated using FDTD and used in the following publication (Gisbert-Quilis et al., 2017).

It is also worth noting that FDTD is the go-to tool to study dynamics, which is not possible beyond trivial situations with frequency-domain solvers. This is particularly interesting when combined with quantum dynamical equations, such as the Maxwell–Bloch equations. This can be implemented in Lumerical using the material-plugin via a C++ script.

1.3.2 *Boundary Element Method*

The boundary element method (BEM) (De Abajo and Howie, 2002; Hohenester and Trügler, 2012) is an excellent tool for studying systems in open domains and calculating quantities such as the optical cross sections or near-fields of individual plasmonic elements. In contrast to finite-difference simulations, only the surface is discretized meaning matrices of order N^2 need to be manipulated as opposed to N^3 for a volume based method. This cuts simulation time for challenging simulations significantly, so where possible, the BEM approach should favoured. The codes available to researchers tend to be frequency domain, such as the excellent Matlab toolbox by A. Trügler and U. Hohenester (Hohenester and Trügler, 2012; Hohenester, 2014; Waxenegger et al., 2015) which I have used extensively in this thesis[79].

[79] The code can be downloaded from `http://physik.uni-graz.at/mnpbem/`

To understand the principles behind the BEM we can consider the quasistatic case and show how the solution to Poisson's equation can be transformed into a surface integral problem. Previously we considered solving Laplace's equations in separate domains separated by discontinuous material interfaces, this is completely equivalent to considering the domain as a whole and solving Poisson's equation, $\nabla^2\phi(\mathbf{r}) = -\rho(\mathbf{r})/\epsilon_0$, where the charge density is localised to an infinitesimal surface layer at the boundaries. This means the scalar potential, in medium i, is given by a surface integral equation on the surface S which separates two media

$$\phi_i(\mathbf{r}) = \oint_S d^2\mathbf{r}' \; G(\mathbf{r}-\mathbf{r}')\sigma(\mathbf{r}') + \phi_{ext}(\mathbf{r}). \tag{1.46}$$

Here $G = \frac{1}{4\pi\epsilon_0}\frac{1}{|\mathbf{r}-\mathbf{r}'|}$ is the Green's functions for the Poisson equation. The BEM equations can be found by applying the EM boundary conditions at the interface, see appendix B. The continuity of the tangential component of the electric field at the surface just means that the surface charge is equal on either side of the boundary. To evaluate the continuity of the normal component of the electric displacement at the surface we need to derive the derivative of the electric potential in the direction normal to the surface

$$\lim_{\mathbf{r}\to\mathbf{s}} \hat{\mathbf{n}}\cdot\nabla\phi(\mathbf{r}) = \partial_n\phi(\mathbf{s}) = \lim_{\mathbf{r}\to\mathbf{s}}\left(\partial_n \oint_S d^2\mathbf{s}'\; G(\mathbf{r}-\mathbf{s}')\sigma(\mathbf{s}') + \partial_n\phi_{ext}(\mathbf{r})\right)$$
$$= \pm 2\pi\sigma(\mathbf{s}) + \oint_S d^2\mathbf{s}'\; F(\mathbf{s},\mathbf{s}')\sigma(\mathbf{s}') + \partial_n\phi_{ext}(\mathbf{s}), \tag{1.47}$$

for details of the last equality see (Trügler, 2011), the $\mathbf{r}\to\mathbf{s}$ involves some thought because of the singularity in the Greens' function and we have defined $F(\mathbf{s},\mathbf{s}') = \hat{\boldsymbol{n}}\cdot\nabla G(\mathbf{s}-\mathbf{s}')$. The surface can then the discretised, with each element having a constant surface charge within. By writing in a compact vector notation and applying the continuity of the normal component of the displacement field, we obtain

$$\boldsymbol{\sigma} = -\left(2\pi\frac{\epsilon_2+\epsilon_1}{\epsilon_2-\epsilon_1}\overline{\mathbf{1}} + \overline{\boldsymbol{F}}\right)^{-1}\partial_n\boldsymbol{\phi_{ext}} \tag{1.48}$$

The main computational cost is the inversion of the matrix, note that the retarded version is more complicated and consequently has larger simulation times (Trügler, 2011). Two pleasing features of the above equation are the decoupling of material and structural properties (described by the matrix $\overline{\boldsymbol{F}}$), and that the excitation (described by the vector $\partial_n\boldsymbol{\phi_{ext}}$) need only be modelled at the surface. This has computational benefits, compared to the FDTD method where one must discretise a volume on the order of λ^3 while simultaneously needing to resolve deep sub-wavelength features; in BEM we need only resolve the sub-wavelength features. It also helps when modelling more exotic incident beams such as twisted light; in FDTD the beam must be propagated before it interacts with the nanostructure, interactions with the simulation boundary can lead to the beam breaking up after a short distance[80]. Because of this, the BEM has recently been used to model the interaction of plasmonic nanostructures with twisted light (Kerber et al., 2017).

[80] This is exacerbated by the rectangular grid in Lumerical, the problem can be solved somewhat by using a cylindrical grid which matches the symmetry of the beam.

An example of a typical BEM calculation is provided by figure 1.15, which shows the extinction, absorption and scattering cross sections of a prolate gold spheroid. The calculation solves the full Maxwell's equations, so includes retardation effects, which is not possible by analytical means. The L-mode is seen to be much stronger than the T-mode, this is because it occurs at longer wavelengths away from interband transitions which heavily damp the T-mode. See also figure 1.12.

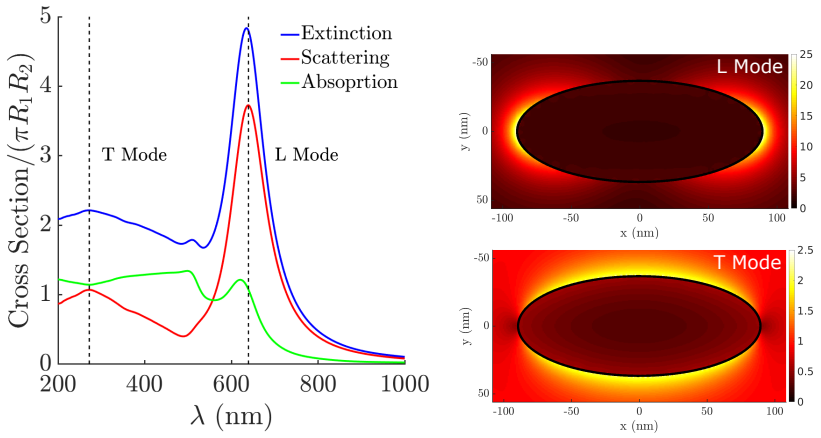


Figure 1.15: BEM calculation of a gold prolate ellipsoid with $R_1 = 90$ nm and a fixed volume equal to that of a 50 nm radius sphere. Indicated by the dotted black lines are the T-mode and L-mode. Dielectric data used for gold from (Johnson and Christy, 1972).

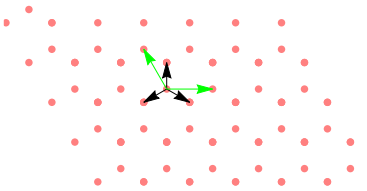
1.3.3 *Generalised Mie Theory Codes*

The GMM theory, an extension of Mie theory to an arbitrary number of spheres, only has limited areas where it can be used, but in these cases it is an extremely powerful method: quick and formally exact, it includes multipolar and retardation effects. Much like ordinary Mie theory for a single sphere, the equations are tedious and dense, so I will not even contemplate presenting the arduous derivation here. It is an old field of research, beginning in the 1930s. Historical details can be found in (Quinten, 2010). It has many obvious areas of applications, such as for studying nanoparticles in colloidal suspensions. The generalisation of Mie theory to multiple scatterers needs the introduction of translation rules for spherical harmonics as now spheres off-coordinate-centre must be considered. Furthermore, the field experienced by a particle is now from the external light *and* the scattered field from all other spheres, this complicates the problem as vector spherical harmonics (VSHs) with $m \neq \pm 1$ will now contribute[81]. The problem can be numerically challenging as the scattered field expansion in VSHs leads to much higher multipoles than are necessary with single particle Mie theory. Several excellent open-source GMM codes are available online (Pellegrini et al., 2007; Mackowski and Mishchenko, 2011).

[81] See appendix F for an introduction to VSHs.

1.4 Polaritonics In Other Materials

While Plasmonics is associated with metal optics, plasmonic type behaviour is in fact found in many different materials as long as they exhibit a negative real part of the permittivity. Here we discuss two such materials of current interest in nanophotonics: graphene and polar dielectrics. At the end a quick overview is given of current materials of interest in nanophotonics.



Graphene's crystal structure can be described by a Bravais lattice with two atoms in a basis. The green arrows show the unit cell vectors and the black arrows show the nearest neighbour vectors.

1.4.1 Graphene

Graphene is a truly 2D material, or at least as close as one can ever get, being one-atom thick, and due to its hexagonal crystal structure supports unique electronic and optical behaviour. This has led to it being one of the hottest research topics in physics over the past decade. When carbon atoms are brought together they can have a unique bonding configuration called sp^2 or trigonal hybridisation. In this case, the small energy splitting between the $2s$ and $2p$ states (~ 4 eV), allows the promotion of a $2s$ electron to a $2p$ state[82], which allows three sp electrons to form a hybrid covalent bond in the plane (Haken and Wolf, 2013). These form the σ band and are responsible for the lattice structural strength in all carbon allotropes, they are also a filled shell and hence form a deep valence band and so can often be ignored when considering the electronic properties[83]. The remaining p orbital points perpendicular to the plane and forms linear combinations with neighbouring p orbitals, this leads to a half-filled π band. The electrons occupying this p orbital can be considered to be delocalised over the entirety of the molecule/crystal. The charge carriers behave as *massless Dirac fermions*, due to the special band structure. The dispersion of the π band meet at a symmetry point, near the corners of the hexagonal Brillouin zone, and the region around this point has a linear dispersion

$$E_k = \hbar v_F k, \tag{1.49}$$

where $v_F \approx c/300$ is the Fermi velocity. This region of the band structure is known as the *Dirac cone* and leads to very different electronic behaviour compared to conventional electron gasses with parabolic dispersion, An accurate analytical expression for the full dispersion can be found using the tight-binding model, see figure 1.16. Undoped graphene is a semi-metal, a hybrid type of material between a metal and an insulator; it has a vanishing density of states at the Dirac point and no bandgap (Kotov et al., 2012).

Despite its novel electronic behaviour, graphene's optical response can be well modelled using the Drude-Lorentz formula. This can be derived from the Lindhard response function, for our

[82] This can be facilitated by the chemical environment; external forces can make the $2s$ and $2p$ states approximately degenerate in energy. In other words, the energy gained from the covalent bond compensates the energy cost of the electron promotion.

[83] At least at the low energies we care about.

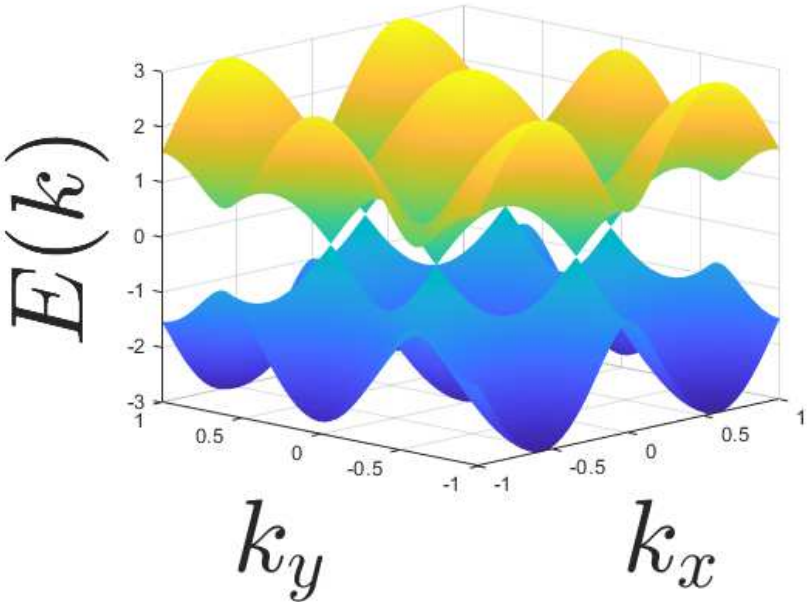
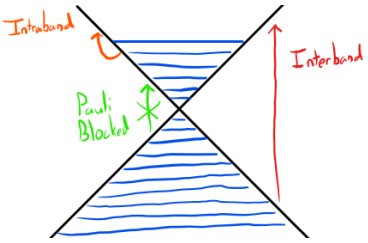


Figure 1.16: Energy-momentum dispersion of graphene calculated within the tight-binding model using the nearest neighbour approximation. The energy is normalised to the hopping energy and the wavenumbers are normalised to $2\pi/a$ where a is the nearest neighbour lattice distance.

purposes we need only the local and the low electronic-temperature approximation ($E_F >> k_bT$), so we can write the 2D dynamical conductivity as (Gonçalves and Peres, 2016)

$$\begin{aligned} \sigma(\omega) &= \sigma_{intra}(\omega) + \sigma_{inter}(\omega) \\ \sigma_{intra}(\omega) &= \frac{4i\sigma_0 E_F}{\pi\hbar(\omega + i\gamma)} \\ \sigma_{inter}(\omega) &= \sigma_0\left[\Theta(\hbar\omega - 2E_F) + \frac{i}{\pi}\log\left(\left|\frac{\hbar\omega - 2E_F}{\hbar\omega + 2E_F}\right|\right)\right] \end{aligned} \tag{1.50}$$

where $\sigma_0 = \frac{\pi e^2}{2h}$ and Θ denotes a step function which is $0(1)$ in the case $\hbar\omega$ is below(above) $2E_F$, physically this describes Pauli blocking and is illustrated in the margin figure. The intraband part, σ_{intra}, can be derived from the semiclassical equations of motion for electrons[84]. A shortcoming of the semiclassical model is that the band index is a constant of motion and so it does not account for interband transitions, so we supplement it with an interband transition model, which can be derived using Fermi's golden rule. The dielectric function can be found simply from (Falkovsky, 2008): $\epsilon(\omega) = \epsilon_\infty + \frac{i\sigma(\omega)}{\epsilon\omega\delta}$, where δ is the thickness of the graphene sheet and is introduced as σ is a 2D conductivity[85]. A plot of the conductivity and permittivity for typical parameters are presented in figure 1.17. This model has been found to work extremely well in plasmonics. Due to the zero bandgap of intrinsic graphene, both the intraband and interband terms can be important to include, depending on the frequency of the incoming light and the Fermi energy of the graphene. It should be noted that there is significant interest in the nonlocal generalisation of the conductivity, a recent experiment has utilised plasmonic excitations to probe the nonlocal properties of the electron gas in graphene (Lundeberg



Possible interband (red) and intraband (orange) transitions for graphene. The mechanism of Pauli blocking is also indicated(green).

[84] This model can be viewed a generalisation of the classical model we presented for the derivation of the Drude model, to the case with arbitrary band structure (Ashcroft and Mermin, 1976).

[85] Note that in simulations it is possible to use a larger thickness than graphene's physical thickness 0.335 nm, as long as it remains much smaller than the wavelength (Vakil and Engheta, 2011). Also note that this requires an anisotropic 3D permittivity model; so for a graphene sheet located in the xy plane $\epsilon_{xx}(\omega) = \epsilon_{yy}(\omega) \neq \epsilon_{zz}$.

et al., 2017).

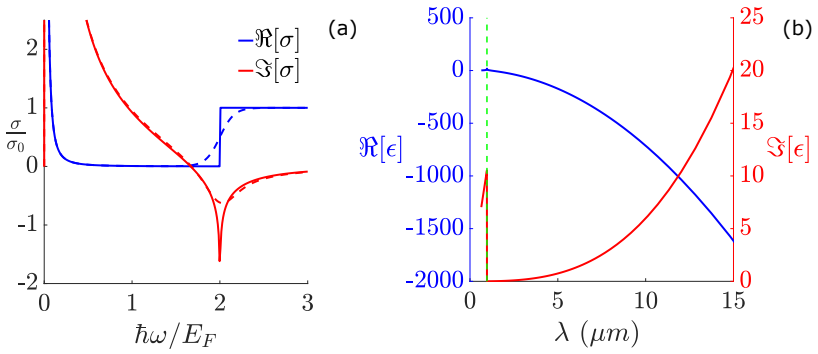


Figure 1.17: (a) The 2D conductivity of a graphene sheet with $E_F = 0.64$ eV and $\gamma = 1$ meV (corresponds to a mobility of 10000 $cm^2V^{-1}s^{-1}$). The solid lines correspond to the zero-temperature limit and the dashed line are for a temperature of 300 K.

For graphene with a large doping, the Drude component will dominate the optical response and so it comes as no surprise to learn that graphene can support plasmons. Further, given graphene's unique band structure, it has a number of advantages compared to traditional metal plasmonics:

- Graphene has very high carrier mobilities due to a low electronic density of states and weak electron-phonon coupling, this leads to low losses and hence large propagation distances for graphene SPPs. Thanks to state-of-the-art encapsulation methods using boron nitride (Wang et al., 2013), the theoretical limit of the sheet resistivity can be approached[86]. For the plots in this section I use a conservative value for the mobility of 10000 $cm^2V^{-1}s^{-1}$, which corresponds to typical value for the mobility of chemical-vapor-deposition-grown graphene.

- Graphene plasmons also exhibit a high field concentration, with SPPs strongly confined to the surface and a plasmon wavelength much smaller than the free-space wavelength[87], this is ideal for enhancing light-matter interactions. Physically this can be understood by considering graphene as the short-range SPP in the limit of an ultra-thin metallic sheet, see appendix C.

- The low concentration of free carriers, relative to metals, allows excitations of SPs in the THz to the mid-IR. In this technologically important spectral range, metal SPPs are essentially light-like and exhibit weak field confinement. Actually, this is only true for simple metal-dielectric interfaces, more complicated geometries can achieve strong EM energy localisation for frequencies well below ω_{SP} with metals. For instance the metal-dielectric-metal interface geometry can support a surface mode which does not have a cut-off for decreasing dielectric-thickness; this allows, at least within the classical approximation, arbitrarily small confinement at long wavelengths by making the dielectric layer very thin (Maier, 2006). Of course this involves potentially complicated structuring of metals so the more simple graphene structure is advantageous.

[86] The minimum limit to the loss is imposed by acoustic phonon scattering.

[87] Plasmon wavelengths 100 times smaller than the free-space wavelength have been demonstrated experimentally (Brar et al., 2013).

- Because of its 2D nature, graphene is much more amendable to doping than 3D metals[88], allowing the electron or hole concentration, and hence the Fermi energy, to be tuned. This can be done via chemical doping (Yan et al., 2013), optically (Ni et al., 2016) or electrostatic gating (Ju et al., 2011; Wang et al., 2013). Optical gaps of up to 2 eV, which corresponds to $E_F \sim 1$eV, can be created using ion-gel gating (Chen et al., 2011). This allows the plasma frequency and loss to be modified, and graphene plasmons can be excited over a wide frequency range from THz to the near-IR (Tielrooij et al., 2015). There is optimism that it may well be possible to extend this into the visible with hybrid structures (Yu et al., 2016; de Vega and García de Abajo, 2017). Further opportunities for tuning of optical properties may be possible via strain engineering (Ni et al., 2014) and application of magnetic fields (Yan et al., 2012b).
- The role of any substrate is hugely important due to the 2D structure of graphene, while undoubtedly a problem a lot of the time, it can also be hugely useful for some applications; a neat example of this was demonstration of using a ferroelectric substrate to permanently modify graphene's plasmonic response (Goldflam et al., 2015).

[88] In a bulk metal subjected to an electrostatic potential, electrons will rearrange themselves on the surface and all electron states will uniformly shift. This lack of relative shift between states means that the Fermi energy cannot be changed relative to the band minimum and hence the volume density of electrons and consequently the optical response are unchanged. In contrast, 2D crystals are effectively just surfaces, so additional electrons on the surface can shift the Fermi energy (Jablan et al., 2013).

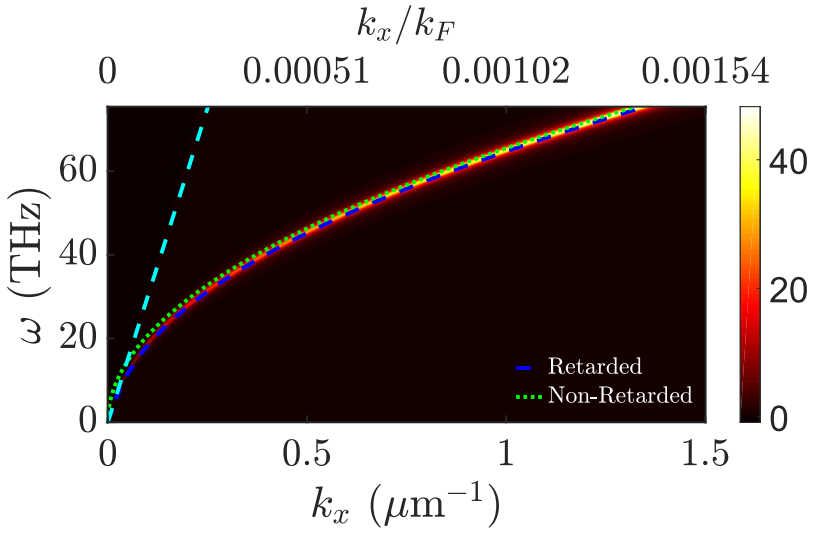


Figure 1.18: The graphene SPP dispersion for $E_F = 0.64$ eV and $\gamma = 1$ meV, superimposed on top of the imaginary part of the reflection Fresnel coefficient for TM polarized light. The blue dashed line shows the full solution to equation 1.51, while the green dotted lines shows the non-retarded approximation, equation 1.52, and the cyan dashed line shows the light line. (b) The real and imaginary part of the corresponding permittivity.

Graphene also has exceptional mechanical properties, having a high tensile strength and stiffness (Booth et al., 2008) and can sustain large deformations (Liu et al., 2007). It also has unique chemical properties as every atom is available for chemical reactions from both above and below. Graphene thus offers an excellent platform for 2D THz/mid-IR nanophotonic circuits and sensors. One should not get too carried away though, the propagation length of graphene SPPs is poor, for instance at IR wavelengths it is at best only a few SPP wavelengths, which may severely limit possible applications (Tassin et al., 2012). Further work is needed to see if these

limitations can be overcome, potentially by suitable structuring and combination with other photonic materials, see chapter 3.

Experimental evidence of graphene plasmons was hampered by the large momentum mismatch between graphene SPPs and incident light. The first confirmations were obtained using EELS (Liu et al., 2008; Liu and Willis, 2010; Koch et al., 2010) and then later complemented by results obtained from scanning near-field microscopy (Fei et al., 2011, 2012; Chen et al., 2012) and optical far-field spectroscopy using structured graphene (Ju et al., 2011). Graphene plasmons have now been demonstrated in a number of different periodically structured geometries including: ribbons (Ju et al., 2011), disks (Yan et al., 2012a,b) and rings (Yan et al., 2012c). Localised graphene plasmons have also been experimentally established (Vasić et al., 2013).

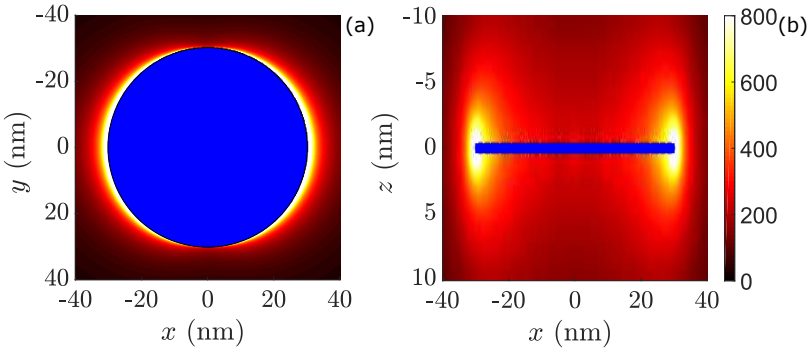


Figure 1.19: Electric FE for a graphene nanodisk shown for (a) top-down and (b) side view. Parameters of $E_F = 0.64$ eV and $\gamma = 1$ meV were used.

In similar fashion to the derivation for SPPs on metal-dielectric interfaces, it is possible to derive the condition for graphene SPPs. This is presented in appendix C, the final result is given by

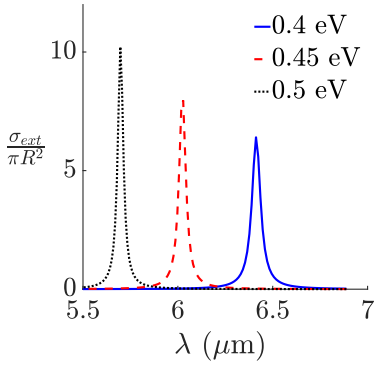
$$\frac{k_{z,1}}{\epsilon_1} + \frac{k_{z,2}}{\epsilon_2} + \frac{i\sigma(\omega)}{\epsilon_0 \omega} = 0, \tag{1.51}$$

where media ϵ_2 and ϵ_1, lie above and below the graphene sheet. This equation is just a generalisation of equation 1.23 to the situation where there is a surface current density in-between two media, it is an implicit equation for ω and k_x and must in general be solved numerically. An analytical expression is possible if we assume the non-retarded limit, $k_x >> \sqrt{\epsilon}\omega/c$, (Jablan et al., 2013)

$$k_x \approx i\frac{\omega(\epsilon_1 + \epsilon_2)\epsilon_0}{\sigma(\omega)}. \tag{1.52}$$

We compare these two expression in figure 1.18 superimposed on the imaginary part of the reflection Fresnel coefficient for TM polarized light. We see that for the parameters chosen, $E_F = 0.64$ eV and $\gamma = 1$ meV, the non-retarded approximation works extremely well away from the light-line. Thus, for large doping, it is usually permissible to neglect retardation effects.

For structured and finite-sized graphene sheets, analytical solutions are not normally possible. Some of the most common systems



The extinction cross section for graphene nanodisks of radius 30 nm and mobility of 10000 $cm^2V^{-1}s^{-1}$, for varying Fermi energy.

to consider is nanoribbons and nanodisks, the results of a BEM calculation for a 30 nm radius graphene nanodisk is shown in figure 1.19. Note that the graphene structure is indicated by the blue region, the internal field is not shown as it is not completely clear how one would treat the internal structure within classical electrodynamics. The FE is on the order 10^2, and in this case approaching 10^3, which is well over an order of magnitude larger than for LSPs in individual metal nanospheres/nanodisks. This compares well to previous values reported in the literature (Thongrattanasiri et al., 2012).

1.4.2 *Polar Dielectrics*

It is well-known that polar dielectric crystals[89] exhibit a high reflectivity in a material-dependent spectral window known as the *reststrahlen band*, which is bounded by the transverse optic (TO) and longitudinal optic (LO) phonon frequencies (Kittel, 1966; Ashcroft and Mermin, 1976), at energies ω_{TO} and ω_{LO} respectively. In crystals with N atoms per primitive cell, the phonon dispersion develops $3N$ branches: $3N-3$ optical and 3 acoustic[90]. The optical phonons correspond to out-of-phase oscillations of the relative position of the atoms and so the transverse variety can couple to EM waves[91]. The permittivity of a polar dielectric around the reststrahlen band is well described by the Lorentz model

$$\epsilon(\omega) = \epsilon_\infty \left(1 + \frac{\omega_{LO}^2 - \omega_{TO}^2}{\omega_{TO}^2 - \omega^2 - i\omega\gamma}\right), \tag{1.53}$$

which clearly reveals that between ω_{TO} and ω_{LO} the real part of the permittivity will go negative and hence can support surface modes[92]. These are known as SPhPs and are coherent oscillations of the ionic charges coupled with external light, and behave in many ways analogously to plasmon polaritons in metals. One of the great benefits compared to plasmonics is their reduced damping rates; this is because scattering time of optical phonons is on the order of picoseconds as compared to $10-100$ femtoseconds for metals (Caldwell et al., 2015). This leads to a small $\Im[\epsilon]$ and hence large quality factors, over an order of magnitude larger than metals, and a tight field confinement that has been demonstrated experimentally (Greffet et al., 2002; Hillenbrand et al., 2002), with quality factors over 100 found and confinement, $\lambda^3_{LSPhP}/V_{eff}$, up to 200 (Caldwell et al., 2013; Chen et al., 2014).

Given the immense interest in combating loss in plasmonics, there is obvious excitement at the possibility of using polar dielectrics for optical *phononics*: a low-loss substitute for plasmonics. While perhaps not quite the holy grail of nanophotonics, for some applications polar dielectrics offer an attractive alternative to metals, albeit over a small frequency window corresponding to the reststrahlen band. This necessarily means that there is only nar-

[89] A material is polar if its crystal structure is constructed with atoms of different electronegativities which ionically bond.

[90] Acoustic phonons are characterised by a linear dispersion at long wavelengths and tend to zero energy. In contrast, optical phonons have a flat dispersion at a finite energy for large wavelengths

[91] Note that the LO frequency appears in equation 1.53 but this does not mean it couples to light, it appears in the same way that the plasmon frequency appears in the Drude model despite the fact that volume plasmons do not couple to light. Mathematically, it pops up from use of the *Lyddane-Sachs-Teller relation*, $\frac{\omega_{LO}^2}{\omega_{TO}^2} = \frac{\epsilon(0)}{\epsilon(\infty)}$, to rewrite the permittivity in accessible parameters (Kittel, 1966; Ashcroft and Mermin, 1976)

[92] This is actually quite a general feature of Lorentz-type permittivities. For example, excitonic materials, such as organic dyes, can demonstrate metallic behaviour over a small range of wavelengths if the oscillator strength is large enough (Núñez Sánchez et al., 2016).

row frequency separation between epsilon-near-zero (at ω_{LO}) (Alu et al., 2007) and epsilon-near-pole (at ω_{TO}) (Molesky et al., 2013) behaviour. Furthermore, similar to the case of graphene plasmonics, phononic resonances occur in the mid-IR/THz, this is a technologically important part of the spectrum that corresponds to the vibrational and rotational transitions of molecules, which can be used to 'fingerprint' them. There is also an important atmospheric transparency window in-between 8 to 12 μm, and blackbody radiation peaks in this zone[93]. There is also the possibility for the coupling with graphene plasmons, which is explored in chapter 3. As noted in section 1.4.1, metal-dielectric structures can be constructed to give large confinements in the IR/THz, using SPhPs one can achieve similar figures of merit with reduced complexity. Also removing metals from the fabrication process can help improve CMOS compatibility (Caldwell et al., 2015).

In this work we will be predominately interested in the case of SiC. It has a diatomic crystal structure and so there is 3 optical phonons: 1 LO and 2 TO. It holds considerable promise for electronic, thermal and optical devices due to excellent mechanical strength and chemical stability, It can be doped and support free carriers, which adds a Drude contribution to equation 1.53 (Dunkelberger et al., 2018), but this will not be considered in this thesis. SiC has a relatively large splitting between the LO and TO phonons, due to a large polarity of the interatomic bonding (Caldwell et al., 2015).

[93] SPhP resonators have been proposed as ideal narrow-band thermal emission sources (Greffet et al., 2002; Schuller et al., 2009).

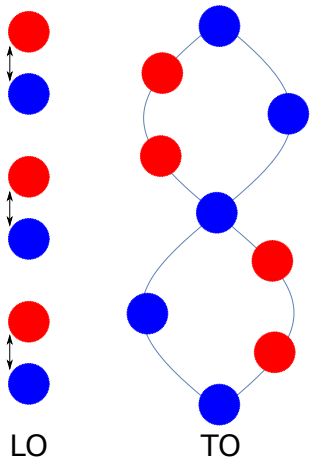


Illustration of the LO and TO phonon modes in a 1D diatomic chain.

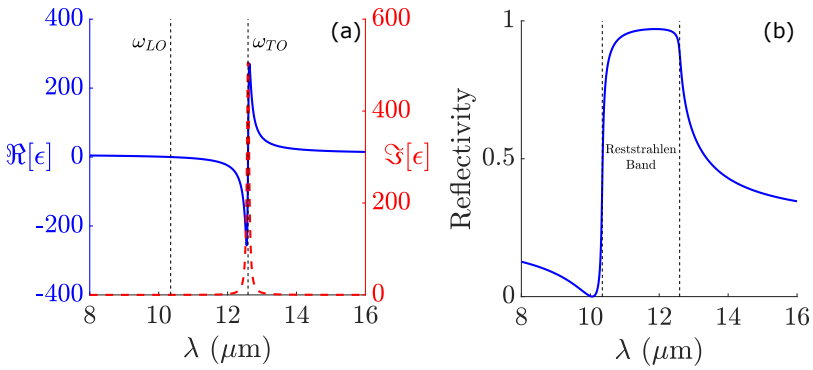


Figure 1.20: The permittivity (a) and the reflectivity (b) of undoped SiC, using equation 1.53 with the parameters $\omega_T = 0.0988$ eV, $\omega_L = 0.120$ eV, $\epsilon_\infty = 6.56$ and $\gamma = 0.00059$ eV (Francescato, 2014).

1.4.3 *Other Promising Polaritonics Materials*

The discovery of plasmonic behaviour in graphene was not only an exciting discovery in itself, it has also paved the way for a wealth of new exotic 2D crystal materials which may provide the building blocks of next-generation ultra-compact nanophotonic devices. Examples include:

- Hexagonal boron nitride is a polar dielectric and has a natural hyperbolic response; it is a van der Waals bonded material meaning that the in-plane and out-of-plane dielectric response will

have different signs over certain wavelength ranges (Jacob, 2014). Its optical properties can be altered by varying the thickness and SPhP wavelengths 25 times smaller than free-space have been demonstrated with much lower losses compared to graphene, allowing propagation lengths of up to 20 times larger (around 5 to 10 μm) (Dai et al., 2014). Away from its *two* reststrahlen bands, it acts as an excellent insulating layer due to a large bandgap of 6 eV (Xia et al., 2014), this has led to many applications as a substrate or encapsulation layer for other 2D materials.

- Single-layer transition metal dichalcogenides[94], such as molybdenum disulfide which is a direct bandgap semiconductor with a gap of $\sim$ 2 eV, exhibit strong and long-lived exciton-polaritons. This is because of the decreased dielectric screening due to the reduced dimensionality. Quantum confinement leads to a strong dependence on number of layers. Furthermore, the lack of inversion centre in the crystal structure and strong spin-orbit coupling leads to valley-selective circular dichroism (Xia et al., 2014). They can support a field confinement over an order of magnitude greater than graphene, albeit with greater loss (Low et al., 2017).
- At the time of writing, the current hot new 2D crystal is black phosphorus, which has a small direct bandgap of around 0.3 eV and thus provides a convenient middle ground between zero bandgap graphene and large bandgap transition metal dichalcogenides (Xia et al., 2014). The unique puckered[95] crystal structure leads to an interesting *in-plane* anisotropy and hence a linear dichroism.
- Other elements in group 14 of the periodic table can also form 2D crystals such as silicene, germanene and stanene (tin). They can be expected to display interesting electronic and optical properties, unfortunately so far they have only been grown on metallic substrates which strongly modify their properties (Castellanos-Gomez, 2016). The development of encapsulation methods should lead to new photonic applications and allow the study of their intrinsic properties.

[94] They have a general formula MX_2, where M is a transition metal and X is a chalcogen, i.e. in group 16 of the periodic table: oxygen, sulphur, etc. They are probably, after graphene, the most studied family of 2D crystals (Castellanos-Gomez, 2016).

[95] The phosphorous atoms bond to one another via sp^3 hybridisation in contrast to sp^2 for graphene (Castellanos-Gomez, 2016).

Particularly exciting is the emerging concept of combining such materials in hybrid structures, which combine the strengths of the constituent materials and (hopefully) negate some of the weaknesses. 2D materials are self-passivated without any dangling bonds, this means 2D crystals with very different lattice constants to be combined (Xia et al., 2014). For example, combining monolayer graphene and multilayer hexagonal boron nitride can lead to tunable, broad-bandwidth and low-loss optical modes (Caldwell et al., 2016).

Also worth mentioning is some more traditional alternatives to metals in plasmonics:

- Doped semiconductors support plasmons over a wide range from the IR to THz via a very tunable carrier density, for small

nanocrystals the addition of only a few electrons can have a huge impact on the LSP resonance (Faucheaux et al., 2014). This dynamic tunability could be very useful for active plasmonic devices. They are also useful systems to study quantum plasmonics and the LSP resonance provides a useful optical probe of semiconductor properties.

- Metal alloys has been proposed as a way to achieve high quality factors and on-demand optical responses (Gong and Leite, 2016).
- Superconductors can exhibit an interesting dynamical response to external EM fields (Robson et al., 2017). They have been used in microwave metamaterials, unfortunately pushing their operation to higher frequencies is unwise as the photons have enough energy to break up the Cooper pairs which give superconductors their unique properties (Singh and Zheludev, 2014). High-temperature superconductors with large bandgaps have been proposed as a potential solutions but seem unlikely to offer superior performance to normal metals (Tassin et al., 2012).

For some applications in nanophotonics one can often do away with plasmonics altogether and use high-refractive-index all-dielectric devices instead. For instance, all-dielectric metamaterials, made from transparent meta-atoms, can achieve an EM response in all the four quadrants[96] (Jahani and Jacob, 2016). Modest FEs are possible, with the advantage of much smaller dissipation. In contrast to plasmonics, enhancement of the magnetic field is also possible (Bakker et al., 2015). Hybrid metal-dielectric devices hold a lot of promise as they can potentially give tailored devices with quality factors ranging from plasmon-like (~ 10) to dielectric-like ($\sim 10^3$) (Yang et al., 2017).

Lastly, it is important to mention the growing use of electron structure methods to predict the suitability of materials for photonic applications. This has recently been demonstrated for topologically insulating quaternary trichalcogenide compounds by using first-principles DFT analysis of the dielectric function (Yin et al., 2017). Another example is the theoretical design of band structure engineered van der Waals layered metals, with suppression of optical losses by reducing the density of states for electron scattering (Gjerding et al., 2017). This emerging combination of fundamental material research with plasmonics can be expected to become increasingly important and will help develop next-generation photonic devices.

[96] The four quadrants refer to the four possible combinations of signs for μ and ϵ. Note that a dielectric metamaterial is fundamentally different than a photonic crystal as it relies on near-field coupling between subwavelength meta-atoms.

2 Quantum Plasmonic Theory

2.1 Why Quantum Theories Of Plasmons Are Necessary

Here we discuss why quantum plasmonics is necessary when considering very small metallic nanoparticles and how current theoretical activity has been motivated by some remarkable experiments over the past decade.

For small nanostructures, which for metals means the characteristic length of the structure is below ~ 5 nm, the electrons will begin to 'feel' their confinement and form a discrete energy spectrum like that of a molecule. This has been observed experimentally (Knight et al., 1984; Zheng et al., 2007). Typically, energy level discreteness can be ignored in plasmonics as the high electron density in metals means very small energy spacing at room temperature (Maier, 2007). This can be contrasted to the case of semiconductor quantum dots, which exhibit quantum effects at larger sizes, relative to metals, due to their much lower electron concentration. Related to this is a second size effect, which originates from the growing importance of the surface[1]. The ratio of the surface to the volume scales with R^{-1}, meaning that for smaller nanoparticles the surface region increasingly dominates the optical response and so it must be modelled accurately for valid results. In particular, the behaviour of the electron density at the surface must be treated correctly and it is not appropriate to use hard wall boundaries[2] that don't allow electron spill-out, as was used in many early theoretical calculations on the problem.

[1] In a quantum picture this is just the boundary conditions for the electronic wavefunctions

[2] i.e. infinite potential well models.

The theory of confined electrons in metals has a long history, arguably starting with Kubo who was the first to explore in detail how thermodynamic properties are modified by the discrete energy level structure of small metal nanoparticles (Kubo, 1962). This helped inspire many similar 'particle-in-a-box' type models (Gor'Kov and Eliashberg, 1965; Rice et al., 1973; Lushnikov and Simonov, 1974; Dasgupta, 1977; Wood and Ashcroft, 1982; Penzar and Šunjić, 1984; Huang and Lue, 1994). In particular, it seems that Rice *et al* were the first to perform a calculation of the polarizability of small particles that took electron screening into account self-consistently (Rice et al., 1973). An electron in a nanostructure or

molecule does not directly experience the applied field, instead it experiences a screened field which is the sum of the external and induced field. This mean-field approach forms the basis of DFT and will feature heavily in our discussions in this chapter.

The current interest in quantum size effects in plasmonics has been a reaction to a number of groundbreaking experiments in the last decade that have probed the quantum plasmonics regime (Scholl et al., 2012; Ciracì et al., 2012; Savage et al., 2012; Raza et al., 2013a; Scholl et al., 2013). In particular, it was around 2012 that quantum plasmonics as a field of research really started to explode, thanks in part to three especially pioneering experiments that deserve special mention:

- Scholl *et al* studied *individual* ligand-free silver nanoparticles from sizes of 20 $\rightarrow$ 2 nm, using EELS (Scholl et al., 2012). They found typical nonlocal effects as the particle size is decreased: a blueshift and a broadening of the SP peak, see figure 2.1(a).
- Ciracì *et al* used chemically deposited sub-nanometre molecular layers to precisely control the separation between gold spheres and a gold film (Ciracì et al., 2012), see figure 2.1(c). This allowed a detailed study of the effects of nonlocality on the FE in small gaps, it was found that the HM can give excellent predictions down to gap-sizes of 1 nm.
- Savage *et al* measured the electrical and optical response between two almost-touching gold tips of an atomic force microscope (Savage et al., 2012), see figure 2.1(b). This allowed sub-nanometre control down to the quantum tunnelling regime where quantum charge transfer plasmons (CTPs) were observed. The quantum-corrected model (Esteban et al., 2012) was used to explain the results. Similar results were found by Scholl *et al* using EELS to study the plasmonic properties of two closely spaced 5 nm radius silver nanospheres, with the separation controlled by the electron beam from a scanning transmission electron microscope (Scholl et al., 2013).

There has been particularly heavy interest in quantum plasmonic dimers due to potential in technological applications and as an ideal testing ground for quantum plasmonic theory. It is possible to distinguish three regimes based on the gap distance: capacitative coupling (classical), a nonlocal regime ($\sim$ 2 $\rightarrow$ 1 nm) and a tunnelling regime ($\lesssim$ 1 nm). Note that the lattice constant of gold and silver is around 0.4 nm, so for any gap less than 1 nm the distinction between separated and merged nanoparticles becomes hazy[3]. At large gap distances, classical electrodynamics works well and predicts a bonding dipole plasmon (BDP) which redshifts for decreasing gap size[4]. The BDP is associated with a large FE for small gaps, and is predicted to increase *ad infinitum* for decreasing gaps by purely classical models. In reality, nonlocal screening will put a limit on the maximum FE at some finite gap

[3] This length scale is dictated by the electron spill-out length.

[4] A BDP arises from the Coulomb interaction coupling two plasmons, leading to charges of an opposite sign on either side of the gap. The structure has an overall dipole moment so couples strongly to light.

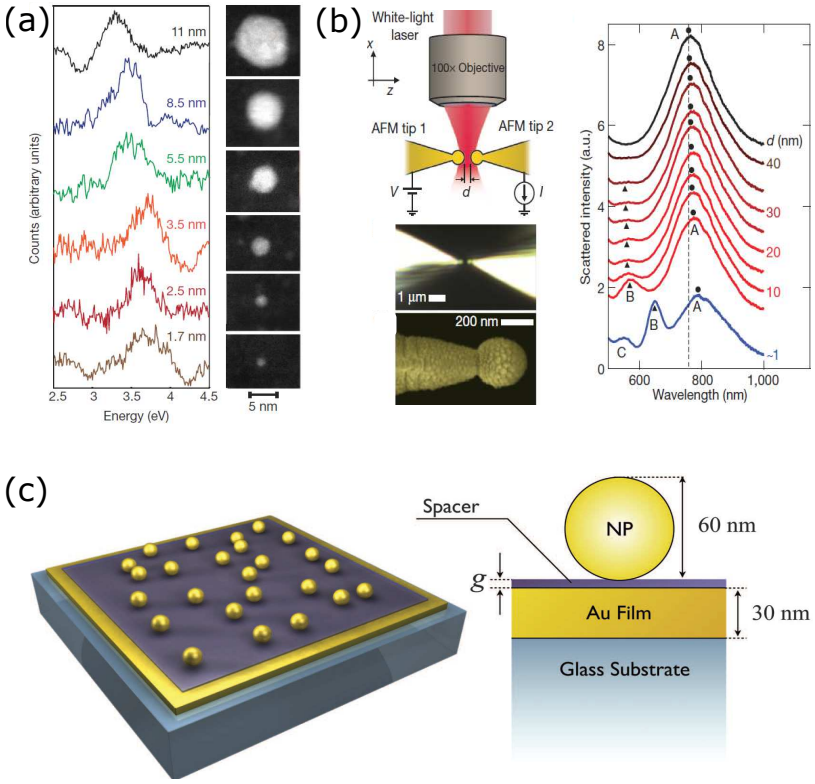


Figure 2.1: (a) Demonstration of quantum size effects for individual silver nanospheres. (b) Demonstration of electronic tunnelling effects between two nearly-touching tips on the plasmonic response. (c) Experimental configuration of gold nanoparticles on a gold film separated by a sub-nanometre thick molecular level which can be used to precisely control the separation. For small gaps nonlocal effects become increasingly important. (a), (b) and (c) reprinted with permission from references (Scholl et al., 2012; Savage et al., 2012; Ciracì et al., 2012) respectively.

size[5]. The appearance of a CTP is the most dramatic change seen in the quantum tunnelling regime, they have been observed experimentally by several groups (Savage et al., 2012; Scholl et al., 2013) and cannot be explained by macroscopic nonlocal models[6]. The redshifting BDP disappears and a CTP appears, which blueshifts with decreasing gap size (Zhu et al., 2016). The CTP is a consequence of electron tunnelling across the gap at optical frequencies, a significant amount of the induced charge can tunnel across leading to a short-circuiting of the gap. Using linker molecules has proved to be a useful method in gap quantum plasmonics to manipulate both the tunnel width and potential height, an intriguing mix of molecular electronics and plasmonics. This was explored by Tan *et al* who demonstrated enhanced tunnelling rates for silver nanocubes bridged by conducting and insulating molecules[7] compared to the equivalent vacuum barrier (Tan et al., 2014). DNA has also been explored experimentally as a linker molecule, with coherent tunnelling observed for up to a 2.8 nm wide gap (Lerch and Reinhard, 2016). Water layers have been predicted to lead to significant tunnelling for up to 2 nm gaps (Teperik et al., 2016). An interesting idea to increase tunnelling distances is to use a strong electric field to enhance the tunnelling rate via Fowler Nordheim tunnelling[8] (Wu et al., 2012). A innovative way of tuning the interparticle distance down to the point of merging using light was devised by Yoon's group, via the process of photooxidative desorption of the molecular linkers (Cha et al., 2014; Jung et al., 2015). Molecules

[5] The finite extent of induced surface charge leads to an effective gap size different than the geometrical gap (Teperik et al., 2013).

[6] The quantum-corrected model is a macroscopic model but as it is an *ad hoc* method, its ability to explain and predict are debatable.

[7] The highest occupied molecular orbital (HOMO)-lowest occupied molecular orbital (LUMO) gap of the molecule dictates the tunnelling rate; the LUMO sits at a lower energy than the vacuum, making tunnelling more energetically favourable. Note that in reference (Tan et al., 2014) they use the quantum-corrected model and identify a CTP peak, this has been demonstrated to be an artificial consequence of using a frequency dependent tunnelling material (Knebl et al., 2016).

[8] Emission of electrons induced by a strong electrostatic field.

placed in the gap can also act as convenient probes of the near-field, for instance allowing a mapping of the FE by SERS (Zhu and Crozier, 2014). Third harmonic generation has also been used to explore the tunnelling limit (Hajisalem et al., 2014). These systems has obvious implications for sensing deep in the quantum limit; molecules differing by only one atom can cause notable shifts of the gap plasmons (Benz et al., 2014). Further applications involves atomic scale plasmonic switches (Lutz et al., 2013; Große et al., 2014; Emboras et al., 2015) and optical rectification (Ward et al., 2010).

It is a feat of modern plasmonics to be able to correctly predict the spectral shift of the SP peak of an individual nanoparticle with decreasing size. It was a problem that long plagued cluster science due to the complexity of many competing mechanisms and the difficulty in comparing theoretical idealisations with experimental reality. For a metallic cluster there are three contributions to the SP resonance shift, which may act in unison or opposition depending on the material:

- Energy level quantisation always leads to a blueshift via the opening of an energy gap at the Fermi surface (Monreal et al., 2013).
- Electron spill-out of valence electrons will lead to a redshift; this is a simple consequence of a reduced electron density in the interior of the nanoparticle which reduces the effective plasma frequency.
- The d-electrons form a polarizable background that screens the valence electron and hence reduces the SP energy, they are also strongly localised meaning screening at the surface is reduced for valence electrons that spill-out. As the surface-to-volume ratio increases, surface screening is reduced and leads to blueshift of the SP.

The last two are together known as dynamical surface screening and can lead to a blue- *or* redshift, depending on the material and the influence of its d-electrons. This physics is contained within a useful quantity known as the *Feibelman's parameter* (Feibelman, 1982), $\delta_F(\omega)$, allowing the plasmon frequency to be written as follows[9]

$$\omega_{SP}(R) = \omega_{SP}(R \to \infty)\sqrt{1 - \frac{\delta_F(\omega)}{R}}. \tag{2.1}$$

In contrast to the classical model within the quasistatic approximation, there is now a dependence on the particles radius R. Whether there is a red- or blueshift depends on the sign of $\Re[\delta_F(\omega)]$.

A fairly clear-cut case is that of the light-alkali metal clusters which experience a simple redshift with decreasing size due to electron spill-out, $\Re[\delta_F(\omega)] > 0$. This has been confirmed experimentally (de Heer et al., 1987; Parks and McDonald, 1989; Reiners et al., 1995; Xia et al., 2009) for sodium, where the transition from molecular to collective behaviour has been mapped out in great detail. In

[9] It may appear at first a little puzzling that the effect of electron spill-out, which is a ground state property, is described within a dynamical model. This is possible because of the sum rules which relate ground state properties to moments of the optical response (Reiners et al., 1995).

contrast, the case for noble metals is more complicated due to the influence of d-shell electrons. For silver it is the reduced screening effect that wins and consequentially a blueshift, $\Re[\delta_F(\omega)] < 0$, is observed with decreasing size (Tiggesbäumker et al., 1993; Raza et al., 2013a). Note that this competition will dictate at which sizes quantum effects will appear; a particularly interesting case is that of mercury clusters that show no strong quantum-size effects down to a very small cluster size, $\Re[\delta_F(\omega)] \approx 0$, so that Mie theory works well even for clusters consisting of tens of atoms (Haberland et al., 1992).

In the following sections, we will now cover the various theoretical models developed in quantum plasmonics to explain the aforementioned experiments and physical phenomena.

2.2 Nonlocality In Plasmonic

QUANTUM THEORIES CAN BE COMPUTATIONALLY INFEASIBLE FOR LARGER SYSTEMS, SO OFTEN BY NECESSITY WE ARE FORCED TO WORK IN A CLASSICAL FRAMEWORK WITH SOME QUANTUM CORRECTIONS. THESE CORRECTIONS ARE LABELLED NONLOCAL EFFECTS. IN SIZE REGIMES WHERE QUANTUM EFFECTS ARE A PERTURBATION ON THE CLASSICAL BEHAVIOUR WE CAN EXPECT THIS TO BE AN ACCURATE APPROACH, THIS IS THE BASIC PHILOSOPHY BEHIND THE HYDRODYNAMIC MODEL (HM). WE WILL ALSO CONSIDER THE EFFECTS OF ELECTRON CONFINEMENT ON THE BROADENING OF THE LSP.

2.2.1 The Need For The Hydrodynamic Model

In section 1.1 we briefly explored the concept of nonlocality, for nanostructures with some small defining length scale such effects can become important. In a fully quantum framework nonlocality emerges quite naturally and one need not give it much thought. On the other hand, to build it in within a macroscopic framework[10] does involve some effort. The HM is one approach to this problem, which is used heavily in quantum plasmonics. Some motivation for this approach is given by the experimental results of (Scholl et al., 2012), shown in figure 2.1(a), where the quantum size effects of a single silver nanosphere are seen to be not so drastic; there is a shift and a broadening. This gives us hope that we could describe the changes within a modified macroscopic approach, this is the guiding principle of the HM. Given its popularity, and the fact it is used in chapter 4, means it is appropriate to explore the HM in some detail.

[10] By this I mean a theory that depends only on macroscopic variables and involves no explicit consideration of electronic wavefunctions.

Briefly, we recap that the nonlocal wave equation in real-space that we need to solve is[11] (Raza et al., 2015)

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{r},\omega) = \left(\frac{\omega}{c}\right)^2 \int d^3r' \; \epsilon(\mathbf{r},\mathbf{r}',\omega)\mathbf{E}(\mathbf{r}',\omega) \tag{2.2}$$

[11] This is obtained by taking the curl of Faraday's law from equations 1.1 and working in the frequency domain.

which is completely general and valid as long as the approximations of macroscopic electrodynamics hold. This is clearly a formidable integro-differential equation to solve and the permittivity $\epsilon(\mathbf{r},\mathbf{r}',\omega)$ is not only difficult to express in closed form, it is also a 6 dimensional quantity in space[12] and so a real pain to handle numerically. The HM provides a convenient approximation by reducing it to a much simpler differential equation[13]. There also exists phenomenological approaches to solving equation 2.2 which may provide significant insight, for instance taking the response function to have a Gaussian form[14] to model nonlocal smearing (Ginzburg and Zayats, 2013).

[12] 7 dimensions with frequency included!

[13] This is not *a priori* obvious as it is not always possible to transform an integral equation into differential form (Arfken, 1985). One would suspect it is only possible in this case because of the drastic approximations in the HM, which does not completely take into account long-range correlation in the electron gas.

[14] The true response function will clearly be more complicated then just depending on the difference $|\mathbf{r}-\mathbf{r}'|$, this will only be valid away from surfaces.

The HM has a long history which begun with Bloch in 1933 (Bloch, 1933) and was applied to plasmonics from the very be-

ginning of the field (Ritchie, 1957). A notable application in the 1970s was from Ruppin who generalised Mie theory for the case of a nonlocal longitudinal and local transverse dielectric function. It was found that metallic spheres can support a family of longitudinal modes above the plasma frequency (Ruppin, 1973, 1975). These additional modes had previously been predicted for thin films (Melnyk and Harrison, 1970) and soon after experimentally demonstrated for silver (Lindau and Nilsson, 1971) and potassium (Anderegg et al., 1971). Long thought to be just an interesting feature of nonlocality, these modes have recently been shown to have potential use in boosting the third-harmonic generation of thin silver wires coated with a third-order nonlinear dielectric (Hu et al., 2018).

The HM has become a popular tool to study plasmonic systems where deviations from classical predictions can be expected, but full quantum calculations are out of the question, this would roughly correspond to the size regime of $2 \rightarrow 10$ nm, but of course this will depend on the material to be studied, the accuracy required and the computational resources available to the researcher(s). Some benefits of the method include:

- The HM quite naturally includes retardation effects, it is just a modified version of the macroscopic Maxwell's equations after all. This is not the case for quantum methods such as TDDFT, which are very much electrostatic in spirit[15].
- Existing electrodynamic codes can be modified to include nonlocality via the inclusion of hydrodynamic terms, although some increase in computational time can be expected due to the need to resolve not only subwavelength but also sub-Fermi wavelength features as well.
- It can be argued that in the Herculean task of running and interpreting the results from a DFT calculation for a large system, that some physical insight is bound to be loss. The HM provides a nice link from the quantum world to the plasmonic which can help build intuition on quantum plasmonic effects. Analytical solutions are even possible in some very simple geometries (Ruppin, 1973, 1975).

[15] There is some work in combining TDDFT and Maxwell's equations using multiscale methods (Yabana et al., 2012).

As noted previously, the HM has had success in explaining experimental result (Ciracì et al., 2012), but it does have some well documented shortcomings that must be kept in mind. These will become apparent as we delve into the theory in the next section

2.2.2 *The Hydrodynamic model*

The key assumption underlying the HM is that the quantum mechanical many-body problem can be expressed in terms of two scalar fields: the electron density, $n(\mathbf{r})$, and a velocity field, $v(\mathbf{r})$. Inspired by DFT, we can guess that the energy of the plasma is a functional of these two fields. The macroscopic EM fields can

then be obtained, in the form of equations of motion, via functional derivatives of the combined light-electron gas Hamiltonian (Boardman, 1982), a derivation is shown in appendix G. These are the *hydrodynamical equations of motion* (Raza et al., 2015)

$$\partial_t n = -\nabla \cdot (n\mathbf{v})$$
$$[\partial_t + \mathbf{v}(\mathbf{r}) \cdot \nabla]\, \mathbf{v}(\mathbf{r}) = \underbrace{-\frac{e}{m}\left(\mathbf{E}(\mathbf{r}) + \mathbf{v}(\mathbf{r}) \times \mathbf{B}(\mathbf{r})\right)}_{\text{LORENTZ FORCE}} \underbrace{-\frac{1}{m}\nabla\left(\frac{\delta F[n](\mathbf{r})}{\delta n(\mathbf{r})}\right)}_{\text{QUANTUM PART}} \underbrace{-\gamma \mathbf{v}(\mathbf{r})}_{\text{DAMPING}}. \tag{2.3}$$

The first equation is simply the continuity equation. In the second equation, the first term of the right hand side is the Lorentz force experienced by the electrons and the third term is a phenomenological term to account for bulk damping. It is the second term that contains the interesting physics, it takes into account the internal kinetic energy and, for a suitable functional $F[n(\mathbf{r})]$ choice, an approximation to the exchange-correlation (XC) energy of the electron gas. The favoured choice for the functional, because of its simplicity, is the Thomas-Fermi functional[16] (Lundqvist and March, 2013)

$$F_{TF}[n](\mathbf{r}) = \frac{3h^2}{10m}\left(\frac{3}{8\pi}\right)^{\frac{2}{3}} \int d^3r'\; n^{\frac{5}{3}}(\mathbf{r}'), \tag{2.4}$$

which ignores the XC contribution. This gives (see appendix G for a short derivation) a *linearised*[17] second term equal to $-\frac{\beta^2 \nabla n_1(\mathbf{r})}{n_0}$, which can be identified as a macroscopic pressure term from the internal kinetic energy of the electrons that will act to homogenize spatial inhomogeneities in the electron gas (Raza et al., 2015). This corresponds to an energy transfer within the system via a non-photonic mechanism and is a source of nonlocality (Raza et al., 2011). In the limit $\beta \to 0$, we recover the usual local equation of motion for electrons. Of course, better approximations for the functional could be chosen, but given the drastic assumptions inherent to the HM, I think it is more sensible to treat β as a phenomenological parameter that best matches experiment. The value for β predicted from the HM, $\beta^2 = \frac{v_F^2}{3}$, is famously incorrect for application in plasmonics. This is because the HM is a low frequency approximation, the velocity $v_F/\sqrt{3}$ is the velocity of sound in hydrodynamics which is consistent with the hydrodynamic regime[18], $\omega << \gamma$, whereas the regime of plasmonics is high frequency, $\omega >> \gamma$ (Boardman, 1982). Electron-electron interactions are unable to establish equilibrium near the plasmon frequency (Giuliani and Vignale, 2005). Using the RPA a value of

$$\beta^2 = \frac{3v_F^2}{5}, \tag{2.5}$$

is obtained. The extra energy cost associated with this higher value of β is due to deviations from the spherical shape of the Fermi surface in a plasma oscillation; in the local equilibrium situation

[16] This is the simplest choice for the non-interacting kinetic energy functional and is based on the kinetic energy density of the non-interacting homogeneous electron gas evaluated at a local point, $n(\mathbf{r})$. It is valid in the limit $\frac{|\nabla n(\mathbf{r})|}{n(\mathbf{r})k_F(\mathbf{r})} << 1$, which is the same limit of applicability as the LDA discussed in section 2.3 (Giuliani and Vignale, 2005).

[17] The density is expanded about the ground state: $n(\mathbf{r}) = n_0 + n_1(\mathbf{r})$.

[18] In the hydrodynamic regime a quasi-particle will suffer many collision in one period of the exciting field and so Fermi-liquid theory will break down. The frequent collisions bring about a local equilibrium and hence it becomes possible to characterize the system by macroscopic properties (Ying, 1974).

assumed by the HM there is only expansion and contraction of the Fermi sphere (Giuliani and Vignale, 2005). β has been generalised to a complex quantity by Mortensen *et al* in the generalized nonlocal optical response (GNOR) model, which was devised to describe the increased broadening of quantum plasmon modes within the HM via the inclusion of a diffusion term (Mortensen et al., 2014). It gives a quasi-quantum theoretical basis for broadening rather than the *ad hoc* Kreibig damping, which is discussed later in this section. In fact the GNOR model predicts the same $1/R$ dependence, unfortunately it predicts, based on bulk parameters, a diffusion constant two orders of magnitude too small (Raza et al., 2015)) and a phenomenological relaxation time must be used to match experiment[19]. It is clear that a better understanding of loss channels at these scales is needed, this will be a recurring theme of this thesis.

Coupling equations 2.3 with Maxwell's equations and taking a linear approximation[20] gives, for a material described by a Drude dielectric function and in the frequency domain (Raza et al., 2015),

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{r},\omega) = \left(\frac{\omega}{c}\right)^2 \left(\epsilon(\omega) + \zeta^2 \nabla(\nabla\cdot)\right) \mathbf{E}(\mathbf{r},\omega). \tag{2.6}$$

Compared to equation 2.2, the above differential equation may be less physically intuitive, but it is a considerable simplification and is computationally more practical as it can be solved using finite element methods quite simply. The nonlocal parameter is given by $\zeta^2 = \frac{\epsilon_\infty \beta^2}{\omega(\omega+i\gamma)}$ and is a purely real quantity in the absence of damping and diffusion processes. By writing the electric field as a sum of a longitudinal and transverse part, $\mathbf{E} = \mathbf{E}_L + \mathbf{E}_T$, it is simple to see that the transverse part is unaffected by the hydrodynamic corrections[21]; the spatial dispersion only affects the longitudinal part of the electric field. The material response for the longitudinal field is described by the longitudinal component of the permittivity tensor, ϵ_L, hence the HM can be classified as a *longitudinal nonlocal response theory* (Wubs, 2015). This is a specific characteristic of the HM, different nonlocal models may influence the transverse response as well; for instance, diffuse surface scattering leads to a nonlocal contribution to the transverse dielectric function in doped quantum dots (Monreal et al., 2015). The presence of longitudinal waves leads to the need for extra boundary conditions, these have been a hot topic of debate in the past and can easily lead to mistakes, as explained in reference (Raza et al., 2011). The exact number depends on the choice of the electron density at the interface and the choice of materials; for a metal-dielectric interface with a step function electron density profile one need only specify the physical reasonable assumption that the normal component of the induced current density goes to zero. This additional need for information is presumably a consequence of the breakdown of the concept of a sharp interface when nonlocal effects are considered. The inclusion of the hydrodynamic term in the wave equation leads to the following:

[19] Perhaps the greatest success of the GNOR model is offering a much more believable explanation for the broadening of the BDP as a dimer gap is decreased, when compared to the quantum-corrected model (Esteban et al., 2012). The GNOR model predicts the loss occurs at the metal boundary in contrast to the quantum-corrected model which predicts it occurring in the gap due to the lossy fictitious material, this is at odds with current understanding of mesoscopic electron transport through a classically forbidden gap as an elastic process (Mortensen et al., 2014).

[20] Note that the HM is almost synonymous with the *linear* HM, in fact the nonlinear hydrodynamic response may be important. For example, very recent experimental work has indicated nonlocality contributes to nonlinear processes that lead to white-light supercontinuum generation from rough thin gold films (Chen et al., 2018).

[21] As $\nabla \cdot \mathbf{E}_T = 0$.

- The longitudinal nonlocality allows the coupling of light to a family of confined longitudinal modes with discrete energies above the plasma frequency (Ruppin, 1975).
- There is a size dependent blueshift arising from the smearing of the induced surface charge at the boundaries over a finite distance of a few Å. This is a consequence of the finite extent of electronic wavefunctions in the metal. Classical electromagnetism has the implicit assumption of point charges and, as a consequence, screening charge is an infinitesimally thin layer.
- The inclusion of diffusion via the GNOR model will lead to a size dependent broadening.

An ideal system to study these effects is a small sphere in the non-retarded limit, which has a simple analytical form[22] in the form of a modified polarizability α_{NL} (Raza et al., 2013b)

$$\begin{aligned} \alpha_{NL}(\omega) &= 4\pi R^3 \frac{\epsilon(\omega) - \epsilon_b\,(1+\delta_{NL})}{\epsilon(\omega) + 2\epsilon_b\,(1+\delta_{NL})} \\ \delta_{NL} &= \frac{\epsilon(\omega)-\epsilon_\infty}{\epsilon_\infty} \frac{j_l(k_{NL}R)}{k_{NL}Rj_l'(k_{NL}R)} \\ k_{NL}^2 &= \left(\omega^2 + i\omega\gamma - \omega_P^2/\epsilon_\infty\right)/\beta^2, \end{aligned} \tag{2.7}$$

where k_{NL} is the nonlocal longitudinal wavevector. In the case of planewave excitation we consider the case $l = 1$. The results for a calculation of a 5 nm radius silver nanosphere are shown in figure 2.2 for the local and nonlocal quasistatic model. The blueshift of the dipole LSP is clearly evident as well as the presence of the higher energy modes, which are only visible on a log scale. It is important to note that the HM can only predict a blueshift of the dipole LSP and is therefore missing some crucial physics of small clusters, as we explored in section 2.1. In fact even the blueshift predicted by the HM may only fortuitously be in agreement with the experimental results for silver nanospheres as it does not take into account dynamical surface screening. Furthermore, it clearly cannot accurately model the electron tunnelling regime of closely gapped dimers.

A positive feature of the HM is that there is a clear road to improving it. Viewed as a long wavelength limit of the theory of the dynamic electron gas, extra terms in the expansion of q can be included as well as better approximate functionals than Thomas-Fermi. Furthermore, the HM is a hard-wall model, so modification to include electron spill-out allows the possibility to describe clusters where the spill-out dominates the small size regime, such as the alkali metals as well as electron tunnelling between dimers. In reality, at a certain point it is often better just to switch allegiances to density functional methods. One promising approach is the quantum HM developed by Yan (Yan, 2015), which uses a more sophisticated functional[23] and includes electron spill-out. While the description of the electron dynamics are not at

[22] Assuming a Drude dielectric function describing the metal.

[23] XC contribution included within the LDA.

high-level as DFT, retardation is included and large nanostructure relevant to plasmonics that are beyond the practical reach of DFT can be tackled.

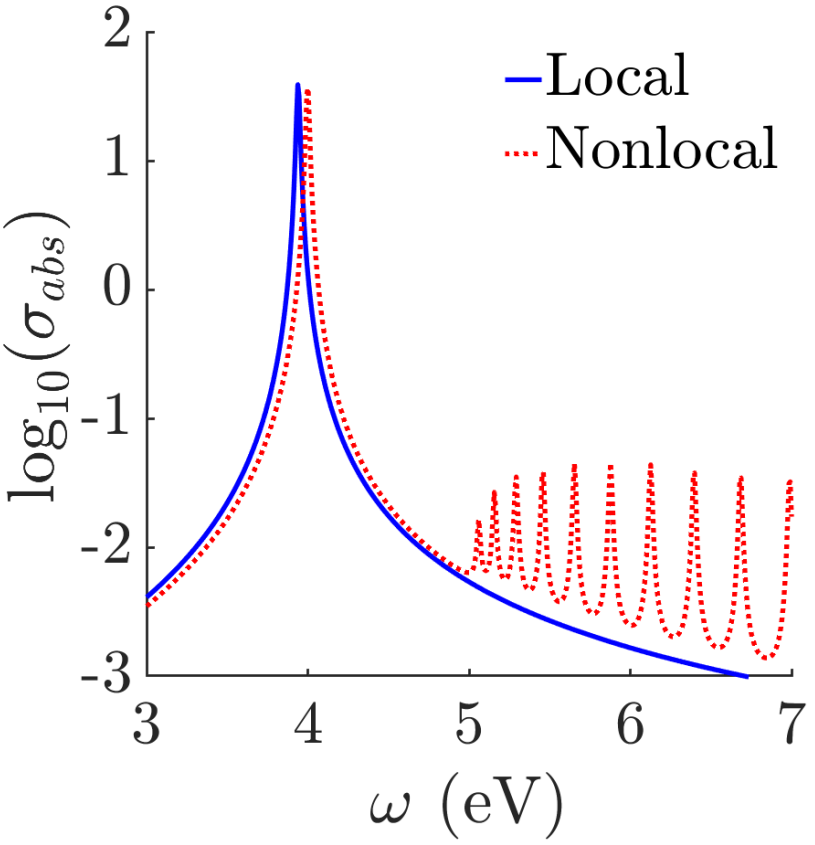


Figure 2.2: The log of the absorption cross section for a 5 nm radius silver sphere within a quasistatic local (blue solid line) and quasistatic HM (red dotted line). The Drude parameters used are $\omega_P = 9.1$ eV, $\hbar\gamma = 0.022$ eV and $\epsilon_\infty = 3.3$.

2.2.3 *The Local Analogue Model*

While numerical implementation of the HM is not the most difficult thing in the world, it would still be nice to include nonlocality within existing local codes. The fact that nonlocal effects tend have rather straightforward effect on optical spectra, i.e. a shift and the appearance of some new modes, may give us heart that there is a straightforward way to model the effects locally. It took the ingenuity of Luo *et al*, who found the inclusion of a fictitious thin[24] dielectric layer, with a suitable permittivity, over a metal surface can mimic nonlocality[25] (Luo et al., 2013). It has been shown to accurately model the LSP blueshift and the smearing of the electric field at interfaces, within the computationally simpler local framework. This method is known as the local analogue model (LAM). The condition for the permittivity of the fictitious layer of thickness Δd is

$$\frac{\epsilon_{fict}(\omega)}{\Delta d} = \frac{\epsilon_b \epsilon(\omega) i k_{NL}}{\epsilon(\omega) - \epsilon_b} \frac{j_l'(k_{NL}R)}{j_l(k_{NL}R)} \approx \frac{\epsilon_b \epsilon(\omega) i k_{NL}}{\epsilon(\omega) - \epsilon_b}, \tag{2.8}$$

[24] Compared to the metals skin-depth.

[25] The real and fictitious system should show the same far- and near-field behaviour for any incident light field.

where the approximation holds if $R >> 1/(ik_{NL})$, i.e. if the radius of curvature is much larger than the Thomas-Fermi wavelength[26]. Note that if this approximation is taken then the bulk longitudinal resonances above $\omega_P/\sqrt{\epsilon_\infty}$ will not appear. We will make use of the LAM incorporated within the BEM in chapter 4. In figure 2.3 we compare the quasistatic model with the LAM for a small 1 nm silver sphere[27]. Note that the LAM includes retardation, in contrast to the quasistatic HM used in the figure, but this is not important at these small sizes.

[26] This will almost always be the case for system of practical interest; the Fermi wavelength is 0.37 nm for silver.

[27] In reality, for a nanosphere so small, a TDDFT calculation is feasible and would offer much more accurate results.

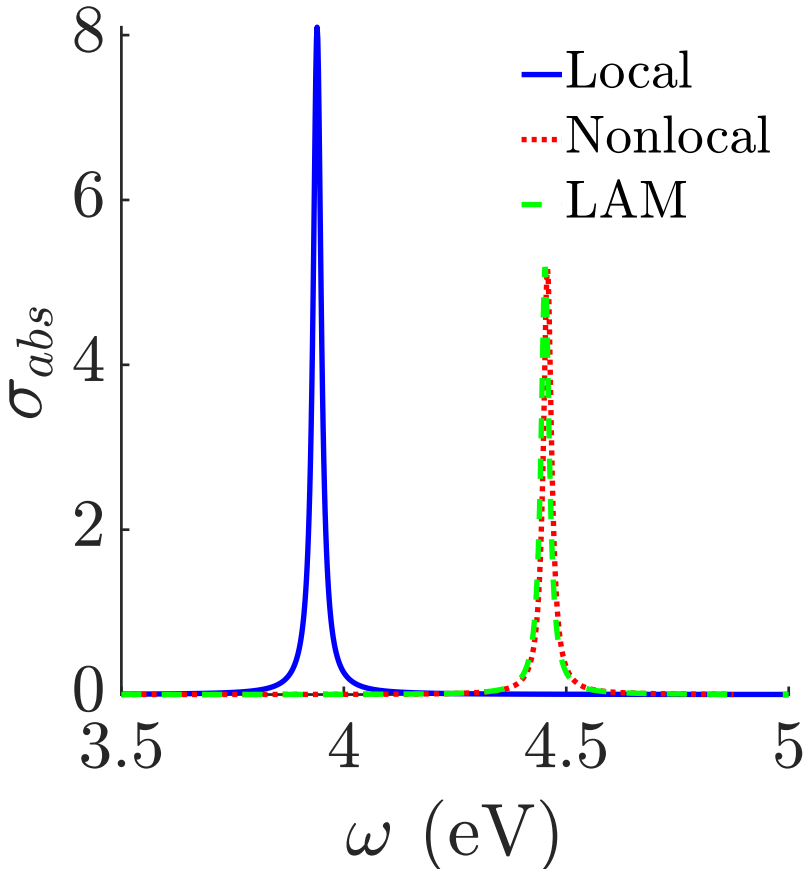


Figure 2.3: Absorption cross section for a 1 nm radius silver sphere within a quasistatic local (blue solid line), quasistatic HM (red dotted line) and the LAM (green dashed line). The Drude parameters used are $\omega_P = 9.1$ eV, $\hbar\gamma = 0.022$ eV and $\epsilon_\infty = 3.3$.

2.2.4 *Electron-Surface Scattering*

When a spherical nanoparticle of radius R is smaller than the electron mean free path there is a surface scattering contribution, known as Kreibig or Landau damping, in addition to the bulk damping γ_0. The two can be simply added thanks to Matthiessen's rule (Ashcroft and Mermin, 1976). The, now size dependent, damping can be written in the form (Kreibig and Vollmer, 2013)

$$\gamma(R) = \gamma_0 + A\frac{v_F}{R}, \tag{2.9}$$

where A is a material-dependent parameter on the order of unity but there exists considerable uncertainty in its value, which is a

consequence of experimental difficulties as well as various size dependent contributions from other sources such as phonon-plasmon coupling (see chapter 5 and (Fitzgerald et al., 2017; Donati et al., 2017)), structural phase transitions (Kreibig, 1978) and chemical interface damping (Hövel et al., 1993; Pinchuk and Kreibig, 2003; Foerster et al., 2017). There have been a number of theoretical (Molina et al., 2002; Lermé et al., 2010; Li et al., 2013) and experimental (Quinten, 1996; Alvarez et al., 1997; Link and El-Sayed, 1999) works exploring this topic. It is important to note that the surface scattering models are only possible for certain systems of high symmetry such as the sphere (Kreibig and Vollmer, 2013) and spherical shell structures (Moroz, 2008). It is not easily generalisable to more complicated geometries, although a possible technique could be to use Monte Carlo simulations, which have been recently used to explore the effect of localisation on the Drude conductivity for confined structures probed at THz frequencies (Cocker et al., 2017). The dependence of surface scattering effects on geometry will be important to understand for ultra-small plasmonic devices, for instance thin metallic nanoshells could offer superior performance for enhancing large FEs; experimental results have demonstrated that 20 nm thick gold nanoshells show no additional broadening (Nehl et al., 2004), which has been backed up by RPA calculations showing Landau damping decreases with decreasing nanoshell thickness (Kirakosyan et al., 2016). Note that I have considered the effects of surface scattering and nonlocality separately but in reality they are two faces of the same coin: a classical representation of a quantum mechanical phenomena. In full quantum simulations, such as TDDFT, it will not be necessary to treat the two separately as both will automatically be included. It is also worth stressing that classical models cannot possibly provide the whole picture of electron scattering in finite systems; quantum effects can lead to either an increase in excitation lifetime via the reduced number of states near the Fermi energy due to confinement or a decrease via the reduction in dynamic screening at the electron surface (Quijada et al., 2010). These discussion are very relevant for calculating FEs and investigating hot-electron dynamics and thermalization (Saavedra et al., 2016) in tiny metal nanoparticles.

2.3 Density Functional Methods In Plasmonics

DENSITY-FUNCTIONAL THEORY (DFT) IS A UNIVERSAL MEAN-FIELD TECHNIQUE FOR ELECTRONIC STRUCTURE CALCULATIONS, IT IS A FORMALLY EXACT APPROACH TO THE STATIC ELECTRONIC MANY-BODY PROBLEM AND CAN BE GENERALISED TO THE DYNAMIC CASE WITH TIME-DEPENDENT DENSITY-FUNCTIONAL THEORY (TDDFT). IT IS THE MOST COMMONLY USED ELECTRONIC STRUCTURE METHOD DUE TO ITS INCLUSION OF EXCHANGE-CORRELATION (XC), RELATIVE SIMPLICITY AND CHEAP COMPUTATIONAL COST COMPARED TO OTHER TECHNIQUES. IN THE PAST DECADE, IT USE HAS INCREASINGLY EXTENDED INTO PLASMONICS; HERE WE DISCUSS THE BASIC UNDERLYING PRINCIPLES AND WHY IT IS USEFUL FOR STUDYING QUANTUM PLASMONICS.

2.3.1 Introduction To DFT

We are interested in solving the N-body, static, non-relativistic, many-body Schrödinger equation within the Born-Oppenheimer approximation[28]

$$\hat{H}\Psi_I(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N) = E_I\Psi_I(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N), \tag{2.10}$$

where I use the notation that $\mathbf{x}$ stands for the combined spatial coordinate $\mathbf{r}$ and spin index σ. Ψ_I is the Ith *antisymmetric*[29] N-electron eigenfunction and is associated with the energy eigenvalue E_I, it describes the correlated motion of the N interacting electrons (Szabo and Ostlund, 2012). In static many-body theory, one is typically interested in the ground state energy $E_0 = \langle\Psi_0|\hat{H}|\Psi_0\rangle$. The wavefunction gives us all the information one could ever want to know about the system, but this is a vast amount of data which soon hits the famous 'exponential wall' for even modest values of N; it is unstorable and impossible to interpret in practise (Ullrich, 2011). A preferable quantity is the ground state electron density which is physically intuitive and depends only on 3 coordinates, rather than $3N$. DFT offers the intriguing opportunity to obtain all the useful information about a many-body system using the ground state density, in principle exactly, without directly solving the many-body Schrödinger equation.

The origin of DFT is with the seminal paper by Hohenberg-Kohn (Hohenberg and Kohn, 1964). It is apparent that the ground state density $n_0(\mathbf{r})$ is a functional of the confining potential $v_0(\mathbf{r})$[30], we can write this notationally as $n_0[v_0](\mathbf{r})$. What is not *a priori* clear is if the converse is also true: can we write $v_0[n_0](\mathbf{r})$? The proof that we can is given by the celebrated *Hohenberg-Kohn theorem*, the cornerstone of DFT. The theorem can be stated as follows:

Theorem (Hohenberg-Kohn Theorem). *In an interacting N-electron*

[28] i.e. we describe the motion of electrons within a field of fixed nuclear point charges, see section 2.3.6.

[29] $\Psi_I(\mathbf{x}_1, ..., \mathbf{x}_j, ..., \mathbf{x}_k, ..., \mathbf{x}_N) = -\Psi_I(\mathbf{x}_1, ..., \mathbf{x}_k, ..., \mathbf{x}_j, ..., \mathbf{x}_N)$, it is anti-symmetric with respect to interchange of *both* the spatial coordinate and spin.

[30] As the potential dictates the form of Schrödinger's equation and hence the N electron wave-functions, it follows that is fixes the ground state density via $n_0(\mathbf{r}) = N\,\Pi_{j=2}^{N}\int d\mathbf{x}_j|\Psi_0(\mathbf{x}, \mathbf{x}_2, ..., \mathbf{x}_N|^2$

*system there exists a **unique** one-to-one correspondence between the confining potential $v_0(\mathbf{r})$ and the ground state density $n_0(\mathbf{r})$ (to within a constant).*

This far-reaching statement is surprisingly easy to prove and proceeds by proof of contradiction (*reductio ad absurdum*). There is more that we can conclude from these considerations: since the kinetic and interparticle interactions are universal, the Hamiltonian must be a functional of the density $\hat{H}[n_0]$. This means all eigenstates, not just Ψ_0, are also density functionals, $\Psi_I[n_0]$. This leads to the conclusion that all ground- and excited state properties of a quantum system are, in principle, determined by the ground state density. Further, Hohenberg and Kohn also deduced that the total-energy functional

$$E[n] = \langle \Psi[n] | \hat{T} + \hat{W} + \hat{V} | \Psi[n] \rangle \,, \tag{2.11}$$

only equals the ground state energy, E_0, for the correct $n_0(\mathbf{r})$, and for any other density will have a *larger* value. This allows the calculation of the density via a minimisation process[31], to this end we write the energy functional in terms of the *universal functional*, $F[n]$, first introduced by (Hohenberg and Kohn, 1964)

$$E[n] = F[n] + \underbrace{\int d^3r \; n(\mathbf{r}) v_0(\mathbf{r})}_{V[n]} = T[n] + W[n] + V[n], \tag{2.12}$$

where $V[n] = \langle \Psi[n] | \hat{V} | \Psi[n] \rangle$ is the functional describing the coupling of the density and confining potential v_0. $T[n]$ and $W[n]$ are the kinetic and electron-electron interaction functionals respectively[32]. The minimisation can be expressed as an Euler equation and written as (Ullrich, 2011)

$$\frac{\delta F[n]}{\delta n(\mathbf{r})} + v_0(\mathbf{r}) = \mu, \tag{2.13}$$

where μ is the familiar chemical potential and plays the role of a Lagrange multiplier to enforce the correct number of electrons: $\int d^3r \; n(\mathbf{r}) = N$.

The above discussions were fruitful but at the moment it is not really of any practical importance; the Hohenberg-Kohn theorem does not provide a way to calculate the ground state density. By rewriting the problem in terms of the Kohn-Sham (KS) equations we can get a computationally practical scheme where the interacting density problem is replaced by an auxiliary system of non-interacting electrons (i.e a single-particle formalism) in an effective *local* potential v_{KS} (the KS potential), which takes into account both Coulomb and XC contributions (Kohn and Sham, 1965). The equations can be derived by considering the relationship between the Euler equation and the single Slater determinant description of a non-interaction electron system ($W[n] = 0$). By rewriting the total-energy functional in a smart way, Kohn and Sham found that

[31] This is akin to standard quantum mechanics where one can deduce the energy by minimisation of the expectation value of the Hamiltonian with respect to the wavefunctions.

[32] This is normally the Coulomb interaction but need not be.

the interacting density can be found by solving a single-particle Schrödinger equation for a system of non-interacting electrons in a confining potential $v_{KS}[n]$, which itself is a functional of the density. The equations must be solved as a set self-consistently, see figure 2.4. We state the equations:

$$\begin{aligned} \left(-\frac{\hbar^2}{2m}\nabla^2 + v_{KS}[n](\mathbf{r})\right)\psi_i[n](\mathbf{r}) &= \varepsilon_i\psi_i[n](\mathbf{r}) \\ n_0(\mathbf{r}) &= \sum_i^N |\psi_i(\mathbf{r})|^2 \\ v_{KS}[n](\mathbf{r}) = v_0(\mathbf{r}) + \underbrace{\frac{e^2}{4\pi\epsilon_0}\int d^3r' \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}}_{v_H[n](\mathbf{r})} &+ \underbrace{\frac{\delta E_{XC}[n]}{\delta n}}_{v_{XC}[n](\mathbf{r})} \end{aligned} \tag{2.14}$$

This is an incredible simplification of the problem, but of course nature does not give us a completely free lunch: we need to construct a suitable expression for the XC energy, E_{XC}. The theory would be exact if the XC energy was explicitly known, but in practise approximate expressions must be used. Luckily the XC contribution is the smallest contribution to the total energy so the error from this approximation is minimised[33]. The non-interacting and Hartree functionals, which make up the largest contribution, are treated exactly in the KS method (Ullrich, 2011). Note that it is possible to write down the many-body wavefunction using the KS orbitals, $\psi_i[n]$, to construct a single Slater determinant, but there is no guarantee this corresponds to the real many-body wavefunction, or even that it is a good approximation. The KS orbitals are simply a mathematical device to construct the density, one should bear this in mind for our later discussions. This also applies to the KS eigenvalues[34], meaning that differences in eigenvalues cannot be simply taken as the excitation energies. We will discuss this further in the next section on TDDFT, which is the appropriate method to obtain accurate excitation energies.

The simplest approximation to the XC functional is the popular LDA[35]

$$\begin{aligned} E_{XC}^{LDA}[n] &= \int d^3r\, u^{LDA}(n(\mathbf{r})) \\ v_{XC}^{LDA} = \frac{\delta E_{XC}^{LDA}[n]}{\delta n}\Big|_{n=n(\mathbf{r})} &= \frac{d}{dn}u^{LDA}(n)\Big|_{n=n(\mathbf{r})}, \end{aligned} \tag{2.15}$$

where $u(n(\mathbf{r}))$ is the XC energy per unit volume of an homogeneous electron gas, *evaluated locally* for $n(\mathbf{r})$ (Giuliani and Vignale, 2005). It was first proposed in the original paper by Kohn and Sham (Kohn and Sham, 1965). The XC energy density can be split into an exchange and correlation part, the exchange part is known exactly and very accurate numerical parametrizations exist for the correlation part (Martin, 2004). One would expect the LDA to only be a successful approximation for a very limited range of systems with slowly varying[36] densities, but it has turned out to be remark-

[33] That is not to say it is not important! It contributes significantly ($\sim$ 100%) to the binding energy of matter and therefore good approximations are needed for 'chemical accuracy' (accuracies within $\sim$ 0.05 eV), long time a holy grail of DFT research (Perdew and Kurth, 2003).

[34] The exception being the HOMO energy which must give the negative of the ionization energy of the system (Ullrich, 2011).

[35] The second line follows from the following identity for functional derivatives: if $F[\phi] = \int d^n x\, f(\phi(x))$ then $\frac{\delta F[\phi]}{\delta\phi(y)} = \partial_\phi f|_{\phi(y)}$.

[36] Here slowly varying is to be measured relative to the most spatially varying (most energetic) electrons in the system, giving the condition $\frac{\nabla n(\mathbf{r})}{n(\mathbf{r})k_F} << 1$ (Giuliani and Vignale, 2005).

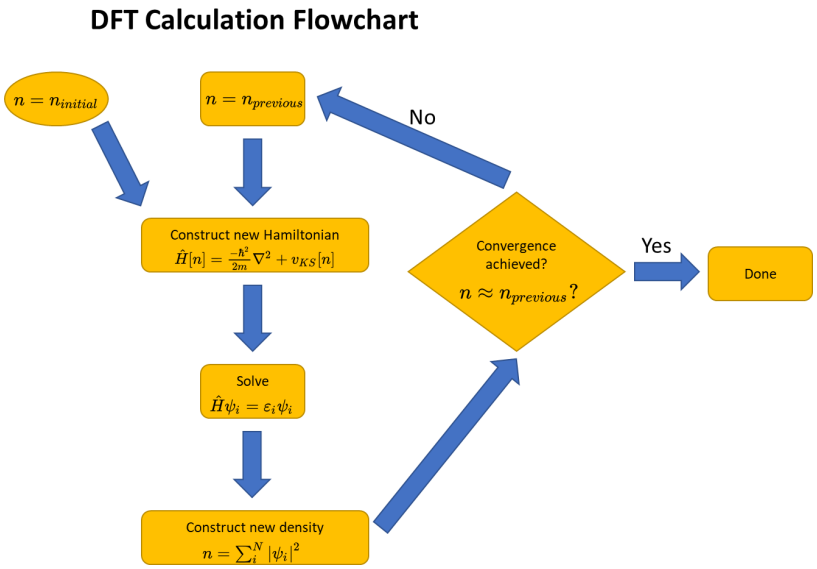


Figure 2.4: Computational scheme for a typical DFT calculation. This figure is a little simplified; the nonlinearity of the KS equations means that if one naively substitutes the density from the previous iteration as the new density for the next iteration, then instabilities can arise. Instead a suitable density-mixing scheme needs to be used (Martin, 2004).

ably successful even for highly inhomogeneous systems such as atoms and molecules. This is partly because it is based on a real physical system (it is exact in the homogeneous limit), meaning it obeys certain sum rules and scaling properties; these result in an error cancellation between overestimated exchange energies and underestimated correlation energies (Perdew and Kurth, 2003). This results in superior accuracy, for most applications, compared to the Hartree-Fock approximation, which treats exchange exactly and neglects correlation. Nevertheless, the LDA has some serious shortcomings, it is not free of self-interaction[37] and so has the wrong asymptotic behaviour, it goes to zero exponentially fast rather than $1/r$. This leads to famous issues with bandgaps in solids and ionisation energies. Furthermore, for physical phenomena where it does work well, accuracy will be around 5% which is not good enough for chemical applications. It does though capture the correct physical trends, especially for simple metals. It is commonly used in solid-state physics and suffices for quantum plasmonics in most cases. Gradient expansions provide a systematic approach to improving upon the LDA but don't work well in practise as they don't satisfy sum rules. Instead, generalized gradient approximations and hybrid functionals (which include a fraction of exact exchange) should be used (Perdew and Kurth, 2003), but we will not consider these any further in this thesis.

A significant further simplification that makes many DFT calculations practical, is the *pseudopotential* concept. A DFT calculation time scales linearly with the number of electrons compared to an exponential scaling for wavefunction based approaches. Even so, this can still be prohibitive for calculations involving elements with many electrons; for example, a typical plasmonics type calculation might involve silver which has a total of 47 electrons per atom, or gold which has 79. Already for a small cluster of 20 atoms this

[37] It is a requirement of the KS scheme that the electron-electron interaction does not contain self-interaction. This results in the condition that the XC energy of each KS orbital cancels with its self-Hartree energy (Ullrich, 2011).

would be a heavy calculation, beyond what is practical on a typical scientific workstation (as of writing in 2018). The pseudopotential idea is based on the observation that many physical properties of atoms and how they form molecules can be understood via the behaviour of the outermost electrons, the core electrons play only a passive role and so by replacing the strong ionic potential by a softer pseudopotential we can hope to still capture the essential physics (Broglia et al., 2013). An added benefit is that the smooth pseudopotential further reduces the computational burden. A good pseudopotential should be transferable between different systems and describe the outermost electrons of interest accurately in different chemical environments. We will make heavy use of pseudopotentials in the calculations presented in chapter 5.

Often it is appropriate to neglect the background ionic structure and use the jellium model for a considerable gain in simplicity, this has been shown to work well for simple metals systems such as alkali clusters[38] and has had success in describing the "supershell" structure observed in experiments (Knight et al., 1984; Brack, 1993). Given that the existence of plasmons is known to not depend on the ionic background, the jellium model has found heavy use in quantum plasmonics, this is discussed in section 2.3.5. In figure 2.5 the density returned via a LDA-DFT calculation for a sodium sphere of 92 valence electrons within the jellium approximation is presented, it shows a pronounced electron spill-out at the cluster boundary as well as density oscillations in the interior[39].

[38] For the jellium approximation to be valid the valence electrons must be strongly delocalised and there must be no directionality to the binding with the ionic cores (Brack, 1993).

[39] These are related to Friedel oscillations, they are similar to the 'Gibbs phenomenon' for Fourier transforms of a discontinuous step function signal. The presence of a sharp Fermi sphere leads to a smallest possible wavelength and means the electron density can not perfectly screen the jellium background.

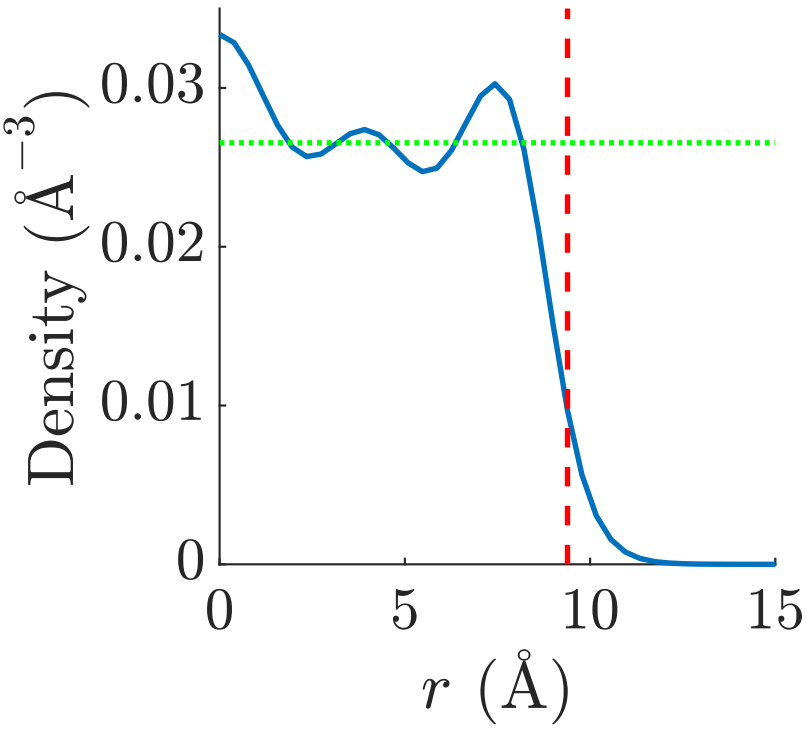


Figure 2.5: The ground state density for the Na_{92} cluster, calculated within LDA-DFT. The green dotted line shows the homogeneous electron gas density for the equivalent Wigner-Seitz radius of $3.93a_0$. The red dashed line shows the radius ($R = r_s N^{1/3}$) of the jellium sphere. This is the input ground state used for the TDDFT result shown in the inset of figure 4.5 in chapter 4.

Finally, it is worth remarking that there is a freedom in which

computational scheme is used for solving the KS equations via the choice of expansion basis for the KS wavefunctions. For periodic systems, expansion in planewaves is most natural and speeds up calculations considerably. For finite systems, planewaves can still be used using a super-cell approach. An expansion in Gaussian functions is common in quantum chemistry codes, this is advantageous as many integrals can be calculated from analytical formulas. It is also possible to not use a basis set and instead represent the wavefunctions on a discrete real-space grid (Andrade et al., 2015), this is the approach we utilise in this thesis. The grid approach is flexible with respect to boundary conditions, simple to discretise, effectively parallelised and competitive on speed with other approaches. Errors can be systematically controlled by changing the grid size. There a few drawbacks: real-space implementations are not variational[40] and the grid can break symmetries artificially (Nogueira et al., 2003). Both issues can be fixed somewhat by choosing a small enough grid. Although we will not explore DFT in great detail within this thesis, all TDDFT calculations need a ground state DFT calculation as input, so an in-depth knowledge of DFT and its capabilities is necessary for quantum plasmonics. There are various open-source DFT codes available online, I have made considerable use of OCTOPUS (Castro et al., 2006; Andrade et al., 2015) and GPAW (Mortensen et al., 2005; Enkovaara et al., 2010) during the PhD.

[40] The energy can be smaller than the real energy.

2.3.2 *Introduction To TDDFT*

TDDFT is an extension of DFT to deal with excitations and time-dependent phenomena. We wish to solve the *time-dependent* many-body Schrödinger equation

$$i\hbar\partial_t\Psi(\mathbf{x}_1,\mathbf{x}_2,...\mathbf{x}_N,t) = \hat{H}\Psi(\mathbf{x}_1,\mathbf{x}_2,...\mathbf{x}_N,t). \tag{2.16}$$

The Hamiltonian has the same form as in the last section but we now include a time dependent potential $v(\mathbf{r},\mathbf{t})$ that couples to the electron density, this could be for example from an impinging laser. This is a very general equation and can describe a huge wealth of potential physical phenomena that is encountered in atomic and molecular physics/chemistry. A particularly relevant situation for us is when the exciting field is weak and we are in the regime of linear response, allowing calculation of optical absorption spectra.

TDDFT is based on a time-dependent extension of the Hohenberg-Kohn theorem called the *Runge–Gross theorem* (Runge and Gross, 1984)

Theorem (Runge–Gross Theorem)**.** *In a time-dependent interacting N-electron system there exists a* ***unique*** *one-to-one correspondence between the external potential $v(\mathbf{r},t)$ and the density $n(\mathbf{r},t)$, i.e. if $v(\mathbf{r},t)$ and $v'(\mathbf{r},t)$ differ by more than a time-dependent function $c(t)$, they cannot produce the same time-dependent density $n(r,t)$.*

The proof proceeds in two steps: first by showing that two different time-dependent potentials must necessarily lead to two different current densities and secondly that this consequently leads to two different time-dependent densities (Ullrich, 2011). A significant consequence of the theorem is that the expectation value of *any* operator is a unique functional of the time-dependent density, this allows calculation of physical observables. In a time-dependent system energy is not a conserved quantity and so it is not possible to build an energy variational principle, instead one can use the action to construct an Euler equation. As in the static case, a non-interacting system of electrons can be considered which reproduces the interacting density $n(\mathbf{r},t)$. This is the time-dependent KS equations

$$\begin{aligned}\left(-\frac{\hbar^2}{2m}\nabla^2 + v_{KS}[n](\mathbf{r},t)\right)\psi_i[n](\mathbf{r},t) &= i\hbar\partial_t\psi_i[n](\mathbf{r},t)\\ n(\mathbf{r},t) &= \sum_i^N |\psi_i(\mathbf{r},t)|^2\\ v_{KS}[n](\mathbf{r},t) = \underbrace{v(\mathbf{r},t)}_{\substack{v_0(\mathbf{r})\\ +v_{ext}(\mathbf{r},t)}} &+ \underbrace{\frac{e^2}{4\pi\epsilon_0}\int d^3r' \frac{n(\mathbf{r}',t)}{|\mathbf{r}-\mathbf{r}'|}}_{v_H[n](\mathbf{r},t)} + v_{XC}[n](\mathbf{r},t).\end{aligned} \tag{2.17}$$

These equations can be solved directly by time propagation of the KS orbitals from an initial DFT calculation to some time t, for instance using a number of different numerical routines in OCTOPUS (Castro et al., 2006; Andrade et al., 2015). As an example, the linear optical absorption spectrum can be obtained by exciting all frequencies using a small delta function kick in time; $v_{ext}(\mathbf{r},t) = -\kappa|\mathbf{r}|\delta(t)$, where κ is a small momentum transferred to all electrons (Castro et al., 2004). This method of obtaining spectra from time-propagation becomes advantageous compared to frequency domain linear response methods for larger systems[41]. These equations can naturally be extended to include nonlinear response and be coupled to molecular dynamics, calculation quality can also be easily checked via energy conservation and sum rules.

[41] Another advantage of this method is that only occupied states need to be calculated. Calculating a high number of unoccupied KS orbital, which can include unbound states, is challenging.

Like with static DFT, the quality of a TDDFT calculation is strongly dependent on the XC kernel used. We have seen that the XC functional must be a complicated quantity, encoding all the subtle correlation effects of an interacting electron gas. On top of that, the time-dependent XC functional must include dynamical effects. This leads to a nonlocal temporal dependence for $v_{XC}(\mathbf{r},t)$, known as a memory dependence (Ullrich, 2011). This means that the XC potential cannot be constructed purely from knowledge of the density at one point in time. The simplest and most common approximation is the adiabatic approximation[42]

$$v_{xc}^{ALDA}(\mathbf{r},t) = \frac{d}{dn}\left(u_{xc}(n)\right)\Big|_{n(\mathbf{r},t)}, \tag{2.18}$$

[42] It was actually first introduced before the formal foundations of TDDFT had been laid down (Zangwill and Soven, 1980).

which ignores memory effects; the static XC functional is evalu-

ated with the instantaneous density $n(\mathbf{r}, t)$. In the limit of a slowly varying (in time) density, this becomes exact. It is commonly combined with the LDA to give the adiabatic local density approximation (ALDA), giving a XC potential local in space *and* time. Its advantages include simplicity and ease of implementation. It is surprisingly successful at predicting accurate excitation energies over a range of different systems (Ullrich, 2011). Of course, the ALDA inherits the weaknesses of the LDA such as the wrong long range spatial behaviour of the XC potential, improvements can often be gained by using better ground state XC potentials and sticking with the adiabatic approximation (Giuliani and Vignale, 2005). Surprisingly, it works well even for high frequency phenomena[43]. Plasmons are known to be well described by TDDFT within the ALDA (Ullrich, 2011), so we need not go beyond this level of approximation. An interesting progression of the work presented later in this thesis in chapter 5 would be to use a dynamic XC kernel with a memory dependence and explore the intrinsic plasmon linewidth due to coupling to multiple eh pairs.

[43] This is partially because the scale of the frequency dependence of v_{XC} is determined by correlated multi-electron excitations. Single electron excitations, which are calculated in a typical TDDFT calculation are lower energy and can therefore treated in an adiabatic limit (Giuliani and Vignale, 2005).

2.3.3 *Linear-Response TDDFT: The Casida Method*

We will construct a linear response theory based on the following two principles:

- For a weak perturbing field the electron density will oscillate around the minimum energy configuration, we can keep only the linear part of the density response. The oscillations will be described by the local environment of the energy surface close to its minimum (Broglia et al., 2013).

- As the density oscillates it causes an induced potential, which in turn acts back on the density. In other words, we must generalise the self-consistent mean-field theory to the dynamic case. We will continue to work in the density functional framework to achieve this.

The great thing about linear response is that we can calculate excited state properties from ground state properties, not only does this make the theory easier it is also a convenient regime for experiments if one wants to probe the ground state[44]. It means one can construct response function from the unperturbed eigenvalues and eigenfunctions; for the Casida method this means we can simply use the KS orbitals returned from a ground state calculation to construct a system of equations that can be solved to return not only the excitation energies, but also a wealth of information on the physical characteristics of the excitations.

[44] The experimentalist usually wants to probe the system not the system modified by the probe.

The linear response theory can be formulated within two complementary pictures: a real-space and an eh representation. The defining equations are best introduced in the more intuitive real-space description and best solved in the eh representation via the

Casida equation. We consider a scalar potential, which couples to the electron density, of the form

$$v(\mathbf{r},t) = v_0(\mathbf{r}) + \Delta v(\mathbf{r},t)\theta(t-t_0), \tag{2.19}$$

where $v_0(\mathbf{r})$ defines the ground state and $\Delta v(\mathbf{r},t)$ is a small time-dependent perturbation switched on at time t_0. The Runge-Gross theorem tells us that we can formally write $n(\mathbf{r},t) = n[v](\mathbf{r},t)$, which we can expand in powers of the perturbation Δv. Keeping to linear order

$$n(\mathbf{r},t) = n_0(\mathbf{r}) + \Delta n(\mathbf{r},t). \tag{2.20}$$

The induced density can be written in the following way within the KS scheme in the frequency domain[45]

$$\begin{aligned}\Delta n(\mathbf{r},\omega) &= \int d^3r' \left.\frac{\delta n[v_{KS}](\mathbf{r},\omega)}{\delta v_{KS}(\mathbf{r}',\omega)}\right|_{v_{KS}[n_0]} \Delta v_{KS}(\mathbf{r}',\omega) \\ &= \int d^3r' \; \chi_{KS}(\mathbf{r},\mathbf{r}',\omega)\Delta v_{KS}(\mathbf{r}',\omega).\end{aligned} \tag{2.21}$$

χ_{KS} is the density-density response function of the non-interacting KS system, the electrons are non-interacting but respond to an effective KS potential

$$\begin{aligned}\Delta v_{KS}(\mathbf{r}',\omega) &= \Delta v(\mathbf{r},\omega) + \frac{e^2}{4\pi\epsilon_0}\int d^3r' \; \frac{\Delta n(\mathbf{r}',\omega)}{|\mathbf{r}-\mathbf{r}'|} + \Delta v_{XC}(\mathbf{r},\omega) \\ &= \Delta v(\mathbf{r},\omega) + \int d^3r' \; \mathfrak{K}(\mathbf{r},\mathbf{r}',\omega)\Delta n(\mathbf{r}',\omega),\end{aligned} \tag{2.22}$$

In the top line, the middle term is simple the frequency-dependent Hartree potential and the final term is the linearised XC potential. $\mathfrak{K} = \mathfrak{K}_H + \mathfrak{K}_{XC}$ is the combined Hartree-XC kernel[46], it is the heart of linear-response theory and controls the level of approximation. These equations reveal that the electrons don't respond to the bare external field Δv, but instead a *total self-consistent* field. In other words, the field experienced by an electron is *screened* by its neighbouring electrons. This is a dynamical mean-field approach; the response of an interacting system is described as a non-interacting system responding to a self-consistent field, which is constructed from global parameters like electron density (Giuliani and Vignale, 2005).

The key observation in constructing a linear response TDDFT is that the non-interacting KS response function is connected to the true interacting density-density response function χ. We can equally write the density response as

$$\Delta n(\mathbf{r},\omega) = \int d^3r' \; \chi(\mathbf{r},\mathbf{r}',\omega)\Delta v(\mathbf{r}',\omega). \tag{2.23}$$

In this representation the interactions are included in the response function not the potential. This means the two response functions

[45] See appendix H for details. As in our discussions in section 1.1, the response function will on depend only on time differences. We assume perturbations are turned on adiabatically from a time in the distant past so that transient contributions can be ignored. It also means that real transitions between states induced by the perturbation are neglected, i.e. the occupation statistics given by the zero-temperature Fermi function is unchanged. Expectation values will change in time only because of changes in the time-dependent wavefunctions.

[46] Where $\mathfrak{K}_H(\mathbf{r},\mathbf{r}') = \frac{1}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}'|}$, note that it is frequency independent.

are related via the following Dyson-like equation[47]

[47] See appendix H for the derivation.

$$\chi(\mathbf{r},\mathbf{r}',\omega) = \chi_{KS}(\mathbf{r},\mathbf{r}',\omega) + \int d^3r_1 d^3r_2\ \chi_{KS}(\mathbf{r},\mathbf{r}_1,\omega)\ \mathfrak{K}(\mathbf{r}_1,\mathbf{r}_2,\omega)\ \chi(\mathbf{r}_2,\mathbf{r}',\omega). \tag{2.24}$$

This can be solved formally, considering the response functions as matrices, to give the interacting response function in terms of the non-interacting response function, which can be constructed from the KS eigenvalues and orbitals, in the following RPA-like form

$$\overline{\chi} = \left(\overline{\mathbf{1}} - \overline{\chi}_{KS}\overline{\mathfrak{K}}\right)^{-1}\overline{\chi}_{KS}. \tag{2.25}$$

In fact, if we set the XC contribution to zero we recover a KS based RPA[48]. The RPA tends to overestimate excitations energies in finite systems, but models plasmonic type excitations well (Ullrich, 2011). The poles of the response functions give the excitations energies of the system, the non-interacting response function, as the name suggests, gives the bare eigenvalue differences[49]. The interacting response function will have poles where the operator $\overline{\mathbf{1}} - \overline{\chi}_{KS}\overline{\mathfrak{K}}$ is not invertible, these are the true excitation energies of the system. The energies of the resulting eh excitations will be renormalised relative to the non-interacting eh pairs. There will also be the formation of new collective modes as a consequence of the electron-electron interaction, which don't appear in the non-interacting spectrum. The physics of quantum plasmons is contained completely within equation 2.25. In general, $\mathfrak{K}_{XC}$ is a retarded response function, but within the ALDA becomes a local function in space and time and hence frequency independent in the frequency domain.

[48] Which can be considered a time-dependent Hartree theory.

[49] In the KS scheme this will be just the KS eigenvalue differences.

An efficient means of solution used in many modern electronic structure codes was introduced by Mark E. Casida (Casida, 1995) by using as a basis the product of occupied and unoccupied orbitals. It is based on earlier methods for time-dependent Hartree-Fock theory (Langhoff et al., 1972) and is a suitable means of calculation for finite systems such as atoms and molecules. It can be considered as a dimension higher than the occupation representation used commonly in physics and is a very natural way to explore excitations: as linear combinations of eh pairs. A pole in the interacting response function means that the system can sustain a response at its excitation frequency in the absence of an external excitation, the Casida method extracts these poles. The Casida equations can be stated as follows[50]:

$$\begin{aligned}\sum_{jb}\left[\delta_{ia,jb}\omega_{bj}^2 + 4\sqrt{\omega_{ai}\omega_{bj}}\mathfrak{K}_{ia,jb}\right]\mathcal{Z}_{jb} &= \Omega_{ia}^2\mathcal{Z}_{ia}\\ \mathfrak{K}_{ia,jb} = \int d^3r\ d^3r'\ \psi_i(\mathbf{r})\psi_a(\mathbf{r})\mathfrak{K}(\mathbf{r},\mathbf{r}')\psi_j(\mathbf{r}')\psi_b(\mathbf{r}')&\\ \omega_{ai} = \varepsilon_a - \varepsilon_i,&\end{aligned} \tag{2.26}$$

[50] See appendix I for a detailed derivation. We assume a spin unpolarized and zero-temperature system, a more general expression can be found in (Martin, 2004).

where I use the standard notation that the indices $\{i,j\}$ refer to occupied orbitals and $\{a,b\}$ to unoccupied. This means the bare

KS energy difference ω_{ai} is a positive quantity, where ε_i is the ith KS eigenvalue. This is an Hermitian form of the equation, which is valid for real KS orbitals and a frequency independent $\mathfrak{K}$. A more general expression can be found in appendix I, but it is this form that is commonly solved in electronic structure codes, such as OCTOPUS. It is an infinite dimension equation so in practise it must be truncated, with a suitable number of unoccupied states chosen. The total eh space will be $N_{pairs} = N_{occ} \times N_{uno}$, the excitations are labelled by the index $I = \{1, 2...N_{pairs}\}$, which stands for the combined indices ia. In the adiabatic limit there will only be as many excitations as the dimensions of the eh space, this means only single excitations[51] are considered. If the kernel has a frequency dependence then it is possible to support more excitations than the dimensions of the eh space; multiple excitations will be included. Fortunately quantum plasmons seem to have a largely singly excited character (Gieseking et al., 2016b) and so are well described within the ALDA. Note that each excitation comes in a pair $\{-\Omega_I, \Omega_I\}$ and corresponds to excitation and de-excitation[52] (stimulated emission) (Ullrich, 2011). In the limit of no interactions, $\mathfrak{K} \to 0$ and the Casida matrix becomes diagonal with the elements giving the bare KS energy differences. Note that the non-interacting system can only support single eh excitations: collective modes cannot appear without electron-electron interactions. ω_{ai} is the 'bare' excitation energy and Ω_{ai} is the 'dressed' excitation energy; the electrons interacts with a cloud of neighbouring electrons effecting how it interacts with an external probe. The dressed electron transitions can be described in terms of the bare KS transitions, using the Casida eigenvectors as weighting coefficients.

The most numerically challenging aspect is construction of the coupling-matrix elements $\mathfrak{K}_{ia,jb}$, which involve a demanding double spatial integral[53]. Once the matrix elements are known, equation 2.26 is simple to solve numerically and returns a set of eigenvectors $\mathcal{Z}_I$ and eigenvalues Ω_I. From these objects, physical quantifies can be derived such as the polarizability, oscillator strength and, most importantly for us, FEs[54]. Note that in the Casida method we do not consider an external potential, only the linearised XC and Hartree potentials. Instead, when calculating physical properties, we 'select' the modes of interest according to the perturbation of interest. For instance, when considering coupling to light, one is interested in the dipole strength function[55].

A couple of useful approximations can be made to the Casida equation. The Tamm-Dancoff approximation (TDA) is when the de-excitations are ignored[56], this a well-known approximation in nuclear physics for the RPA (Fetter and Walecka, 2012) and has been adopted in TDDFT (Hirata and Head-Gordon, 1999). This results in a simpler eigenvalue problem and has some theoretical advantages as it analogous to the configuration interaction singles (CIS) method. It allows the simple transfer of quantum chemistry concepts, such as the collectivity index discussed in chapter 5, to

[51] We speak of single excitations if an expansion of the excited many-body wavefunction Ψ_I is dominated by a singly excited Slater determinant. A double excitation is a many-body excited state dominated by a doubly excited Slater determinant, and so on. A general excitation will have contributions from a mixture of different multiply excited Slater determinants (Szabo and Ostlund, 2012).

[52] This can be seen from $-\Omega_I = -\Omega_{ia} = +\Omega_{ai}$.

[53] Note though that if we use an adiabatic XC kernel we need only calculate it once, usually one would already have this quantity from a regular DFT calculation.

[54] See the next section and appendix L.

[55] Higher order moment can be considered. For instance, the transition quadrupole moment was recently computed to explore quadrupole quantum plasmons in small gold nanorods (Gieseking et al., 2016a).

[56] So only particle-hole (ph) elements are considered, hole-particle (hp) elements are neglected.

TDDFT. The results are often close to the corresponding full TDDFT values as de-excitation tends to have a small contribution. Another approach is possible if the bare KS response function poles are well separated and only differ slightly from the true excitation energies. In which case, an expansion about the pole is possible to significantly simplify the equations (Petersilka et al., 1996). The coupling-matrix elements away from the diagonal tend to be small compared to the diagonals, reflecting the decreasing overlap of energetically separated orbitals, justifying this approach for some systems (Ullrich, 2011). For a spin-unpolarized system, this single-pole approximation gives (Ullrich and Turkowski, 2008)

$$\Omega_{\pm}^2 = \omega_{ai}^2 + 4\omega_{ai}\mathfrak{K}_{ia,ia}, \tag{2.27}$$

which can be further simplified by application of the TDA: $\Omega = \omega_{ai} + 2\mathfrak{K}_{ia,ia}$. This can be viewed as the TDA for a two-level system. Equation 2.27 is instructive as it nicely shows the role of interaction on the excitations energy. By splitting the coupling matrix into its Hartree and XC part, one can observe that the Hartree term blueshifts the excitation energy relative to the bare KS transition while the XC redshifts[57]. As expected, the shift induced by Coulomb effects is much larger than then the XC. The Hartree term is often called a "depolarization shift" in solid state physics. While these two approximations are useful for theoretical considerations, much of the time there is no considerable extra effort to solving the full eigenvalue equation and this is what we will do in this thesis.

[57] There is an interesting case of spin plasmons where the Hartree shifts cancel out and the resulting collective modes are due to XC effects. They are consequently redshifted relative to single-particle excitations (Ullrich and Turkowski, 2008)

2.3.4 *Extracting Physical Observables From The Casida Equations*

Assuming the Casida equation has been solved, the task turns to extracting physical quantities of interest from the returned eigenvectors and eigenvalues. Transitions to excited states are characterised by physical observables such as: dynamic polarization, current flow, bond breaking and formation, etc. These come in two flavours: local quantities[58] which allow visualisation of the excitation and integrated global quantities that are usually linked to measurements, such as optical cross sections.

We begin with the very useful transition density, $\Delta n^I(\mathbf{r})$, which gives the density fluctuation of the Ith Casida excitation. It is most useful for building a picture of excitation as it gives the spatial profile of the density oscillation, it is also necessary for calculating the FE as shown in appendix L. It is defined as follows for real orbitals[59] (Yasuike et al., 2011)

$$\begin{aligned}\Delta n^I(\mathbf{r}) = \langle\Psi_0|\sum_i^N \delta(\mathbf{r}-\mathbf{r}_i)|\Psi_I\rangle &= \sqrt{2}\sum_{ia}\psi_i(\mathbf{r})\psi_a(\mathbf{r})\left(\mathcal{X}_{ia}^I + \mathcal{Y}_{ia}^I\right) \\ &= \sqrt{2}\sum_{ia}\psi_i(\mathbf{r})\psi_a(\mathbf{r})\sqrt{\frac{\omega_{ai}}{\Omega_I}}\mathcal{Z}_{ia}^I.\end{aligned} \tag{2.28}$$

[58] There is also interest in nonlocal quantities such as the transition density matrix and the recently introduced "particle-hole map" (Li and Ullrich, 2015).

[59] There are a few different definition floating about the literature, sometimes without the $\sqrt{2}$ and also sometimes with a minus rather a plus. This version seems to agree with the implementation of OCTOPUS.

$\mathcal{X}_{ia}$ and $\mathcal{Y}_{ia}$ are the eigenvectors from the more general anti-Hermitian form of the Casida, they are directly related to the excitation and de-excitation component of the transition density matrix. The details can be found in appendix I. In the limit that one single excitation completely dominates the response then we can expect the transition density to give the observable induced density, $\Delta n(\mathbf{r}, \omega)$. On a more abstract level, the transition density is the diagonal of the transition density matrix in a spatial representation[60] The integral over all space must be zero for density conservation, and the square, $|\delta n^I(\mathbf{r})|^2$, measures the probability that an electron and hole are simultaneously at position $\mathbf{r}$ (Plasser et al., 2014).

From the transition density it is simple to define the transition dipole moment[61] in the direction $q = \{x, y, z\}$

$$\begin{aligned} \mu_q^I = \langle \Psi_0 | \hat{\mu}_q | \Psi_I \rangle &= \sqrt{2} \sum_{ia} \mu_{q,ia} \left(\mathcal{X}_{ia}^I + \mathcal{Y}_{ia}^I \right) \\ &= \sqrt{2} \sum_{ia} \mu_{q,ia} \left(\sqrt{\omega_{ai}} \mathcal{Z}_{ia}^I \right), \end{aligned} \tag{2.29}$$

where $\mu_{q,ia}$ is the KS transition matrix elements given by: $\mu_{q,ia} = \int d^3r \; \psi_i(\mathbf{r}) q \psi_a(\mathbf{r})$. This makes intuitive sense, the dipole moment of an excitation is the weighted sum of the dipole moments of constituent KS transitions. The transition dipole moment characterises how 'bright' the excitations is, for other moments we would simply replace the KS dipole matrix elements with the relevant quantity.

The dimensionless oscillator strength for a non-interacting system is (Fox, 2002)

$$\mathfrak{f}_{ai} = \frac{2m}{3\hbar} \omega_{ai} \sum_{q=x,y,z} |\, \langle \psi_a | q | \psi_i \rangle \,|^2. \tag{2.30}$$

From which it is natural to define the oscillator strength for excitation I in an interacting system

$$\mathfrak{f}^I = \frac{2m}{3\hbar} \Omega_I \sum_{q=x,y,z} |\mu_q^I|^2 \tag{2.31}$$

Oscillator strengths were introduced in spectroscopy before the advent of quantum theory, they measure the relative strength of absorption and emission lines. The oscillator strengths are convenient for three reasons:

- They are the residues of the mean polarizability (Ullrich, 2011), $\alpha(\omega) = \frac{1}{3} \operatorname{Tr}[\overline{\alpha}(\omega)] = \sum_I \frac{f_I}{(\omega + i\eta)^2 - \Omega_I^2}$.
- They satisfy the well-known Thomas-Reiche-Kuhn sum rule[62] $\sum_{I=1}^{\infty} f_I = N$, where N is the number of electrons considered in the calculation. This gives a convenient check of TDDFT calculations, in particular on the unoccupied state truncation chosen for a Casida calculation.
- The absorption spectrum can be understood in terms of the *dipole strength function*, which is related to the oscillator strength

[60] Looking at the off-diagonal elements provides a way to visualise eh pair dynamics (Li and Ullrich, 2011), for instance to understand charge separation and recombination in excitonic molecules (Plasser and Lischka, 2012). A detailed explanation on obtaining the transition density matrix from the Casida eigenvectors, with and without de-excitation, is given by (Etienne, 2015).

[61] The transition moment is a vector quantity given geometrically by a line joining the centre of charge of the region of induced electron density gain (i.e. negative charge accumulation), with the centre of charge of the region of induced electron density loss (i.e regions of positive charge depletion, it alternatively can be seen as a region of hole density gain).

[62] Also known as the f-sum rule.

by summing up the oscillator strengths at each frequency, $S(\omega) = \sum_I f_I \delta(\omega - \Omega_I)$. To visualise the spectrum, the delta functions are artificially broadened by convolving with Lorentzian functions of width η. The total area under the spectrum will give the number of electrons considered due to the Thomas-Reiche-Kuhn sum rule, $\int d\omega\ S(\omega) = N$.

The dipole strength function is related to the photoabsorption cross section and imaginary part of the polarizability

$$S(\omega) = \frac{2m\omega}{\hbar e^2} \frac{\Im[\alpha(\omega)]}{\pi}. \tag{2.32}$$

Absorption in a quantum system is a dissipation process induced by electronic excitations and can be identified with the imaginary component of the induced density. For an excitation wavelength much larger than the quantum system it is possible to take the dipole approximation and write the polarizability tensor element (Ullrich, 2011)

$$\begin{aligned} \alpha_{qp}(\omega) &= e^2 \int d^3r\ d^3r'\ q\ \chi(\mathbf{r}, \mathbf{r}', \omega)\ p' \\ &= \frac{e^2}{\hbar} \sum_I^{\infty} \frac{2\Omega_I \langle \Psi_0 | \sum_{n=1}^N \hat{q}_n | \Psi_I \rangle \langle \Psi_I | \sum_{n=1}^N \hat{p}_n | \Psi_0 \rangle}{(\omega + i\eta)^2 - \Omega_I^2}, \end{aligned} \tag{2.33}$$

where the Lehman representation of the density-density response function is used. The average polarizability is found by taking the trace, from which the oscillator strengths can be identified.

2.3.5 *The Use Of Density Functional Methods In Plasmonics*

The study of plasmons in TDDFT predates the current boom in plasmonics research. Given that plasmons are well modelled within the LDA and RPA, they were an obvious phenomena of interest in early TDDFT studies. A nice example of this sort of work is provided by Vasiliev *et al* who explored ultra-small sodium clusters/molecules from Na_2 to Na_8 (Vasiliev et al., 2002). TDDFT was also immediately used to explain experimental results in cluster science (Brack, 1993). An early pioneer in applying DFT to small metallic clusters was Ekardt, who self-consistently calculated the effective mean potential and electron density within the LDA and jellium approximation (Ekardt, 1984a). The model was soon extended to the dynamic case, within a frequency domain description, to explore the polarizability (Ekardt, 1984b) and photoabsorption (Ekardt, 1985). Another notable work is from Yabana and Bertsch who studied the excitation spectra of sodium and lithium clusters, as well as C_{60} molecules, in the time-domain (Yabana and Bertsch, 1996).

Some of earliest work that kick-started current interest in quantum plasmonics came from Nordlander's group who used frequency-domain techniques to study nanoshells (Prodan and Nordlander, 2001, 2002), spherical dimers (Zuloaga et al., 2009)

and nanorods (Zuloaga et al., 2010). These studies were performed within the jellium approximation and by making use of suitable geometrical symmetries a large number of conduction electrons could be considered, up to 2.5×10^4 (Prodan and Nordlander, 2002). Time-domain simulations have also been used on jellium spheres, revealing the presence of surface and core plasmons (Townsend and Bryant, 2011). Since then there has been increasing interest in moving beyond the jellium approximation and including atomic structure explicitly (Zhang et al., 2014; Barbry et al., 2015; Yang et al., 2016). As interest in smaller quantum systems, such as atomic wires (Fitzgerald et al., 2017) and polycyclic aromatic hydrocarbons (Manjavacas et al., 2013; Cui et al., 2016), is rising, there is clearly a need to move beyond the jellium model in some cases. But, despite its simplicity, the jellium model continues to be used heavily where possible as it often captures the essential physics and can drastically cut simulation costs. For instance, gold jellium nanospheres have been used to study plasmon-induced water splitting completely within the TDDFT framework (Yan et al., 2016). A number of reviews are available on the subject of modern quantum plasmonics (Varas et al., 2016; Zhu et al., 2016; Morton et al., 2011), including my own modest contribution (Fitzgerald et al., 2016). Particularly interesting recent developments in quantum plasmonics include active quantum plasmonics (Marinica et al., 2015), quantum plasmonic metamaterials (Sementa et al., 2014) and quantum induced nonlinearities (Marinica et al., 2012).

2.3.6 *Phonons In Density Functional Perturbation Theory*

So far in this chapter we have assumed the Born-Oppenheimer approximation and neglected ionic motion, in this limit the large mass difference between the ionic mass and electron mass justifies the decoupling of electronic and ionic motion[63]. We will find that relaxing this approximation to allow electron-phonon coupling opens up a new loss channel relevant to ultra-small quantum plasmonics. This is becoming an increasingly topical subject, having been explored for polycyclic aromatic hydrocarbons (Cui et al., 2016) and later in chapter 5 we discuss it for sodium atomic chains (Fitzgerald et al., 2017). A possible further application could be to study quantum plasmonic structures under attosecond pulsed illumination, where nonadiabatic effects are known to be important (Nisoli et al., 2017). This could lead to plasmon induced bond formation and breaking in clusters.

Calculations of phonon modes can be efficiently done within the density functional framework, with the Hohenberg and Kohn theorem once again showing us the way (Hohenberg and Kohn, 1964). The lattice dynamics of a molecule[64] are determined by the following Schrödinger equation within the Born-Oppenheimer

[63] This can be understood as follows: the small electron mass means the time scale of electronic motions is much faster than ionic motion. Therefore the electrons can adapt instantaneously to the ionic positions (Broglia et al., 2013).

[64] These discussions also apply for infinite extent crystals.

approximation (Baroni et al., 2001)

$$\left(-\frac{\hbar^2}{2}\sum_I \frac{1}{M_I}\frac{\partial^2}{\partial \mathbf{R}^2} + E(\mathbf{R})\right)\Phi(\mathbf{R}) = \mathfrak{E}\Phi(\mathbf{R}), \tag{2.34}$$

where $\{R_I\}$ is the set of all nuclear coordinates and $\{M_I\}$ is their respective mass. The dynamical properties of the lattice is dictated by the eigenvalues $\mathfrak{E}$ and the eigenfunctions, $\Phi(\mathbf{R})$. The energy $E(\mathbf{R})$ is known as the *Born-Oppenheimer energy surface*, it is the ground state energy of a system of interacting electrons in a fixed nuclei Coulombic field (Baroni et al., 2001). It is the 'glue' that holds matter together and dictates bond formation and is a key quantity, for instance allowing one to find the equilibrium geometry of a molecule via the vanishing of the force, $F_I = \frac{\partial E(\mathbf{R})}{\partial \mathbf{R}_I} = 0$. The vibrational frequencies are simply found by the eigenvalues of the Hessian of $E(\mathbf{R})$

$$\det\left|\frac{1}{\sqrt{M_I M_J}}\frac{\partial^2 E(\mathbf{R})}{\partial \mathbf{R}_I \partial \mathbf{R}_J} - \omega^2\right| = 0. \tag{2.35}$$

By using the well-known Hellmann-Feynman theorem (Feynman, 1939), it is possible to obtain the Hessian by calculating the ground state electron density[65], $n_{\mathbf{R}}(\mathbf{r})$, as well as the linear response to geometrical distortion of the nuclear geometry, $\frac{\partial n_{\mathbf{R}}(\mathbf{r})}{\partial \mathbf{R}_I}$. An efficient procedure to obtain the linear response is given by density-functional perturbation theory (DFPT). The mathematics is similar to the linear response theory discussed in the last section and is well explained in the authoritative review by Baroni *et al* (Baroni et al., 2001). In chapter 5 DFPT is used to calculate the phonon normal modes of a sodium chain.

[65] It has a parametric dependence on the nuclei coordinates, which I have indicated by a subscript.

Intermission

The theory has now been covered and the textbook style of this work will now retire over to the real substance of this thesis. I present here three projects I covered during my PhD.

3 Hybrid Plasmon-Phonon Modes In Graphene Plasmon Cavities Made With Silicon Carbide

There is a strong interest in pushing the operation of plasmonic devices to the lower frequencies of mid-IR to THz, which is a fertile ground for spectroscopy. This is challenging to do with metals as they begin to behave as perfect conductors and exhibit weak electric field confinement in this spectral region. An auspicious route is to move away from metals and look at graphene plasmonics and phonon-polaritonics with polar dielectrics. As these excitations can be engineered to occur at similar frequencies, there is the intriguing prospect to couple the two and excite hybrid plasmonic-phononic excitations, which potentially can combine the advantages of each type of mode. In this chapter, we explore this possibility with a versatile graphene covered SiC grating geometry.

3.1 Introduction And System Setup

We consider a monolayer graphene sheet suspended upon an undoped SiC grating, the schematic is shown in figure 3.1. Only TM polarized light at normal incidence is considered and we assume that the SiC layer below the grating is infinitely thick to switch off the transmission channel and make the analysis simpler. The abrupt breaking of symmetry in the SiC plane leads to the excitation of both a graphene SPP and a localised surface phonon polariton (LSPhP) in the SiC. For certain excitation frequencies and geometrical parameters, we can expect a coupling between the two, this could be advantageous for photonic applications with the aim to combine the advantages of graphene plasmonics (large field confinement and tunability) with SiC phononics (low loss), both of which were discussed in section 1.4. It is also a convenient geometry as it allows excitation of graphene plasmons at normal

incidence; the large momentum mismatch between graphene SPPs and normal incident light means this can be challenging.

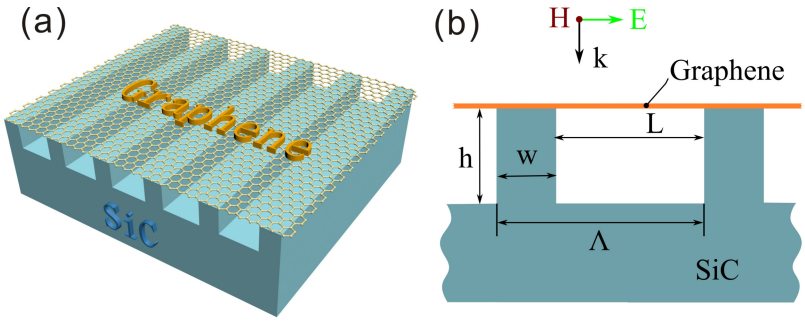


Figure 3.1: Schematic of the SiC grating and graphene device. (a) Three-dimensional illustration of the structure. (b) Cross sectional view with the geometrical parameters indicated: Λ is the grating period, w is the grating ridge width, L is the grove width and h is the grating height. The incident light is normal incident with TM polarization and the structure is considered to be immersed in air.

There has been a lot of interest in the coupling of surface phonons and graphene plasmons; this is because, as a 2D layer, graphene must usually be supported by a substrate that can have a large impact on its optical properties. This is particularly the case for the IR response of graphene located on typical polar dielectric substrates such as SiC (Liu and Willis, 2010; Koch et al., 2010) and SiO_2 (Fei et al., 2011; Yan et al., 2013; Zhu et al., 2014), which is dominated by the coupling of substrate surface phonons with graphene plasmons. There have also been demonstrations for graphene+hBN heterostructures (Brar et al., 2013, 2014; Dai et al., 2015; Woessner et al., 2015; Barcelos et al., 2015) and thin layers of surface-absorbed polymers (Li et al., 2014). These effects can be either desirable or a nuisance. As an example of an application, the interaction of the LSPs of a gold antenna with the surface phonons of a SiO_2 coating has been explored recently, it was found that the strong-coupling between the modes leads to a transparency window (Huck et al., 2016). Another noteworthy example is the demonstration of strong-coupling between a nanometre scale split-ring resonator planar metamaterial with the phonon modes of a thin SiO_2 layer (Shelton et al., 2011). It was shown that the coupling allowed control of the FEs spatial location, giving a maximum field in the dielectric layer rather than the metal surface, in the strong-coupling regime.

In this chapter we report on a numerical study, using finite element analysis with COMSOL multiphysics, on the absorption properties and near-field distribution of the graphene + SiC grating system in the mid-IR. Away from the LSPhP resonance, the SiC cavity induces a cavity effect in the graphene and allows excitation of graphene plasmon standing waves. This is a novel way to structure the photonic response of graphene using the near-field coupling, without direct structuring of the graphene itself. Near the LSPhP resonance we find anti-crossing points in the absorption spectrum, providing evidence of strong-coupling between the two modes. The graphene + SiC cavity geometry thus acts as a convenient means to observe strong-coupling between plasmon and phononic polaritons *and* as a potential platform to observe strongly coupled light-matter

states with deposited molecules.

Strong-coupling is a popular current topic in modern plasmonics. If the coupling between two subsystems is strong enough, then the dynamics is qualitatively modified from the original subsystems; two new hybrid modes arise which are the eigenmodes of the coupled system[1]. In the time domain, this corresponds to *coherent* periodic energy transfer between the subsystems at a rate faster than *both* subsystems dissipative dynamics[2]. In the frequency domain, at the point of energy crossing between the two strongly interacting modes a characteristic splitting is seen, known as Rabi splitting or anti-crossing, each branch is associated with a hybrid mode. Strong-coupling can therefore be defined as the situation when this splitting is observable[3] (Törmä and Barnes, 2014). The model system of strong-coupling is two coupled classical harmonic oscillators[4], which nicely demonstrates the physics of energy level splitting; a brief overview is presented in appendix J.

In nanophotonics, strong-coupling has been explored in the context of interactions between different nanophotonic modes (Zhao and Zhang, 2015; Gubbin et al., 2016) and the interaction of nanophotonic modes with quantum systems[5] (Schlather et al., 2013; Törmä and Barnes, 2014; Zengin et al., 2015; Chikkaraddy et al., 2016). Light-matter interactions tend to be weak because of the large difference in spatial localization between electronic states and photons. This is not always a problem; in spectroscopy one wants to measure the transitions between energy levels, not modify the energy levels themselves. But there is interest in exploring mixed light-matter states, especially single-emitters strongly coupled to the light-field. These could form the building blocks for quantum information systems, such as atomic-scale switches and logic gates. They can operate at the deep quantum limit where addition or subtraction of a single atom or photon can fundamentally change the system state (Khitrova et al., 2006). This is also known as cavity quantum electrodynamics (QED), so-called because the initial developments in this area were based on high-finesse cavities (Raizen et al., 1989), i.e. the quantum emitter is coupled to a cavity mode. Achieving strong-coupling is experimentally challenging, there are two approaches that can be taken:

- High quality factor cavities can be used. This ensures a photon makes many round trips to interact with the same emitter, boosting the interaction. It can be achieved with dielectric microcavities and is often combined with cryogenic temperatures to reduce loss rates in both the cavity and emitter. This is the traditional approach of cavity QED with single emitters (Raizen et al., 1989), the first demonstration of single-atom strong-coupling to cavity modes was done in 1992 (Thompson et al., 1992).

- Another approach is to minimise the size mismatch between light and quantum systems. This is the plasmonic approach; plasmons are used to concentrate the light field to well below the

[1] Note that plasmon and phonon polaritons themselves can be considered as hybrid modes arising from strong-coupling.

[2] Hence the presence of strong-coupling is dictated by the relative values of the coupling strength to the damping rate.

[3] If the linewidths of the uncoupled modes are larger than the Rabi splitting, then it will not be possible to observe it experimentally.

[4] It turns out that for our purposes the classical formalism is equivalent to the quantum mechanical Jaynes-Cummings picture (Zengin et al., 2015).

[5] This can be either atoms, molecules, clusters or quantum dots.

diffraction limit, $\sim \lambda^3$, which is the minimum possible volume of dielectric cavities. Recently, single-molecule strong-coupling was demonstrated in ultra-small plasmonic cavities (Chikkaraddy et al., 2016).

Both approaches have their strengths and weaknesses: dielectric cavities have superb quality factors but diffraction-limited field confinement, plasmonics boasts great confinement and can operate at ambient conditions, but suffers from low quality factors. It will then not come as a surprise that both graphene and polar dielectrics have been touted as a potential means of achieving the holy grail of strong-coupling physics: low loss and high confinement. By combining the strengths of the two building blocks, the graphene + SiC cavity geometry is potentially a suitable platform for strongly coupling light to trace amounts of deposited molecules.

The graphene layer is modelled using the local RPA in the zero-temperature limit[6]. At the frequencies and levels of doping that we consider, the interband transitions can be ignored. The Drude expression, equation 1.50, can be simply substituted into the dispersion relation, equation 1.52, to get

[6] Details can be found in section 1.4.

$$k_x = k_{SPP} = \frac{\pi\hbar^2\epsilon_0(\epsilon_1+\epsilon_2)}{e^2 E_F}\left(1+\frac{i}{\omega\tau}\right)\omega^2, \tag{3.1}$$

where ϵ_1 and ϵ_2 are the permittivities below and above the sheet. Throughout, we use a conservative mobility of $\mu = 10000\ \mathrm{cm}^2/(\mathrm{Vs})$. There is no doubt that the results presented in this chapter could be improved by considering higher quality graphene sheets.

The numerical results discussed in this section were obtained by Dr Ke Li and Xiaofei Xiao. The theoretical analysis and manuscript preparation was performed by myself and I was a joint-first author of the publication (Li et al., 2017) that this chapter is based on. The figures used throughout this chapter are either reproduced or adapted from this publication, with permission given by the American Chemical Society.

3.2 *Graphene Plasmon Standing Waves*

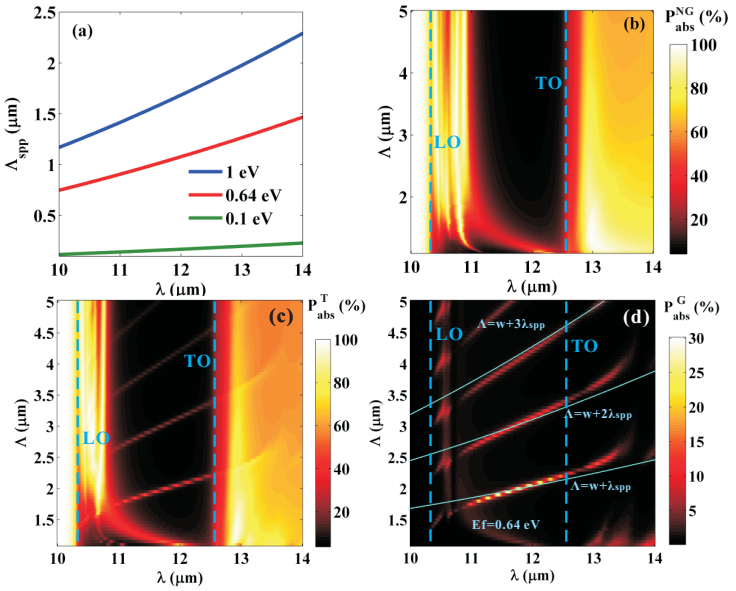
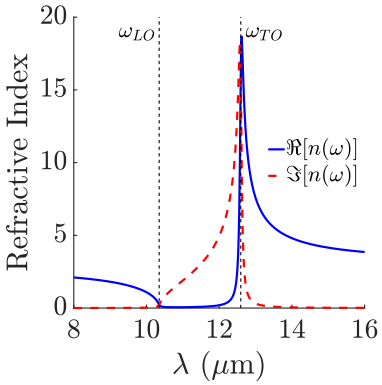


Figure 3.2: (a) The graphene SPP wavelength against excitation wavelength for three different chemical potentials. (b) Total absorption (P^{NG}_{abs}) against grating period (Λ) and incident wavelength (λ) for a bare SiC grating with no graphene. (c) Total absorption (P^{T}_{abs}) for the hybrid SiC + graphene structure. (d) Absorption in just the graphene layer (P^{G}_{abs}). The blue continuous lines indicate a fit to a Fabry-Pérot model with a phase shift of $-\pi$. For (b), (c) and (d) we set $w = h = 1$ μm.

We begin by estimating how large we should set the period of the grating structure, for this we must consider the graphene SPP wavelength at the excitation wavelengths of interest, i.e. located within the reststrahlen band of SiC. We need spectral overlap between the plasmonic and phononic excitations for coupling. Using equation 3.1 we can plot the SPP dispersion, shown in figure 3.2(a), and find that the plasmon wavelength is on the order of $10 \rightarrow 100\times$ smaller than the free-space wavelength, depending on the chemical potential of the graphene. This reveals the large field confinement and momentum mismatch that must be overcome for efficient coupling of light and plasmons. The dispersion is shown for three experimentally realistic Fermi energies, revealing the large tunability that makes graphene ideal for building active nanophotonic devices. Unless otherwise specified, we will set $E_F = 0.64$ eV in this chapter. Using figure 3.2(a), we can see that we should consider grating periods around $10\times$ smaller than the excitation wavelength.

We now compare the absorption for different grating periods of the SiC grating, with and without the graphene layer. We fix $w = h = 1$ μm, so changing the period corresponds to changing the cavity width L. Starting with the bare SiC grating with no graphene, shown in figure 3.2(b), we see that within the reststrahlen band, indicated by the dashed blue lines, there is a number of absorption peaks just above the LO wavelength and above that a region of low absorption. The region of high absorption corresponds to the excitation of LSPhPs, made possible by the structuring of



The refractive index of SiC, note that the real part becomes very small in the reststrahlen band leading to $R \sim 1$. Comparison should be made with figure 1.2(b) for the case of a metal below the plasma frequency.

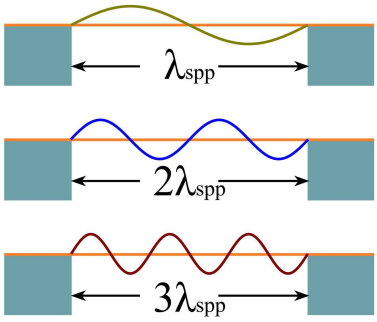


Illustration of cavity standing waves in the graphene sheet.

the SiC. This should be compared to figure 1.20 where no peaks in absorption are seen within the reststrahlen band. At longer wavelengths, but still within the reststrahlen band, the SiC acts as a perfect electrical conductor; it behaves 'metallic' and hence is a very good reflector. Outside of the reststrahlen band we see some strong absorption, it is a very similar behaviour to that seen for a flat SiC interface.

Next we consider the addition of a graphene layer on top of the SiC grating and study the absorption. We analyse two relevant quantities; the total absorption, figure 3.2(c), and the absorption purely in the graphene layer, figure 3.2(d). The most apparent difference, when compared to figure 3.2(b), is the appearance of absorption peaks in the previously non-absorbing part of the reststrahlen band. They only appear for certain cavity widths so we may take a guess that they are standing wave modes, their appearance in both 3.2(c) and (d) would suggest in the graphene. This is confirmed by looking at the electric near-field, shown in figure 3.3(a). The modes occur at integer numbers of the ratio L/λ_{SPP} and the near-field shows they correspond to a discrete number of graphene SPP wavelengths fitting in the cavity width, this is shown in figure 3.3(a) for different values of w at a fixed wavelength. This can be understood by noting that SiC acts as a very good mirror in this region of the reststrahlen band and screens out external fields very effectively. Thus, the SiC enforces the field to zero at the cavity boundaries, leading to the excitation of only antisymmetric (odd) plasmon modes[7], this is because a normal-incident field excites SPPs at the SiC edges that have an opposite phase and can only constructively interfere for odd modes (Du et al., 2014). For non-normal light it is possible to excite symmetric plasmon standing waves in the cavity as the symmetry of the light-matter system in the x-direction is broken. The small peak to the far left of figure 3.3(a) is present even without the graphene and is due to near-field coupling between the SiC slabs for small cavities. Note that the FE near the graphene is rather modest, on the order of 10. This is far below the maximum achievable with graphene[8] so there is clear room for improvement of the geometry in this respect.

It is worth emphasising that the modes we explore are not diffraction modes, the momentum to excite the SPP is from the cavity edges not from the periodicity. Proof of this is provided by two means:

- The modes are not reproduced by the diffraction equation $k_{SPP} = N\frac{2\pi}{\Lambda}$, where N is the diffraction order.
- The peak position does not change for varying Λ (for a fixed cavity length L), as illustrated nicely in figure 3.3(a).

A grating is only considered because it simplifies the numerical calculations and corresponds to a probable experimental setup. Similar results would be obtained for an isolated cavity.

[7] This is specified with respect to the plane of symmetry at $x = 0$ and refers to symmetry of the electric field. The magnetic field profile in comparison will be symmetric.

[8] For example, see figure 1.19.

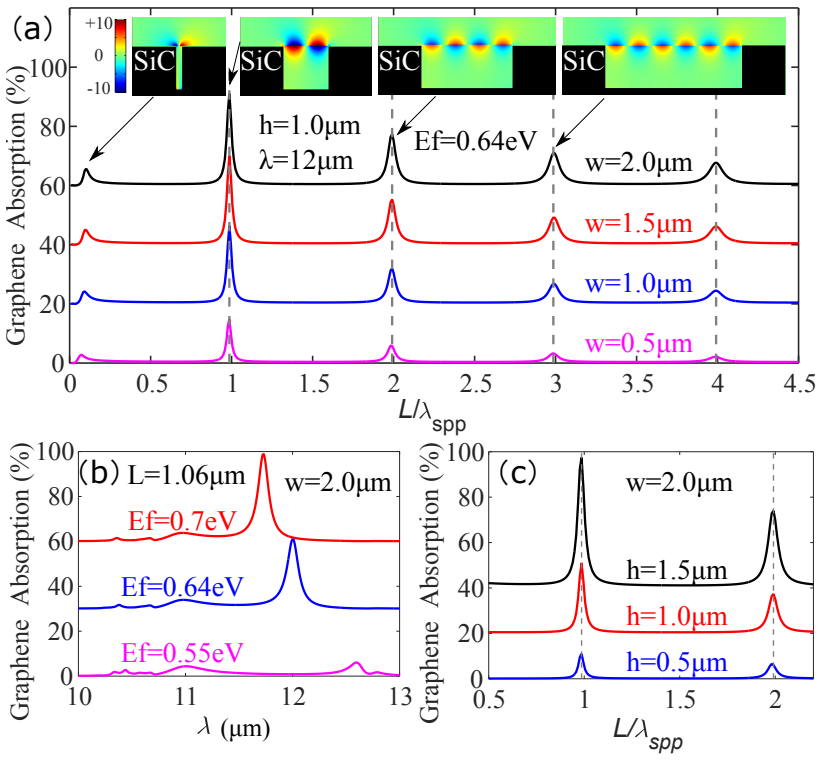


Figure 3.3: (a) Absorption in the graphene layer at different geometrical parameters of the SiC grating. The peaks in absorption appear at integer values of L/λ_{SPP}, indicating the presence of graphene plasmon standing waves. For each peak the corresponding electric field profile is shown. (b) Absorption in the graphene layer against excitation wavelength for various Fermi energies (for fixed parameters $L = 1.06$ μm, $w = 2.0$ μm and $h = 1.0$ μm). (c) Absorption in the graphene layer at different geometrical parameters L/λ_{SPP} for different cavity heights h (with $E_F = 0.64$ eV and $w = 2.0$ μm).

To further confirm our intuition that the repeated absorption feature in figure 3.2(d) is because of the excitation of standing waves, we have used a simple Fabry-Pérot model based on the following equation

$$\delta\phi + \Re[k_{SPP}(\omega)]L = m\pi, \tag{3.2}$$

where $\delta\phi$ is a phase shift experienced at the point of contact between the graphene and SiC, and m is an integer > 0. Taking the phase shift to be $-\pi$, which is suitable for a perfect reflector boundary condition[9], we find a good agreement with the simulation results around the TO frequency (the right blue dashed line in figure 3.2(d)). The fit is not perfect and deviates away from the TO frequency, which we can attribute to a frequency dependence of the phase shift. How 'perfect' a reflector the SiC is depends on frequency, as it becomes less reflecting the SPP field can increasingly penetrate into the SiC and accumulate extra phase on reflection. This is an interesting feature of using SiC that cannot be replicated with metals at these frequencies as they will behave as perfect conductors over the entire frequency range considered in this chapter (Zhao and Zhang, 2015). It opens up the intriguing possibility to manipulate boundary conditions of graphene plasmon resonators. Another interesting possibility would be to consider closely separated graphene plasmonic resonators in a SiC environment where the near-field coupling between the two would have a strong frequency dependence.

[9] In comparison, a phase shift of $\approx \pi/4$ is found for free-standing graphene nanoribbons in vacuum (Nikitin et al., 2014; Du et al., 2014). The phase shift indicates how far the plasmonic field penetrates into the outside material and can be considered in terms of an effective length of the resonator, larger than the geometric length (Novotny, 2007).

This system is very suitable for a practical setup. It is easily excited, tunable due to the presence of the graphene and also exhibits the highest FE on the graphene layer which is convenient for potential applications for molecular sensing. In an experiment it would be more natural to explore this system by scanning the excitation wavelength rather than the geometrical parameters, which would usually be fixed for a sample. This is demonstrated in figure 3.3(b), with parameters $L = 1.06$ μm, $w = 2.0$ μm and $h = 1.0$ μm, for various chemical potentials to illustrate the tunability of the system. We find that as the Fermi energy is increased, the absorption peak blueshifts[10] and grows in intensity. This makes physical sense as a larger chemical potential means an increased carrier concentration which leads to a greater restoring force and hence a blueshifted plasmon frequency. The increased carrier concentration also leads to a larger oscillator strength and hence the observed growth in intensity. The tunability is an important feature of the system as bare SiC exhibits only a limited intrinsic tunability (Dunkelberger et al., 2018).

[10] Note that there is a small mistake in (Li et al., 2017) where this is incorrectly stated as a redshift.

It is interesting to explore the role of the cavity height, in figure 3.3(c) the absorption for various cavity heights h is shown. The resonance position is found to be independent of h, this is a consequence of exploring large heights ($h > 500$ nm) so that the SPP depends only on the material directly above and below the graphene layer; in this case air. For shorter cavities the plasmonic near-field would interact with the SiC at the cavity bottom and a more complicated dependence of the spectral position on h can be expected. In contrast, the absorption percentage in the graphene is dependent on h, growing with increasing height, this is due to moving towards the condition for a Fabry-Pérot resonance in the vertical direction. Recently it has been demonstrated that complete absorption can be achieved for certain cavity heights when this condition is met[11] (Xiao et al., 2018).

[11] Physically, this corresponds to the reflection from the graphene sheet being completely cancelled by the reflected wave from the cavity bottom (Xiao et al., 2018).

It is interesting to consider the appropriateness of the cavity as a platform for inducing strong light-matter interactions, for instance for molecules deposited on the graphene layer. A suitable characterising quantity is the ratio of the quality factor over the volume, Q/V, also known as the Purcell factor. To calculate a volume we consider a square cavity, with sides of length L and $w = 1$ μm, the out-of-plane length is taken to be the out-of-plane decay length, δ_z of the plasmon. For the case of large confinement $k_{SPP} >> \omega\epsilon/c$, therefore the out-of-plane wavevector is given by[12] $k_z \approx k_{SPP}$ and therefore

$$\delta_z(\omega) = \frac{1}{\Re[k_z(\omega)]} \approx \frac{e^2 E_F}{(\epsilon_1 + \epsilon_2)\epsilon_0 \pi \hbar^2 \omega^2}. \tag{3.3}$$

[12] This depends on the convention for the out-of-plane wavenumber: $e^{ik_z z}$ or $e^{-k_z z}$. This gives $k_z \approx ik_{SPP}$ or $k_z \approx k_{SPP}$ respectively in the non-retarded regime. In reference (Li et al., 2017) the first convention was used meaning the imaginary part of the inverse gives the decay length, while in this work the second convention is used meaning the real part of the inverse gives the decay length.

For 12 μm incident wavelength this distance corresponds to 0.17 μm. Note that the decay length is heavily frequency dependent and can be further manipulated via the chemical potential and dielectric environment. To obtain the quality factor we use the

estimation $Q \approx \omega/\Delta\omega$, where $\Delta\omega$ is the linewidth obtained by a Lorentzian fit. We find a value of 96, this is larger than typical plasmonic values, for instance $Q = 16$ in reference (Chikkaraddy et al., 2016), but far smaller than in dielectric systems such as photonic crystal nanocavities, which are on the order $10^4 \rightarrow 10^5$ (Yoshie et al., 2004). At an excitation wavelength 12 μm, we find $Q/V \approx 8 \times 10^5/\lambda_0^3$, which is comparable to values obtained for other graphene nanoresonators reported in the literature (Brar et al., 2013). To give some more context for these numbers, our value compares well with different dielectric systems commonly used to observe Rabi splitting of single quantum objects, such as photonic crystal cavities, micropillars and microdisks. As these systems, at smallest, have confining volumes on the order $(\lambda_0/n)^3$, the Purcell factor $(Q/V)(\lambda_0/n)^3 \sim Q$, which for the aforementioned systems can reach $10^4 \rightarrow 10^5$ (Yoshie et al., 2004; Khitrova et al., 2006). The Purcell factor for our system falls somewhat short of the impressive value of $Q/V \sim 40 \times 10^6/(\lambda_0/n)^3$ reported by Chikkaraddy *et al*, which can be considered state-of-the-art. This was achieved by an ultra-small cavity volume of $\sim$ 36 nm^3 using a nanoparticle-on-mirror geometry (Chikkaraddy et al., 2016). Of course, our system has the redeeming quality that it should be much simpler to construct, it also operates in a different spectral range. Phononic and graphene cavities provide an interesting middle ground in-between dielectric and metal-based plasmonic cavities, whether they can eventually outperform the traditional alternative remains an open question.

3.3 *Strong-Coupling Between Graphene Plasmons And Cavity Surface Phonon Polaritons*

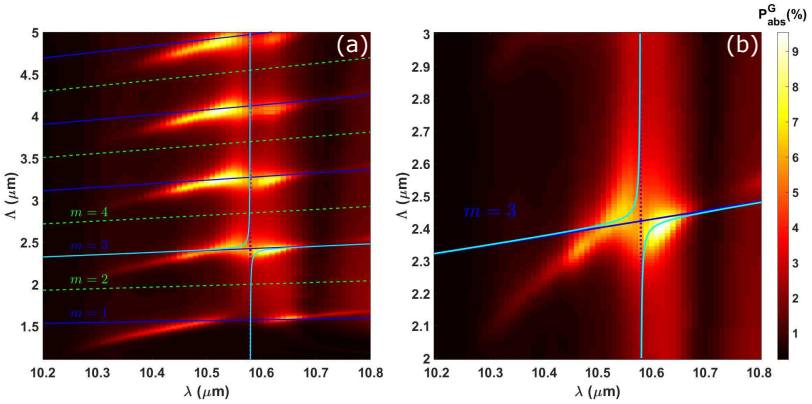


Figure 3.4: (a) Zoom in of the absorption in the graphene layer (for $w = h = 1$ μm) highlighting Rabi splitting. (b) Further zoom in for $m = 3$ mode. The purple dotted line indicates the LSPhP wavelength and the blue solid line shows the graphene SPP dispersion using the Fabry-Pérot model with a phase shift of -0.35π. The cyan line shows the Rabi splitting of the hybrid mode obtained using a coupled harmonic oscillator model (equation 3.4).

Finally for this chapter, we explore another interesting feature revealed in 3.2(d) and shown in more detail via a zoom-in in figure 3.4. We can observe a set of anti-crossing points in the absorption spectrum, this is the manifestation of strong-coupling between a set of graphene plasmons and an LSPhP mode in the SiC. The coupling can be intuitively understood using a simple model of two coupled harmonic oscillators, see appendix J. The particular shape of the Rabi splitting seen in figures 3.4(a) and (b) is due to the dispersive graphene SPP and the SiC cavity LSPhP at a fixed wavelength of 10.6 μm. Entering the strong-coupling regime is highly dependent on the damping of the two subsystems, therefore the small linewidths of the graphene plasmons and SiC LSPhPs should be favourable for the observation of Rabi splitting[13]. To quantify the coupling we apply a fit to the numerical results using the sum of two Lorentzian functions, this is shown figure 3.5. The peak separation (Rabi splitting), $\hbar\omega_R$, is found to be 1 meV and can be used within a coupled harmonic oscillator model

$$E_{\pm}(\omega) = \frac{\hbar\omega_{SPP} + \hbar\omega_{SPhP}}{2} \pm \frac{1}{2}\sqrt{(\hbar\omega_R)^2 + (\hbar\omega_{SPP} - \hbar\omega_{SPhP})^2}, \tag{3.4}$$

which has been used heavily in plasmonics (Schlather et al., 2013) and phonon-polaritonics (Gubbin et al., 2016). The LSPhP energy is taken to be fixed at a value of 0.12 eV and the graphene SPP energy is given by the combined dispersion relations 3.1 and 3.2

$$\omega_{SPP} = \sqrt{\frac{(m\pi - \delta\phi)\, e^2 E_F}{\pi\hbar^2\epsilon_0\,(\epsilon_1 + \epsilon_2)}}. \tag{3.5}$$

The phase shift $\delta\phi$ is difficult to determine and will play the role of a fitting parameter, we will also take $\epsilon_1 = \epsilon_2 = 1$. To obtain a good fit for the whole family of graphene plasmon modes displayed in figure 3.4, a phase shift of $\delta\phi = -0.35\pi$ is chosen[14]. This indicates

[13] It is important to note that near to the LO frequency, graphene SPPs can become lossy due to interactions with substrate phonons. In the cavity system a substantial amount of the graphene is free standing and not in contact with the substrate, which should negate this.

[14] Note that there is a misprint in (Li et al., 2017), which states that the fit is for $m = 2$: consequently the phase shift quoted of -1.35π is incorrect by a difference of π. This does not effect any of the physics discussed in the paper and the fitting procedure is correct.

that at these frequencies the SiC is far from a perfect reflector and allows significant field penetration.

Interestingly, the near-field plots displayed in figure 3.5 show the different behaviour of the hybrid mode, depending on the excitation frequency. At lower frequencies, the mode is more phonon-polariton-like and the field is concentrated at the cavity bottom. At higher frequencies, the mode is graphene-plasmon-like and the field is concentrated in the graphene layer, it is a similar field profile to the standing-wave modes discussed previously. Therefore, over a small range of frequencies, one has a huge control over the spatial distribution of the near-field. This is a big advantage of hybrid modes brought about by strong-coupling, such behaviour has obvious possible application for plasmon based switches. Further control of the relative contribution of each constituent mode could be achieved by manipulating the cavity width and graphene doping.

To quantify the strength of the coupling, we can compare the relative size of the linewidths of the constituent modes (W_1 and W_2) to the splitting D. For Rabi splitting to be experimentally visible, which can be taken as a practical definition of strong-coupling (Törmä and Barnes, 2014), we need $D \gtrapprox W$. Taking the largest linewidth, we get the ratio $\frac{D}{W} \approx 1$, which gives some justification for considering the interaction of the two modes as strong-coupling. Of course it is just on the threshold and there is some arbitrariness to the definition of strong-coupling. Further optimisation of the geometry can be expected to improve this figure.

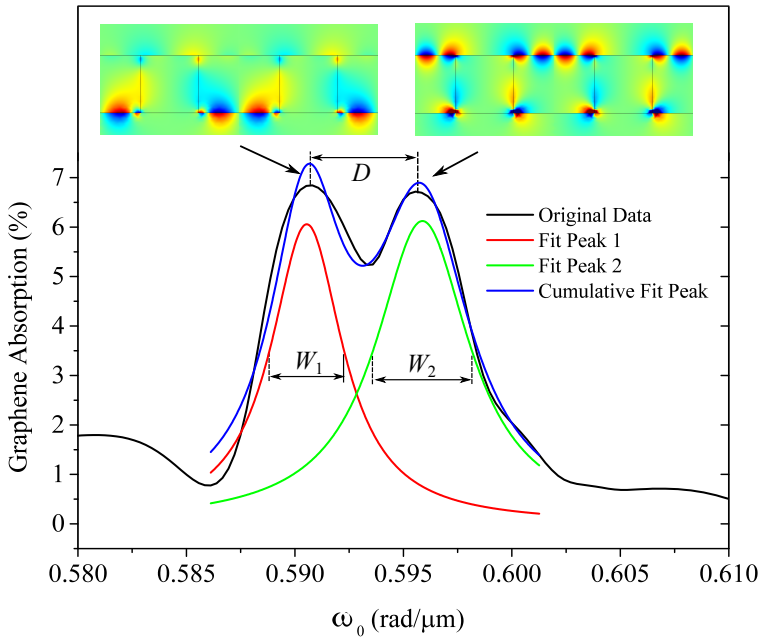


Figure 3.5: Spectrum for parameters $\Lambda = 2.45\ \mu$m and $L = 1.45\ \mu$m. A sum of two Lorentzians is used to fit the peaks, the full width at half maximum of the two is $W_1 = 0.004$ rad/μm and $W_2 = 0.005$ rad/μm. The peak distance is $D = 0.005$ rad/μm.

4 The Ultimate Limits Of Field Enhancement In Plasmonic Nanolenses

The nanolens, a finite chain of self-similar spheres, is one of the most familiar structures in plasmonics and was proposed as a means to deliver massive FE and confinement. The reality has been less impressive, with experiments not demonstrating any significant advantages compared to other less complex structures such as the nanosphere dimer. Here we pass a critical eye over the nanolens and explore boosting the FE by extending the large size difference between constituent spheres that is critical for strong nanolensing, via two strategies: by going small into the quantum limit, and going large by considering operation in the mid-IR to THz using polar dielectrics.

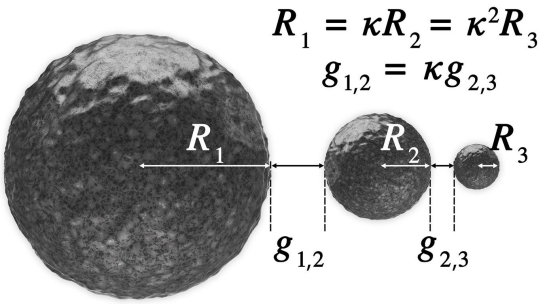


Figure 4.1: The nanolens geometry considered in this chapter, illustrated for a three-particle nanolens. The geometry is completely specified by the three parameters R_3, $g_{2,3}$ and κ.

4.1 Introduction To The Nanolens

The plasmonic cascade nanolens is a widely studied structure in plasmonics, offering potentially huge FEs. The design was first proposed in 2003 by Li *et al* and has fuelled a large body of further work, both experimental (Bidault et al., 2008; Kneipp et al., 2008; Ding et al., 2010; Kravets et al., 2010; Höppener et al., 2012;

Coluccio et al., 2015; Das et al., 2016; Heck et al., 2017; Lloyd et al., 2017; Heck et al., 2018) and theoretical (Li et al., 2006; Dai et al., 2008; Pellegrini et al., 2016), in the subsequent years since the initial publication (Li et al., 2003). The highly confined fields could have applications in nano-localised and controllable chemistry (Cortés et al., 2017), sensing (Mayer and Hafner, 2011), energy conversion (Atwater and Polman, 2010), nonlinear plasmonics (Kauranen and Zayats, 2012) and spectroscopies with light (Le Ru and Etchegoin, 2008). They also support large field gradients useful for exciting non-dipolar transitions. One of the ultimate aims in modern photonics is controllable and reliable surface-based single-molecule spectroscopies, the cascade effect is one of the most promising routes towards this.

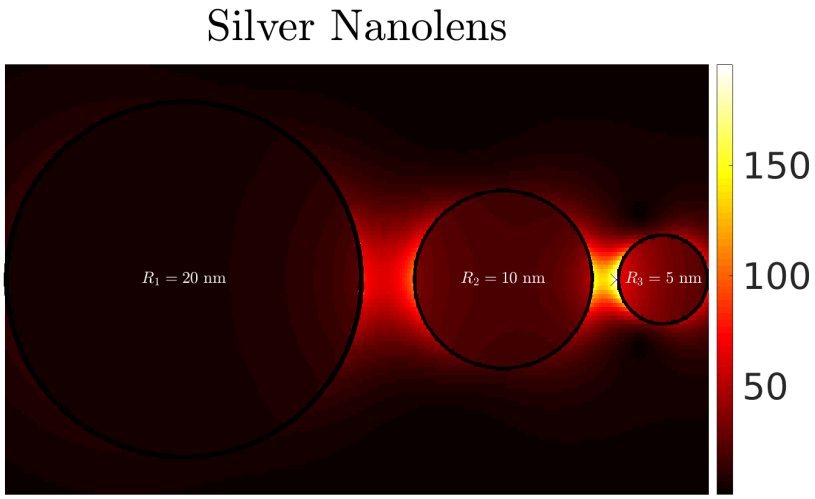


Figure 4.2: Electric field map for a silver nanolens of $R_3 = 5$ nm, $g_{2,3} = 3$ nm and $\kappa = 2$. The blue cross is 0.3 nm away from the surface of the smallest nanosphere and indicates the point we record the electric field value at for the rest of this chapter.

The basic idea behind the nanolens is very simple. A single plasmonic sphere excited at its LSP dipole resonance will enhance the field on the order of the quality factor Q. If a smaller sphere of the same material is placed in the plasmonic near-field, and it is chosen small enough that it does not significantly perturb the larger nanosphere, then it will experience a field Q and will have its own plasmonic near-field on the order Q^2. This of course can be repeated again and again; for N progressively smaller nanospheres we can, in principle, achieve a FE on the order Q^N. For comparison, the plasmonic dimer with a small gap can achieve a FE on the order of Q^2. This though cannot be the whole truth as the field map of a nanolens at resonance, an example is shown in figure 4.2, reveals the field is enhanced most strongly in the smallest gap on the surface of the smallest sphere; it is not a symmetrical dipole field distribution around the smallest sphere. A deeper understanding can be gained by considering higher-order plasmonic coupling between spheres, as illustrated in figure 4.3 for an asymmetrical dimer. In practise, the largest nanosphere will induce a large field gradient. The next largest nanosphere will not feel a constant planewave excitation but instead a spatially inhomogeneous driving field; this

allows the excitation of higher order plasmonics modes, see the discussions in section 1.2.3. Higher order modes are characterized by a stronger confinement along the polarization direction. Thus, exciting higher order modes is an excellent way to confine and enhance electric fields and is the secret to nanolensing; the self-similar structure allows the efficient excitation of very high-order plasmon modes.

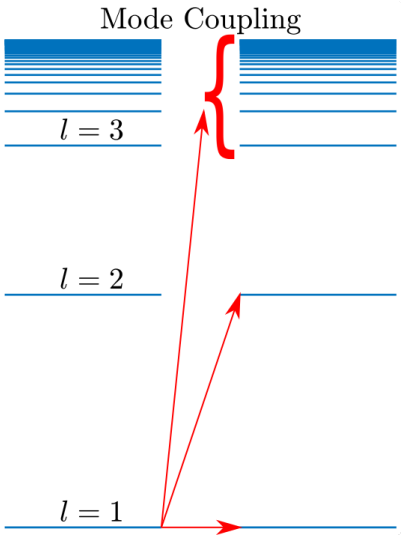


Figure 4.3: Illustration of the plasmon mode coupling between an asymmetrical pair of nanospheres. The larger sphere is assumed large enough that the smaller sphere does not significantly perturb it, but small enough that only the dipole mode is excited by the external field.

The original paper introducing the nanolens (Li et al., 2003), and our discussions above, were based on the electrostatic limit. It was quickly realised that retardation would limit the FE considerably for constituent spheres larger than a few tens of nanometres (Li et al., 2006; Dai et al., 2008). Ideally one would stay within the electrostatic approximation but a large size difference between spheres leads to stronger nanolensing, this would suggest an optimum size for a nanolens. To explore this in more detail we define the geometry as shown in figure 4.1 for a three-particle nanolens. For a N particle nanolens we label the smallest nanosphere radius R_N, the next smallest is labelled R_{N-1} and has the radius κR_N, and so on till the largest nanosphere $R_1 = \kappa^{N-1} R_N$. We also define the gap sizes in a similar fashion: the smallest gap is $g_{N-1,N}$, the next smallest gap is $g_{N-2,N-1} = \kappa g_{N-1,N}$ and so on till the largest gap $g_{1,2} = \kappa^{N-2} g_{N-1,N}$. We choose the same ratio, κ, for both the radii and gap sizes for simplicity and to allow us to explore different geometries via a single parameter, but there is no guarantee that this will give the largest possible FEs. A large size ratio is characterized by $\kappa >> 1$ and leads to strong nanolensing.

We begin by considering a silver nanolens and vary κ from $\kappa = 1$ (all spheres the same size) to $\kappa = 4$ (the largest sphere has a radius 16 times larger than the smallest). The smallest nanosphere, R_3, is chosen to be 5 nm and the smallest gap, $g_{2,3}$, is 3 nm, this is similar to the parameters used by (Dai et al., 2008). We measure the FE a distance of 0.3 nm away from the smallest nanoparticle within the gap between the two smallest nanospheres, this point

is indicated by a blue cross in figure 4.2. It is of course a rather arbitrary point to record and was chosen to be close to the surface of the smallest sphere, where the largest FE will be, while being on the limit of a physically meaningful distance in the macroscopic picture. The results are shown in figure 4.4, for varying values of κ. It is convenient to extract the maximum FE from the near-field spectrum for each κ and plot it, as is done in figure 4.4(a). It should be stressed that these FE may correspond to different spectral positions as the near-field resonance may shift with changing geometry. In figures 4.4(b) and (c) the spectra are shown for $\kappa = 2$ and $\kappa = 4$ respectively. We find already at small κ of around 2 the retarded solution is reaching an optimum value, above which retardation limits the FE. In contrast, the electrostatic calculation shows a monotonic increase with κ and shows no upper limit to the FE. Studying the spectra we observe that the nanolensing results in a single intense peak, this breaks down for the retarded result in figure 4.4(c) (for $\kappa = 4$), where multiple peaks can be seen. We found in general that the presence of a single peak is a good indicator of nanolensing. Interestingly, due to the difficulty in creating nanolenses, experiments have been limited to small κ; the results in figure 4.4 suggest that there is little reason to consider larger κ and that the rather modest[1] FE values obtained in experiments to date with metal nanolenses are unlikely to be significantly improved. Therefore, in this chapter we explore two approaches to boost the nanolensing effect:

[1] Relative to the FEs obtained using simpler geometries such as dimers.

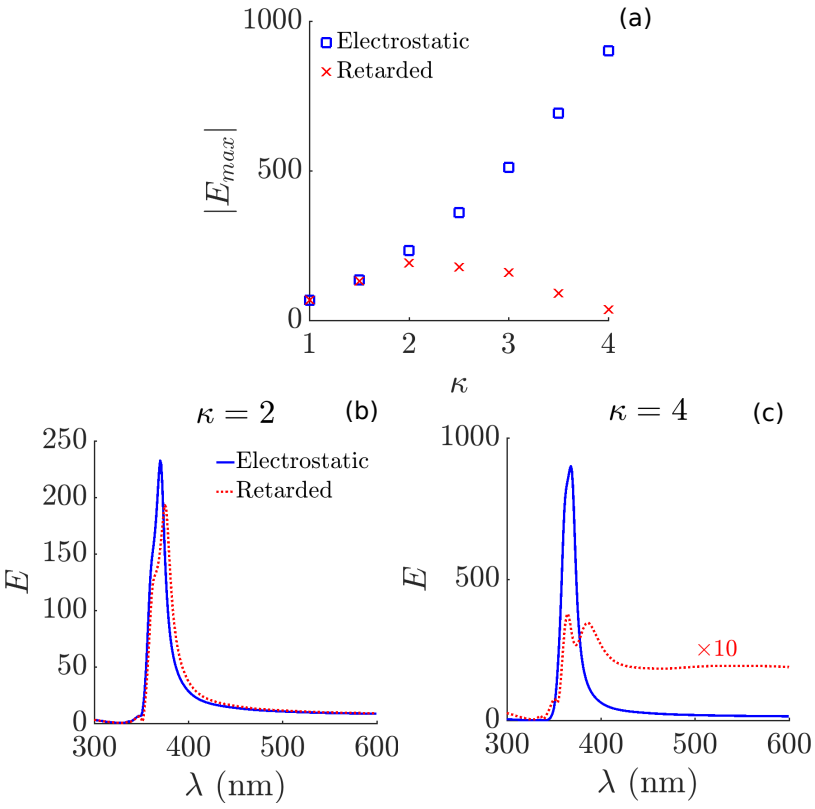


Figure 4.4: (a) The maximum electric field, for a sweep of the geometrical parameter κ, of a silver nanolens with $R_3 = 5$ nm and $g_{2,3} = 3$ nm, calculated within the electrostatic approximation (blue squares) and full retarded calculation (red crosses) using the BEM. (b) The electric field for a $\kappa = 2$ and (c) $\kappa = 4$ nanolens. Note in (c) that the retarded result (red dotted line) is multiplied by 10 to display clearly on the same scale as the electrostatic result (blue solid line). Dielectric data used was from (Johnson and Christy, 1972).

- Our first approach is to decrease the size of the smallest sphere so it is possible to remain within the electrostatic approximation for a larger range of κ. To accurately model nanolenses in this size regime it becomes important to include nonlocal effects. We perform an in depth study of sodium, which due to its similarities with the free electron gas is an excellent test metal, and consider the role of nonlocality in terms of spatial dispersion and surface scattering separately.
- The second approach is to ditch metals and use materials with longer wavelength resonances; this results in a larger range of κ being within the electrostatic approximation, assuming the material can be nanostructured on the same small scale as metals. Furthermore, metals are famously lossy, using materials such as polar dielectrics that support high quality factor resonances should offer superior nanolensing.

The results discussed in this chapter were published in the following paper (Fitzgerald and Giannini, 2018). Some of the figures used in this chapter are either reproduced or adapted from this publication, with permission given by the American Chemical Society.

4.2 Going Small: The effect Of Nonlocality On The Nanolens

4.2.1 Going Small Within The Classical Model

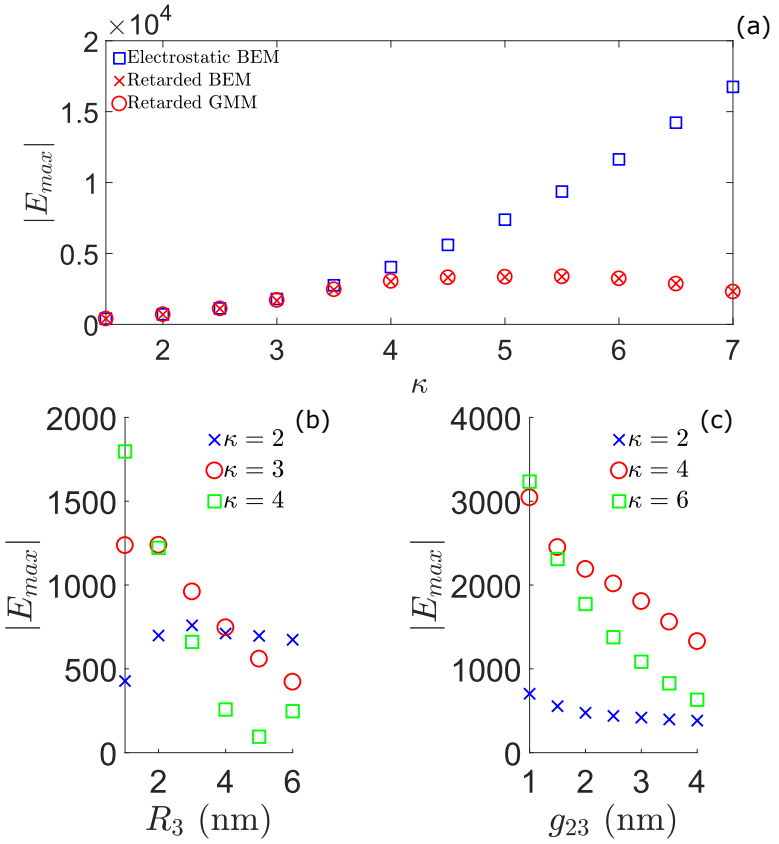


Figure 4.5: (a) Maximum FE for a sodium nanolens of $R_3 = 0.94$ nm and $g_{2,3} = 1$ nm, for varying κ. For the full retarded calculation (red markers) both the BEM (crosses) and GMM (circles) results are shown, showing excellent agreement. (b) Maximum FE for varying smallest nanosphere radii R_3 and (c) smallest gap size $g_{2,3}$. Dielectric function used was a Drude approximation with parameters $\omega_P = 6.05$ eV and $\gamma = 0.0268$ eV.

We begin by displaying the results for a κ sweep of a sodium nanolens with geometrical parameters $R_3 = 0.94$ nm and $g_{2,3} = 1$ nm, within the classical model, see figure 4.5(a). The rather specific radius for R_3 corresponds to a closed shell Na_{92} cluster, the gap size of 1 nm is about as small as one can go before quantum tunnelling starts to become important[2]. The results for figure 4.5(a) are similar to what we saw for figure 4.4, although now we see good agreement between the electrostatic approximation and the full retarded solution up to a larger κ of 4 due to the smaller size of the structure. We also find larger FEs, with a maximum value of over 3000 around $\kappa = 5$ predicted by the retarded calculation. This nicely illustrates that within the classical model, the 'going small' strategy works. In reality we must include quantum effects at these sizes for accurate results. We also show the results for a separate GMM calculation (Pellegrini et al., 2007) (red circles) at each κ and find excellent agreement with the BEM simulations (red crosses), confirming the accuracy of our calculations.

To gain further insight into the nanolens, in figure 4.5(b) we explore the role of the smallest nanoparticle size, R_3, on the max-

[2] We ignore electron tunnelling completely in the following discussions, it could be included within the classical framework using the quantum-corrected model (Esteban et al., 2012). Tunnelling will only limit the FE via short circuiting and hence there would be no benefit of working with gap sizes where it is relevant.

imum possible FE, for three different values of κ. As one would expect, the largest FEs are achieved for simultaneously large κ and small R_3; this is a consequence of ignoring nonlocal effects and will not be the case once quantum effects are included. Most interesting, for some values of R_3 the largest FE is not achieved for the largest κ; for instance, at $R_3 = 5$ nm, the $\kappa = 2$ nanolens has a maximum FE 2.7× larger than for $\kappa = 4$. A larger κ is not always best! Similarly, figure 4.5(c) reveals that at larger gap sizes it may sometimes be preferable to use a smaller κ, although the overall trend is that a small $g_{2,3}$ and large κ give the biggest FEs. Together these two results highlight the complicated interference effects captured by the full Maxwell's equations and missing from the simple electrostatic approximation. In particular, they reveal that one should be careful in modelling the nanolens; it is always possible to pick a lucky/unlucky nanolens geometry leading to constructive/destructive interference effects. This is why we have made the effort to model nanolens geometries over an extensive range of parameters.

4.2.2 *Going Small Within The Hydrodynamic Model*

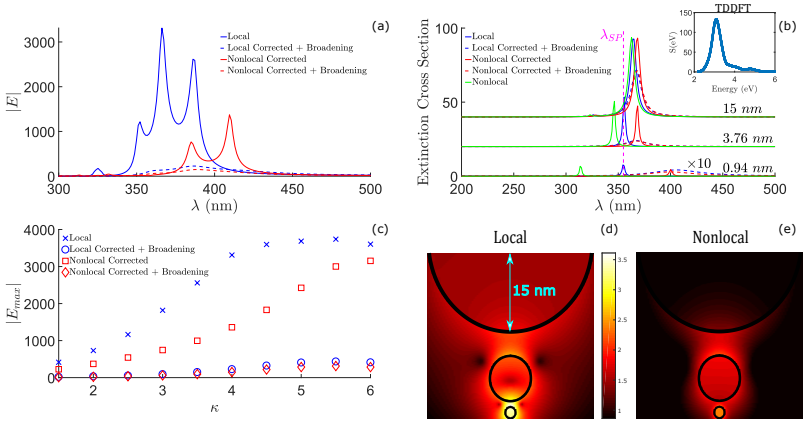


Figure 4.6: (a) FE for a nanolens ($R_3 = 0.94$nm, $g_{2,3} = 1$ nm and $\kappa = 4$) calculated with the four models described in the main text. (b) Extinction cross section of the individual nanospheres for the four different models, also shown is the result for the non-corrected nonlocal model (in green). The local and nonlocal models with broadening included are multiplied by 10 for clarity for the smallest nanosphere. The inset shows the TDDFT result for the Na_{92} cluster. (c) The maximum FE for the four models over a range of κ. (d) and (e) show the field profile, for $\kappa = 4$, within the local and nonlocal models at resonance (logarithmic scale).

Our findings from figure 4.5 revealed that it is desirable to choose a small R_3 and a large κ for massive FEs, we now explore how this breaks down with nonlocal effects included. Clearly it is infeasible to run a full quantum calculation dues to the large size differences between the constituent nanospheres, so we make a number of approximations and include quantum effects via the HM within the LAM and calculated using the BEM[3]. Unfortunately, while the HM captures the smearing of the electric field at the metal surface well, it predicts the wrong shift of the LSP for alkali metals (see sections 2.1 and 2.2). The small size quantum effects of alkali metals is dominated by an electron spill-out induced redshift. To model this we used the spherical jellium model within the LDA-TDDFT formalism, which provides a good description for closed-shell clusters, to describe the optical response of the Na_{92} smallest

[3] The artificial dielectric thickness used in the LAM throughout this section was set to $R/200$ and its accuracy was confirmed by comparison with a coated Mie theory code (Bohren and Huffman, 2008).

nanosphere[4]. This provides accurate results within a few tenths of an eV in comparison to experimental results (Xia et al., 2009). Thus, the nonlocal result can be 'corrected' following the prescription of (Teperik et al., 2013). Performing the TDDFT calculation allows an estimation of the electron spill-out, which is not possible relying on the HM, and allows modelling of nanolenses down to the quantum limit.

[4] The DFT calculation used a jellium density determined by sodium's Wigner-Seitz radius of 2.08 Å.

In figure 4.6(a) we show the FE for a sodium nanolens of $R_3 = 0.94$ nm, $g_{2,3} = 1$ nm and $\kappa = 4$ for four different models to illustrate the role of the different nonlocal contributions.

- The 1st is a purely local model, as used in the last section.
- The 2nd model is local but with a surface scattering contribution included in the damping, as well as an electron spill-out correction to the plasma frequency provided by the TDDFT calculation. For the Na_{92} smallest sphere we use an experimental value of 0.42 eV (Xia et al., 2009) for the damping rather than the Kreibig formula, equation 2.9, which will breakdown at this small scale. Sodium is well described by a free electron gas and as $\omega_P >> \gamma$, we may take the plasmon linewidth to be given by γ.
- The 3rd model includes nonlocality with *both* an electron spill-out correction and a correction to fix the incorrect value of the Feibelman parameter for alkali clusters (Teperik et al., 2013; Monreal et al., 2013) (see section 2.1), but *no* surface scattering. The plasma frequency can be corrected as follows: $\omega_{SP} = \omega_{SP}^{nonloc}(1 - \frac{\Delta}{R} - \frac{\delta}{R})$, where Δ is the position of the induced charge relative to the edge, R is the sphere radius and δ is the spill-out length (Teperik et al., 2013). The local results can also be corrected via $\omega_{SP} = \omega_{SP}^{loc}(1 - \frac{\delta}{R})$, which is implemented for model 2. Both quantities are on the order of an Å in size and so we will take $\delta \approx \Delta$ and fit the value using the TDDFT simulation for the Na_{92} cluster. A good fit is given by a value of $\delta = 0.12$ nm, this agrees reasonably well with the experimental value of 0.145 nm (Reiners et al., 1995). Note that the δ parameter is calculated from the single Na_{92} cluster and that value is used for the correction for *all* the spheres in the nanolens.
- The 4th model includes nonlocality with the Feibelman parameter correction *and* the surface scattering contribution to the damping.

From consideration of these four models we can explore the contribution from the various small-size effects on the FE. In the local model (model 1, blue solid line), we find a number of peaks due to a complicated plasmon hybridisation between the three particles. The presence of a number of peaks suggests retardation is already limiting the nanolensing effect, this is an interesting size regime to explore as both quantum and retardation effects are important to include. Interestingly, compared to the individual nanosphere

response, the large FE response is more broadband; over a wavelength range of about 50 nm a large FE of over 1000 is possible. We find that the nonlocal model with no broadening (model 3, red solid line) leads to a redshift (due to electron spill-out) and a reduction in the maximum FE to about 40% of the local result, there is also a smoothing out of the number of peaks visible. For both model 2 and 4 there is a drastic reduction in the FE by over a factor of 10 due to the Kreibig damping. This reduction means that the nanolens, in this small size regime, gives no benefit over ordinary individual nanospheres and dimers, which offer FEs on the order of Q and Q^2 respectively (Sun et al., 2011). We can conclude that it is the increased damping via electron-surface scattering that severely limits the cascade FE rather than the nonlocal induced shift of the resonances.

To gain further understanding of these results, we analyse the optical response of each individual nanosphere. The extinction cross section of the individual nanospheres in the $\kappa = 4$ system, for each of the four models, is shown in figure 4.6(b). Also shown for comparison is the non-corrected nonlocal model (green solid line). For the largest particle ($R = 15$ nm), somewhat surprisingly, retardation is already important and redshifts the resonance relative to the quasistatic LSP resonance indicated by the purple dashed line, at smaller wavelengths we can see a weak higher order mode beginning to form. For all three sizes the non-corrected nonlocal model incorrectly predicts a blueshift. The corrected models, in contrast, show a redshift as a consequence of electron spill-out, naturally this shift decreases for larger nanoparticles. The inset of 4.6(b) shows the jellium TDDFT result for Na_{92}, where S is the dipole strength function and we use an artificial 0.1 eV broadening. Close to 3 eV a prominent LSP is present, for lower broadening it is possible to see that in fact the peak is fragmented via interactions with single electron excitations. Also present is a Bennett plasmon and the volume plasmon at higher energies (Varas et al., 2016), these modes are not included in the local and nonlocal models. The SPP predicted by TDDFT is redshifted compared to the classical result due to the electron spill-out.

The max FE for the four different models, over a range of κ, is shown in figure 4.6(c). We see that the findings from figure 4.6(a) continue over a range of κ. Interestingly, at larger κ the Kreibig damping's role in decreasing the FE is increased. This is a consequence of the smaller particle not contributing to the cascade effect, whilst the larger two spheres begin to behave purely classical; the system begins to behave as a classical asymmetric dimer and we can expect FEs on the order of Q^2 rather than Q^3 if electron-surface scattering is not included. Also shown is a logarithmic plot of the field at resonance, for $\kappa = 4$, within the local (model 1) and nonlocal corrected (model 4) in figure 4.6(d) and 4.6(e) respectively. The local result shows that the largest FE is near the smallest sphere in the gap $g_{2,3}$, this agrees with what was found in the original work

on the nanolens (Li et al., 2003). In contrast, the field map for the nonlocal model shows only a small FE in the gap with the largest fields found within the smallest nanosphere, this illustrates the breakdown of the cascade effect.

From these results, it seems that going smaller is not an effective strategy for building plasmonic cascade nanolenses. The decrease in the quality factor of the smaller nanospheres in the nonlocal limit, due to Kreibig damping, overrides any benefit of working within the electrostatic limit. Recently, similarly drastic reductions from nonlocality (up to 7 times) have been shown for the fluorescence enhancement of a dipole near a gold nanoparticle (Jurga et al., 2017). We emphasise that the inclusion of electron tunnelling, a negligible effect at these gap sizes, would only further limit the FE.

4.3 Going Large: The Polar Dielectric Nanolens

We have now seen the detrimental effects of retardation and nonlocality on nanolensing. To explore different strategies for achieving large FEs, we next consider the role of the nanolens material. In figure 4.7(a) we perform a sweep of κ for typical plasmonic metals: silver, gold and aluminium[5]. We also consider the polar dielectric SiC excited within its Reststrahlen band, where it can support LSPhPs, see section 1.4.2. The smallest radius and gap are set to 5 nm and 3 nm respectively. We find that gold, the most commonly used plasmonic metal, is not suitable for nanolensing, with the lowest FEs of the four materials. This is a consequence of interband transitions close the LSP resonance, leading to a low quality factor. Possible improvement could come from structuring the nanoparticles to shift the LSP resonance towards lower frequencies, for instance using ellipsoid or rod shaped nanoparticles. Aluminium is known to have a good quality factor but its high energy plasma frequency means even very small nanospheres become comparable to the excitation wavelength, consequently retardation heavily limits nanolensing. Thus, for larger κ, aluminium does not perform much better than gold. Silver rules supreme amongst the plasmonic metals, offering a good quality factor[6] and is hit by retardation less severely than aluminium because of its lower energy plasma frequency. For larger κ though it does not offer much better nanolensing than gold or aluminium. In contrast, SiC offers much improved performance compared to metals, especially at larger κ. This is due to a high quality factor and reduced retardation loss compared to the metals. It shows a monotonic increase in FE with increasing κ: the hallmark of nanolensing. The lack of retardation effects can be confirmed by comparison to an electrostatic calculation; the results are identical, which is a consequence of the low energy of the LSPhP mode. Using equation 1.27, the electrostatic conditions holds approximately up to 1 micron for SiC at the LSPhP resonance; for the $\kappa = 4$ nanolens the largest sphere is only 80 nm and so the condition is very well fulfilled. This result highly suggests that the nanolens concept is most viable in the mid-IR/THz regime where the electrostatic regime holds over a much wider range of nanosphere sizes.

To emphasise the difference in the wavelength operation between metal and polar dielectric nanolenses, the FE spectra for the four materials, with $\kappa = 4$, is shown in figure 4.7(b). Wavelength is displayed on a logarithmic scale and the FE for SiC is divided by 10 to fit it on the same scale as the other materials. Note that while polar dielectrics are of zero benefit if we wish to achieve large FEs in the visible, luckily the longer wavelength regime of the mid-IR/THz is technologically important, see section 1.4 and chapter 3. Similar results in this spectral range with plasmons may be possible by using graphene and doped semiconductors. The FE values obtained in this section should be taken as upper bounds

[5] For gold and silver we use experimental data from (Johnson and Christy, 1972). For aluminium a Drude fit with $\omega_P = 15.3$ eV and $\hbar\gamma = 0.598$ eV is used.

[6] The quality factor of silver is approximately an order of magnitude larger than gold and twice that of aluminium at the plasmon frequency (Caldwell et al., 2015). Note that the maximum FE is highly sensitive to the exact material data used; this was shown by Pellegrini *et al*, where they demonstrated that the FE can change by a factor of 5, depending on the experimental dielectric function used (Pellegrini et al., 2016). See appendix A for a short discussion on the different experimental datasets available and the displeasing disparity between them.

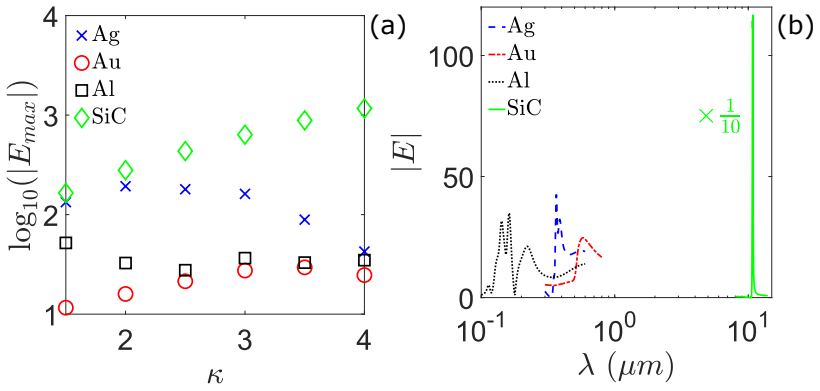
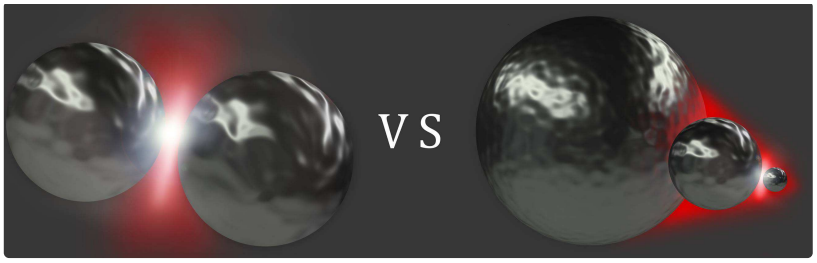


Figure 4.7: (a) Maximum FE, in log scale, for typical plasmonic metals: silver, gold and aluminium, as well as the polar dielectric SiC. (b) FE spectra for the $\kappa = 4$ geometry. The SiC result is divided by 10 to allow easy visualisation on the same plot. The nanolens geometry in both plots is $R_3 = 5$ nm and $g_{2,3} = 3$ nm.

as no nonlocal or surface scattering corrections were used. We expect such effects to be small and should compare favourably to metals; but it is an area in desperate need for research, especially as experiments start to explore ever smaller structures (Caldwell et al., 2013). The flat dispersion of optical phonons in the long wavelength limit means a low group velocity on the order of $\sim 10^4$ m/s. This gives, despite the long scattering time of $\sim 10^2$ ps, a short mean free path on the order of 10 nm. Furthermore, quantum corrections can be expected to be negligible due to the absence of free carriers in the polar dielectrics; we assume no doping and ignore any free carrier contribution.



4.4 *The Nanolens Vs The Dimer*

To rigorously test the cascade effect in SiC nanolenses, we compare against the more usual symmetric dimer geometry. The dimer is one of the most heavily studied systems in plasmonics, providing the simplest example of plasmonic coupling (Nordlander et al., 2004) that leads to a well-documented intense FE in the centre of the gap (Hao and Schatz, 2004). It has been recently suggested by Pellegrini *et al* that the nanolens holds no significant advantage over a symmetric dimer of equal or less total volume (Pellegrini et al., 2016). This was a very surprising result, but the authors only tested for a single nanolens geometry. As we have seen in preceding sections, the nanolens is a non-trivial system and a range of geometries should be studied before conclusions are made. In this section we extend the study of (Pellegrini et al., 2016) for silver over a range of κ and compare against SiC. For a fair comparison, we enforce the dimer volume to be equal to the equivalent nanolens, for a given κ, so that $R_{dimer} = ((R_1^3 + R_2^3 + R_3^3)/2)^{1/3}$. The dimer gap is set equal to the smallest gap of the corresponding nanolens and the field is measured in the middle where the largest FE is found. Of course there is some arbitrariness when defining a suitable comparison between the nanolens and dimer; we have chosen to fix the volume as many physically relevant quantities, such as the dipole moment and extinction cross section, depend on it. Furthermore, if the volume is not fixed then the difference in volume between the two will change with κ, meaning an analysis based on changing κ would be much less meaningful.

In figure 4.8 the maximum FE over a range of κ is shown for a silver nanolens and the equivalent dimer for: (a) the electrostatic approximation and (b) the full retarded solution. We find that the results of Pellegrini *et al* hold for all the geometries considered in the electrostatic approximation and all, expect at $\kappa = 5$, in the full retarded calculation. Our work strongly suggests that, in the case of silver, **the nanolens geometry holds no significant advantage over the dimer system for exciting strong FEs**. Given that silver is a superior plasmonic metal, as shown in figure 4.7, we would expect these results to hold for *all* metals in the visible. This result suggests that the large amount of work in search for large FEs with nanolenses built of noble metals may be wasted effort, when

much simpler dimers are preferable. Most surprising is that the dimer remains superior within the electrostatic approximation as well, see figure 4.8(a). Therefore, it is not retardation effects that limit nanolensing, instead it seems silver is too lossy to build an effective nanolens. To analyse this in more detail, we construct a toy model for a Drude metal based on silver, with a variable quality factor. Assuming that $\omega >> \gamma$, we may approximate $Q \approx \frac{\omega}{\gamma}$ and so we need only change the damping parameter in the Drude model. The results are shown in figure 4.9, we plot the maximum FE of the dimer minus the nanolens (ΔE_{max}). Positive numbers correspond to the dimer outperforming the nanolens and are indicated by the blue region in the plot, negative numbers correspond to the nanolens outperforming the dimer and are indicated by the cream region. The results show that the nanolens geometry, for these particular parameters, is desirable only when Q is very large , $Q \gtrsim 800$, which is beyond what is achievable in plasmonics in the visible regime by over an order of magnitude[7] and is only just within the reach of the best of polar dielectrics at much lower frequencies; SiC has $Q \sim 900$ at the LSPhP resonance (Caldwell et al., 2015). These results suggest that the metal-based cascade nanolens is yet another victim of the large inherent loss in plasmonics. Alternative materials suitable for constructing nanolens could be high index dielectrics that exhibit very large quality factors[8] or hybrid dielectric-metal systems where the loss can be modified over orders of magnitude from metal- to dielectric-like.

[7] The quality factor of silver at the LSP resonance is ~ 30.

[8] But unfortunately offer only modest FEs $\lesssim 5$.

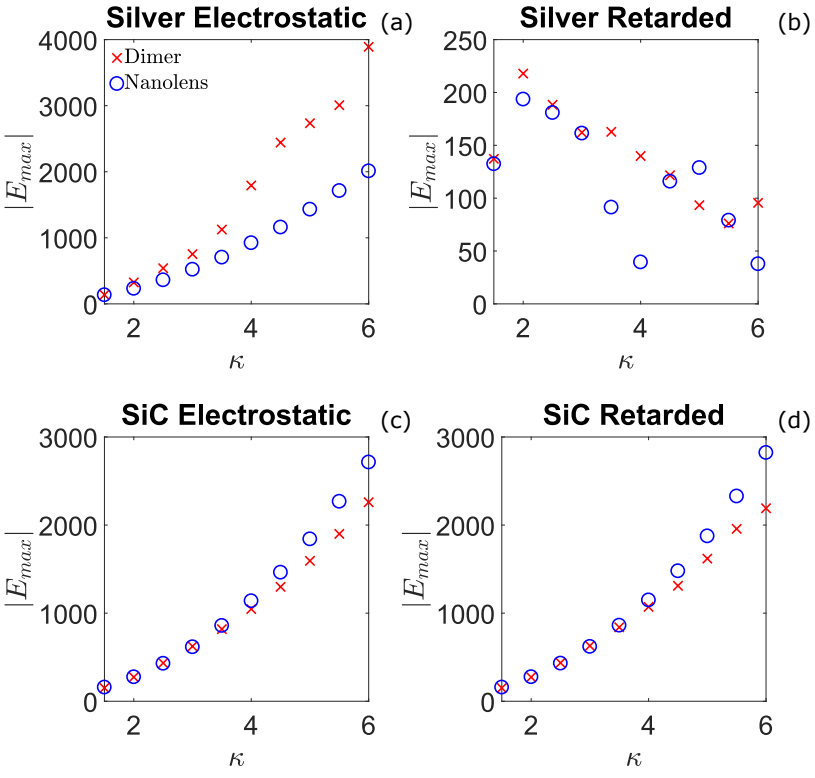


Figure 4.8: Maximum FE of a silver nanolens and dimer against κ within the electrostatic approximation (a) and full retarded solution (b). (c) and (d) are the same but for SiC. The geometry parameters in all the plots is $R_3 = 5$ nm and $g_{2,3} = 3$ nm.

Returning to figure 4.8, we find, in contrast to the case of silver, that the SiC nanolens outperforms the dimer in both the electrostatic approximation, figure 4.8(c), and the full solution to Maxwell's equations, figure 4.8(d). The nanolens become increasingly superior to the dimer for larger κ, at $\kappa = 6$ the nanolens geometry leads to a 29% larger FE compared to the dimer. The suitability of SiC for nanolensing is a consequence of being well within the electrostatic limit at these sizes, combined with a large quality factor. The former being confirmed by the electrostatic results being approximately equal to the retarded results.

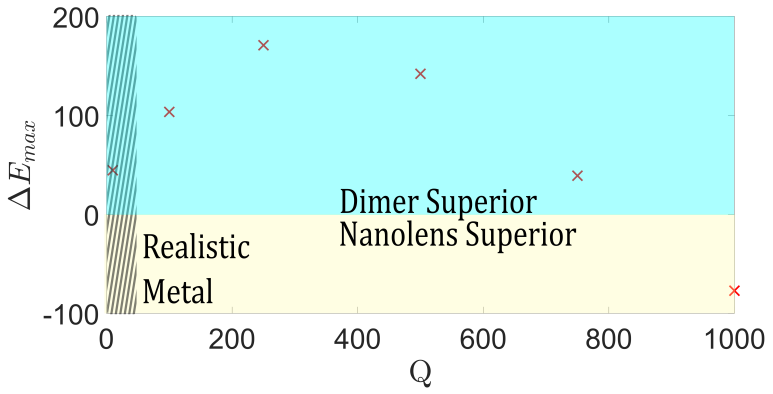


Figure 4.9: The maximum FE of a dimer minus that of a nanolens, for a fictitious Drude metal with variable Q. Indicated in blue are the regions where the dimer is superior and cream are the region where the nanolens is superior. The hashed region indicates the region of realistic Q values for metals. The geometry parameters are $\kappa = 4$, $R_3 = 5$ nm and $g_{2,3}$ nm.

At even larger κ we can expect the nanolens to be further superior than shown in figure 4.8, where we restrict ourselves to $\kappa \leq 6$. Unfortunately the BEM employed here become increasingly difficult to converge as the size ratio between the constituent spheres increases. We have therefore performed additional calculations using the multi-sphere GMM code by Mackowski and Mishchenko (Mackowski and Mishchenko, 2011), which is able to achieve convergence for challenging geometries up to a $\kappa = 10$ nanolens, shown in figure 4.10. The nanolens geometry has a substantial 82% improvement of the maximum FE compared to the equivalent volume dimer for $\kappa = 10$. The nanolens become increasingly superior for larger κ, continuing the trend from the BEM calculation shown in 4.8(d). The drop off in FE increment with increasing κ for the dimer is presumably due to growing retardation loss, for large enough κ a similar drop off can be expected for the nanolens. Thus, by using SiC structures we can achieve massive FEs approaching 10^4, which corresponds to an intensity enhancement of 10^8, further demonstrating that polar dielectrics are a suitable material for constructing extreme-cascade nanophotonic devices. An immediate potential application is for deterministic surface-enhanced infrared absorption spectroscopy (SEIRAS), where the molecular IR signal scales as $\sim |E|^2$ (Yang et al., 2018). The spectral information shown in figure 4.11(a) for $\kappa = 10$ reveals that the strongest FE occurs for a narrow single peak, again revealing this to be a hallmark of strong nanolensing. This is in contrast to the dimer, which has multiple

peaks in the spectrum due to mode hybridisation (Nordlander et al., 2004). To confirm that the response of the system is down to a material resonance rather than a pure geometric resonance, we have calculated the FE for a silver nanolens in the same wavelength regime as for the SiC nanolens ($8 \rightarrow 14$ μ m), where silver acts as a perfect electrical conductor. Only a small FE on the order of 10 is found.

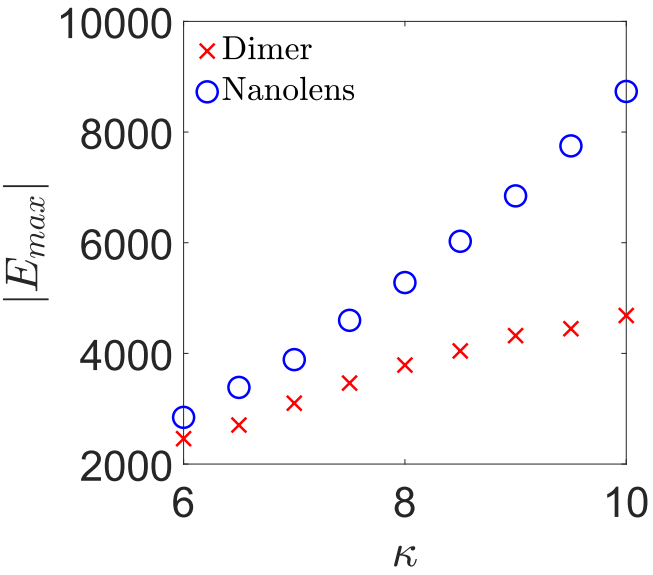


Figure 4.10: (a) Maximum FE of a SiC nanolens and dimer against κ for an extended range. The geometry parameters are $R_3 = 5$ nm and $g_{2,3} = 3$ nm.

So far we have stuck to the three-particle nanolens, as this corresponds to the most common type of nanolens studied in experiments and will remain in the nanolensing regime for the largest range of κ. Higher FEs can be achieved with a larger number of elements, although at the cost of added complexity to build. In the original work of Li *et al* it was shown that a symmetric nanolens increases the FE by a factor of 2 (Li et al., 2003). $N > 3$ particle devices have been experimentally demonstrated (Coluccio et al., 2015) and lead to improved SERS intensity compared to the three-particle nanolens (Das et al., 2016). In figure 4.11(b) we show the result for a nanolens built of four spheres compared to an equivalent dimer. The smallest sphere radius, R_4, and gap, $g_{3,4}$, are 5 nm and 3 nm respectively, and $\kappa = 4$. We find that the four-particle nanolens can achieve a maximum FE of almost double the equivalent dimer, again demonstrating the effectiveness of the cascade effect in SiC devices. Higher numbers of nanospheres of course could be considered and might well lead to even larger improvements. We have also explored four-particle silver nanolenses for low κ and observed no significant advantage compared to the equivalent dimer system. Interestingly, to observe a consistent advantage for the silver nanolens we had to consider extreme cases such as $R_4 = 1$ nm and $g_{3,4} = 3$ nm, in the range $1 < \kappa < 2.5$. Of course, these are unrealistic cases as nonlocal effects, which are not included in these calculations, would destroy the nanolensing in reality.

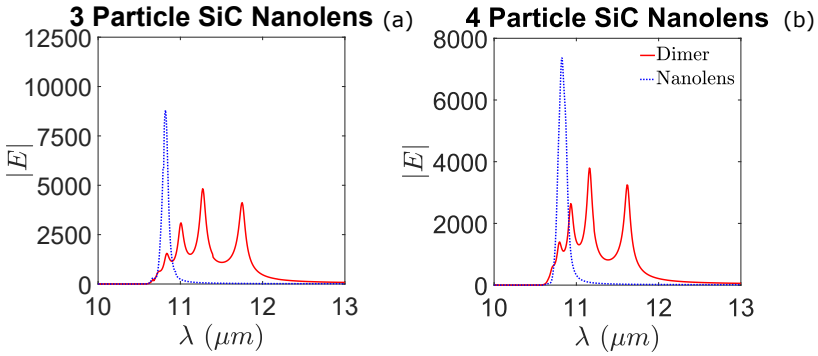


Figure 4.11: FE for a (a) three-particle SiC nanolens with geometry $R_3 = 5$ nm, $g_{2,3} = 3$ nm and $\kappa = 10$ (b) four-particle SiC nanolens with geometry $R_4 = 5$ nm, $g_{3,4} = 3$ nm and $\kappa = 4$. Both are compared to the equivalent volume dimer.

As a final remark for this chapter, we note that the geometries previously shown, whilst demonstrating extreme FE, may be difficult to produce for experimental demonstration due to the large κ and small gaps. These results should be taken as an indication of the ultimate achievable FEs in polar dielectrics, although we emphasise that further geometrical optimisation is certainly possible! To demonstrate a more realistic device, we consider a nanolens with both the smallest nanosphere radius and gap to be 10 nm and limit ourselves to $\kappa = 4$. The results are shown in figure 4.12(a) and (b), we find for these geometries that the maximum FE is 52% and 116% larger for the three- and four-particle nanolens geometry, as compared to the equivalent dimer, respectively. The value of the FE is, of course, lower than the results shown in figure 4.11 due to the larger gap, but we see an improved performance of the nanolens relative to the dimer for more modest κ. A wider spacing leads to a lower number of higher order plasmon modes being excited and a consequent drop in the field concentration near the smallest nanosphere. It seems by picking an extremely small gap in the previous calculations, we inadvertently favoured the dimer. The results are clear evidence that *experimentally realistic* SiC devices can utilise the cascade effect to achieve large FEs, beyond what is achievable with metal-based devices.

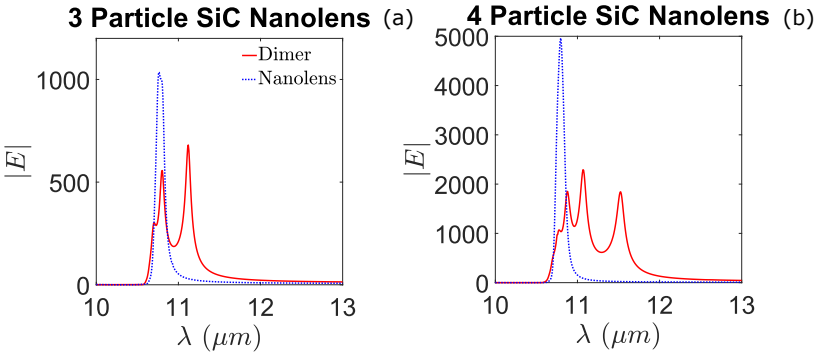


Figure 4.12: FE for a (a) three- and (b) four-particle SiC nanolens. Both with smallest nanosphere radius and gap equal to 10 nm and $\kappa = 4$. Both are compared to the equivalent volume dimer.

5 The Ultimate Small Size Limit Of Plasmonic Nanoantennas

Given the huge technological rewards of miniaturisation in electronics over the past century, it is of obvious interest to consider shrinking down plasmonics components to their ultimate small size limit. This raises intriguing theoretical questions on the nature of plasmons in few-electron systems. In this chapter we explore the quantum nanoantenna concept by considering the few-atom limit of a plasmonic nanorod. A new measure of plasmonicity is defined based on the collectivity of optical transitions, this allows the precise and quantitative identification of quantum plasmons. We then consider the FE of these modes by taking into account loss induced by plasmon-phonon coupling.

5.1 Introduction: The Ultra Small Size Limit Of Plasmonics

As was discussed in chapter 2, some recent experiments have probed such tiny nanoparticles and gaps that classical electrodynamics, and even quasi-quantum models like the HM, will fail to model the plasmonic behaviour correctly. Instead, the optical response of the constituent electrons must be treated using *ab initio* methods such as TDDFT. This will automatically take into account electron spill-out, nonlocality, and energy-level discreteness, although at much higher computational cost. It is also becoming increasingly apparent that the atomic structure may be important to explicitly include rather than relying on simplifications such as the jellium model (Zhang et al., 2014). Few-atom systems have been shown to behave similar to classical plasmonic components (Bursi et al., 2014) and can even be surprisingly well described by classical models (Sinha-Roy et al., 2017; Urbieta et al., 2018). This has led to the emerging concept of 'quantum plasmonic nanoantennas' (Fitzgerald and Giannini, 2017). Within this picture, plasmonic nanoparticles behave like large molecules, and the plasmon, which will capture much of the spectral weight in

the near- and far-field if optically active, is built out of multiple discrete electronic transitions mediated by the Coulomb interaction. To explore this classical-quantum transition we employ, what has become a standard 'workhorse' of quantum plasmonics, the single-atom-thick chain (Gao and Yuan, 2005; Yan et al., 2007; Yan and Gao, 2008; Bernadotte et al., 2013; Yasuike et al., 2011; Piccini et al., 2013; Huang et al., 2014; Rossi et al., 2015). It represents the ultimate aspect ratio limit of the plasmonic nanorod. Their small size could lead to added tunability and novel quantum plasmonic behaviour as well as provide an ideal case study, which is computationally practical, for understanding how plasmons arise within a quantum picture. We consider a sodium nanolens as the valence electrons in alkali metals are known to be well modelled by a free electron gas. This also enables a simple choice of the LDA for the XC functional, although it is important to stress that XC effects do not effect the physics of plasmons in a great fashion; only providing a small quantitative spectral shift. Our insights gained could have been achieved by ignoring XC contribution, but seeing as its inclusions leads to no significant additional computational cost it might as well be included. We ignore substrate effects and enforce a linear geometry with a constant atomic spacing of 3.08Å [1], so it is a rather artificial system. Only one electron (3s) per atom is explicitly considered with the remainder modelled by a norm-conserving Troullier-Martins pseudopotential (Troullier and Martins, 1991). There have been experimental demonstrations of single-atom-thick atomic chains, in particular via the precise manipulation of single gold atoms on a substrate (Nilius et al., 2002),

[1] This value is taken from a relaxed Na_2 structure (Vasiliev et al., 2002). In reality the spacing will likely depend on the number of atoms and a linear chain may not even be the lowest energy configuration (Bergara et al., 2003). It would also be very sensitive to the specific substrate chosen.

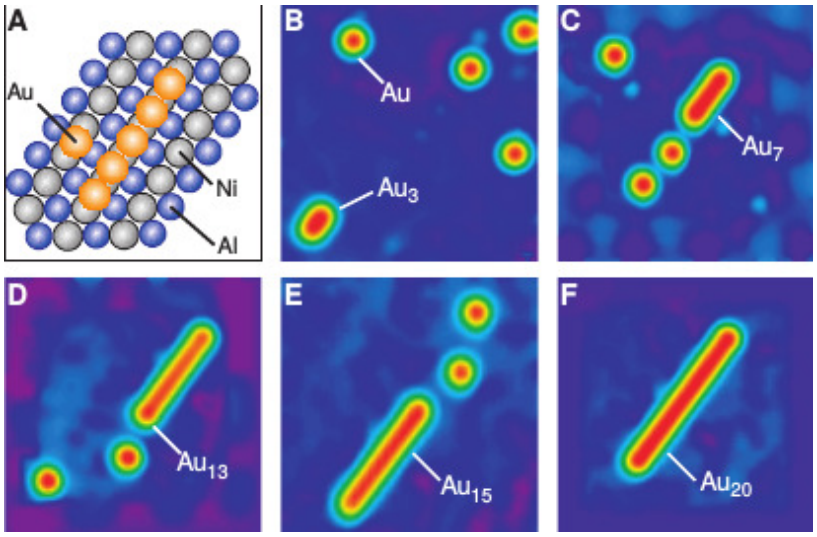


Figure 5.1: Scanning tunnelling microscopy images of the manipulation of single gold atoms on a NiAl substrate using the tip of the microscope. From (Nilius et al., 2002), reprinted with permission from the AAAS

It is important to have clear and robust methods to analyse the results from electronic structure calculations, which may return a large amount of data and many different excitations. The surface of the nanoparticle allows mixing of plasmons and single-particle excitations (Landau damping), by breaking translational invariance.

This means identification of peaks in excitation spectra can be challenging. In the quantum plasmonic literature, often peaks with a large dipole strength function are claimed to be plasmonic (Gao and Yuan, 2005; Yan et al., 2007; Yan and Gao, 2008) using the argument that this indicates collectivity. This can be erroneous, as the dipole strength function gives a measure of the dipole moment and hence how strongly the excitation interacts with light (how 'bright' it is). A single eh pair separated over a large distance can give a large dipole moment but it would not be sensible to describe it as plasmonic, it would be more usually described as excitonic[2]. Thus, we need more sophisticated methods to identify plasmons in small metal clusters. Some examples from the quantum plasmonics literature include the Coulomb scaling method (Bernadotte et al., 2013), electron density rearrangement (Paul and Balanarayan, 2018), the plasmonicity index (Bursi et al., 2016; Zhang et al., 2017) and analysis of the electron population dynamics (Townsend and Bryant, 2014; Ma et al., 2015; Rossi et al., 2017).

Collectivity is one of the most intuitive and defining characteristics of a plasmon and distinguishes it from excitonic excitations, which may superficially appear similar. It is therefore an interesting route to quantify collectivity to identify plasmonic excitations. In the next section and elaborated in appendix K, we employ quantum chemistry methods using data which can be obtained from the standard output of any electronic structure code[3], to define a new index that is simply the product of the collectivity index (Martin, 2003; Plasser et al., 2014) and the dipole strength. This provides a simple and convenient method to quantify how plasmon-like a molecular excitation is, at no extra cost to a TDDFT calculation.

Once one has found a cluster supports quantum plasmons, the question is naturally: are they good for anything practical? Given the theme of this thesis, naturally we would enquire about the potential FEs achievable from these systems. Unfortunately it is not straightforward to extract the FE from a typical TDDFT calculation due to the rather artificial treatment of loss. A system calculated in DFT will typically be a *closed* system, for instance if we perturb the system with a delta function kick then the consequent coherent motion will continue forever; the time evolution remains unitary at all times. This is a consequence of the TDDFT calculation not taking into account the interaction of the excited subspace with the surrounding universe nor the photon field[4]. Therefore loss is normally added 'by hand' in the form a parameter η. It is possible to derive and solve the Casida equation without any loss, the excitations are returned at precise energies. Typically the results are then convolved with Lorentzian functions of width η. In a time-propagation calculation the linewidth is equally artificial, it is determined by the inverse of the total propagation time[5]. At best this value can be taken as a phenomenological value from experiment, but for many systems where experimental data does not exist, such as the atomic chains commonly studied in quantum

[2] An exciton is a Coulombically bound eh pair, they mediate a large number of different light-matter phenomena in inorganic and organic systems. Typically they are associated with semiconductors and insulators as they are heavily screened in bulk metals and consequently short-lived. Experimentally this leads to a suppression of low-frequency scattering in favour of the single plasmon peak (Contreras et al., 1985). In metal clusters the screening is reduced and the lower density of states around the Fermi level can lead to long-lived excitons (Peng et al., 2015).

[3] As long as the transition density matrix can be calculated.

[4] A neat way to explore the former is to split a TDDFT calculation into two subsystems: one of which is excited externally and the other unperturbed apart from its interaction with the induced field of the excited subsystem. A subsystem of a TDDFT calculation can exhibit non-unitary time evolution. This has been used to explore exciton decay in silver nanowires (Peng et al., 2015).

[5] Note though that in both cases the area under the peaks does have a physical meaning, it is the oscillator strength of the excitation. The total area of the spectrum must give the total number of active electrons in the simulation via the Thomas-Reiche-Kuhn sum rule, see section 2.3.4.

plasmonics, instead a 'reasonable' value must be plucked from thin air. One suspects it is often a value that gives the nicest looking plot! This is obviously displeasing as loss can be crucial to include, for example when calculating the contributions to the Raman signal in SERS for molecules close to metal nanoparticles, both the FE and chemical interactions[6] depend strongly on the excited state linewidths (Jensen et al., 2008). Therefore, to analyse the atomic chain in terms of a plasmonic component, we need to identify the main loss channels at these size scales. They come in two flavours:

- Landau damping; as a number of single eh or quasi-collective modes start to merge at the same energy they become indistinguishable. The width of the merged mode gives a measure of the Landau damping. From the other point of view, as a nanoparticle gets smaller the plasmon can couple to eh excitation due to momentum transfer from the surface. This fragments the mode and increases the width. (Lermé et al., 2010).

- Phonon-electron/plasmon coupling; each Casida excitation has an intrinsic linewidth, η, given by coupling to the environment. In molecular physics it is well-known that electronic transitions are broadened by electron-phonon coupling (Fox, 2002). Vibronic transitions typically occur in the IR spectral region, and so on the scale of electron transitions in molecules, typically in the UV and visible, form a continuum of energy levels around each electron state. When an electron is excited to a higher electron state, this will simultaneously create vibrational quanta. This leads to a spread of possible energies that the system can absorb light at and hence broadens the excitation spectrum.

[6] Charge transfer between the molecule and the metal.

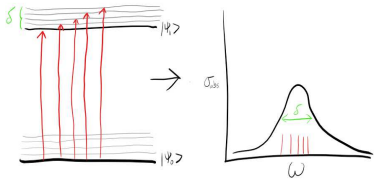


Figure 5.2: Illustration of electron-vibron coupling in a molecule. Note that the absorption need not have a Lorentzian shape.

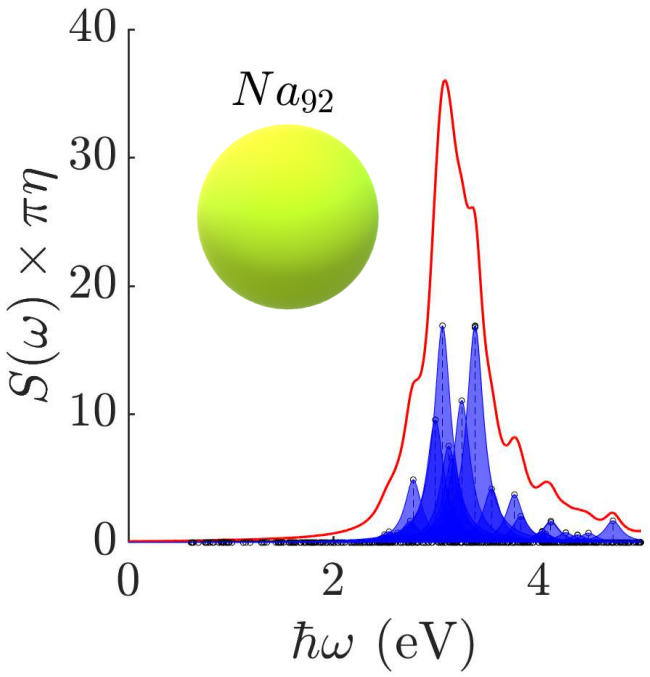


Figure 5.3: The Casida spectrum for a Na_{92} cluster modelled with the jellium approximation and ALDA-TDDFT. The black lines show the oscillator strength for each excitation, note they all have a degeneracy of 3 due to the spherical symmetry. The blue shaded curves show the Lorentzian curves of linewidth $\eta = 0.1$ eV arising from the broadening of the oscillator strength. The red curve shows the resulting dipole strength function multiplied by $\pi\eta$ for ease of appearance.

These two damping pathways are nicely illustrated for the Na_{92} cluster that we explored in chapter 4, the results for a Casida calculation are presented in figure 5.3. The blue shaded curves show the result of the artificial broadening, $\eta = 0.1$ eV, to the oscillator strengths, which are shown by the black lines. The resulting dipole strength function is shown by the red line. Observe that it has an intrinsic width determined by the merging excitations; if the broadening of the Casida excitations is large enough then they cannot be distinguished in experiment and one would observe a single plasmon peak with a linewidth determined by the so-called Landau damping. Of course in the simulation results the magnitude of the Landau damping will have a dependence on the broadening chosen, the choice of η dictates how fragmented the LSP peak appears as well as how visible secondary modes such as the volume plasmon are. We find for atomic chains that Landau damping does not occur; quantum plasmons at these sizes are well defined Casida excitations and there is no overlapping of excitations. We therefore estimate how phonons couple to electronic excitations in a separate calculation using DFPT (Baroni et al., 2001) to calculate the phonon normal modes of the sodium chains. It is found that high FEs, comparable to what is achievable in classical plasmonic systems, are possible in atomic chains. This opens up the exciting possibility of designing quantum antennas that display plasmonic behaviour in the deep quantum regime. It is also provides a new theoretical framework in which to consider molecular interactions with light, borrowing concepts from classical plasmonics.

The following results discussed in this chapter were published in the following paper (Fitzgerald et al., 2017). Some of the figures used in this chapter are either reproduced or adapted from this publication, with permission given by the American Physical Society.

5.2 Quantum Plasmonics Of The Sodium Chain

5.2.1 The Electronic Ground State

Before proceeding to the excitation spectrum, a DFT ground state calculation is required to obtain the KS orbitals as well as the ground state density $n_0(\mathbf{r})$. This is done using the open-source code OCTOPUS (Castro et al., 2006; Andrade et al., 2015). All densities, wavefunctions and potentials are represented on equally spaced real-space grids with the simulation box built via a union of spheres centred on each sodium atom. A radius of 12 Å and a grid spacing of 0.3 Å is chosen as it gives a good compromise between accuracy and speed. As unoccupied orbitals need to be calculated for the Casida method, the grid is chosen to be finer than is necessary for an occupied ground state calculation. The ground state density is shown in figure 5.4(a) and the various potentials in 5.4(b), the positions of the sodium atoms are indicated by the green dots. In particular, note the difference and similarities between the confining potential v_0 (blue line) and the self-consistent KS potential (red line), the electron-electron interaction screens the confining potential reducing the potential seen by the electrons. This is especially important outside the chain where the KS potential strongly screens the ionic charge, this reduces the electron spill-out significantly compared to the non-interacting case. Notice also the magnitude of the Hartree potential (yellow line) compared to the much smaller XC potential (purple line).

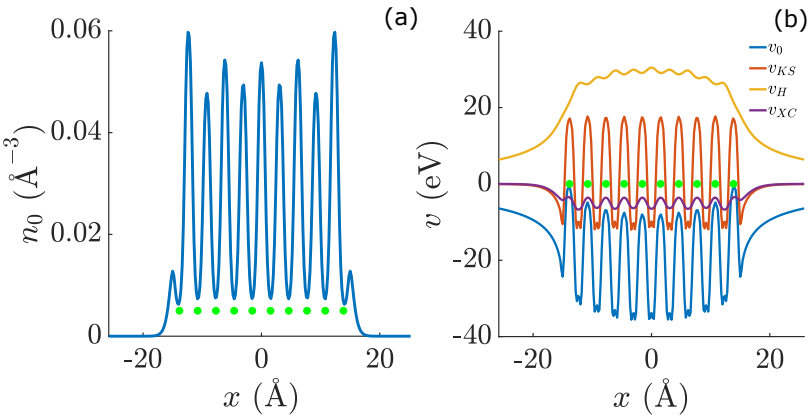


Figure 5.4: (a) Ground state density of Na_{10} chain. The atom positions are indicated by the green circles. (b) The corresponding confining potential, KS potential, Hartree potential and XC potential.

5.2.2 Casida Calculation

Once the KS orbitals and eigenvalues are known then the Casida equation can be solved (equation 2.26) to give the Casida eigenvalues and eigenvectors. Calculating the dipole strength function[7] gives a useful visualisation of the excitation spectrum, which is presented in figure 5.5(a). The red line shows the full Casida spectrum and the blue dashed line, for comparison, shows the excitation

[7] See section 2.3.4 for details.

spectrum with interactions switched off[8], i.e the Casida equation is solved but with the Hartree and XC contribution to the Casida matrix switched off. I call this the "non-interacting" spectrum but it must be emphasised that it is based on a DFT ground state with interactions included, only dynamical interactions are neglected. As we know that the existence of plasmons relies on the Hartree interaction, looking at the non-interacting spectrum can reveal clues to identifying plasmons, this can be viewed as a simplified version of the Coulomb scaling method (Bernadotte et al., 2013). The transition densities, presented at the bottom of figure 5.5, identify the three most distinct modes: the longitudinal mode (L-mode), transverse end mode (TE-mode) and transverse central mode (TC-mode) that were first identified by Yan *et al* (Yan et al., 2007) and have been studied for other materials such as gold (Piccini et al., 2013), graphene nanoribbons (Cocchi et al., 2012) and carbon chains (Broglia et al., 2013). The L-mode and both T-modes have a dipolar character, along the long length (x direction) and the short lengths (y and z direction) of the chain respectively. A broadening of $\eta = 0.0228$ eV is used for each Casida excitation, which is obtained by considering phonon-electron coupling and is discussed in the next section. The area under the curve must give the number of valence electrons according to the Thomas-Reiche-Kuhn sum rule, considering the spectrum up to 6 eV as well as 105 unoccupied states satisfies the sum rule to over 96%, this confirms the quality of the basis set truncation. The T-modes are higher in energy than the L-mode due to the greater electron confinement along the short length of the chain, alternatively in a classical plasmon picture this can be considered as due to a greater restoring force from the positive background. The L-mode is the lowest optically active mode, same as is found for a classical nanorod (Bryant et al., 2008).

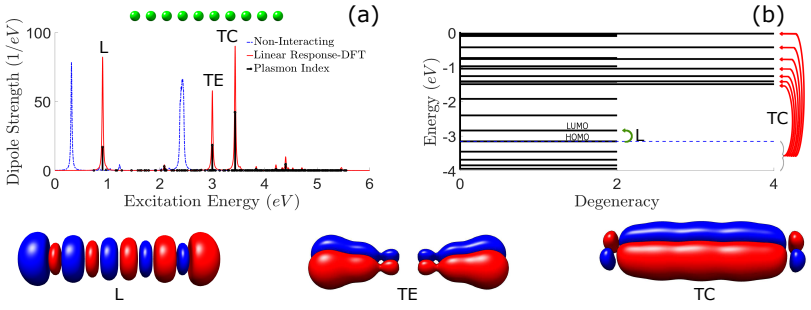


Figure 5.5: (a) Dipole strength function of a Na_{10} chain with the transition densities for the L-mode, TE-mode and TC-mode. The vertical black bars indicate the collectivity index multiplied by the oscillator strength for every Casida excitation and reveals the most plasmonic peaks. The 3D transition density plots are created with UCSF Chimera (Pettersen et al., 2004). (b) Transitions between KS states that make up the L-mode and TC-mode.

An elementary way to explore the nature of the modes is to vary the geometry and see the effect it has on them. In figure 5.6(a), a sweep of the chain length is shown revealing that dipole strength intensity grows for the L-mode and TC-mode but saturates for the TE-mode[9]. We can immediately conclude that the TE-mode cannot possibly be a plasmon as it does not survive taking the macroscopic limit, it seems to be a consequence of the atomic nature of the chain end; we have found it does not appear in simplified models where

[8] Excitations are given by the energy difference of KS levels, i.e. the poles of the KS density-density response function.

[9] Note that in figure 5.6(a) the dipole strength function is normalised to the electron number, so an increase in excitation intensity with increasing chain length corresponds to a constant normalised peak intensity.

the atomic structure is ignored. Note that the accumulation in oscillator strength of the L-mode and TC-mode with increasing length does not necessarily mean an increase in electron collectivity, as has been erroneously claimed sometimes in the literature.

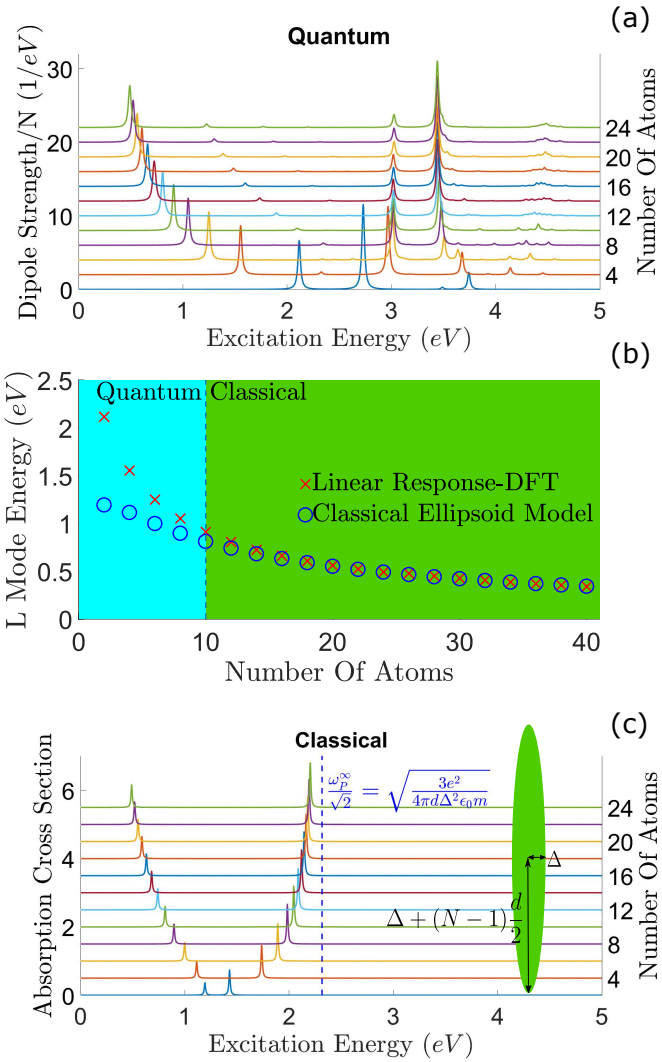


Figure 5.6: (a) The dipole strength function (divided by the number of valence electrons) for chain lengths ranging from 2 to 24 atoms. (b) Fit for the L-mode energy using a classical ellipsoid model. The best fit for the longest chain size of $N = 40$ is given by $\Delta = 1.45d$. (c) The normalised absorption cross section (averaged over 3 dimensions) calculated using the classical ellipsoid model. The dashed blue line indicates the transverse plasmon frequency for an infinite wire $\frac{\omega_p^\infty}{\sqrt{2}}$ where ω_p^∞ is the limit of the plasma frequency for an infinite number of atoms.

Further insight into the nature of the modes is given by the Casida eigenvectors, which reveal how the excitations can be constructed from KS transitions. The L-mode is found to be predominately made up of the HOMO-LUMO gap transition, this is the transition between states $5 \rightarrow 6$ and is illustrated in figure 5.5(b). This has been checked by repeating the Casida calculation with all states frozen apart from states 5 and 6, which recreates the L-mode excitation[10]. The L-mode can therefore be identified to be equivalent to the left-hand side peak of the non-interacting spectrum in figure 5.5(a) but blueshifted by electron screening. This can be further confirmed by employing the collectivity index, introduced in detail in appendix K. In agreement with Yasuike *et al* (Yasuike et al., 2011), we find that the L-mode has a collectivity index of 1. All together, this would leads us to conclude that the L-mode is a long-range charge transfer excitation[11] rather than a plasmon.

[10] Apart from a small depolarization shift

[11] The L-mode is very similar to high absorption excitations seen in conjugated polymers with high persistence lengths (Vezie et al., 2016)

What complicates matters is that the L-mode has been identified as plasmon-like in previous works because of it scaling with the Coulomb strength, (Bernadotte et al., 2013) and large plasmonicity index (Bursi et al., 2016). The role of the Coulomb interaction is illustrated by the fact the L-mode is the second lowest Casida excitation[12], this hints that the Coulomb interaction does play an important role but is not able to couple different transitions (Piccini et al., 2013). The L-mode thus exhibits a duality of plasmonic and low collectivity behaviour, we believe it is sensible to label it as a "proto-plasmon": a mode peculiar to the extreme geometry of the atomic chain and precursor to the classical longitudinal plasmon of a rod/ellipsoid.

[12] The first is a dark mode and so is not visible in figure 5.5(a).

Compared to the L-mode, the TC-mode has a more complicated structure; it is made up of multiple KS orbital transitions, which can be gleaned from the Casida eigenvectors. This already suggests that the TC-mode is much more collective and is confirmed by a collectivity index of 4.6. This tells us that virtually all the electrons take part in this excitation[13]. The explanation for this is the large number of transitions which have a change in the quantum number associated with the y and z direction (the short lengths of the chain), these transitions are visible as the right-hand side peak in the non-interacting spectrum shown in 5.5(a). These transitions have the same symmetry and are nearly degenerate in energy (because the length is much longer than the radius, states with different quantum numbers associated in the x direction are close in energy), which has been previously identified as a condition for single-particle excitations to couple and form collective modes (Guidez and Aikens, 2014). The effect is further increased by a degeneracy of two associated with the equivalence of the y and z direction, breaking this symmetry would decrease the collectivity of transverse modes. The electron-electron interactions merge these transitions into a single blueshifted peak, this is illustrated in figure 5.5(b), note that there are contributions from higher energy transitions to unoccupied states for the TC-mode not shown in figure. Hence, it is necessary to include a large number of unoccupied states for an accurate description of high energy T-modes[14].

[13] Due to spin degeneracy, the collectivity index should be multiplied by two to get in terms of electron number.

Also shown in figure 5.5(a) is our newly introduced plasmon index indicated by the vertical black lines[15]. The collectivity index, while useful for identifying collective modes, does not by itself hold any information on how the mode couples to light. If we are interested in bright plasmons then the mode must also possess a strong dipole moment to couple with light, we therefore define a simple plasmon index by multiplying the dipole strength function[16] and the collectivity index. The index identifies the TC-mode as the most plasmon-like excitation due to its combined large dipole moment and collectivity. The TE-mode and L-mode are identified as moderately plasmonic. Interestingly, the second-order L-mode, with an energy of 2.09 eV and just visible on figure 5.5(a), has a much larger collectivity of 3.7 than the first-order L-mode, although

[14] It is important to pick a number of unoccupied states such that there is an equal number of transitions associated with the y and z direction, else the excitation degeneracy associated with this symmetry can be lifted from the truncation.

[15] For convenience we divide by the number of electrons multiplied by $\pi\eta$. This comes from the normalisation of the Lorentzian which is convolved with the Casida oscillator strengths to obtain the continuous spectrum.

[16] Of course if one was not interested in dipole plasmons then the index can be suitably modified. For instance quadrupole modes could be identified using the quadrupole strength function.

it is not identified as very plasmonic by the index due to a small dipole moment.

Figure 5.6(a) the results of a sweep over different chain lengths, from 2 to 24 atoms[17] is shown. How the excitation energy depends on the chain length can aid our understanding, the low computational cost of the sodium chain affords us this luxury of a thorough geometry sweep not possible for larger systems. The L-mode redshifts with increasing length and can be understood either classically as a decrease in the restoring force on the oscillating electrons, or quantum mechanically as a decrease in the HOMO-LUMO gap. This similarity between collective charge oscillations and single eh pairs has been noted elsewhere for a Na_2 molecule (Kuemmel and Brack, 2001) and perhaps should not be too surprising given both need to have the same symmetry to couple strongly to light. The TE-mode and TC-mode modes do not shift above a chain length of about 6 atoms, as the chain width is kept constant. We find that the L-mode does not become more collective as the chain size is increased, the largest chain checked was a Na_{60} chain and again a collectivity of 1 was found for the L-mode[18]. The low energy absorption spectrum for Na_{60} is shown in figure 5.7, revealing many higher order L-modes which have increasingly higher collectivities. This is related to the increased number of eh transitions able to contribute at higher energies. As the chain size is increased these modes bunch together and in the infinite limit give the dispersion curve of a plasmon in a 1D homogeneous electron gas (Bernadotte et al., 2013). In the limit of an infinite chain the restoring force provided by the ends of the chain will go to zero, this is reflected in the dispersion by the long wavelength limit of the 1D longitudinal plasmon going to zero energy[19].

[17] In steps of two so all chains have an even number of valence electrons to ensure the system is always spin unpolarized for simplicity.

[18] Interestingly it is now the 5th lowest energy excitation again indicating the role of the Coulomb interaction.

[19] This should be contrasted for the dispersion of plasmons in 2D and 3D homogeneous electron gas which goes to a finite energy in the long wavelength limit (Giuliani and Vignale, 2005).

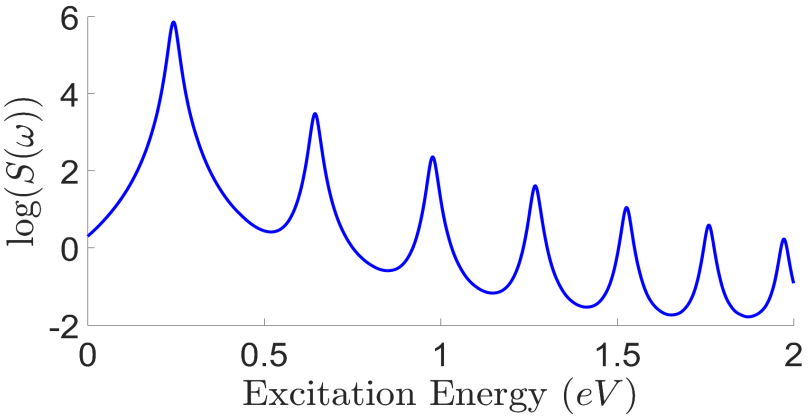


Figure 5.7: The log of the dipole strength function for a Na_{60} chain, showing the higher order longitudinal plasmon modes. A total of 60 occupied states and 80 unoccupied states were considered.

To achieve further understanding, we try and identify the Casida excitations with the corresponding classical modes, see appendix M for further details. This is perhaps the simplest and crudest method for identifying a plasmon in a quantum system. We choose a prolate spheroid as the classical comparison, which has two plasmon modes that can be expressed analytically: a longitudinal

and two degenerate transverse. The plasma frequency is assumed to be dependent on the minor semiaxis, which will depend on the amount of electron spill-out and we will label as Δ, via the electron concentration, this reduces the number of fitting parameters[20]. The L-mode energy is calculated for chain sizes up to 40 atoms, for the longer chains, which can be expected to behave the most classical, we fit for Δ and get a value of $1.45 \times d$, the fit is shown in figure 5.6(b). For larger chains the classical model works surprisingly well, but below about 10 atoms the model starts to fail and predicts too low excitation energies. This can be attributed to quantum effects and a failure of the ellipsoid approximation. For these fitting parameter the TC-mode energies predicted are too low by about an eV, unsurprisingly highlighting that the classical model breaks down for such small length scales; the larger energy in the quantum model is probably due to the strong quantum confinement in the transverse direction. The classical model also predicts the incorrect behaviour for the small energy shift of the TC-mode with chain length; the classical calculation shows a slight blueshift with increasing chain length for small chain sizes before becoming constant, while the quantum model predicts a redshift. In figure 5.6(c), the corresponding classically calculated absorption cross section is shown. We can see similarities with the Casida spectrum in 5.6(a), showing that the identification of the TC-mode as the quantum limit of the classical T-mode seems to be correct.

[20] Similar analysis was performed in reference (Yan and Gao, 2008) they found a surprisingly good agreement with the classical ellipsoid model by fitting both the shorter semiaxis Δ and using a bulk plasma frequency of 3.83 eV. In fact, whilst the fit of the L-mode looks good they seem to have incorrectly fitted to the TE-mode rather than the TC-mode.

5.2.3 *Calculating The Field Enhancement*

The local FE of an excitation can be found from the induced density[21] $\Delta n(\mathbf{r}, \omega)$, which can be obtained from the real part of the density matrix (Casida, 1995). Close to an excitation the induced density can be written in the following revealing form for the Ith excitation

$$\Delta n(\boldsymbol{r}, \omega_I) = -\frac{\Delta n^I(\boldsymbol{r})}{i\eta} \boldsymbol{\mu}^I \cdot \boldsymbol{E}. \tag{5.1}$$

In appendix L we show a derivation of this equation. It has also been derived previously, in a slightly different manner (Cocchi et al., 2012). The FE can be obtained within the electrostatic approximation by calculating the induced electric field, $\Delta \boldsymbol{E}(\boldsymbol{r}, \omega)$, by solving the Poisson equation[22].

$$FE = \frac{|\Delta \boldsymbol{E}(\boldsymbol{r}, \omega) + \boldsymbol{E}_{ext}|}{|\boldsymbol{E}_{ext}|}. \tag{5.2}$$

Equation 5.1 is a very intuitive equation that can be understood as follows:

- The transition density $\Delta n^I(\boldsymbol{r})$ gives the spatial profile of the excitation, it gives a 'snapshot' of the density oscillation and tells us how the excitation moves charge around the system.
- The factor $\boldsymbol{\mu}^I \cdot \boldsymbol{E}$ measures the projection of the external field

[21] Not to be confused with the transition density $\Delta n^I(\mathbf{r}, \omega)$, although they are closely connected, see section 2.3.4.

[22] This is solved on the same real-space grid as the TDDFT calculation.

on the transition dipole moment and tells us how strongly the excitation couples to light, i.e. whether it is a bright mode.

- The response is purely imaginary as the electrons oscillate out of phase at the resonance.
- It depends inversely on the broadening parameter η. As expected, the higher the loss the lower the FE.

Clearly the loss parameter dictates the FE, and so a careful choice is necessary if we want to estimate the FE possible in quantum plasmonics. Using DFPT, the vibrational normal modes of the Na_{10} chain are calculated. In total there are $3N = 30$ modes, with only the top 9 modes corresponding to nonzero energies. In figure 5.8(a) we show the excitation energies and illustrate the corresponding motion of the sodium atoms, which all correspond to longitudinal motion only. We take the maximum energy of 0.0228 eV and use it as the linewidth η in the Casida calculations. Unfortunately there is no experimental values to test it against, but we note that it is comparable to exciton spectral linewidths observed for carbon nanotubes at room temperature (Jones et al., 2005). It is about an order of magnitude lower than experimental values for 3D sodium clusters Na_8 (Chris Wang et al., 1990) and Na_{20} (Xia et al., 2009), this could be down to the different geometry or from extra loss channels present in these experiments. This is the simplest method to include electron-phonon coupling and should give an upper-bound on the true value. More detailed plasmon-phonon coupling models will exhibit richer spectral features, as recently shown for aromatic hydrocarbons (Cui et al., 2016). For increasing chain length the L-mode energy reduces while the phonon energy increases, this suggests for longer chains there could be an interesting strong-coupling regime between the two.

The maximum FE found using the phonon-induced line broadening is 80.5 for the L-mode and 78.3 for the TC-mode, which corresponds to intensity enhancements of approximately 6400. These values are comparable to the FEs in larger plasmonic nanostructures. The electric fields spatial profile is also qualitatively very similar to that seen for larger plasmonic nanoantennas although with obvious deviations near the sodium atoms. The calculated FEs are over two times larger than found in the gap region between two coupled naphthalene polyacenes (Bursi et al., 2014), there the authors used a broadening of 14 meV which is smaller than our calculated value for the sodium chain, suggesting even at these length scales metals are efficient materials for exciting large FEs. From a practical point of view, whilst the maximum FEs are pleasing, the sharp fall off of the induced field in space, due to the small number of electrons, would limit any potential plasmonic applications. Any molecules placed in these high field regions would feel a strong electronic interaction with the antenna; the back-coupling, via chemical and EM interaction[23], would significantly perturb the antenna. Further studies of strongly coupled quantum plasmonic antenna + molecule

[23] On these tiny scales it gets increasingly difficult/pointless to distinguish the two.

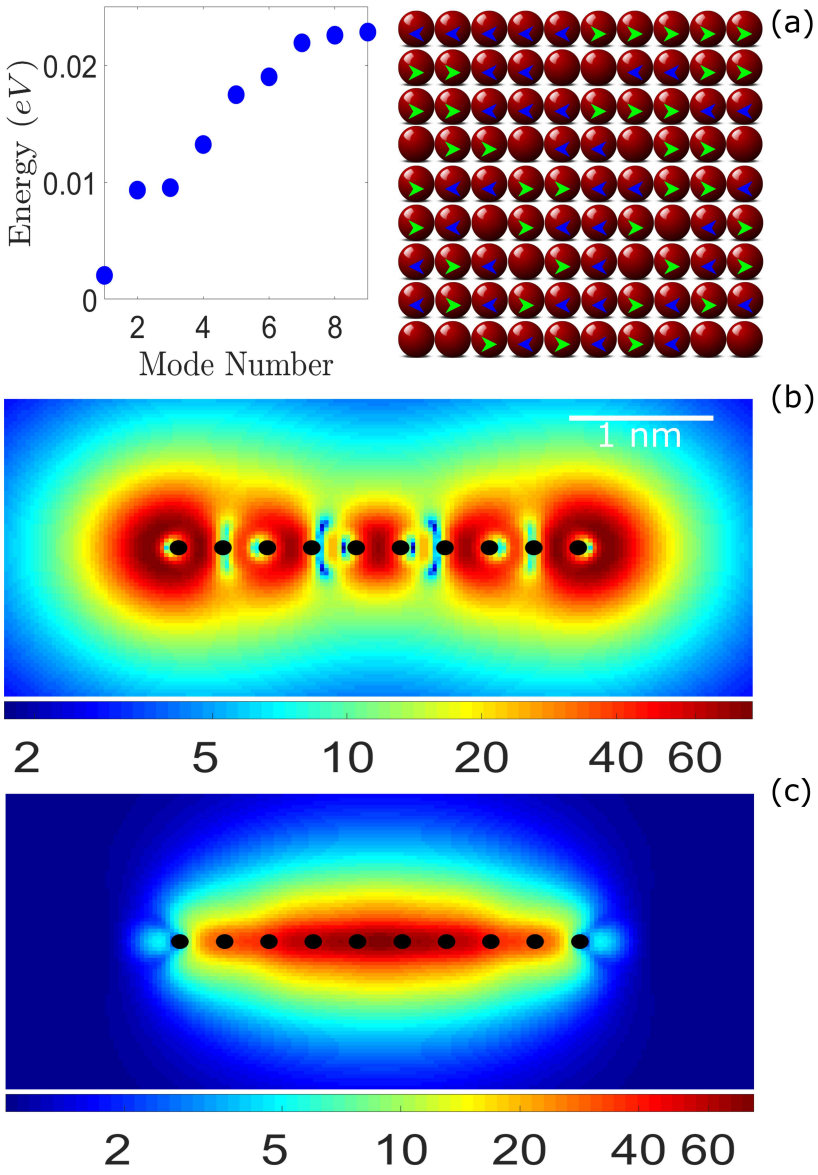


Figure 5.8: (a) The phonon normal mode energies and corresponding atomic motion of the Na_{10} chain. (b) The FE calculated using equation 5.1, in $\log_{10}$ scale, of the L-mode. (c) And TC-mode. The black dots represent the sodium atom positions and the white bar indicates a length 1 nm.

systems are necessary. More modest FEs are possible further away from the nanoantenna; the TC-mode supports FEs of around 5 at about 2.2 bond lengths away which suggests it may be possible to alter the photonic environment with minimal electronic coupling between the chain and a small absorbing molecule. Fortunately, as we saw in section 1.2.3 and appendix E, small nanoparticles support extremely large field gradients. These could be useful in exciting higher order multipole excitations (Alabastri et al., 2016; Rivera et al., 2016) and nonlinear effects (Cox and De Abajo, 2014; Yamaguchi and Nobusada, 2016) in small molecules and atoms. As an example, the TC-mode has a maximum gradient enhancement $\nabla|\Delta\boldsymbol{E} + \boldsymbol{E}_{ext}|/|\boldsymbol{E}_{ext}|$ of 17.1 a_0^{-1} in the z direction for the slice $x = y = 0$. Finally, the system if ever experimentally realised would most likely be built on a substrate, which could have a large effect on the optical response and in particular the FE. It is not included here to avoid clouding the essential physics and keep the calculation simple, but the inclusion of a substrate would be an interesting direction of future research.

5.3 *Quantum Plasmonics Of The Sodium Chain Dimer*

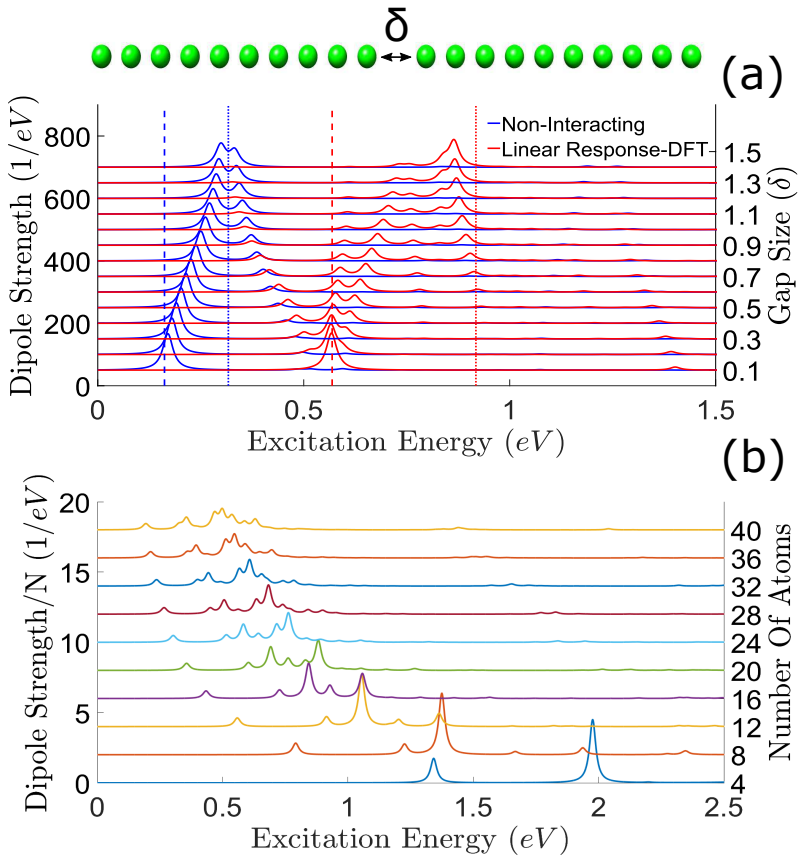
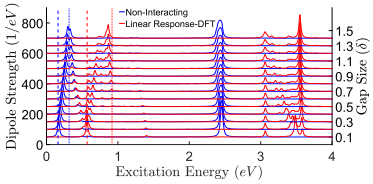


Figure 5.9: (a) The dipole strength function of a Na_{20} chain with a varying gap width, δ, in the centre. δ is varied from 0 to $1.4d$, with a value of $1d$ corresponding to a missing atom in the centre. The dashed (dotted) blue lines show the Na_{20} (Na_{10}) L-mode energies for the non-interacting case and the red lines for the interacting. (b) Dipole strength function, divided by the number of valence electrons, for varying chain sizes from 4 to 40 atoms with a constant gap in the middle of $1d$.

Now we consider two Na_{10} chains placed end to end, or alternatively taking a single Na_{20} chain and separating it in the middle. Remarkably, given the amount of interest in dimers within classical plasmonics, there has been no excited state structural analysis of gap plasmons in ultrasmall metallic systems that we are aware of. In figure 5.9(a) we show the low energy spectra as the gap length is changed by enlarging the middle sodium bond (δ) up to $1.4d$, $\delta = 0$ corresponds to an unchanged Na_{20} chain and $\delta = d$ to a Na_{20} chain with a missing atom in the centre. We resist considering larger gaps as the LDA is not suitable to consider long range van der Waals interactions that would begin to dominate. The non-interacting electronic spectra can be simply understood as a splitting of the KS energy levels due to the gap and the creation of symmetric and anti-symmetric pairs of eigenstates. The large difference between the non-interacting and interacting spectra highlights the crucial role of electron-electron interactions. Turning to the interacting spectra, the transverse modes[24] are relatively unchanged from the presence of a gap as is the case for classical nanorod dimers (Jain et al., 2006), there is though some interference effects between the TC-mode and TE-mode for small gaps. The spectra for the longitudinal modes are much more interesting, they show the formation of a number



The full excitation spectra for a Na_{20} chain with varying gap width.

[24] For clarity they are not shown in figure 5.9(a), the full spectra up to 4 eV is shown above in the margin figure.

of new longitudinal modes for gap sizes $0.4d \rightarrow 1.4d$, which we will label as "tunnelling" or "gap" modes as they involve the transfer of charge across the gap, due to electron tunnelling, every optical cycle. They are qualitatively similar to the low energy CTPs that appear for nearly touching dimers (Zhu et al., 2016). We find that they tend to have a low collectivity compared to the T-modes, but have a larger collectivity than the L-mode of a single chain, indicating that they arise as a consequence of multiple eh pairs excited across the length of the total system. At larger separation distances, $> 1.2d$, the excitation spectra are dominated by a longitudinal BDP, which originates from the electrostatic coupling of the two chains. It is redshifted from the Na_{10} L-mode that would be present with no coupling.

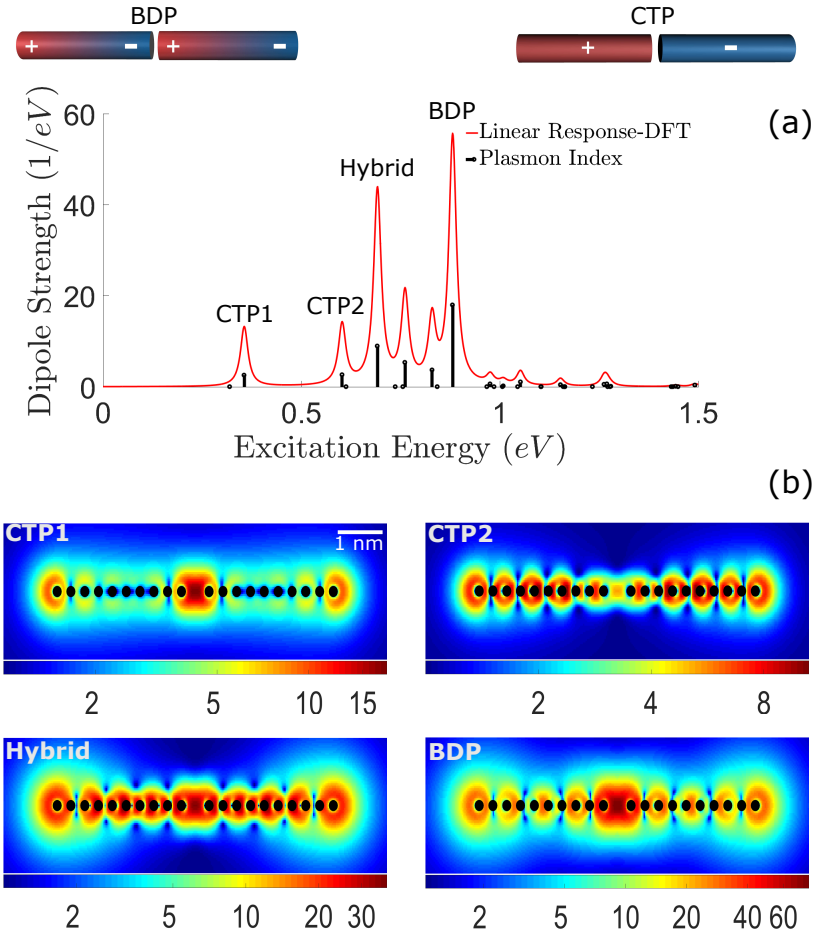


Figure 5.10: (a) The dipole strength function of a Na_{20} chain with a gap of $1d$ (i.e. one atom removed) opened in the centre. (b) The FEs, in $\log_{10}$ scale, of the CTP1, CTP2, hybrid and BDP modes.

To help identify all the peaks at intermediate gap sizes, we will employ our knowledge of macroscopic plasmonics. We will use a practical definition of CTPs: each chain should have a different sign of induced charge (see the top right hand side image of figure 5.10) and the excitation intensity should go to zero for large gap widths. The BDP must have a dipole induced charge distribution for each chain (see the top left hand side image of figure 5.10) and the excitation intensity will not go to zero for large gap widths, but

rather merge with the L-mode of a single Na_{10} chain eventually[25]. At small gap sizes, where electron tunnelling is most important, two CTPs dominate the Casida spectra. The lowest energy CTP, we will label it as CTP1, redshifts and loses intensity as the gap width increases, it becomes negligible above a gap size of about $1d$. The second CTP, which we label CTP2, also loses intensity as the gap size increases. As expected, both CTP modes merge with the Na_{20} L-mode in the limit $\delta \to 0$. We identify a single BDP that begins to become apparent around $\delta = 0.7d$ and grows in intensity for increasing gap size, it is the highest energy tunnelling mode and is close in energy to the single Na_{10} chain L-mode. In-between the CTPs and the BDP there are three more complicated modes which we believe are best explained as hybrid modes; they have a similar charge distribution as the BDP, although the change in the potential over the gap region is smaller, which is due to charge transfer neutralising the induced charge build-up either side of the gap. They die out as the gap length is widened so cannot be identified as BDPs.

[25] This would be at very wide gaps and will not be seen in our simulations.

To further aid identification of the various modes, we have also considered holding the gap width fixed and sweeping the chain length, see figure 5.9(b). We consider only the lower energy L-modes, allowing chain sizes up to 40 atoms to be considered[26]. We only look at dimers which have a total number of atoms that are multiples of 4, this is to make sure that the separated chains each have an even number of valence electrons to ensure the validity of unpolarized spin calculations. It is found that CTPs are present for very small chains of only a few atoms, highlighting that they do not involve many electron transitions and originate from the interaction of electronic wavefunctions near the HOMO-LUMO gap. Interestingly, the longitudinal tunnelling modes keep a roughly constant normalised dipole strength intensity, highlighting that they scale with system size; this suggests they would continue to exist in the macroscopic limit.

[26] Lower energies means not as many unoccupied states need to be calculated.

To explore the possible FE and further confirm our mode labelling, we consider a specific gap width of $\delta = d$ for a Na_{20} dimer, this is equivalent to one extra bond length in the centre of a N_{20} chain. In figure 5.10, the low energy excitation spectrum (a) and the FE of four of the tunnelling modes (b) is shown. We find that at this gap size our plasmon index predicts the hybrid mode and BDP to be far more plasmonic than the CTP1 and CTP2. Furthermore, the calculated FE in figure 5.10(b) shows maximum values of 38.2 and 80.1 for hybrid and BDP modes respectively, compared to 18 and 10.6 for the CTP1 and CTP2. The large plasmonicity for the BDP is found to come from its large dipole moment rather than a large collectivity, although of the four modes analysed it has the largest collectivity of 3.29, comparing to values close to 2 for the other three modes. The BDP mode is also the most interesting out of the four in terms of potential applications as it exhibits the largest FE in the gap region, where potentially a small atom could be placed

and its emission and absorption properties modified. These atomic chain dimers act as the ultimate small size limit of picocavities. Of course for such a tiny molecular system, even an individual atom can strongly back-couple to the cavity and heavily modify the achievable FEs. Further research is needed to see how small plasmonics cavities can realistically be pushed to, the methods developed in this chapter are a first step in this direction.

Classification of the various peaks in the Casida spectrum is actually easiest using the induced electric potential. The nature of the modes is not apparent from the FEs in 5.10(b). The induced potential can be calculated from the induced density by solving Poisson's equation $\nabla^2(\Delta\phi(\boldsymbol{r},\omega)) = e\frac{\Delta n(\boldsymbol{r},\omega)}{\epsilon_0}$, the results of this calculation for the six major low-energy tunnelling modes is shown in figure 5.11. By comparing to the potential for classical plasmonic dimers, see the top of figure 5.10(a), it is seen that the CTP1 mode is most CTP-like and the modes become increasingly BDP-like with increasing energy. Both the CTP1 and BDP show a fantastic similarity with their classical analogue. This is a really pleasing example of how plasmonics can be used for understanding of an electronic excitation spectrum. This should be compared to the analysis of the KS orbitals, which in this case is complicated and not very illuminating.

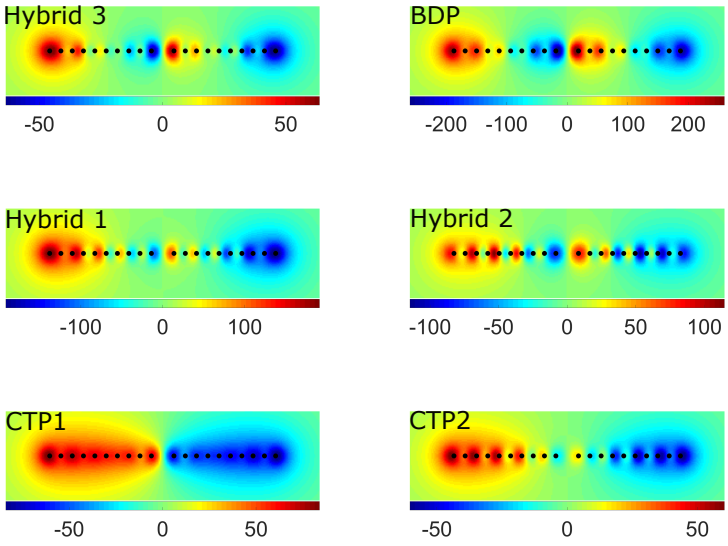


Figure 5.11: The induced electric potential for the six tunnelling modes considered in the text. 'Hybrid 1' corresponds to the lowest energy and 'Hybrid 3' to the largest energy hybrid modes shown in figure 5.10 between the CTP2 and BDP.

5.4 *The Quantum Nanoantenna*

Finally, we give some justification for considering these molecular systems as quantum nanoantennas, i.e. antennas operating in the quantum regime (Fitzgerald and Giannini, 2017). An antenna *efficiently* converts EM energy from the far- to near-field, or vice versa. Ultra-small systems are seen as poor antennas, in fact one of the major application of plasmonic nanoantennas is to boost the light-matter interaction of nearby small quantum systems such as molecules and quantum dots. We propose quantum plasmonic antennas as ideal middle ground between a classical plasmonic antenna and a molecule/quantum dot, we justify this as follows:

- We are often interested in designing antennas that are efficient for their size, by considering excitation collectivity we can design molecular systems with the biggest possible cross section. For instance, the size could be dictated by other requirements such as catalytic behaviour where smaller nanoparticles are often favourable (Gomes Silva et al., 2010), an 'all-in-one' nanoparticle with certain desired plasmonic and chemical properties may have an optimum size within the quantum plasmonic limit.

- Small size comes with added benefits such as tunability. For example, it has been shown for polycyclic aromatic hydrocarbons that the addition/removal of a single electron can reversibly switch on/off a molecular plasmon excitation (Lauchner et al., 2015). Quantum plasmonic devices will work within the compromise of a large cross section and tunability.

- We have seen in chapter 3 that small volumes are necessary for plasmonic-based strong-coupling. Ultra-small nanoantennas could be favourable for designing strongly mixed light-matter states, which could have applications in cavity-QED.

- Using the analogy of plasmonic antennas can help understand complicated quantum systems. It is well-known that atoms/molecules can be considered as ultra-small antennas. In the last section we demonstrated that the complicated Casida spectra from the sodium chain dimer can be nicely interpreted within a plasmonic picture. Furthermore, the analysis of FEs seems to be largely ignored in electronic structure calculations, its worth has recently been demonstrated in TDDFT calculations of water splitting by ultra-small gold nanospheres (Yan et al., 2016).

- Quantum nanoantennas do not necessarily have to act in isolation, they could be coupled to larger classical antennas to allow manipulation of light over a huge range of size scales. They could even be arranged periodically to form a quantum metamaterial; the small cross section of each individual element can be compensated by arranging the meta-molecules closely. This is

similar to the case of a crystal solid which, despite the small size of the constituent atoms, can interact strongly with light.

For these reasons I believe it is sensible to push the antenna concept down to the smallest possible length scales and that quantum plasmonic nanoantennas look set to be a rich area of research for many years to come.

6 *Some Conclusions*

It is customary to conclude with some closing words: we have taken a journey across five orders of magnitude of length scales (~ 1Å $\rightarrow 10$ μm), frequencies from the UV to the mid-IR and have used a wealth of different theoretical and numerical techniques. We have considered three very different theoretical projects:

- In chapter 3 we looked at a very practical graphene + SiC grating system, relevant to current nanophotonic experiments. The theory is based on well-known classical electrodynamics and modelled using the familiar FDTD method, despite this the outcomes are surprising due to the material-dependent physics contained in the dielectric functions and the complexities of the near-field interaction between the graphene and the grating. A novel cavity effect was found, with the structured SiC allowing graphene plasmon standing waves to be excited without direct structuring of the graphene itself. This could provide an easy to construct, tunable, high quality factor and small-volume cavity ideal to detect and study trace amount of molecules deposited on the graphene surface. Strong-coupling phenomena between phonon polaritons and graphene plasmons was also explored, revealing further opportunities for manipulation of the near-field and fundamental macroscopic electrodynamics research.

- In chapter 4 we explored the extreme limits of macroscopic plasmonic theory, performing a numerical study of the plasmonic cascade nanolens to scrutinize the achievable FEs in metal-based plasmonics, and offering an exciting alternative with phonon-polaritonics; with maximum FEs up to 10000 as a tantalising possibility. This is theory that will hopefully spur on experimental work and divert seemingly wasted energy away from metal nanolenses. SiC nanolenses designed on the scales envisioned in this thesis would be challenging to make, probably beyond current state-of-the-art experimental techniques, but I am confident that they are achievable structures over the next couple of decades. Perhaps this work can help provide some of the motivation in exploring these challenging regimes.

- In chapter 5 we unshackled our imaginations and considered a system that is perhaps not of immediate practical relevance but is an ideal 'toy model' for quantum plasmonics: the sodium chain. By using theoretical techniques borrowed from quantum

chemistry, namely excited state structural analysis, new methods for identifying and understanding quantum plasmons based on collectivity were developed. This is not a well-known method in plasmonics[1] and complements existing quantum plasmonic theory. We have also translated classical plasmonic concepts down to these ultra-small scales, introducing the FE into electronic structure calculations; not only does this allow us to study the plasmonic response of molecular structures, it also reveals a new way of analysing the optical response of molecules which seems to be ignored in molecular physics and chemistry. For example, our methods are relevant to the huge field of research on light-emitting polymers with applications in organic optoelectronics (Fox, 2002). Furthermore, we have introduced the concept of gap plasmonics for closely separated molecular dimers; using knowledge from classical systems helps interpret the complicated excitation spectra. It nicely illustrates how electronic structure can benefit from plasmonics and that the relationship between the two is not just a one-way street.

[1] An arguably not even in physics in general!

The underlying concept unifying all three projects is the collective motion of the constituent charge carriers of the materials: electrons and optical phonons. The hope is that this thesis has provided a revealing glimpse into some of the diverse research in nanophotonics. Finally, as it is pleasing to come full circle, it is observed that Archimedes' burning lens was ultimately unsuccessful at repelling the Romans and he met his demise via a quick-to-anger soldier. Let us hope that the potential nanophotonic applications discussed in this thesis have more luck!

Figure 6.1: The Death of Archimedes. Painted in 1815 by Thomas Degeorge.

A Some Remarks On Experimental Dielectric Functions

Traditionally, for silver and gold, the experimental data used in plasmonics is either from Johnson and Christy (Johnson and Christy, 1972) or Palik (Palik, 1998). There are though criticisms of this data and questions of whether the plasmonic community should rely so heavily on these particular data sets. In the case of silver, the Johnson and Christy data, which has been used for some of the results in this current work, has been criticised for giving a too optimistic value for the quality factor[1], which cannot be matched by experiment (Jiang et al., 2016). In contrast, the Palik data underestimates the quality factor. Particularly relevant to this thesis, it has been shown by Pellegrini *et al* that the theoretical maximum FE of a plasmonic cascade nanolens can change by a factor of 5, depending on the experimental dielectric function used (Pellegrini et al., 2016).

[1] Hence predicting quantities like the plasmonic FE to be too large.

To compound the uncertainty, it has been claimed that many plasmonic components used in experiment have limited performance due to incorrect deposition methods (McPeak et al., 2015). In figure A.1 we show a comparison of the quantity $-\Re[\epsilon]/\Im[\epsilon]$ for the two commonly used data sets along with some recent experimental data, taken over the wavelength region most relevant to plasmonics (Wu et al., 2014; McPeak et al., 2015; Jiang et al., 2016). While the results seem to converge to a common value at shorter wavelengths, even the more recent experimental results shows a displeasing discrepancy with one another at longer wavelengths. This uncertainty in the dielectric function, combined with the differences in the idealised geometrical structures used in theory with imperfect real nanoparticles created in the lab, means the comparison of theory with experiment can be difficult and the bar of what constitutes good agreement between the two is somewhat lower than in other areas of physics. It is most definitely something that the plasmonics community needs to improve on in the future. An interesting discussion of the dissipative loss in the case of graphene is given in (Tassin et al., 2013).

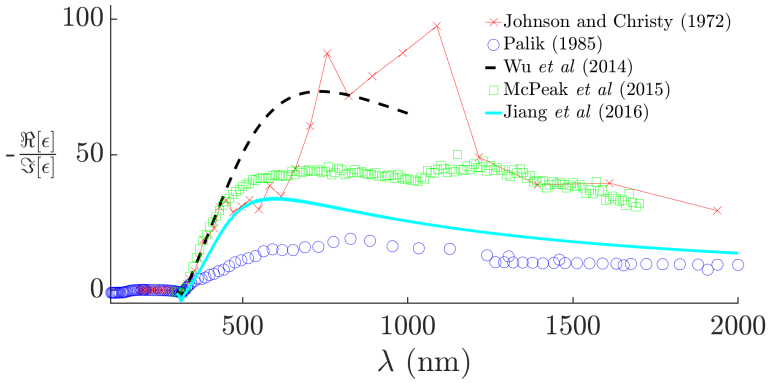


Figure A.1: Comparison of the different experimental dielectric functions described in the text.

B Macroscopic Electromagnetic Boundary Conditions

Central to many of our discussions in chapter 1 is the use of boundary conditions at the interface between different dielectric regions. The use of sharp interfaces is a pillar of classical electrodynamics and so, for perspicuity, the textbook derivation of the electric and magnetic field boundary conditions is presented here. The EM fields will be discontinuous in the presence of surface charges and currents. This can be understood by considering the integral form of Maxwell's equations (equations 1.1)

$$\begin{aligned} \oint_S \mathbf{D} \cdot d\mathbf{A} &= Q_f \\ \oint_S \mathbf{B} \cdot d\mathbf{A} &= 0 \\ \oint_l \mathbf{E} \cdot d\mathbf{l} &= -\partial_t \int_S \mathbf{B} \cdot d\mathbf{A} \\ \oint_l \mathbf{H} \cdot d\mathbf{l} &= I_f + \partial_t \int_S \mathbf{D} \cdot d\mathbf{A}, \end{aligned} \tag{B.1}$$

where Q_f and I_f are the free charge and current contained within the closed integrals, respectively. The first two integrals are over closed surfaces and can be manipulated for our purposes by choosing the famous *Gaussian pillbox* located on the surface which extends slightly into either dielectric region, see figure B.1. If we let the thickness go to zero then the contributions of the edge are zero and the exact shape we choose for the pillbox doesn't matter. For an idealised surface charge, the volume charge contribution will be zero. If we take the area of the pillbox to be A and the normal of the surface points from region 1 into 2

$$\left(\mathbf{D}^{(2)} - \mathbf{D}^{(1)}\right) \cdot \mathbf{A} = \sigma_f A, \tag{B.2}$$

where we have take into account that the only contribution to the enclosed charge can be from a surface contribution, $Q_f = \sigma_f A$. Therefore, the normal component of the electric displacement is discontinuous by an amount given by the surface charge density σ_f. The magnetic field at the boundary is evaluated in a similar fashion but will always be continuous due to the apparent absence of magnetic charge in the universe.

The second pair of integrals are evaluated using a very thin

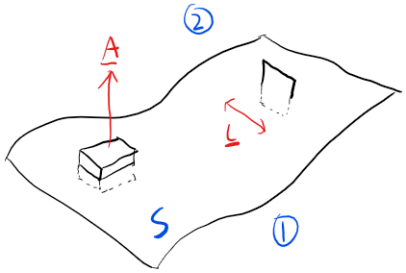


Figure B.1: Illustration of the Gaussian pillbox and the Amperian loop used to derive equations B.7.

Amperian loop, illustrated in figure B.1. In the limit of infinitesimal width, the magnetic flux through the loop will go to zero and hence the parallel component of the electric field ($\mathbf{L}\cdot\mathbf{E}$) must be continuous

$$\left(\mathbf{E}^{(2)}-\mathbf{E}^{(1)}\right)\cdot\mathbf{L}=\lim_{A\to 0}\partial_t\int_S\mathbf{B}\cdot d\mathbf{A}=0. \tag{B.3}$$

In a similar fashion the 4th equation is evaluated, again the integral on the right hand side will go to zero as the width is decreased, but the enclosed current contribution will not

$$\left(\mathbf{H}^{(2)}-\mathbf{H}^{(1)}\right)\cdot\mathbf{L}=I_f. \tag{B.4}$$

In the limit of infinitesimal width the volume current will not contribute, but a surface current can. The contributing surface current will be in the area perpendicular to both $\mathbf{L}$ and $\hat{\mathbf{n}}$

$$I_f=\mathbf{J}_f\cdot\hat{\mathbf{n}}\times\mathbf{L}=(\mathbf{J}_f\times\hat{\mathbf{n}})\cdot\mathbf{L}. \tag{B.5}$$

This means that

$$\left(\mathbf{H}^{(2)}-\mathbf{H}^{(1)}\right)\times\hat{\mathbf{n}}=\mathbf{H}_{\parallel}^{(2)}-\mathbf{H}_{\parallel}^{(1)}=(\mathbf{J}_f\times\hat{\mathbf{n}})\times\hat{\mathbf{n}}=-\mathbf{J}_f. \tag{B.6}$$

So the four macroscopic electrodynamic boundary conditions in all their glory are:

$$\boxed{\begin{aligned} D_{\perp}^{(2)}-D_{\perp}^{(1)}&=\sigma_f, \quad B_{\perp}^{(2)}-B_{\perp}^{(1)}=0\\ \mathbf{E}_{\parallel}^{(2)}-\mathbf{E}_{\parallel}^{(1)}&=0, \quad \mathbf{H}_{\parallel}^{(2)}-\mathbf{H}_{\parallel}^{(1)}=-\mathbf{J}_f. \end{aligned}} \tag{B.7}$$

A very relevant geometrical specification of these relations is that of the boundary conditions for the surface of a sphere. If $\hat{\mathbf{r}}$ is the spherical radial vector then

$$\begin{aligned} \left(\epsilon_2\mathbf{E}^{(2)}-\epsilon_1\mathbf{E}^{(1)}\right)\cdot\hat{\mathbf{r}}&=\frac{\sigma_f}{\epsilon_0}\\ \left(\mathbf{B}^{(2)}-\mathbf{B}^{(1)}\right)\cdot\hat{\mathbf{r}}&=0\\ \left(\mathbf{E}^{(2)}-\mathbf{E}^{(1)}\right)\times\hat{\mathbf{r}}&=0\\ \left(\mathbf{H}^{(2)}-\mathbf{H}^{(1)}\right)\times\hat{\mathbf{r}}&=-\mathbf{J}_f. \end{aligned} \tag{B.8}$$

C Surface Plasmon Polaritons In Planar Geometries

We will consider flat interfaces in the xy plane and simplify the equations by considering wave propagation in the x direction, in the y direction both the dielectric environment and the EM fields have no spatial dependence. This reduces the problem down to $2D$ and in fact the x direction is particularly simple due to it planewave form, so it is effectively a $1D$ problem. The dielectric function has a spatial dependence of the form $\epsilon(z)$ and we will assume we have a number of well defined dielectric environments dived by sharp interfaces[1], described mathematically by step functions. This allows us to split up the problem into separate regions in which we solve Helmholtz equation, as per the discussion of section 1.2. We can write the electric and magnetic field as follows

$$\partial_z^2 \begin{bmatrix} \mathbf{E}(z) \\ \mathbf{H}(z) \end{bmatrix} + (k_0^2 \epsilon - k_x^2) \begin{bmatrix} \mathbf{E}(z) \\ \mathbf{H}(z) \end{bmatrix} = 0. \tag{C.1}$$

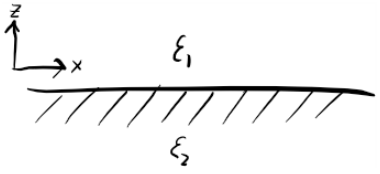
Using Ampere's and Faraday's laws from equations 1.1 gives us the following six coupled differential equation

$$\begin{aligned} \partial_z E_y = -i\omega\mu_0 H_x, \quad \partial_z H_y = i\omega\epsilon_0\epsilon E_x \\ \partial_z E_x - ik_x E_z = i\omega\mu_0 H_y, \quad \partial_z H_x - ik_x H_z = -i\omega\epsilon_0\epsilon E_y \\ ik_x E_y = i\omega\mu_0 H_z, \quad ik_x H_y = -i\omega\epsilon_0\epsilon E_z. \end{aligned} \tag{C.2}$$

These equations are the starting point for the transfer matrix method, which can be seen as a generalisation of the derivations presented in this appendix. The set of equations have two distinct set of solutions: transverse magnetic (TM) and transverse electric, we are interested in the TM modes.

The first problem we tackle in this appendix is that of a single interface between two infinite half-spaces, we wish to derive the SPP condition that is analysed in section 1.2.2. The approach we take is to use an ansatz for the solution field profile of H_y, i.e. it must show oscillatory behaviour in the x direction and exponential decay in the z direction. Then we use the following two equations

[1] The EM boundary conditions of sharp interfaces are discussed in appendix B.



The problem geometry for a single interface between two dielectric media.

from equations C.2 to calculate the electric field components

$$E_x^{(j)} = -\frac{i}{\omega\epsilon_0\epsilon_j}\partial_z H_y^{(j)}$$
$$E_z^{(j)} = -\frac{k_x}{\omega\epsilon_0\epsilon_j}H_y^{(j)}, \tag{C.3}$$

where the index j denotes the different dielectric regions, in this case $j = \{1, 2\}$. This leads to

$$\mathbf{E}(x,z) = A_j[ik_{z,j}\,\mathrm{sgn}(z), 0, -k_x]\frac{1}{\omega\epsilon_0\epsilon_j}e^{ik_x x}e^{-\mathrm{sgn}(z)k_{z,j}z}$$
$$\mathbf{H}(x,z) = A_j[0,1,0]e^{ik_x x}e^{-\mathrm{sgn}(z)k_{z,j}z}, \tag{C.4}$$

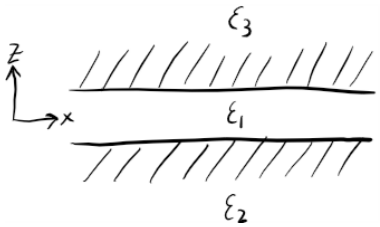
where $\mathrm{sgn}(z)$ is the sign function and is -1 for $z < 0$ and $+1$ for $z > 0$. A_j is a constant, defined in each half-space, that is determined by continuity of H_y

$$H_{x,1}(z=0) = H_{x,2}(z=0),$$
$$\therefore A_1 = A_2. \tag{C.5}$$

The wavenumber perpendicular to the surface, $k_{z,j}$, is determined by the continuity of the tangential electric field

$$E_{x,1}(x, z=0) = E_{x,2}(x, z=0),$$
$$\therefore \frac{k_{z,2}(\omega)}{k_{z,1}(\omega)} = -\frac{\epsilon_2(\omega)}{\epsilon_1(\omega)}, \tag{C.6}$$

which is just equation 1.23 from section 1.2.2, the SPP dispersion can be obtained as detailed there.



The problem geometry for a thin film.

Another simple and insightful problem to tackle is two interfaces, i.e. a film of material in-between two infinite half-spaces. If the film is very thick then the problem is trivial and one can just use the solution above for each interface, but when the film is thin the SPPs at each interface can couple, leading to new physics (Dionne et al., 2005). The problem is set up as before but now $j = \{1, 2, 3\}$, where index 1 denotes the middle layer. The solutions for regions 2 and 3 are of the same form as in the previous problem, but the ansantz for region 1 is a combination of a decaying and growing exponentials[2]

[2] The x dependence is left out in the following equation as it is just an unimportant phase that will cancel out.

$$\left.\begin{aligned} H_y^{(3)} &= Ae^{-k_{z,3}z} \\ E_x^{(3)} &= \frac{Aik_{z,3}}{\omega\epsilon_0\epsilon_3}e^{-k_{z,3}z} \\ E_z^{(3)} &= -\frac{Ak_x}{\omega\epsilon_0\epsilon_3}e^{-k_{z,3}z} \end{aligned}\right\} \text{Region 3}$$

$$\left.\begin{aligned} H_y^{(2)} &= Be^{+k_{z,2}z} \\ E_x^{(2)} &= -\frac{Bik_{z,2}}{\omega\epsilon_0\epsilon_2}e^{+k_{z,2}z} \\ E_z^{(2)} &= -\frac{Bk_x}{\omega\epsilon_0\epsilon_2}e^{+k_{z,2}z} \end{aligned}\right\} \text{Region 2} \qquad \text{(C.7)}$$

$$\left.\begin{aligned} H_y^{(1)} &= \left(Ce^{+k_{z,1}z} + De^{-k_{z,1}z}\right) \\ E_x^{(1)} &= \left(-Ce^{+k_{z,1}z} + De^{-k_{z,1}z}\right)\frac{ik_{z,1}}{\omega\epsilon_0\epsilon_1} \\ E_z^{(1)} &= \left(Ce^{+k_{z,1}z} + De^{-k_{z,1}z}\right)\frac{k_x}{\omega\epsilon_0\epsilon_1} \end{aligned}\right\} \text{Region 1}$$

Applying continuity of E_x and H_y at the interfaces located at $+a$ and $-a$ leads to four coupled equations

$$\begin{aligned} H_y^{(3)}(a) = H_y^{(1)}(a) \quad &\rightarrow \quad Ae^{-ik_{z,3}a} = \left(Ce^{+k_{z,1}a} + De^{-k_{z,1}a}\right) \\ H_y^{(2)}(-a) = H_y^{(1)}(-a) \quad &\rightarrow \quad Be^{-ik_{z,2}a} = \left(Ce^{-k_{z,1}a} + De^{+k_{z,1}a}\right) \\ E_x^{(3)}(a) = E_x^{(1)}(a) \quad &\rightarrow \quad \frac{Ak_{z,3}}{\epsilon_3}e^{-ik_{z,3}a} = \frac{k_{z,1}}{\epsilon_1}\left(-Ce^{+k_{z,1}a} + De^{-k_{z,1}a}\right) \\ E_x^{(2)}(-a) = E_x^{(1)}(-a) \quad &\rightarrow \quad -\frac{Bk_{z,2}}{\epsilon_2}e^{-ik_{z,2}a} = \frac{k_{z,1}}{\epsilon_1}\left(-Ce^{-k_{z,1}a} + De^{+k_{z,1}a}\right), \end{aligned} \qquad \text{(C.8)}$$

which can be solved systematically by eliminating the factors $Ae^{-ik_{z,3}a}$ and $Be^{-ik_{z,2}a}$, and then C and D to obtain the following condition for a surface mode (Maier, 2007)

$$e^{-4k_{z,1}a} - \frac{\left(\frac{k_{z,1}}{\epsilon_1} + \frac{k_{z,3}}{\epsilon_3}\right)}{\left(\frac{k_{z,1}}{\epsilon_1} - \frac{k_{z,3}}{\epsilon_3}\right)}\frac{\left(\frac{k_{z,1}}{\epsilon_1} + \frac{k_{z,2}}{\epsilon_2}\right)}{\left(\frac{k_{z,1}}{\epsilon_1} - \frac{k_{z,2}}{\epsilon_2}\right)} = 0. \qquad \text{(C.9)}$$

As a check we find in the limit $a \to \infty$ the result for a single interface is recovered. The coupling of the SPPs leads to an energy splitting between two modes[3], we analyse this in the case of an insulator-metal-insulator geometry, i.e. a thin metal film surrounded by two dielectric half-spaces. For simplicity we will consider a symmetrical environment so $\epsilon_3 = \epsilon_2$. In figure C.1 a contour plot of the solution to equation C.9, in ω-k_x space, for a 30 nm thick silver film is shown[4]. Blue regions indicate where the expression is satisfied and reveal the dispersion of the non-radiative thin-film SP modes. The higher energy mode ω^+ is often labelled as the odd mode[5] and the lower energy mode ω^- is called the even mode (Maier, 2007). ω^+ asymptotically approaches the single-interface SP energy from above and ω^- from below, the amount of splitting between the two modes depends inversely on the film thickness. The natural length scale to measure the thickness is the

[3] This energy splitting between interacting modes is a common theme in this thesis, see chapter 3 in particular.

[4] There are actually four solutions to equation C.9, but two correspond to physically insignificant 'leaky' modes, they are to the left of the light line and, like the bound modes, come in a symmetric and antisymmetric pair (Johnstone et al., 1990). They are ignored in the plot.

[5] E_x is an antisymmetric function with respect to $z = 0$. It can be a little confusing as sometimes the modes are labelled according to the symmetry of H_y and E_z (Alù and Engheta, 2006), which will be *opposite* to the symmetry of E_x. Because of this, I will not use this labelling.

skin depth of the metal, for a thickness much greater than the skin depth the SPPs will not couple. In the limit of small thickness, the higher energy mode, because it has a zero in the tangential field due to the destructive interference between the coupled SPPs, has only a small amount of the EM field present in the metal; this leads to a long-range propagation length. Hence it is often labelled as the *long-range* SPP. In contrast for the lower energy mode, the field is strongly confined to the metal layer but the price paid is a very short propagation length, it is often labelled as the *short-range* SPP (Maier, 2007).

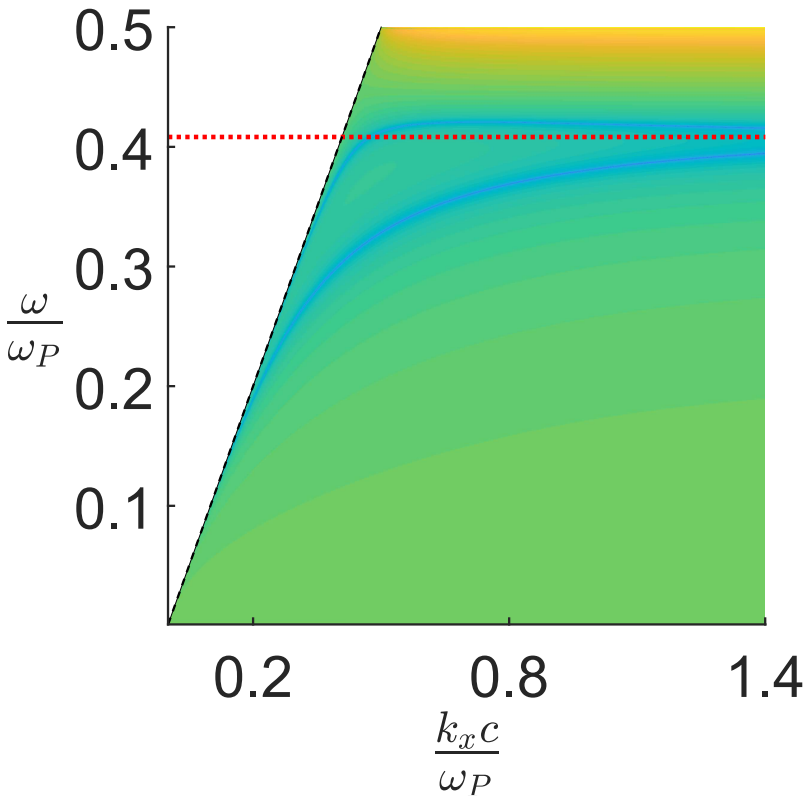


Figure C.1: The roots of equation C.9 for a 30 nm thick silver film in air. The blacked dashed line indicates the light line and the red dotted lined indicates the SP resonance for a single interface. The silver is described by a Drude formula with parameters $\omega_P = 8.9$ eV, $\hbar\gamma = 0.0366$ eV and $\epsilon_\infty = 5$.

As well as providing an interesting system to understand SPP coupling, the thin metal film in the limit $a \to 0$ provides a route to understanding graphene SPPs. We can write the dielectric function as $\epsilon(\omega) = 1 + \frac{i\sigma_V(\omega)}{\epsilon_0\omega}$, where σ_V is a *volume* conductivity. This expression is simple to derive from combining Ampere's law with Ohm's law

$$\begin{aligned} \nabla \times \mathbf{B}(\mathbf{r},\omega) &= \mu_0 \left(\sigma(\mathbf{r},\omega) - i\omega\epsilon_0\epsilon(\mathbf{r},\omega)\right) \mathbf{E}(\mathbf{r},\omega) \\ &= -i\omega\mu_0\epsilon_0\tilde{\epsilon}(\mathbf{r},\omega)\mathbf{E}(\mathbf{r},\omega), \end{aligned} \tag{C.10}$$

where $\tilde{\epsilon}$ is a redefined permittivity that includes the effect of the conductivity (we will relabel it as simply ϵ). The volume conductivity can be rewritten as a surface conductivity σ_S, which is more suitable for graphene, using $\sigma_S(\omega) = 2a\sigma_V(\omega)$. We consider the short-range SPP and expand for small $k_{z,1}a$. Using equation C.9, it

is possible to write the following equation for the short-range SPP in the case of equal sub- and superstrates $\epsilon_2 = \epsilon_3$ (Maier, 2007)

$$\coth(k_{z,1}a) = -\frac{k_{z,2}\epsilon(\omega)}{k_{z,1}\epsilon_2} \tag{C.11}$$

Expanding to first-order in $k_{z,1}a$ gives (Vakil and Engheta, 2011)

$$k_{z,2} \approx k_x \approx i\frac{2\epsilon_0\epsilon_2\omega}{\sigma(\omega)}, \tag{C.12}$$

which gives the graphene dispersion given by equation 1.52 in section 1.4.1 (Jablan et al., 2013).

Perhaps a more elegant derivation for the graphene SP condition is to include it as a boundary condition via a surface current, this means there is a discontinuity in the magnetic field at the interface[6]

[6] I pick the surface normal to point into region 2.

$$H_y^{(2)}(z=0) - H_y^{(1)}(z=0) = [\mathbf{J}\times\hat{\mathbf{n}}]_y = -J_x = -\sigma E_x^{(2)}(z=0). \tag{C.13}$$

If we now repeat our analysis of a single interface, but including this discontinuity, we find that the constants A_j are related as follows

$$A_1 = A_2(1 + \frac{i\sigma k_{z,2}}{\omega\epsilon_0\epsilon_2}), \tag{C.14}$$

which means continuity of the tangential electric field now gives

$$\begin{aligned} -A_1\frac{i\sigma k_{z,1}}{\epsilon_1} &= A_2\frac{i\sigma k_{z,2}}{\epsilon_2} \\ \therefore -(1 + \frac{i\sigma k_{z,2}}{\omega\epsilon_0\epsilon_2})\frac{i\sigma k_{z,1}}{\epsilon_1} &= \frac{i\sigma k_{z,2}}{\epsilon_2}, \end{aligned} \tag{C.15}$$

which, on multiplying by $\frac{\epsilon_1\epsilon_2}{k_{z,1}k_{z,2}}$, gives the graphene SP condition

$$\frac{k_{z,1}}{\epsilon_1} + \frac{k_{z,2}}{\epsilon_2} + \frac{i\sigma(\omega)}{\epsilon_0\omega} = 0. \tag{C.16}$$

D Derivation Of The Quasistatic Condition

This simple derivation is mainly taken from (Bohren and Huffman, 2008) and gives a rough measure of when retardation effects can be expected to influence the optical response of a nanoparticle, it is included as it reveals some interesting physics of light propagation in materials. For a constant electric field over a sphere excited by a planewave of wavevector k, then we need $e^{ikR} \approx 1$, in other words there is no phase accumulation over the length of the nanoparticle. This requires that the decay over the radius[1] of the nanoparticle is negligible and hence $e^{-\Im[k]R} \approx 1$, meaning

[1] Or for a non-spherical object some characteristic length.

$$\Im[k]R = \Im[n(\omega)]R\frac{\omega}{c} << 1, \quad \text{Condition 1}$$

This condition is often stated as the justification for using the electrostatic approximation. Actually, a second condition is also necessary and arises from the requirement that the time for a signal to pass across the nanosphere is small compared to the time variation of the incident light; this means an almost static polarization will be established in the nanoparticle in a short time compared to the period (van de Hulst, 1957). The field changes over a time characterised by $\frac{1}{\omega} = \tau$, the signal propagation time t should be less than this and can be written as[2], $t = R/v_p = \Re[n(\omega)]R/c$, therefore

[2] This assumes the frequency is located away from a region of strong absorption in the material and that the phase, v_p, and group velocity, v_g, can be taken to be equal, i.e. a region of *normal* dispersion rather than *anomalous* (Jackson, 2007).

$$\frac{\omega\Re[n(\omega)]R}{c} << 1, \quad \text{Condition 2.}$$

Combining these two inequalities gives $\Re[n(\omega)]\Im[n(\omega)]\frac{R^2\omega^2}{c^2} = \Im[\epsilon(\omega)]\frac{R^2\omega^2}{2c^2} << 1$, giving the final inequality statement for the quasi-static limit:

$$R << \frac{\sqrt{2}c}{\omega\sqrt{\Im[\epsilon(\omega)]}}. \tag{D.1}$$

The legitimacy of this expression is rather suspect due to our assumption that $v_g = v_p$, it shouldn't really be used for metals near the plasma frequency where significant dispersion leads to a substantial difference between the two velocities.

E Derivation Of The LSP Resonance In The Quasistatic Approximation

This derivation is a well-known and given in many textbooks[1], we include it here for completeness. Given the spherical symmetry, it is sensible to expand the potentials inside and outside of a homogeneous and isotropic sphere of radius R in spherical harmonics

$$\phi_i(r,\theta,\phi) = \sum_{l=1}^{\infty}\sum_{m=-l}^{+l} Y_{l,m}(\theta,\phi) f_i(r),$$

where $Y_{l,m}(\theta,\phi)$ is a spherical harmonic function, the coordinate axis is chosen to be centred on the middle of the sphere and the region label is given by i. Our task is to find the radial functions, $f_i(r)$, and, from the boundary conditions connecting the potential in the two regions, find the LSP resonance condition. The form of the radial function is well-known and can be found by subbing the above expansion into the Laplace equation and solving the resulting radial differential equation (Jackson, 2007)

$$f_i(r) = A_i r^l + \frac{B_i}{r^{l+1}}, \tag{E.1}$$

where A_i and B_i are constant in each region determined by boundary conditions. By demanding that the solution is well behaved at $r = 0$ and $r = \infty$, we can conclude

$$\begin{aligned} f_{in}(r) &= A r^l \\ f_{out} &= f_{ext} + \frac{B}{r^{l+1}}, \end{aligned} \tag{E.2}$$

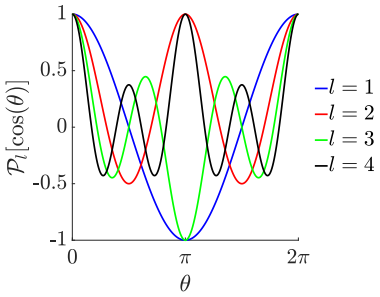
where we have introduced f_{ext} defined in an expansion of the external potential, $\phi_{ext} = \sum_{l=1}^{\infty}\sum_{m=-l}^{+l} Y_{l,m}(\theta,\phi) f_{ext}(r)$.

We now consider a general multipolar excitation of the form[2]

$$\phi_{ext} = -E_0 r^l \mathcal{P}_l[\cos(\theta)], \tag{E.3}$$

for $l = 1$ this is just a planewave excitation directed along the z direction. Larger l can be considered to be terms in an expansion of a general excitation field, in reality one would expand in spherical harmonics but we ignore the azimuthal angle here for simplicity. Enforcing boundary continuity of the tangential electric field

[1] At least for planewave excitation case, see (Jackson, 2007; Maier, 2007; Novotny and Hecht, 2012), amongst others.



The first 4 Legendre functions.

[2] The Legendre polynomials $\mathcal{P}_l$ form a complete set on a finite closed interval $\{-1, 1\}$.

component and normal electric displacement at $r = R$ gives the following conditions on the expansion coefficients (taking $\epsilon_{in} = \epsilon(\omega)$ and $\epsilon_{out} = \epsilon_b$)

$$\begin{aligned} B_l(\omega) &= E_0 R^{2l+1} \frac{l(\epsilon(\omega) - \epsilon_b)}{l\epsilon(\omega) + (l+1)\epsilon_b} = E_0 \beta_{out,l}(\omega) R^{2l+1} \\ A_l(\omega) &= -E_0 \frac{(2l+1)\epsilon_b}{l\epsilon(\omega) + (l+1)\epsilon_b} = -E_0 \beta_{in,l}(\omega), \end{aligned} \tag{E.4}$$

where I have introduced two spectral functions $\beta_{out,l}(\omega)$ and $\beta_{in,l}(\omega)$. The expansion coefficients can immediately be used to obtain the *total* potential inside and outside the nanoparticle:

$$\begin{aligned} \phi_{in}(\mathbf{r}, \omega) &= -E_0 \beta_{in,l}(\omega) r^l \mathcal{P}_l[\cos(\theta)] \\ \phi_{out}(\mathbf{r}, \omega) &= \underbrace{-E_0 r^l \mathcal{P}_l[\cos(\theta)]}_{\text{External Field}} + \underbrace{E_0 \beta_{out,l}(\omega) \frac{R^{2l+1}}{r^{l+1}} \mathcal{P}_l[\cos(\theta)]}_{\text{Induced Field}}. \end{aligned} \tag{E.5}$$

The denominator of the spectral functions gives the resonance condition for an SP mode in a sphere within the electrostatic approximation

$$-\frac{\epsilon(\omega)}{\epsilon_b} = \frac{l+1}{l}. \tag{E.6}$$

For the particular case of planewave excitation, where only the dipole mode is excited ($l = 1$), we get the textbook result:

$$\begin{aligned} \phi_{in}(\mathbf{r}, \omega) &= -E_0 \frac{3\epsilon_b}{\epsilon(\omega) + 2\epsilon_b} r\cos(\theta) \\ \phi_{out}(\mathbf{r}, \omega) &= -E_0 r\cos(\theta) + E_0 \left(\frac{\epsilon(\omega) - \epsilon_b}{\epsilon(\omega) + 2\epsilon_b} \right) \frac{R^3}{r^2} \cos(\theta). \end{aligned} \tag{E.7}$$

The electric field can be found as follows in spherical coordinates (from now on the spatial and frequency dependence is implied)

$$\mathbf{E}_i = -\nabla \phi_i = -\partial_r(\phi_i)\hat{\mathbf{r}} - \frac{1}{r}\partial_\theta(\phi_i)\hat{\boldsymbol{\theta}}, \tag{E.8}$$

and we will need the following identity for the derivative of a Legendre polynomial

$$\begin{aligned} \partial_\theta \mathcal{P}_l[\cos(\theta)] &= -\sin(\theta) \frac{d}{d\cos(\theta)} \mathcal{P}_l[\cos(\theta)] \\ &= \frac{l+1}{\sin(\theta)} \left[\mathcal{P}_{l+1}[\cos(\theta)] - \cos(\theta)\mathcal{P}_l[\cos(\theta)] \right]. \end{aligned} \tag{E.9}$$

We find

$$\begin{aligned} \mathbf{E}_{in} &= E_0 \beta_{in,l} r^{l-1} \left[l\mathcal{P}_l[\cos(\theta)]\hat{\mathbf{r}} + \partial_\theta \mathcal{P}_l[\cos(\theta)]\hat{\boldsymbol{\theta}} \right] \\ \mathbf{E}_{out} &= E_0 r^{l-1} \left[l\mathcal{P}_l[\cos(\theta)]\hat{\mathbf{r}} + \partial_\theta \mathcal{P}_l[\cos(\theta)]\hat{\boldsymbol{\theta}} \right] \\ &+ E_0 \beta_{out,l} \frac{R^{2l+1}}{r^{l+2}} \left[(l+1)\mathcal{P}_l[\cos(\theta)]\hat{\mathbf{r}} - \partial_\theta \mathcal{P}_l[\cos(\theta)]\hat{\boldsymbol{\theta}} \right]. \end{aligned} \tag{E.10}$$

These expressions simplify considerably in the case of planewave

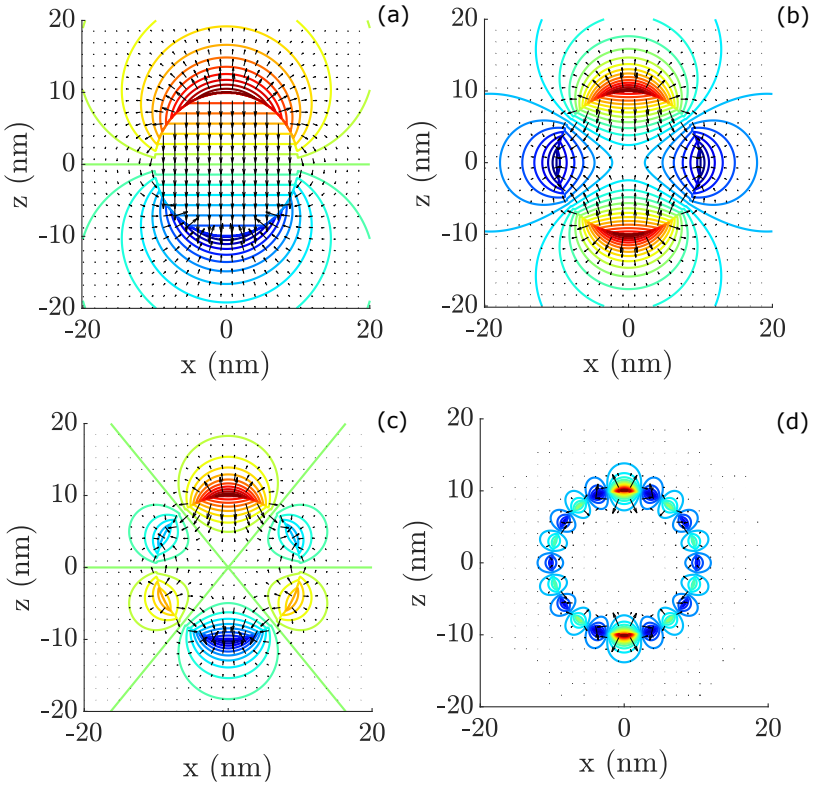


Figure E.1: The induced electric field vector plot superimposed on a contour plot of the electric potential (real part), for the (a) dipole, (b) quadrupole, (c) octupole and (d) $l = 10$ resonance of a 10 nm radius silver nanoparticle described by the dielectric function from (Johnson and Christy, 1972).

excitation

$$\begin{aligned} \mathbf{E}_{in} &= E_0 \beta_{in,1} \left(\cos(\theta)\hat{\mathbf{r}} - \sin(\theta)\hat{\boldsymbol{\theta}}\right) = \mathbf{E}_0 \beta_{in,1} \\ \mathbf{E}_{out} &= \underbrace{E_0 \left(\cos(\theta)\hat{\mathbf{r}} - \sin(\theta)\hat{\boldsymbol{\theta}}\right)}_{\mathbf{E}_0} + E_0 \beta_{out,1} \frac{R^3}{r^3} \left(2\cos(\theta)\hat{\mathbf{r}} + \sin(\theta)\hat{\boldsymbol{\theta}}\right), \end{aligned} \tag{E.11}$$

where I have used $\cos(\theta)\hat{\mathbf{r}} - \sin(\theta)\hat{\boldsymbol{\theta}} = \hat{\mathbf{z}}$. It turns out that the outside electric field can be written in terms of the field of a dipole centred at $r = 0$, this may not be so clear from the expression above. The clue is given by the form of the external potential, by writing it in terms of the dipole moment of the sphere $p = \alpha E_0 = 4\pi\epsilon_0\epsilon_b R^3 \left(\frac{\epsilon(\omega)-\epsilon_b}{\epsilon(\omega)+2\epsilon_b}\right) E_0$, meaning

$$\phi_{out} = \phi_{ext} + \frac{\mathbf{p}\cdot\hat{\mathbf{r}}}{4\pi\epsilon_0\epsilon_b r^2} = \phi_{ext} + \frac{\mathbf{p}\cdot\mathbf{r}}{4\pi\epsilon_0\epsilon_b r^3}, \tag{E.12}$$

this is the well-known dipole potential valid for any orientation and $\mathbf{r}$ is the radial vector. The gradient can be found as follows

$$\begin{aligned} \nabla\left(\frac{\mathbf{p}\cdot\mathbf{r}}{r^3}\right) &= \mathbf{p}\cdot\mathbf{r}\nabla\left(\frac{1}{r^3}\right) + \frac{\nabla(\mathbf{p}\cdot\mathbf{r})}{r^3} \\ &= \frac{\mathbf{p} - 3\hat{\mathbf{r}}(\mathbf{p}\cdot\hat{\mathbf{r}})}{r^3}, \end{aligned} \tag{E.13}$$

where I have used $\nabla\left(\frac{1}{r^3}\right) = -\frac{3}{r^4}\nabla(\mathbf{r}) = -\frac{\hat{\mathbf{r}}}{r^4}$, which gives us equations 1.32.

The surface charge which leads to these induced fields is given by the discontinuity in the normal component of the polarization $\sigma = -\hat{\mathbf{n}} \cdot (\mathbf{P}_{in} - \mathbf{P}_{out})$ and so can be found from the radial components of the electric field. For simplicity we set $\epsilon_b = 1$ so that there is no polarization outside of the sphere ($\chi_b = 0$), in which case

$$\begin{aligned} \sigma &= -\hat{\mathbf{n}} \cdot \mathbf{P}_{in}|_{r=R} = -\epsilon_0(\epsilon(\omega) - 1)E_{in,r}|_{r=R} \\ &= \epsilon_0 E_0 \frac{l(2l+1)(1-\epsilon(\omega))}{l\epsilon(\omega) + (l+1)} R^{l-1} \mathcal{P}_l[\cos(\theta)]. \end{aligned} \tag{E.14}$$

The gradient of the electric field is an interesting quantity to look at. As it is a gradient of a vector field, it is a tensor quantity

$$\nabla \mathbf{E}_{ij} = \partial_{x_i} E_j = -\frac{\partial^2 \phi}{\partial_{x_i} \partial_{x_j}}. \tag{E.15}$$

The tensor, in Cartesian coordinates, will clearly be traceless as Laplace's equation holds. An interesting component to consider is the change in the radial component in the radial direction $\nabla \mathbf{E}_{rr}$. For $\theta = 0$ this is along the direction of maximum FE. It can be written as follows

$$\partial_r E_{out,r} = -E_{max,l}(l+2)\mathcal{P}_l[\cos(\theta)] \frac{R^{2l+1}}{r^{l+3}}, \tag{E.16}$$

where $E_{max,l}$ is the maximum field (at $\theta = 0$). For planewave excitation we recover the result from reference (Alabastri et al., 2016)

$$\partial_r E_{out,r}|_{l=1} = -\frac{3E_{max,l}\cos(\theta)}{r} \left(\frac{R}{r}\right)^3. \tag{E.17}$$

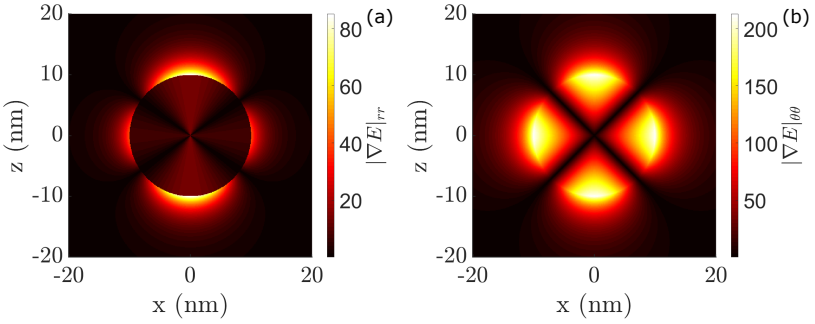


Figure E.2: The electric field gradient in the (a) radial and (b) azimuthal direction for the quadrupole mode at resonance of a 10 nm radius silver nanoparticle described by the dielectric function from (Johnson and Christy, 1972).

F Mie Theory Derivation

Like the quasistatic derivation presented in the preceding appendix, there are many textbook derivations of Mie theory; the most famous probably being from (Bohren and Huffman, 2008). I repeat it here as there are a myriad of different notations around and so it is nice to see the full derivation in one place. Our task is to solve the Helmholtz equation for a spherical region described by dielectric function ϵ_{in} and the space outside described by ϵ_{out}, the resulting expressions will be valid for *any* sphere size. Given the spherical symmetry of the problem, one should expand the electric and magnetic field in terms of VSHs: $\boldsymbol{\mathcal{M}}(r,\theta,\phi)$ and $\boldsymbol{\mathcal{N}}(r,\theta,\phi)$. If the dielectric function commutes with the angular momentum operator, i.e. it a radial function, we can gain a bucketload of insight and elegance by making use of an operator formalism; completely analogous to what is done is quantum mechanics[1]. We define the angular momentum operator as follows:

$$\hat{\mathbf{L}} = \frac{\mathbf{r}\times\nabla}{i} = \mathbf{r}\times\hat{\mathbf{p}} = \frac{1}{i}\left[-\frac{\hat{\boldsymbol{\theta}}}{\sin(\theta)}\partial_\phi + \hat{\boldsymbol{\phi}}\partial_\theta\right], \tag{F.1}$$

where $\hat{\mathbf{p}} = -i\nabla$ is the *linear* momentum operator. The fundamental theorem of vector calculus, also known as the Helmholtz decomposition/theorem, states that any smooth vector field which decays to zero at infinity can be decomposed into an irrotational and solenoidal fields; this will apply to all physically meaningful fields. Mathematically, this can be stated as $\mathbf{V} = -\nabla\phi + \nabla\times\mathbf{A}$. We made use of this in section 1.1 when discussing the two classes of solutions to the wave equation; longitudinal and transverse. A non-obvious consequence of the Helmholtz theorem is that *any* vector field can be expanded in terms of $\hat{\mathbf{L}}$

$$\mathbf{V} = \nabla\psi_1 + \hat{\mathbf{L}}\psi_2 + \frac{\nabla\times\hat{\mathbf{L}}}{i}\psi_3, \tag{F.2}$$

the derivation is not presented here and can be found elsewhere (Low, 2008). The angular momentum operator is useful as the spherical harmonics are an eigenfunction of its square

$$\begin{aligned}\hat{\mathbf{L}}^2 Y_{lm}(\theta,\phi) &= \hat{\mathbf{L}}\cdot\hat{\mathbf{L}}Y_{lm}(\theta,\phi)\\ &= -\left[\frac{1}{\sin(\theta)}\partial_\theta\left(\sin(\theta)\partial_\theta\right) + \frac{1}{\sin^2(\theta)}\partial_\phi^2\right] = l(l+1)Y_{lm}(\theta,\phi),\end{aligned} \tag{F.3}$$

[1] It must stressed that even though the mathematics is very similar to the quantum mechanical calculation of a central potential problem (Merzbacher, 1970), this is a completely classical derivation.

where the spherical harmonics are defined as

$$Y_{lm}(\theta,\phi) = \underbrace{\sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}}_{C_{lm}} P_{lm}[\cos(\theta)]e^{im\phi}, \tag{F.4}$$

which has the correct normalisation to ensure $\int d\Omega\, Y_{l_1,m_1}(\theta,\phi)Y_{l_2,m_2}(\theta,\phi) = \delta_{l_1,l_2}\delta_{m_1,m_2}$, where $d\Omega = d\phi d\theta \sin(\theta)$. Note that P_{lm} is an *associated* Legendre function and is distinct from the Legendre polynomials $\mathcal{P}_l$ that we used in the previous Appendix[2].

[2] Many authors use the same symbol P for both Legendre polynomials, I find this can be quite confusing. The two are related (for non-negative m) via $P_{lm}(x) = (1-x^2)^{m/2}\frac{d^m\mathcal{P}_l(x)}{dx^m}$. For $m = 0$ the two are equivalent $P_{l,m=0} = \mathcal{P}_l$. For $m < 0$ we can use the formula $P_{l,-m}(x) = (-1)^m\frac{(l-m)!}{(l+m)!}P_{lm}(x)$.

We will write the electric field in the form suggested by equation F.2

$$\mathbf{E} = \nabla\psi^{(L)} + \hat{\mathbf{L}}\psi^{(M)} + \frac{1}{ik}\nabla\times\hat{\mathbf{L}}\psi^{(E)}, \tag{F.5}$$

where three scalar functions are introduced: a longitudinal $\psi^{(L)}$, magnetic $\psi^{(M)}$ and electric $\psi^{(E)}$. One can deduce

$$\begin{aligned}\nabla\cdot\mathbf{D} = 0 &= \nabla\cdot\left[\epsilon(\mathbf{r})\left(\nabla\psi^{(L)} + \hat{\mathbf{L}}\psi^{(M)} + \frac{\nabla\times\hat{\mathbf{L}}}{ik}\psi^{(E)}\right)\right]\\ &= \epsilon(r)\nabla^2\psi^{(L)} + \frac{\partial_r\epsilon(r)}{r}\left(\mathbf{r}\cdot\nabla\psi^{(L)} + \frac{\mathbf{r}\cdot\nabla\times\hat{\mathbf{L}}}{ik}\psi^{(E)}\right)\\ &= \epsilon(r)\nabla^2\psi^{(L)} + \frac{\partial_r\epsilon(r)}{r}\left(\mathbf{r}\cdot\nabla\psi^{(L)} + \frac{\hat{\mathbf{L}}^2}{k}\psi^{(E)}\right),\end{aligned} \tag{F.6}$$

where I have assumed that the dielectric function is isotropic and hence $\nabla\epsilon(r) = \partial_r\epsilon(r)\hat{\mathbf{r}}$ and have used the identity $\hat{\mathbf{r}}\cdot\nabla\times\hat{\mathbf{L}} = -\hat{\mathbf{L}}^2/(ir)$ which can be simply proven by writing out the curl explicitly and comparing to equation F.3. This equation reveals that the longitudinal and electric scalar functions are coupled depending on the spatial dependence of the dielectric function[3]. If we simplify matters for now by taking the dielectric function to be a step function, as we did in the previous appendix, then we can conclude

[3] This has obvious interesting implications if one considers electron spill-out, which could be described via a spatially dependent permittivity.

$$\nabla^2\psi^{(L)} = 0, \tag{F.7}$$

in each section which can be solved to give:

$$\psi^{(L)} = \frac{1}{4\pi\epsilon_0}\int d^3r'\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \tag{F.8}$$

This reveals the longitudinal scalar function, in the case of piecewise dielectric functions, is the instantaneous coulomb potential of the free charge and so we are free to set it to zero. This term will be nonzero in the context of nonlocal Mie theory (Ruppin, 1973, 1975), see section 2.2.2. It is possible to deduce that both the electric and magnetic scalar fields obey the *scalar* Helmholtz equation, from calculating the projection of $\hat{\mathbf{L}}$ and $\hat{\mathbf{L}}\times\nabla$ on the *vector* Helmholtz equation for the electric field. This is fairly simple although tedious

and the details can be found in (Low, 2008)

$$\nabla^2 \begin{bmatrix} \psi^{(M)} \\ \psi^{(E)} \end{bmatrix} + k^2 \begin{bmatrix} \psi^{(M)} \\ \psi^{(E)} \end{bmatrix} = 0. \tag{F.9}$$

This is a very useful result; it means the difficult task of solving a differential equation for vector quantities can be repackaged as a scalar differential equation. What's more, for a spherical system the solution is well-known and the scalar functions can be expanded in a series of spherical Bessel functions and spherical harmonics

$$\psi^{(\nu)} = \sum_{l=1}^{+\infty} \sum_{m=-l}^{l} \mathfrak{C}_{lm}^{(\nu)} \psi_{lm}^{(\nu)}(r,\theta,\phi) = \sum_{l=1}^{+\infty} \sum_{m=-l}^{l} \mathfrak{C}_{lm}^{(\nu)} z_l(kr) Y_{lm}(\theta,\phi), \tag{F.10}$$

where $\nu = \{M, E\}$, $\mathfrak{C}_{lm}^{(\nu)}$ are expansion coefficients to be determined and $z_l(kr)$ represents one of the four types of spherical Bessel function[4]. We can therefore write the VSH expansion of the electric field, using equation F.5, as

$$\begin{aligned} \mathbf{E} &= \sum_{lm} \mathfrak{C}_{lm}^{(M)} \overbrace{z_l(kr)\hat{\mathbf{L}} Y_{lm}(\theta,\phi)}^{i\boldsymbol{\mathcal{M}}_{lm}} + \mathfrak{C}_{lm}^{(E)} \overbrace{z_l(kr) \frac{\nabla \times \hat{\mathbf{L}} Y_{lm}(\theta,\phi)}{ik}}^{\boldsymbol{\mathcal{N}}_{lm}} \\ &= \sum_{lm} a_{lm} \boldsymbol{\mathcal{M}}_{lm}(r,\theta,\phi) + b_{lm} \boldsymbol{\mathcal{N}}_{lm}(r,\theta,\phi). \end{aligned} \tag{F.11}$$

both inside and outside the sphere, where

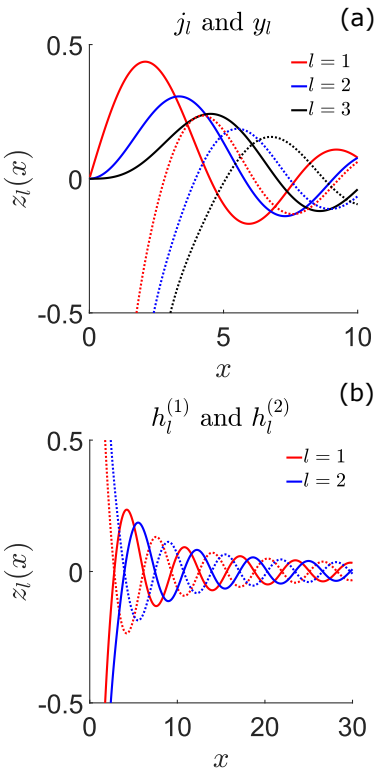
$$\begin{aligned} \boldsymbol{\mathcal{M}}_{lm}(r,\theta,\phi) &= \nabla \times (\mathbf{r}\psi_{lm}) \\ \boldsymbol{\mathcal{N}}_{lm}(r,\theta,\phi) = \frac{\nabla \times \boldsymbol{\mathcal{M}}_{lm}(r,\theta,\phi)}{k} &= \frac{\nabla \times \nabla \times (\mathbf{r}\psi_{lm})}{k}. \end{aligned} \tag{F.12}$$

and we define the expansion coefficients $a_{lm} = i\mathfrak{C}_{lm}^{(M)}$ and $b_{lm} = \mathfrak{C}_{lm}^{(E)}$. We have taken the long and scenic route to arrive at the VSHs introduced by (Bohren and Huffman, 2008) at the start of their derivation. They call $\psi(r,\theta,\phi)$ the *generating function* and the vector $\mathbf{r}$ is chosen as the *pilot* vector to assure that $\boldsymbol{\mathcal{M}}$ is everywhere tangential on a spherical surface $|\mathbf{r}| = const$. They are both zero for $l = 0$. For reference, the components of the VSHs can be written as follows

$$\begin{aligned} \boldsymbol{\mathcal{M}}_{lm}(\rho,\theta,\phi) &= 0\hat{\mathbf{r}} + ime^{im\phi} \frac{C_{lm} P_{lm}}{\sin(\theta)} z_l^{(i)} \hat{\boldsymbol{\theta}} - e^{im\phi} C_{lm} (\partial_\theta P_{lm}) z_l^{(i)} \hat{\boldsymbol{\phi}} \\ \boldsymbol{\mathcal{N}}_{lm}(\rho,\theta,\phi) &= l(l+1) e^{im\phi} C_{lm} P_{lm} \frac{z_l^{(i)}}{\rho} \hat{\mathbf{r}} + e^{im\phi} C_{lm} (\partial_\theta P_{lm}) \frac{\partial_\rho \rho z_l^{(i)}}{\rho} \hat{\boldsymbol{\theta}} \\ &\quad + ime^{im\phi} C_{lm} P_{lm} \frac{\partial_\rho \rho z_l^{(i)}}{\rho} \hat{\boldsymbol{\phi}}, \end{aligned} \tag{F.13}$$

where $\rho = kr$. By taking the real and imaginary part of these expressions it is possible to get the even and odd functions solutions used in (Bohren and Huffman, 2008). Now we have quite rigor-

[4] Although there are four types of spherical Bessel functions, we will need only two: the spherical Bessel function of the first kind j_l, which is finite at the origin as opposed to the spherical Bessel function of the second kind y_l, and the spherical Hankel function of the first kind $h_l^{(1)} = j_l + iy_l$ which gives the correct radial behaviour at infinity of an *outgoing* spherical wave. The second spherical Hankel function $h_l^{(2)} = j_l - iy_l$ represents an *incoming* wave at infinity.



(a) The first three spherical Bessel functions of the 1st (solid line) and 2nd (dashed line) type. (b) The first two spherical Hankel functions of the 1st (solid line) and 2nd (dashed line) kind.

[5] For the normalisation condition of $\boldsymbol{\mathcal{N}}$ I have used the identity $(\mathbf{A}\times\mathbf{B})\cdot(\mathbf{C}\times\mathbf{D}) = (\mathbf{A}\cdot\mathbf{C})(\mathbf{B}\cdot\mathbf{D}) - (\mathbf{B}\cdot\mathbf{C})(\mathbf{A}\cdot\mathbf{D})$.

[6] The ladder operators $\hat{L}_+ = \hat{L}_x + i\hat{L}_y$ and $\hat{L}_- = \hat{L}_x - i\hat{L}_y$ are introduced, they are well-known from quantum mechanics (Merzbacher, 1970).

ously setup the problem, we proceed to the substance of the Mie theory derivation. We need to specify the excitation field and hence the field inside and outside the sphere, as well as the continuity conditions of the fields at the sphere surface.

We begin by specifying the external field. The VSHs are the EM *normal modes* of a sphere, in general a scattered field will be a superposition of normal modes with a weighting give by a pair of coefficients a_{lm} and b_{lm} (Bohren and Huffman, 2008). The coefficients will depend on the form of the exciting field, we will consider a $\hat{\mathbf{x}}$ polarized planewave excitation with the propagation axis aligned along $\hat{\mathbf{z}}$, which allows us to use the following useful identity (Jackson, 2007):

$$\begin{aligned}\mathbf{E}_{ext}(\mathbf{r}) &= E_0 e^{ikz}\hat{\mathbf{x}} = E_0 e^{ikr\cos(\theta)}\hat{\mathbf{x}}\\ &= \hat{\mathbf{x}}E_0\sum_l i^l(2l+1)j_l(kr)\mathcal{P}_l(\theta)\\ &= \hat{\mathbf{x}}E_0\sum_l i^l\sqrt{4\pi(2l+1)}j_l(kr)Y_{l,0}(\theta).\end{aligned} \tag{F.14}$$

To calculate the expansion of the incident field we project the above equation onto $\boldsymbol{\mathcal{M}}_{lm}$ and $\boldsymbol{\mathcal{N}}_{lm}$, this is most elegantly done by moving into a more abstract operator formalism (Cambiasso, 2017); we define $|\boldsymbol{\mathcal{M}}_{k,lm}\rangle = \hat{\mathbf{L}}\,|k,lm\rangle$ and $\psi_{lm}(k\mathbf{r}) = \langle\mathbf{r}|k,lm\rangle$. We begin with the ortho-normalisation conditions[5]

$$\begin{aligned}\langle\boldsymbol{\mathcal{M}}_{k',l'm'}|\boldsymbol{\mathcal{M}}_{k,lm}\rangle &= \int d\Omega\,\left(\boldsymbol{\mathcal{M}}^*_{k',l'm'}\cdot\boldsymbol{\mathcal{M}}_{k,lm}\right)\\ &= \langle k',l'm'|\hat{\mathbf{L}}^\dagger\cdot\hat{\mathbf{L}}|k,lm\rangle = l(l+1)\delta_{l'm',lm}\,\langle k',k\rangle\\ \langle\boldsymbol{\mathcal{N}}_{k',l'm'}|\boldsymbol{\mathcal{N}}_{k,lm}\rangle &= \frac{1}{k'k}\,\langle k',l'm'|\left(\nabla\times\hat{\mathbf{L}}\right)^\dagger\cdot\left(\nabla\times\hat{\mathbf{L}}\right)|k,lm\rangle\\ &= \frac{1}{k'k}\,\langle k',l'm'|\left(\hat{\mathbf{p}}^\dagger\cdot\hat{\mathbf{p}}\right)\left(\hat{\mathbf{L}}^\dagger\cdot\hat{\mathbf{L}}\right) - \left(\hat{\mathbf{L}}\cdot\hat{\mathbf{p}}\right)\underbrace{\left(\hat{\mathbf{p}}\cdot\hat{\mathbf{L}}\right)}_{=0}|k,lm\rangle\\ &= \frac{1}{k'k}\,\langle k',l'm'|\hat{\mathbf{p}}^2\hat{\mathbf{L}}^2|k,lm\rangle = \frac{k'}{k}l(l+1)\delta_{l'm',lm}\,\langle k',k\rangle\\ \langle\boldsymbol{\mathcal{M}}_{k',l'm'}|\boldsymbol{\mathcal{N}}_{k,lm}\rangle &= \frac{1}{k'k}\,\langle k',l'm'|\underbrace{\hat{\mathbf{L}}^\dagger\cdot\left(\nabla\times\hat{\mathbf{L}}\right)}_{=0}|k,lm\rangle = 0,\end{aligned} \tag{F.15}$$

$\langle k',k\rangle$ is to be understood as $z^{(1)*}_{l'}(k')z^{(2)}_l(k)$ where $z^{(1)}_l$ and $z^{(2)}_l$ represent Bessel functions which need not be the same type. We can thus write the expansion coefficients of the planewave expansion as follows[6]

$$\begin{aligned}\frac{\langle\boldsymbol{\mathcal{M}}_{k,lm}|\mathbf{E}_{ext}\rangle}{\langle\boldsymbol{\mathcal{M}}_{k,lm}|\boldsymbol{\mathcal{M}}_{k,lm}\rangle} &= E_0\sum_{l'=1}^{\infty}\frac{i^{l'}\sqrt{4\pi(2l'+1)}}{l(l+1)|j_l(kr)|^2}\,\langle k,lm|\overbrace{\hat{\mathbf{L}}\cdot\hat{\mathbf{x}}}^{=\hat{L}_x}|k,l'0\rangle\\ &= E_0\sum_{l'=1}^{\infty}\frac{\langle k|k\rangle\,i^{l'}\sqrt{4\pi(2l'+1)}}{l(l+1)|j_l(kr)|^2}\,\langle lm|\frac{\hat{L}_+ + \hat{L}_-}{2}|l'0\rangle\\ &= \frac{E_0 i^l}{2}\frac{j_l(kr)}{z_l(kr)}\sqrt{\frac{4\pi(2l+1}{l(l+1)}}\left(\delta_{m,+1}+\delta_{m,-1}\right).\end{aligned} \tag{F.16}$$

The b_{lm} coefficient could be found in a similar way using $\boldsymbol{\mathcal{N}}_{k,lm}$ projected onto the electric field $b_{lm} = \frac{\langle \boldsymbol{\mathcal{N}}_{k,lm} | \mathbf{E}_{ext} \rangle}{\langle \boldsymbol{\mathcal{N}}_{k,lm} | \boldsymbol{\mathcal{N}}_{k,lm} \rangle}$, but actually it is easier to use the magnetic field. The trick is to use the cyclic property of the VSHs under the curl operator (Bohren and Huffman, 2008)

$$\boldsymbol{\mathcal{M}}_{lm} = \frac{\nabla \times \boldsymbol{\mathcal{N}}_{k,lm}}{k}, \tag{F.17}$$

to write

$$\mathbf{B}(\mathbf{r}, \omega) = \frac{\nabla \times \mathbf{E}}{i\omega} = \frac{k}{i\omega} \sum_{lm} a_{lm} \boldsymbol{\mathcal{N}}_{k,lm} + b_{lm} \boldsymbol{\mathcal{M}}_{k,lm}, \tag{F.18}$$

and and the planewave expansion for the magnetic field is

$$\mathbf{B}_{ext}(\mathbf{r}) = \frac{E_0}{i\omega} \nabla \times \left(e^{ikz} \hat{x} \right) = \frac{E_0 k}{\omega} e^{ikz} \hat{\mathbf{y}} = \frac{E_0 k}{\omega} \hat{y} \sum_l i^l \sqrt{4\pi(2l+1)} j_l(kr) Y_{l,0}(\theta). \tag{F.19}$$

so we can solve for the b_{lm} coefficient

$$\begin{aligned} b_{lm} = \frac{i\omega}{k} \frac{\langle \boldsymbol{\mathcal{N}}_{k,lm} | \mathbf{B}_{ext} \rangle}{\langle \boldsymbol{\mathcal{M}}_{k,lm} | \boldsymbol{\mathcal{M}}_{k,lm} \rangle} &= i E_0 \sum_{l'=1}^{\infty} \frac{i^{l'} \sqrt{4\pi(2l'+1)}}{l(l+1)|z_l(kr)|^2} \langle k, lm | \overbrace{\hat{\mathbf{L}} \cdot \hat{\mathbf{y}}}^{=\hat{L}_y} | k, l'0 \rangle \\ &= i E_0 \sum_{l'=1}^{\infty} \frac{\langle k|k \rangle \, i^{l'} \sqrt{4\pi(2l'+1)}}{l(l+1)|j_l(kr)|^2} \langle lm | \frac{\hat{L}_+ - \hat{L}_-}{2i} | l'0 \rangle \\ &= \frac{E_0 i^l}{2} \frac{j_l(kr)}{z_l(kr)} \sqrt{\frac{4\pi(2l+1}{l(l+1)}} \left(\delta_{m,+1} - \delta_{m,-1} \right). \end{aligned} \tag{F.20}$$

The appearance of the $m = \pm 1$ components in the expansion is a consequence of linearly polarized planewaves being a superposition of two circularly polarized waves. For light fields with more complicated spatial structures we can expect contributions from other m components. We can therefore write the electric field of a planewave as

$$\mathbf{E}_{ext} = \sum_l i^l \sqrt{\frac{4\pi(2l+1}{l(l+1)}} \left[\left(\frac{\boldsymbol{\mathcal{M}}_{l,+1} + \boldsymbol{\mathcal{M}}_{l,-1}}{2} \right) + \left(\frac{\boldsymbol{\mathcal{N}}_{l,+1} - \boldsymbol{\mathcal{N}}_{l,-1}}{2} \right) \right]. \tag{F.21}$$

This can be written in terms of the odd and even VSHs used by Bohren and Huffman using $\boldsymbol{\mathcal{M}}_{l,+1} + \boldsymbol{\mathcal{M}}_{l,-1} = 2i\boldsymbol{\mathcal{M}}^{odd}_{l,+1}$ and $\boldsymbol{\mathcal{N}}_{l,+1} - \boldsymbol{\mathcal{N}}_{l,-1} = 2\boldsymbol{\mathcal{N}}^{even}_{l,+1}$ to give

$$\mathbf{E}_{ext} = \sum_l E_l \left(\boldsymbol{\mathcal{M}}^{odd}_{l,+1} - i\boldsymbol{\mathcal{N}}^{even}_{l,+1} \right), \tag{F.22}$$

where I define $E_l = i^{l+1} \sqrt{\frac{4\pi(2l+1}{l(l+1)}}$. This is, apart from a difference in normalisation constant, the same result as Bohren and Huffman (Bohren and Huffman, 2008). Compare how concise this derivation is compared to the arduous journey through an agglomeration of

identities presented in the aforementioned reference.

Now we have the form of the external field in a suitable representation, we need the boundary conditions at the sphere boundary to put the final piece in the puzzle. First we need to write the fields in each of the two dielectric regions, we will do this in terms of the scalar functions

$$\begin{aligned}
\begin{bmatrix}\psi^{(M,ext)}\\ \psi^{(E,ext)}\end{bmatrix} &= \sum_l \begin{bmatrix}\mathfrak{C}^{(M,ext)}_{l,\pm1}\\ \mathfrak{C}^{(E,ext)}_{l,\pm1}\end{bmatrix} j_l(k_{out}r)Y_{l,\pm1}(\theta,\phi) \quad \text{External field}\\
\begin{bmatrix}\psi^{(M,int)}\\ \psi^{(E,int)}\end{bmatrix} &= \sum_l \begin{bmatrix}\mathfrak{C}^{(M,int)}_{l,\pm1}\\ \mathfrak{C}^{(E,int)}_{l,\pm1}\end{bmatrix} j_l(k_{in}r)Y_{l,\pm1}(\theta,\phi) \quad \text{Internal field}\\
\begin{bmatrix}\psi^{(M,sca)}\\ \psi^{(E,sca)}\end{bmatrix} &= \sum_l \begin{bmatrix}\mathfrak{C}^{(M,sca)}_{l,\pm1}\\ \mathfrak{C}^{(E,sca)}_{l,\pm1}\end{bmatrix} h_l(k_{out}r)Y_{l,\pm1}(\theta,\phi) \quad \text{Scattered field.}
\end{aligned} \tag{F.23}$$

The boundary conditions, written in terms of the scalar functions, are (Low, 2008; David and García de Abajo, 2011)

$$\begin{aligned}
\psi^{(M,int)}(R) &= \psi^{(M,ext)}(R) + \psi^{(M,sca)}(R)\\
\partial_r(r\psi^{(M,int)}(r))|_R &= \partial_r(r\psi^{(M,ext)}(r))|_R + \partial_r(r\psi^{(M,sca)}(r))|_R\\
\epsilon_{in}\psi^{(E,int)}(R) &= \epsilon_{out}\left(\psi^{(E,ext)}(R) + \psi^{(E,sca)}(R)\right)\\
\partial_r(r\psi^{(E,int)}(r))|_R &= \partial_r(r\psi^{(E,ext)}(r))|_R + \partial_r(r\psi^{(E,sca)}(r))|_R.
\end{aligned} \tag{F.24}$$

Starting with the magnetic coefficients, we can write

$$\begin{aligned}
j_l(\theta_{out}) + t_l^M h_l(\theta_{out}) &= A_l^M j_l(\theta_{in})\\
[\theta_{out}j_l(\theta_{out})]' + t_l^M[\theta_{out}h_l(\theta_{out})]' &= A_l^M[\theta_{in}j_l(\theta_{in})]',
\end{aligned} \tag{F.25}$$

where $t_l^M = \mathfrak{C}^{(M,sca)}_{l,\pm1}/\mathfrak{C}^{(M,ext)}_{l,\pm1}$ and $A_l^M = \mathfrak{C}^{(M,int)}_{l,\pm1}/\mathfrak{C}^{(M,ext)}_{l,\pm1}$. This pair of equations can be solved to get both the scattering and internal coefficients

$$\begin{aligned}
t_l^M &= \frac{-\theta_{in}j_l(\theta_{out})j_l'(\theta_{in}) + \theta_{out}j_l(\theta_{in})j_l'(\theta_{out})}{\theta_{in}h_l(\theta_{out})j_l'(\theta_{in}) - \theta_{out}j_l(\theta_{in})h_l'(\theta_{out})}\\
A_l^M &= \frac{\theta_{out}h_l(\theta_{out})j_l'(\theta_{out}) - \theta_{out}j_l(\theta_{out})h_l'(\theta_{out})}{\theta_{in}h_l(\theta_{out})j_l'(\theta_{in}) - \theta_{out}j_l(\theta_{in})h_l'(\theta_{out})}.
\end{aligned} \tag{F.26}$$

where I have introduced the notation $\theta_\nu = k_\nu R = k_0 R\sqrt{\epsilon_\nu}$, $\nu = \{in, out\}$. Similarly, we can derive the electric scattering and internal coefficients

$$\begin{aligned}
t_l^E &= \frac{-m^2 j_l(\theta_{in})[\theta_{out}j_l(\theta_{out})]' + j_l(\theta_{out})[\theta_{in}j_l(\theta_{in})]'}{m^2 j_l(\theta_{in})[\theta_{out}h_l(\theta_{out})]' - h_l(\theta_{out})[\theta_{in}j_l(\theta_{in})]'}\\
A_l^E &= \frac{j_l(\theta_{out})[\theta_{out}h_l(\theta_{out})]' - j_l(\theta_{out})[\theta_{out}j_l(\theta_{out})]'}{m^2 j_l(\theta_{in})[\theta_{out}h_l(\theta_{out})]' - h_l(\theta_{out})[\theta_{in}j_l(\theta_{in})]'},,
\end{aligned} \tag{F.27}$$

where $m = \sqrt{\frac{\epsilon_{in}}{\epsilon_{out}}} = \frac{n_{in}}{n_{out}} = \frac{k_{in}}{k_{out}}$. Note the dominator for each pair of coefficients must be the same as it gives the resonance condition for that particular electric or magnetic mode when set to zero.

These coefficients need to be linked to physical quantities that can be measured in experiments. By considering the power flow though an imaginary sphere with a radius going to infinity[7], which contains the real sphere, it is possible to derive the following relations for the scattering and extinction cross section (Bohren and Huffman, 2008)

$$\begin{aligned} \sigma_{sca} &= \frac{2\pi}{k^2}\sum_{l=1}^{\infty}(2l+1)(|t_l^M|^2+|t_l^E|^2) \\ \sigma_{ext} &= \frac{2\pi}{k^2}\sum_{l=1}^{\infty}(2l+1)\Re(t_l^M+t_l^E). \end{aligned} \tag{F.28}$$

The Mie coefficients have also recently been used to study the ultrafast dynamics of metal and dielectric spheres driven by ultrashort Gaussian pulses, revealing a rich temporal behaviour of the LSP and magnetic resonances even in the linear regime (Lazzarini et al., 2017).

[7] So the near-field and intermediate-field terms go to zero.

G Derivation Of The Hydrodynamic Model

G.1 Derivation Of The Thomas-Fermi Pressure Term

Starting from the Thomas-Fermi function, equation 2.4, the following functional derivative identity is used (Ullrich, 2011)

$$\frac{\delta}{\delta\phi(y)}\int d^n x f(\phi(x)) = \partial_\phi f(\phi(y)), \tag{G.1}$$

to get

$$\frac{\delta F_{TF}[n]}{\delta n} = \frac{5}{3}n^{2/3}(\mathbf{r}). \tag{G.2}$$

This gives

$$-\frac{1}{m}\nabla\frac{\delta F_{TF}[n]}{\delta n} = -\frac{\hbar^2}{2m^2}(3\pi^2)^{2/3}\nabla n^{2/3}(\mathbf{r}). \tag{G.3}$$

If we assume that the density is only slightly perturbed from its ground state n_0, then we can expand $n(\mathbf{r}) = n_0 + n_1(\mathbf{r}) + ..$ and keeping to first-order gives

$$\begin{aligned} -\frac{\hbar^2}{3m^2}(3\pi^2)^{2/3}\frac{\nabla n_1(\mathbf{r})}{n_0^{1/3}} &= -\frac{\hbar^2}{3m^2}(3n_0\pi^2)^{2/3}\frac{\nabla n_1(\mathbf{r})}{n_0} \\ &= \frac{v_F^2}{3}\frac{\nabla n_1(\mathbf{r})}{n_0}, \end{aligned} \tag{G.4}$$

where I have used (Kittel, 1966) $v_F = \frac{\hbar}{m}\left(3\pi^2 n_0\right)^{1/3}$ and we can now introduce $\beta^2 = v_F^2/3$.

G.2 Derivation Of The Hydrodynamic Equations Of Motion

The Hamiltonian for the hydrodynamic electron gas + light system is (Boardman, 1982)

$$\begin{aligned} H = \frac{m}{2}\int d^3r\, n(\mathbf{r},t)\left|\nabla\psi(\mathbf{r},t) - \frac{e\mathbf{A}(\mathbf{r},t)}{m}\right|^2 \\ +\frac{1}{2}\frac{e^2}{4\pi\epsilon_0}\int\int d^3r\, d^3r'\,\frac{n(\mathbf{r},t)n(\mathbf{r}',t)}{|\mathbf{r}-\mathbf{r}'|} - \int d^3r\, u_0(\mathbf{r})n(\mathbf{r},t) + F[n(\mathbf{r},t)], \end{aligned} \tag{G.5}$$

where the velocity field is written in terms of the velocity potential, which must be gauge dependent via $\nabla\psi(\mathbf{r},t) = -\mathbf{v}(\mathbf{r},t) + e\mathbf{A}(\mathbf{r},t)/m$.

From now on I will suppress the spatial and time dependence of the macroscopic fields, unless confusion can arise. The terms can be understood as follows:

- The first term describes the hydrodynamic kinetic energy, $\frac{m}{2}\int d^3r\, n|\nabla\psi|^2$, as well as the coupling to the external light field described by a vector potential. We will work in the Coulomb gauge.

- The second term represents the classical Coulomb energy, it tells us the energy price of displacing the electron from its equilibrium density profile n_0.

- The third term is the electrostatic potential, u_0, from the neutralising positive background, normally taken to be jellium. It is negative as it represents a confining potential.

- The fourth term is the universal functional valid for any number of electrons and any external potential. It gives the internal kinetic energy of the electron gas as well as XC contributions.

The equations of motion can be found by minimisation of the action (Lundqvist and March, 2013)

$$\delta S = \delta \int dt L = \delta \int dt \int d^3r\, m(\partial_t\psi)n - H[n, \nabla\psi] \tag{G.6}$$

with respect to *both* canonically conjugate variables, n and $m\psi$, to get *two* equations of motion. We first vary the velocity potential $\psi \to \psi + \delta\psi$ to get the continuity equation

$$\partial_t n = \nabla\cdot(n\nabla\psi)\,. \tag{G.7}$$

This is obtained by two simple integration by parts to isolate the $\delta\psi$, from which it is possible to deduce that the integrand of the time integral must be zero for the action to be stationary. The variation with respect to n is even simpler and one obtains the Bernoulli equation[1]

$$\partial_t\psi = \frac{1}{2}|\nabla\psi - \frac{e\mathbf{A}}{m}|^2 + \underbrace{\frac{e^2}{4\pi\epsilon_0 m}\int d^3r' \frac{n}{|\mathbf{r}-\mathbf{r}'|}}_{=-\frac{e}{m}\phi} - \frac{u_0}{m} + \frac{1}{m}\frac{\delta F[n]}{\delta n}, \tag{G.8}$$

[1] Also sometimes known as the Euler equation.

where we have noted that the retarded potential is generated by electron-electron interactions. This equation can be brought to the more recognisable form of equations 2.3 by taking the gradient and using

$$\frac{1}{2}\nabla\left|\nabla\psi - \frac{e\mathbf{A}}{m}\right|^2 = \frac{1}{2}\nabla|\mathbf{v}|^2 = \mathbf{v}\cdot\nabla\mathbf{v} + \underbrace{\mathbf{v}\times\nabla\times\mathbf{v}}_{=\frac{e}{m}\mathbf{v}\times\mathbf{B}} \tag{G.9}$$

to get

$$\partial_t \mathbf{v} + \mathbf{v}\cdot\nabla\mathbf{v} = \underbrace{+\frac{e}{m}\mathbf{A} + \frac{e}{m}\phi}_{-\frac{e}{m}\mathbf{E}} - \frac{e}{m}\mathbf{v}\times\mathbf{B} + \frac{1}{m}\nabla u_0(\mathbf{r}) - \frac{1}{m}\frac{\delta F[n,\mathbf{v}]}{\delta n}$$
$$\therefore m\left(\partial_t + \mathbf{v}\cdot\nabla\right)\mathbf{v} = -e(\mathbf{E} + \mathbf{v}\times\mathbf{B}) + \nabla u_0(\mathbf{r}) - \nabla\frac{\delta F[n,\mathbf{v}]}{\delta n}, \tag{G.10}$$

which on adding a phenomenological damping term gives the second equation of 2.3, although now with the positive background explicitly shown. A closer look at the above equation reveals that it is an expression for the balance of forces on a fluid element where the $\nabla\frac{\delta F[n,\mathbf{v}]}{\delta n}$ term acts equivalently to the usual pressure gradient term in fluid dynamics. Note that the derivation presented here is more general that many other derivations which assume an irrotational fluid (Ying, 1974; Lundqvist and March, 2013) and so includes retardations effects.

It is often convenient to linearize the HM equations, to do this we make a small expansion around the equilibrium solution n_0 at $t = 0$:

$$\begin{aligned} n &= n_0 + n_1 \\ \mathbf{v} &= 0 + \mathbf{v}_1, \end{aligned} \tag{G.11}$$

where n_1 is the first-order change in density and we take the equilibrium velocity to be zero $\mathbf{v}_0 = 0$ at time zero, this must necessarily mean that the zero-order magnetic field is zero (Toscano et al., 2015). Keeping to first-order gives for the two HM equations

$$\begin{aligned} m\partial_t\mathbf{v}_1 &= -e(\mathbf{E}_0 + \mathbf{E}_1) + \nabla u_0(\mathbf{r}) - \nabla\frac{\delta F[n,\mathbf{v}]}{\delta n}\Big|_0 - \nabla\frac{\delta F[n,\mathbf{v}]}{\delta n}\Big|_1 \\ \partial_t n_1 &= -\nabla\cdot(n_0\mathbf{v}_1)\,. \end{aligned} \tag{G.12}$$

Combining orders gives two sets of equations. First: the zeroth-order condition, which gives the equilibrium density condition

$$-\frac{\delta F[n,\mathbf{v}]}{\delta n}\Big|_0 + e\phi_0 + u_0(\mathbf{r}) = \mu, \tag{G.13}$$

and defines the chemical potential. Second: the first-order conditions

$$\begin{aligned} \partial_t\mathbf{J} &= -\omega_P^2\epsilon_0\mathbf{E}_1 - \frac{\rho}{m}\nabla\frac{\delta F[n,\mathbf{v}]}{\delta n}\Big|_1 \\ \partial_t\rho &= -\nabla\cdot\mathbf{J}, \end{aligned} \tag{G.14}$$

where I have defined the charge $\rho = en_1$ and current density $\mathbf{J} = n_0 e\mathbf{v}$. This pair of equations can be solved along with the wave equation $\nabla\times\mathbf{E}_1 + \frac{1}{c^2}\partial_t^2\mathbf{E}_1 = -\mu_0\partial_t\mathbf{J}_1$ to give a self-consistent scheme to obtain the first-order fields (Toscano et al., 2015).

H Derivation Of The Microscopic Density-Density Response Function

In this appendix we will use the self-consistent field approach to linearize the Liouville equation and obtain the non-interacting density-density response function (Ehrenreich and Cohen, 1959). This is a very simple and physically transparent derivation, we use a single-particle formalism and take into account Fermionic statistics via the use of the single-particle density matrix. We consider a single-particle Hamiltonian

$$\hat{H}(\mathbf{r},t) = \hat{H}_0 + \hat{H}_1(\mathbf{r},t) \tag{H.1}$$

where $\hat{H}_1(\mathbf{r},t)$ is a small perturbation, it can be the *self-consistent potential* or just be the external field. $\hat{H}_0$ is the unperturbed Hamiltonian and we assume that we can calculate it eigenstates and eigenvalues: $\hat{H}_0 |i\rangle = \varepsilon_i |i\rangle$. The *single-particle Liouville equation* is given by (Feynman, 1972)

$$i\hbar \frac{\partial \hat{\rho}}{\partial t} = [\hat{H}, \hat{\rho}] \tag{H.2}$$

$\hat{\rho}$ is the single-particle density operator and we assume we can split it into a perturbed and unperturbed part given by $\hat{\rho}_0$ and $\hat{\rho}_1$ respectively. The unperturbed density operator is given by[1]

$$\hat{\rho}_0 = \frac{1}{1 + e^{-\frac{(\hat{H}-\mu)}{k_B T}}} = \sum_i f_i |i\rangle \langle i| . \tag{H.3}$$

[1] It is presented in the basis of the unperturbed eigenstates. It can be seen to be the sum of projection operators $|i\rangle \langle i|$, weighted by the Fermi function.

f_i is the probability the system is in state $|i\rangle$ and we can identify it with the Fermi function $f(\varepsilon_i) = \frac{1}{1+e^{-\frac{(\varepsilon_i-\mu)}{k_B T}}}$. We proceed to solve equation H.2 by writing it in the form

$$i\hbar \frac{\partial(\hat{\rho}_0 + \hat{\rho}_1)}{\partial t} = [\hat{H}_0, \hat{\rho}_0] + [\hat{H}_1, \hat{\rho}_0] + [\hat{H}_0, \hat{\rho}_1] + [\hat{H}_1, \hat{\rho}_1]. \tag{H.4}$$

We throw away second-order terms and can cancel out the zeroth-order Liouville equation to get

$$i\hbar \frac{\partial \hat{\rho}_1}{\partial t} = [\hat{H}_1, \hat{\rho}_0] + [\hat{H}_0, \hat{\rho}_1], \tag{H.5}$$

and expand the operators $\hat{H}_1$ and $\hat{\rho}_1$ in the eigenstates of $\hat{H}_0$

$$\begin{aligned}\hat{H}_1 &= \sum_{ij} |j\rangle \langle j|\hat{H}_1|i\rangle \langle i| \\ \hat{\rho}_1 &= \sum_{ij} |j\rangle \langle j|\hat{\rho}_1|i\rangle \langle i| .\end{aligned} \tag{H.6}$$

We assume the time dependence of the induced density change is exactly the same as the perturbation $e^{-i\omega t}$, this further justifies throwing away the second-order term in equation H.2 as it will have a time dependence of $e^{-2i\omega t}$. Working through the calculation reveals the elements of the density matrix

$$\langle j|\hat{\rho}_1|i\rangle = \frac{f_i - f_j}{\hbar\omega + \varepsilon_i - \varepsilon_j + i\eta} \langle j|\hat{H}_1|i\rangle , \tag{H.7}$$

where a positive imaginary infinitesimal, η, has been introduced. This reduces instabilities in numerical calculations and is equivalent to averaging the response by a Lorentzian function with width η, i.e. it is an artificially added loss. Physically, it corresponds to an adiabatic 'switch on' of the perturbation; the perturbation intensity goes to zero very slowly as $t \to -\infty$. This enforces that the system does not have an infinite amount of energy before $t = 0$ and, when the system reaches a steady state, that the energy supplied by the external field is matched by the dissipation into the environment (Giuliani and Vignale, 2005).

We will use of the nice property that the expectation value of an operator in a statistical ensemble is given by the trace of the operator with the density operator[2]

[2] I include a factor of $1/N$ due to the use of the Fermi functions in the density matrix.

$$\langle \hat{O} \rangle = \frac{1}{N} Tr[\hat{\rho}\hat{O}] = \sum_j \langle j|\hat{\rho}\hat{O}|j\rangle = \frac{1}{N} \sum_{ij} \langle j|\hat{\rho}|i\rangle \langle i|\hat{O}|j\rangle . \tag{H.8}$$

To calculate the induced density $\Delta n(\mathbf{r})$ we take the trace of the 1st order density matrix with the number density operator $\hat{n} = \sum_I^N \hat{n}_I$, where N is the number of electrons and $\hat{n}_I$ is the single-electron density operator. Since we are considering an independent electron approximation we can simply consider the single-electron number operator and multiply by N, which will kill the normalisation factor in equation H.8. By projection onto the position eigenstates, $\hat{\mathbf{r}}|\mathbf{r}\rangle = \mathbf{r}|\mathbf{r}\rangle$, we obtain

$$\begin{aligned}\langle i|\hat{n}(\hat{\mathbf{r}})|j\rangle = \int d^3r' \; \overbrace{\langle i|\hat{n}|\mathbf{r}'\rangle}^{\psi_i^*(\mathbf{r}')\delta(\mathbf{r}-\mathbf{r}')} \overbrace{\langle \mathbf{r}'|j\rangle}^{\psi_j(\mathbf{r}')} \\ = \psi_i^*(\mathbf{r})\psi_j(\mathbf{r}),\end{aligned} \tag{H.9}$$

where I have used the useful identity (Economou, 1983) $\langle \mathbf{r}|\hat{O}(\mathbf{r})|i\rangle = O^*(\mathbf{r})\psi_i(\mathbf{r})$, since a function of an operator shares the same eigenstates as the operator, i.e. $\hat{O}(\hat{\mathbf{r}})|\mathbf{r}\rangle = O(\mathbf{r})|\mathbf{r}\rangle$. The other matrix is similarly tackled, we will consider a scalar perturbation Δv that

couples to the electron density

$$\hat{H}_1 = \int d^3r' \; \Delta v(\mathbf{r}) \hat{n}(\mathbf{r}') = \sum_I^N \Delta v(\mathbf{r}_I) \tag{H.10}$$

Since we are working in linear response, perturbations just add up and so with no loss of generality we can consider just one and ignore the sum over I (Giuliani and Vignale, 2005). Again inserting the identity $\hat{\mathbf{1}} = \int d^3r' \, |\mathbf{r}'\rangle \langle \mathbf{r}'|$ into the matrix element gives us

$$\Delta n(\mathbf{r}, \omega) = \int d^3r' \sum_{ij}^{\infty} \frac{f_i - f_j}{\hbar\omega - \hbar\omega_{ji} + i\eta} \psi_j^*(\mathbf{r}')\psi_i(\mathbf{r}')\psi_i^*(\mathbf{r})\psi_j(\mathbf{r})\Delta v(\mathbf{r}'), \tag{H.11}$$

where $\hbar\omega_{ji} = \varepsilon_i - \varepsilon_j$. By using the definition of the density-density response function as the functional derivative of the induced density relative to the perturbation $\chi_0(\mathbf{r}, \mathbf{r}', \omega) = \frac{\delta \Delta n(\mathbf{r},\omega)}{\delta \Delta v(\mathbf{r}',\omega)}$, we get

$$\boxed{\chi_0(\mathbf{r}, \mathbf{r}', \omega) = \sum_{ij}^{\infty} \frac{f_i - f_j}{\hbar\omega - \hbar\omega_{ji} + i\eta} \psi_j^*(\mathbf{r}')\psi_i(\mathbf{r}')\psi_i^*(\mathbf{r})\psi_j(\mathbf{r}).} \tag{H.12}$$

If we take the perturbation to be the self-consistent KS field, $\Delta v = \Delta v_{KS}$ and identify the single-particle orbitals with the KS orbitals, then by comparing to equation 2.21 we can identify the non-interacting KS response function, $\chi_0 = \chi_{KS}$. We can clearly see the pole structure; excitations are given by the bare KS energy differences, $\omega = \omega_{ji}$. By playing around with the Fermi functions and assuming real orbitals, we can write the response function in a couple of different instructive ways

$$\begin{aligned}\chi_0(\mathbf{r}, \mathbf{r}', \omega) &= \sum_{i=1}^{N} \sum_{j=1}^{\infty} \left[\frac{1}{-\hbar\omega_{ji} + \hbar\omega + i\eta} + \frac{1}{\hbar\omega_{ji} + \hbar\omega + i\eta} \right] \psi_j(\mathbf{r}')\psi_i(\mathbf{r}')\psi_i(\mathbf{r})\psi_j(\mathbf{r}) \\ &= \sum_{i=1}^{N} \sum_{j=N}^{\infty} \left[\frac{2\hbar\omega_{ji}}{-\hbar\omega_{ji}^2 + (\hbar\omega + i\eta)^2} \right] \psi_j(\mathbf{r}')\psi_i(\mathbf{r}')\psi_i(\mathbf{r})\psi_j(\mathbf{r}),\end{aligned} \tag{H.13}$$

where the last equality follows from $f_j - f_i = f_j(1 - f_i) - f_i(1 - f_j)$ and then relabelling. The last form is most physically revealing; it tells us that we need only consider initial states below the Fermi surface and final states above.

Equation H.12 is valid for a system of independent particles described by a simple product state of one-particle wavefunctions $\Psi_0(\mathbf{r}_1, \mathbf{r}_2, ...\mathbf{r}_N) = \psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2)...\psi_N(\mathbf{r}_N)$. We can derive a more general expression by considering the time-dependent Schrödinger equation and evaluating it with time-dependent perturbation theory, $|\Psi_J\rangle$ are the eigenstates of the unperturbed many-body

Hamiltonian. This is a simple derivation (Harl, 2008) and gives

$$\chi_0(\mathbf{r},\mathbf{r},\omega) = -\sum_{J\neq 0}\left[\frac{\langle\Psi_J|\hat{n}(\mathbf{r}')|\Psi_0\rangle\,\langle\Psi_0|\hat{n}(\mathbf{r})|\Psi_J\rangle}{E_J - E_0 - (\omega+i\eta)} + \frac{\langle\Psi_0|\hat{n}(\mathbf{r}')|\Psi_J\rangle\,\langle\Psi_J|\hat{n}(\mathbf{r})|\Psi_0\rangle}{E_J - E_0 + (\omega+i\eta)}\right]. \tag{H.14}$$

Equation H.12 can then be derived from this equation by considering an excited state of the configuration $\Psi_I(\mathbf{r}_1,\mathbf{r}_2,...\mathbf{r}_{i\to j},...\mathbf{r}_N)$ and noting that $\langle\Psi_0|\hat{n}(\mathbf{r})|\Psi_I\rangle = \psi_i^*(\mathbf{r})\psi_j(\mathbf{r})$.

We can write the full density-density response function $\chi(\mathbf{r},\mathbf{r}',\omega)$ in terms of the KS density-density response function by noting that

$$\Delta n(\mathbf{r},\omega) = \int d^3r'\chi(\mathbf{r},\mathbf{r}',\omega)\Delta v(\mathbf{r}',\omega) = \int d^3r'\chi_{KS}(\mathbf{r},\mathbf{r}',\omega)\Delta v_{KS}(\mathbf{r}',\omega) \tag{H.15}$$

and performing a first-order expansion of the KS potential

$$v_{XC}[n](\mathbf{r},\omega) \approx v_{XC}[n_0](\mathbf{r}) + \overbrace{\int d^3r' \underbrace{\frac{\delta v_{xc}[n](\mathbf{r},\omega)}{\delta n(\mathbf{r}',\omega)}\Big|_{n_0}}_{\mathfrak{K}_{XC}(\mathbf{r},\mathbf{r}',\omega)} \Delta n(\mathbf{r}',\omega)}^{\Delta v_{XC}(\mathbf{r},\omega)} \tag{H.16}$$

$$\Delta v_{KS}(\mathbf{r},\omega) = \Delta v(\mathbf{r},\omega) + \Delta v_H(\mathbf{r},\omega) + \Delta v_{XC}(\mathbf{r},\omega).$$

The different contributions to the KS potential can be expressed using the chain rule

$$\begin{aligned}\Delta v_H(\mathbf{r}',\omega) &= \int d^3r_1 \int d^3r_2\, \frac{\delta v_H(\mathbf{r}',\omega)}{\delta n(\mathbf{r}_1,\omega)}\,\frac{\delta n(\mathbf{r}_1,\omega)}{\delta \Delta v(\mathbf{r}_2,\omega)}\,\Delta v(\mathbf{r}_2,\omega)\\ &= \int d^3r_1 \int d^3r_2\, \mathfrak{K}_H(\mathbf{r}',\mathbf{r}_1,\omega)\,\chi(\mathbf{r}_1,\mathbf{r}_2,\omega)\,\Delta v(\mathbf{r}_2,\omega)\\ \Delta v_{XC}(\mathbf{r}',\omega) &= \int d^3r_1 \int d^3r_2\, \mathfrak{K}_{XC}(\mathbf{r}',\mathbf{r}_1,\omega)\,\chi(\mathbf{r}_1,\mathbf{r}_2,\omega)\,\Delta v(\mathbf{r}_2,\omega).\end{aligned} \tag{H.17}$$

By combining these equations and relabelling some of the variables it is simple to derive the Dyson equation relating the two response functions

$$\begin{aligned}\chi(\mathbf{r},\mathbf{r}',\omega) = \chi_{KS}(\mathbf{r},\mathbf{r}',\omega) +\\ \int d^3r_1 \int d^3r_2\, \chi_{KS}(\mathbf{r},\mathbf{r}_1,\omega)\,\underbrace{(\mathfrak{K}_H(\mathbf{r}_1,\mathbf{r}_2,\omega) + \mathfrak{K}_{XC}(\mathbf{r}_1,\mathbf{r}_2,\omega))}_{\mathfrak{K}(\mathbf{r}_1,\mathbf{r}_2,\omega)}\,\chi(\mathbf{r}_2,\mathbf{r}',\omega).\end{aligned} \tag{H.18}$$

I Derivation Of The Casida Equations

This derivation is a slight modification of the original by Mark E. Casida (Casida, 1995), it is based on the single-particle transition-density matrix that we derived in appendix H. As we saw, the density matrix elements are linked to the *effective* potential experienced by the electrons

$$\Delta\rho_{kl} = \langle k|\hat{\rho}_1|l\rangle = \frac{f_l - f_k}{\hbar\omega - \hbar\omega_{kl}}\Delta v_{KS,kl}, \tag{I.1}$$

the induced density can be written as

$$\Delta n(\mathbf{r},\omega) = \sum_{kl}^{\infty} \langle k|\hat{\rho}_1|l\rangle \langle l|\hat{n}|k\rangle = \sum_{kl} \Delta\rho_{kl}\psi_k(\mathbf{r})\psi_l^*(\mathbf{r}), \tag{I.2}$$

thus the transition density gives a convenient representation of the induced density in a configuration representation. The effective potential is the sum of the external potential and the induced KS field, which takes into account screening and XC effects, it can therefore be linked to the density matrix[1]

$$\Delta v_{KS}(\mathbf{r},\omega) = \int d^3r' \Re(\mathbf{r},\mathbf{r}',\omega)\Delta n(\mathbf{r},\omega). \tag{I.3}$$

[1] Note that the linear response of the KS density matrix gives the true response of the density but not the true density matrix.

It can be written in a configuration representation by noting that $|r\rangle$ is an eigenstate of $\Delta\hat{v}_{KS}$

$$\begin{aligned} \Delta\hat{v}_{KS} &= \sum_{kl}^{\infty} \langle k|\Delta\hat{v}_{KS}|l\rangle \, |k\rangle \langle l| \\ \therefore \Delta v_{KS}(\mathbf{r},\omega) = \langle r|\Delta\hat{v}_{KS}|r\rangle &= \sum_{kl}^{\infty} \langle k|\Delta\hat{v}_{KS}|l\rangle \, \psi_k(\mathbf{r})\psi_l^*(\mathbf{r}). \end{aligned} \tag{I.4}$$

where the matrix elements can be written as

$$\begin{aligned} \Delta v_{KS,kl}(\omega) = \langle k|\Delta\hat{v}_{KS}|l\rangle &= \int d^3r \ \langle k|\Delta\hat{v}_{KS}|\mathbf{r}\rangle \langle \mathbf{r}|l\rangle \\ &= \int d^3r \ \Delta v_{KS}(\mathbf{r},\omega)\psi_k^*(\mathbf{r})\psi_l(\mathbf{r}). \end{aligned} \tag{I.5}$$

This gives the following matrix equation

$$\begin{aligned}\Delta v_{KS,kl}(\omega) &= \int d^3r\, d^3r'\, \psi_k^*(\mathbf{r})\psi_l(\mathbf{r})\aleph(\mathbf{r},\mathbf{r}',\omega)\Delta n(\mathbf{r}',\omega) \\ &= \sum_{mn}^{\infty}\int d^3r\, d^3r'\, \psi_k^*(\mathbf{r})\psi_l(\mathbf{r})\aleph(\mathbf{r},\mathbf{r}',\omega)\psi_m(\mathbf{r}')\psi_n^*(\mathbf{r}')\Delta\rho_{mn} \\ &= \sum_{mn}^{\infty}\aleph_{kl,mn}\Delta\rho_{mn},\end{aligned} \tag{I.6}$$

where I have defined the two-body coupling matrix

$$\aleph_{kl,mn} = \int d^3r\, d^3r'\, \psi_k^*(\mathbf{r})\psi_l(\mathbf{r})\aleph(\mathbf{r},\mathbf{r}',\omega)\psi_m(\mathbf{r}')\psi_n^*(\mathbf{r}'), \tag{I.7}$$

calculation of this quantity can constitute the main computational effort of the Casida method. The density matrix can therefore be written as

$$\Delta\rho_{kl} = \frac{f_l - f_k}{\hbar\omega - \hbar\omega_{kl}}\left(\Delta v_{kl} + \sum_{mn}^{\infty}\aleph_{kl,mn}\Delta\rho_{mn}\right). \tag{I.8}$$

The Casida approach involves the calculation of eigenmodes of the system, these are characterised by a finite response of the system in the limit of vanishing perturbation, i.e. it corresponds to the poles of the interacting response function χ. With this in mind, we set the external perturbation to zero and rearrange to get

$$\sum_{mn}^{\infty}\left(\delta_{kl,mn}\hbar\omega_{mn} + (f_m - f_n)\aleph_{kl,mn}\right)\Delta\rho_{mn} = \hbar\omega\Delta\rho_{kl}. \tag{I.9}$$

We now play around with the indices and switch to the conventional notation of i, j for occupied orbitals and a, b for unoccupied, the summation $\sum_{ia}$ is now understood to mean $\sum_{i=1}^{N_F}\sum_{a=N_F+1}^{\infty}$ and spin is included simply by a factor of 2. This allow the Casida equation to be written in the following conventional super-matrix form[2] (Hirata and Head-Gordon, 1999)

$$\begin{bmatrix}\boldsymbol{\mathcal{A}} & \boldsymbol{\mathcal{B}} \\ \boldsymbol{\mathcal{B}}^* & \boldsymbol{\mathcal{A}}^*\end{bmatrix}\begin{bmatrix}\boldsymbol{\mathcal{X}} \\ \boldsymbol{\mathcal{Y}}\end{bmatrix} = \Omega\begin{bmatrix}\mathbf{1} & 0 \\ 0 & -\mathbf{1}\end{bmatrix}\begin{bmatrix}\boldsymbol{\mathcal{X}} \\ \boldsymbol{\mathcal{Y}}\end{bmatrix} \tag{I.10}$$

where Ω is the values of frequency at which I.9 holds and hence are eigenvalues. The matrix elements of $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathcal{B}}$ are

$$\begin{aligned}\mathcal{A}_{ia,jb} &= \delta_{ia,jb}\omega_{bj} + \aleph_{ia,jb} \\ \mathcal{B}_{ia,jb} &= \aleph_{ia,jb}.\end{aligned} \tag{I.11}$$

Mathematically, this is an infinite-dimensional anti-Hermitian eigenvalue problem; taking the Hermitian conjugate leads to the same eigenvalue equation but with the opposite sign of the eigenvalue, the solutions come in pairs of $\boldsymbol{\mathcal{X}}_I, \boldsymbol{\mathcal{Y}}_I, \Omega_I$ and $\boldsymbol{\mathcal{X}}_I^*, \boldsymbol{\mathcal{Y}}_I^*, -\Omega_I$ (Ullrich, 2011). The eigenvectors are related to the transition density matrix and allows us to extract a wealth of information about excitations, as demonstrated in section 2.3.4.

[2] Some references differ by a minus sign (Ullrich, 2011).

I.1 Alternative Representation Of The Casida Super Matrix Equation

If the KS orbitals are real[3] and the XC kernel is frequency independent, then the matrices $\mathcal{A}$ and $\mathcal{B}$ are real and equation I.10 can be recast as

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ -\mathcal{B} & -\mathcal{A} \end{bmatrix} \begin{bmatrix} \mathcal{X} \\ \mathcal{Y} \end{bmatrix} = \Omega \begin{bmatrix} \mathcal{X} \\ \mathcal{Y} \end{bmatrix}. \tag{I.12}$$

This is equivalent to the following pair of coupled equations

$$\begin{aligned} (\mathcal{X} - \mathcal{Y}) &= \frac{(\mathcal{A} + \mathcal{B})}{\Omega} (\mathcal{X} + \mathcal{Y}) \\ (\mathcal{X} + \mathcal{Y}) &= \frac{(\mathcal{A} - \mathcal{B})}{\Omega} (\mathcal{X} - \mathcal{Y}). \end{aligned} \tag{I.13}$$

It follows that a particularly useful transformation to the supermatrix equation can be made, which was first noted for the RPA (McCurdy and Cusachs, 1971) and then extended to TDDFT (Bauernschmitt and Ahlrichs, 1996). We apply a unitary transformation on the Casida equation

$$\begin{aligned} \mathbf{U} &= -\frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{1} & \mathbf{1} \\ \mathbf{1} & -\mathbf{1} \end{bmatrix} \\ \therefore \boldsymbol{U} \begin{bmatrix} \mathcal{A} & \mathcal{B} \\ -\mathcal{B} & -\mathcal{A} \end{bmatrix} \boldsymbol{U}^{\dagger} &= \begin{bmatrix} 0 & \mathcal{A} - \mathcal{B} \\ \mathcal{A} + \mathcal{B} & 0 \end{bmatrix}. \end{aligned} \tag{I.14}$$

This is convenient as it is halves the dimensions of the original equation, but unfortunately it is not Hermitian and gives complex eigenvalues. This is solved by taking the square of both sides to, giving two equations

$$\begin{aligned} (\mathcal{A} - \mathcal{B})(\mathcal{A} + \mathcal{B})(\mathcal{X} + \mathcal{Y}) &= \Omega^2 (\mathcal{X} + \mathcal{Y}) \quad (1) \\ (\mathcal{A} + \mathcal{B})(\mathcal{A} - \mathcal{B})(\mathcal{X} - \mathcal{Y}) &= \Omega^2 (\mathcal{X} - \mathcal{Y}) \quad (2) \\ &. \end{aligned} \tag{I.15}$$

We need only solve one of these equations to obtain the eigenvalues and eigenvectors, again revealing we have reduced the dimensionality. Choosing the first equation, (1), we need to make one more transformation to make the problem Hermitian, multiplying both sides by[4] $(\mathcal{A} - \mathcal{B})^{-1/2}$ gives

$$\begin{aligned} &(\mathcal{A} - \mathcal{B})^{-1/2} (\mathcal{A} - \mathcal{B})(\mathcal{A} + \mathcal{B})(\mathcal{X} + \mathcal{Y}) \\ &= (\mathcal{A} - \mathcal{B})^{1/2} (\mathcal{A} + \mathcal{B})(\mathcal{A} - \mathcal{B})^{1/2} (\mathcal{A} - \mathcal{B})^{-1/2} (\mathcal{X} + \mathcal{Y}) \\ &= \Omega^2 (\mathcal{A} - \mathcal{B})^{-1/2} (\mathcal{X} + \mathcal{Y}), \end{aligned} \tag{I.16}$$

[3] This is normally the case in molecular calculations and is always possible if the ground state Hamiltonian has time-reversal symmetry

[4] A quick note that it is important to understand that the square root of a matrix $\sqrt{\mathbf{M}}$ in this case denotes the principal square root and means $\sqrt{\mathbf{M}}\sqrt{\mathbf{M}} = \mathbf{M}$.

which leads us to identify (Bauernschmitt and Ahlrichs, 1996)

$$\boxed{\begin{aligned} \boldsymbol{\mathcal{C}}_+ \boldsymbol{\mathcal{Z}}_+ &= \Omega_+^2 \boldsymbol{\mathcal{Z}}_+ \\ \boldsymbol{\mathcal{C}}_+ &= (\boldsymbol{\mathcal{A}}-\boldsymbol{\mathcal{B}})^{1/2}(\boldsymbol{\mathcal{A}}+\boldsymbol{\mathcal{B}})(\boldsymbol{\mathcal{A}}-\boldsymbol{\mathcal{B}})^{1/2} \\ \boldsymbol{\mathcal{Z}}_+ &= \mathfrak{c}(\boldsymbol{\mathcal{A}}-\boldsymbol{\mathcal{B}})^{-1/2}(\boldsymbol{\mathcal{X}}+\boldsymbol{\mathcal{Y}})?? \end{aligned}} \tag{I.17}$$

This is now in Hermitian form and note that the vector $(\boldsymbol{\mathcal{A}}-\boldsymbol{\mathcal{B}})$ is a very simple diagonal matrix with the KS energy differences ω_{ai} as its elements. I have added a normalisation factor of $\mathfrak{c}$, which I will elucidate later on. In index form we recover equation 2.26.

There actually exists another, completely equivalent, representation of equations I.10 as revealed in a very nice paper by Luzanov and Zhikol (Luzanov and Zhikol, 2010). Let us start again from equations I.15, but now choose the second equation , (2), and multiply both sides by $(\boldsymbol{\mathcal{A}}+\boldsymbol{\mathcal{B}})^{-1/2}$. This gives

$$\begin{aligned} (\boldsymbol{\mathcal{A}}+\boldsymbol{\mathcal{B}})^{1/2}(\boldsymbol{\mathcal{A}}-\boldsymbol{\mathcal{B}})(\boldsymbol{\mathcal{A}}+\boldsymbol{\mathcal{B}})^{1/2}(\boldsymbol{\mathcal{A}}+\boldsymbol{\mathcal{B}})^{-1/2}(\boldsymbol{\mathcal{X}}-\boldsymbol{\mathcal{Y}}) \\ = \omega^2(\boldsymbol{\mathcal{A}}+\boldsymbol{\mathcal{B}})^{-1/2}(\boldsymbol{\mathcal{X}}-\boldsymbol{\mathcal{Y}}), \end{aligned} \tag{I.18}$$

leading to

$$\boxed{\begin{aligned} \boldsymbol{\mathcal{C}}_- \boldsymbol{\mathcal{Z}}_- &= \Omega_-^2 \boldsymbol{\mathcal{Z}}_- \\ \boldsymbol{\mathcal{C}}_- &= (\boldsymbol{\mathcal{A}}+\boldsymbol{\mathcal{B}})^{1/2}(\boldsymbol{\mathcal{A}}-\boldsymbol{\mathcal{B}})(\boldsymbol{\mathcal{A}}+\boldsymbol{\mathcal{B}})^{1/2} \\ \boldsymbol{\mathcal{Z}}_- &= \mathfrak{c}(\boldsymbol{\mathcal{A}}+\boldsymbol{\mathcal{B}})^{-1/2}(\boldsymbol{\mathcal{X}}-\boldsymbol{\mathcal{Y}}). \end{aligned}} \tag{I.19}$$

The two set of equations are spectrally equivalent, $\Omega_- = \Omega_+ = \Omega$, but the Hermitian form of the Casida equations is not unique: $\boldsymbol{\mathcal{Z}}_- \neq \boldsymbol{\mathcal{Z}}_+$ (Luzanov and Zhikol, 2010). This poses problems when trying to reduce the TDDFT equations to a CIS like form, which is often necessary for excited state structural analysis. For most 'normal' molecular systems this non-uniqueness does not seem to cause too much trouble as $\boldsymbol{\mathcal{Z}}_- \approx \boldsymbol{\mathcal{Z}}_+$, so different CIS like representations are similar (Luzanov and Zhikol, 2010). The two set of eigenvectors can be related as follows[5]

[5] The second equality follows from the first equation of equations I.13.

$$\begin{aligned} \boldsymbol{\mathcal{Z}}_- &= \sqrt{\Omega}(\boldsymbol{\mathcal{A}}+\boldsymbol{\mathcal{B}})^{-1/2}(\boldsymbol{\mathcal{X}}-\boldsymbol{\mathcal{Y}}) \\ &= \sqrt{\Omega}(\boldsymbol{\mathcal{A}}+\boldsymbol{\mathcal{B}})^{-1/2}\left(\frac{(\boldsymbol{\mathcal{A}}+\boldsymbol{\mathcal{B}})(\boldsymbol{\mathcal{X}}+\boldsymbol{\mathcal{Y}})}{\Omega}\right) \\ &= \frac{(\boldsymbol{\mathcal{A}}+\boldsymbol{\mathcal{B}})^{1/2}}{\Omega}(\boldsymbol{\mathcal{A}}-\boldsymbol{\mathcal{B}})^{1/2}(\boldsymbol{\mathcal{A}}-\boldsymbol{\mathcal{B}})^{-1/2}\sqrt{\Omega}(\boldsymbol{\mathcal{X}}+\boldsymbol{\mathcal{Y}}) \\ &= \frac{(\boldsymbol{\mathcal{A}}+\boldsymbol{\mathcal{B}})^{1/2}(\boldsymbol{\mathcal{A}}-\boldsymbol{\mathcal{B}})^{1/2}}{\Omega}\boldsymbol{\mathcal{Z}}_+. \end{aligned} \tag{I.20}$$

Next we discuss the issue of normalisation, which is especially important when implementing these equations in a Casida code. Normally the normalisation factor $\mathfrak{c}$ does not seem to be included, but it is important as it enforces the normalisation $|\boldsymbol{\mathcal{Z}}_\pm| = 1$. To see this we note that the eigenvectors $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Y}}$ have the following

unusual normalisation

$$(\boldsymbol{\mathcal{X}}+\boldsymbol{\mathcal{Y}})(\boldsymbol{\mathcal{X}}-\boldsymbol{\mathcal{Y}}) = |\boldsymbol{\mathcal{X}}|^2 - |\boldsymbol{\mathcal{Y}}|^2 = 1, \tag{I.21}$$

meaning[6]

$$\begin{aligned} |\boldsymbol{\mathcal{Z}}_+|^2 &= \mathfrak{c}^2(\boldsymbol{\mathcal{A}}-\boldsymbol{\mathcal{B}})^{-1}(\boldsymbol{\mathcal{X}}+\boldsymbol{\mathcal{Y}})(\boldsymbol{\mathcal{X}}+\boldsymbol{\mathcal{Y}}) \\ &= \mathfrak{c}^2(\boldsymbol{\mathcal{A}}-\boldsymbol{\mathcal{B}})^{-1}\frac{(\boldsymbol{\mathcal{A}}-\boldsymbol{\mathcal{B}})(\boldsymbol{\mathcal{X}}-\boldsymbol{\mathcal{Y}})}{\Omega}(\boldsymbol{\mathcal{X}}+\boldsymbol{\mathcal{Y}}) \\ &= \frac{\mathfrak{c}^2}{\Omega}\left(|\boldsymbol{\mathcal{X}}|^2 - |\boldsymbol{\mathcal{Y}}|^2\right). \end{aligned} \tag{I.22}$$

[6] The second equality follows from the second equation of equations I.13.

Therefore for the correct normalisation we require

$$\mathfrak{c} = \sqrt{\Omega}. \tag{I.23}$$

The above procedure can be repeated for $\boldsymbol{\mathcal{Z}}_-$ and the same result is found.

Finally for this appendix, we discuss how to derive the eigenvectors $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Y}}$ from $\boldsymbol{\mathcal{Z}}_+$ which is the quantity usually returned by electronic structure codes, for example OCTOPUS. We begin by writing $\boldsymbol{\mathcal{X}}$ in terms of $\boldsymbol{\mathcal{Z}}_+$ and $\boldsymbol{\mathcal{Z}}_-$ by cancelling out $\boldsymbol{\mathcal{Y}}$

$$\begin{aligned} \boldsymbol{\mathcal{Y}} &= \boldsymbol{\mathcal{X}} - \frac{\sqrt{\boldsymbol{\mathcal{A}}+\boldsymbol{\mathcal{B}}}}{\sqrt{\Omega}}\boldsymbol{\mathcal{Z}}_- \\ \therefore \boldsymbol{\mathcal{Z}}_+ &= \frac{\sqrt{\Omega}}{\sqrt{\boldsymbol{\mathcal{A}}-\boldsymbol{\mathcal{B}}}}\left(2\boldsymbol{\mathcal{X}} - \frac{\sqrt{\boldsymbol{\mathcal{A}}+\boldsymbol{\mathcal{B}}}}{\sqrt{\Omega}}\boldsymbol{\mathcal{Z}}_-\right) \\ \therefore \boldsymbol{\mathcal{X}} &= \frac{\sqrt{\boldsymbol{\mathcal{A}}-\boldsymbol{\mathcal{B}}}}{2\sqrt{\Omega}}\boldsymbol{\mathcal{Z}}_+ + \frac{\sqrt{\boldsymbol{\mathcal{A}}+\boldsymbol{\mathcal{B}}}}{2\sqrt{\Omega}}\boldsymbol{\mathcal{Z}}_- \end{aligned} \tag{I.24}$$

we repeat this but now solve for $\boldsymbol{\mathcal{Y}}$

$$\begin{aligned} \boldsymbol{\mathcal{X}} &= \boldsymbol{\mathcal{Y}} + \frac{\sqrt{\boldsymbol{\mathcal{A}}+\boldsymbol{\mathcal{B}}}}{\sqrt{\Omega}}\boldsymbol{\mathcal{Z}}_- \\ \therefore \boldsymbol{\mathcal{Z}}_+ &= \frac{\sqrt{\Omega}}{\sqrt{\boldsymbol{\mathcal{A}}-\boldsymbol{\mathcal{B}}}}\left(2\boldsymbol{\mathcal{Y}} + \frac{\sqrt{\boldsymbol{\mathcal{A}}+\boldsymbol{\mathcal{B}}}}{\sqrt{\Omega}}\boldsymbol{\mathcal{Z}}_-\right) \\ \therefore \boldsymbol{\mathcal{Y}} &= \frac{\sqrt{\boldsymbol{\mathcal{A}}-\boldsymbol{\mathcal{B}}}}{2\sqrt{\Omega}}\boldsymbol{\mathcal{Z}}_+ - \frac{\sqrt{\boldsymbol{\mathcal{A}}+\boldsymbol{\mathcal{B}}}}{2\sqrt{\Omega}}\boldsymbol{\mathcal{Z}}_- \end{aligned} \tag{I.25}$$

I now introduce the notation

$$\begin{aligned} \boldsymbol{\mathcal{D}}_+ &= \frac{\sqrt{\boldsymbol{\mathcal{A}}+\boldsymbol{\mathcal{B}}}}{2\sqrt{\Omega}} \\ \boldsymbol{\mathcal{D}}_- &= \frac{\sqrt{\boldsymbol{\mathcal{A}}-\boldsymbol{\mathcal{B}}}}{2\sqrt{\Omega}}. \end{aligned} \tag{I.26}$$

and we get our final result

$$\boxed{\begin{aligned} \boldsymbol{\mathcal{X}} &= \boldsymbol{\mathcal{D}}_-\boldsymbol{\mathcal{Z}}_+ + \boldsymbol{\mathcal{D}}_+\boldsymbol{\mathcal{Z}}_- \\ \boldsymbol{\mathcal{Y}} &= \boldsymbol{\mathcal{D}}_-\boldsymbol{\mathcal{Z}}_+ - \boldsymbol{\mathcal{D}}_+\boldsymbol{\mathcal{Z}}_- \end{aligned}}. \tag{I.27}$$

It is only by making use of the non-uniqueness of the Hermitian

transformation to derive the spectrally equivalent set of equations, which gives the same number of equations as unknowns, that allows us to extract the values for $\mathcal{X}$ and $\mathcal{Y}$ from $\mathcal{Z}_+$ or $\mathcal{Z}_+$.

J Coupled Harmonic Oscillator Model

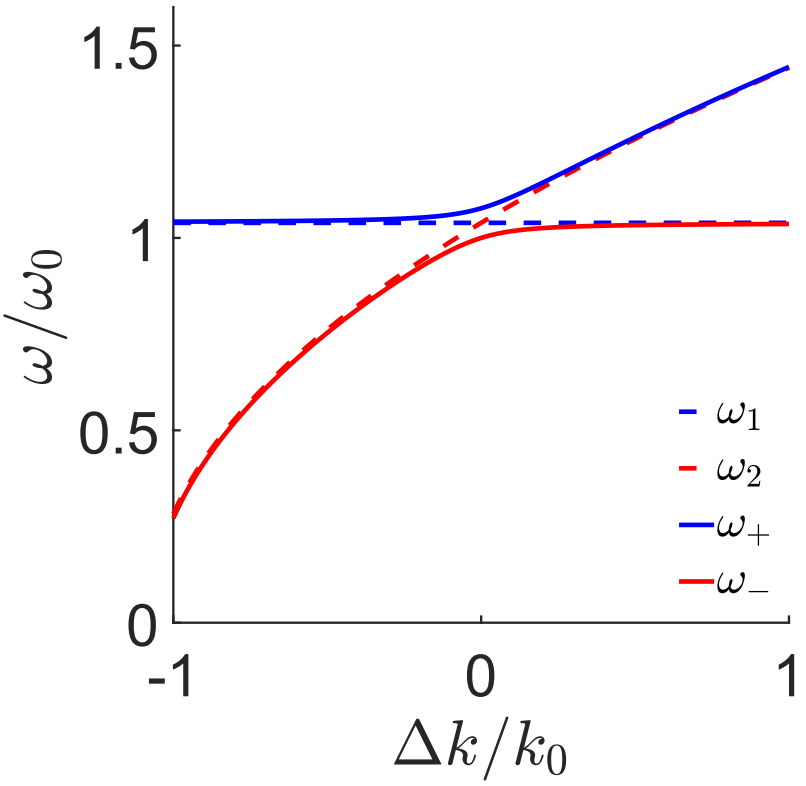


Figure J.1: A plot of the coupled oscillator eigenfrequencies (solid lines), given by equation J.3, and the uncoupled eigenfrequencies (dashed lines). Parameters used are $k_1 = k_0$, $k_2 = k_0 + \Delta k$, $m_1 = m_2 = 1$ and $\kappa = 0.08 k_0$.

Two coupled harmonic oscillators is the paradigm model of strong-coupling. The exchange of energy between two resonantly matched subsystems leads to a modified energy spectrum which differs from the subsystem oscillator modes. The model can be constructed within a classical, semi-classical or fully quantum framework (Törmä and Barnes, 2014), for simplicity we will consider the purely classical case (Novotny, 2010). We can quite generally write down the following set of differential equations for a set of N coupled harmonic oscillators

$$m_i \ddot{x}_i + k_i x_i + \sum_{j \neq i}^{N} \kappa_{ij} \left(x_i - x_j \right) = 0. \tag{J.1}$$

where k_i is the 'spring constant' of the ith oscillator and κ_{ij} is the coupling constant between the ith and jth oscillator. If we take $N = 2$ and $\kappa_{ij} = \kappa_{ji} = \kappa$, we can write the solution by Fourier

transforming and setting the following determinant to zero

$$\begin{vmatrix} -m_1\omega^2 + (k_1 + \kappa) & -\kappa \\ -\kappa & -m_2\omega^2 + (k_2 + \kappa) \end{vmatrix} = 0, \tag{J.2}$$

which has two solutions

$$\omega_{\pm}^2 = \frac{\omega_1^2 + \omega_2^2}{2} \pm \frac{1}{2}\sqrt{(\omega_1^2 - \omega_2^2)^2 + 4\Gamma^2\omega_1\omega_2}, \tag{J.3}$$

$\omega_1 = \sqrt{\frac{k_1+\kappa}{m_1}}$, $\omega_2 = \sqrt{\frac{k_2+\kappa}{m_2}}$ and $\Gamma = \frac{\sqrt{\kappa/m_1}\sqrt{\kappa/m_2}}{\sqrt{\omega_1\omega_2}}$. These are the eigenfrequencies of hybrid modes sometimes called *dressed states*. An example of the anti-crossing phenomenon for a non-dispersive mode of energy ω_1 and dispersive mode of energy $\omega_2 = \sqrt{\frac{1.08k_0+\Delta k}{m_2}}$ is shown in figure J.1.

The inclusion of damping via a dissipative term $\gamma_i\dot{x}_i$ in the defining differential equation, leads to complex frequencies (Novotny, 2010). The relative strength of the damping parameters γ_i compared to the coupling parameter Γ indicates whether Rabi splitting will be visible. This is demonstrated in figure J.2.

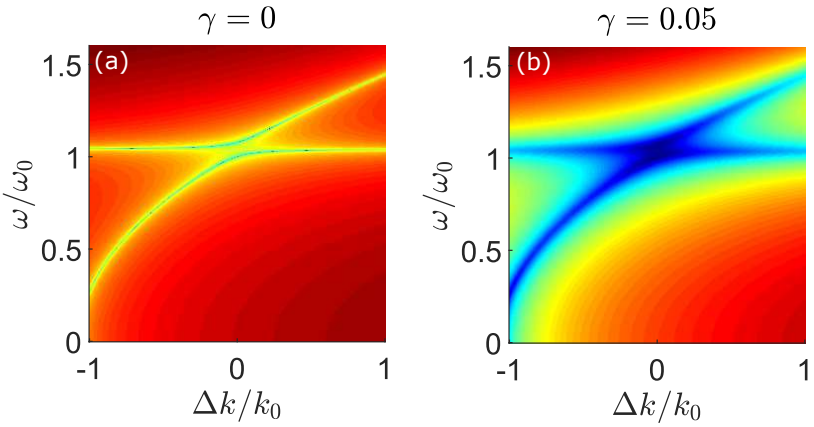


Figure J.2: A plot of the coupled oscillator determinant, equation J.2, with blue regions showing zeros and hence solutions to equation J.1 for two oscillators with different damping. Parameters used are $k_1 = k_0$, $k_2 = k_0 + \Delta k$, $m_1 = m_2 = 1$ and $\kappa = 0.08k_0$.

K Calculating Excitation Collectivity

We will work in the TDA, see section 2.3.3 for a discussion on this approximation. It is convenient as it allows simple identification of the transition density matrix, $\mathcal{X}_{ia} = \Delta\rho_{ia}$. There is no hp-ph coupling meaning it is possible to identify the excited state wavefunction as an expansion of single excited Slater determinants, this is the CIS picture commonly used in quantum chemistry (Etienne, 2015). We will borrow ideas from this field to construct a useful measure of excited state collectivity.

Inspired by ordinary quantum mechanics, where diagonalization of an operator gives us the eigenstates and eigenvalues, we may enquire can we do the same for the transition density matrix to derive a 'natural' representation of the excitation. By obtaining an excitation in its 'purest' and most compact form, we can hope to obtain quantitative measure of its character that could otherwise be masked by the particular basis set we use, i.e. the KS orbitals obtained from the ground state calculation. For example, an unsuitable set of ground state orbitals can lead to a 'false collectivity'. Furthermore, if there is no dominant configuration in a set of Casida excitation amplitudes, then interpretation of the excited state is difficult[1]. A diagonalization of the transition density matrix is in general not possible; it has the dimensions $N_{occ} \times N_{uno}$ and in a typical Casida calculation $N_{uno} > N_{occ}$, so it is not a square matrix. Instead we must use a generalisation of diagonalization for non-square matrices called a singular-value decomposition, which works for matrices of *any* shape. This orbital transformation was first proposed by Amos and Hall (Amos and Hall, 1961). This is a powerful technique for solving sets of simultaneous linear equations where the set of equations is not equal to the number of unknowns, for instance this is a common situation when fitting data (Riley et al., 2006).

[1] To make matters even worst, we should not forget the dubious physical reality of the KS orbitals.

Following (Martin, 2003), we write the transition density matrix in the following form[2]

$$\Delta\boldsymbol{\rho} = \boldsymbol{\mathcal{U}}\boldsymbol{\mathcal{S}}\boldsymbol{\mathcal{V}}^{\dagger}, \tag{K.1}$$

[2] Note this is just a generalisation of writing a diagonalizable square matrix in the form $\mathbf{A} = \mathbf{PDP}^{-1}$, where $\mathbf{D}$ is a diagonal matrix, with the eigenvalues of $\mathbf{A}$ as its elements, and $\mathbf{P}$ is an invertible matrix with the eigenvectors of $\mathbf{A}$ as its columns.

where $\boldsymbol{\mathcal{U}}$ and $\boldsymbol{\mathcal{V}}$ are *unitary* matrices of dimension $N_{occ} \times N_{occ}$ and $N_{uno} \times N_{uno}$. The middle matrix, $\boldsymbol{\mathcal{S}} = \boldsymbol{\mathcal{U}}^{\dagger}\Delta\boldsymbol{\rho}\boldsymbol{\mathcal{V}}$ has the following

diagonal form (Yasuike et al., 2011)

$$\boldsymbol{\mathcal{S}} = \underbrace{\begin{bmatrix} \lambda_1 & 0 & \dots & 0 & 0 & \\ 0 & \lambda_2 & \dots & 0 & 0 & \\ \vdots & \vdots & \ddots & \vdots & \vdots & \mathbf{0} \\ 0 & 0 & \dots & \lambda_{N_{occ}-1} & 0 & \\ 0 & 0 & \dots & 0 & \lambda_{N_{occ}} & \end{bmatrix}}_{N_{uno}} \Bigg\} N_{occ}\,. \tag{K.2}$$

It is simple to show that

$$\begin{aligned} \Delta\boldsymbol{\rho}\Delta\boldsymbol{\rho}^\dagger &= \boldsymbol{\mathcal{U}}\boldsymbol{\mathcal{S}}\boldsymbol{\mathcal{V}}^\dagger\boldsymbol{\mathcal{V}}\boldsymbol{\mathcal{S}}^\dagger\boldsymbol{\mathcal{U}}^\dagger \\ &= \boldsymbol{\mathcal{U}}\boldsymbol{\mathcal{S}}^2\boldsymbol{\mathcal{U}}^\dagger, \end{aligned} \tag{K.3}$$

thus the unitary transformation via $\boldsymbol{\mathcal{U}}$ diagonalizes the matrix $\Delta\boldsymbol{\rho}\Delta\boldsymbol{\rho}^\dagger$ and contains the eigenvectors as columns. Similarly it can be shown

$$\Delta\boldsymbol{\rho}^\dagger\Delta\boldsymbol{\rho} = \boldsymbol{\mathcal{V}}\boldsymbol{\mathcal{S}}^2\boldsymbol{\mathcal{V}}^\dagger. \tag{K.4}$$

Note that $\Delta\boldsymbol{\rho}\Delta\boldsymbol{\rho}^\dagger$ and $\Delta\boldsymbol{\rho}^\dagger\Delta\boldsymbol{\rho}$ will have different dimensions of $N_{occ} \times N_{occ}$ and $N_{uno} \times N_{uno}$ respectively. The first N_{occ} eigenvalues of the two matrices, $\lambda_1, ...\lambda_{N_{occ}}$, are identical and the remaining eigenvalues, $\lambda_{N_{occ}+1}, ...\lambda_{N_{uno}}$, of the larger matrix $\Delta\boldsymbol{\rho}^\dagger\Delta\boldsymbol{\rho}$ are zero (Dreuw and Head-Gordon, 2005). The eigenvalues have the following interesting properties

$$\begin{aligned} &0 \leq \lambda_i \leq 1 \\ &\sum_{i=1}^{N_{occ}} \lambda_i = 1. \end{aligned} \tag{K.5}$$

The $N_{occ} \times N_{uno}$ transitions that defined the excitation has now been reduced to N_{occ} ph amplitudes, each hole in the occupied space has a partner particle in the virtual space. The importance of each ph is measured by the value of the eigenvalue λ_i. Note that the above transformation, when applied to the full Casida equations, is slightly modified due to the presence of de-excitation. This leads to the sum $\sum_{i=1}^{N_{occ}} \lambda_i$ deviating from 1, the magnitude of the deviation will give a measure of the importance of de-excitation (Martin, 2003). The transition orbitals, ϕ, can be calculated via the following unitary transformation

$$\begin{aligned} [\phi_1, \phi_2, ..\phi_{N_{occ}}] &= [\psi_1, \psi_2, ..\psi_{N_{occ}}]\boldsymbol{\mathcal{U}} \\ [\phi'_1, \phi'_2, ..\phi'_{N_{uno}}] &= [\psi'_1, \psi'_2, ..\psi'_{N_{uno}}]\,\boldsymbol{\mathcal{V}}. \end{aligned} \tag{K.6}$$

Note that a unitary transformation will leave the determinant constructed from the orbitals unchanged (Amos and Hall, 1961).

It is usual for electronic transitions to be expressible as a single ph pair in the natural transition orbital picture (Dreuw and Head-Gordon, 2005), even if they appear as highly mixed in the molecular orbital basis; this can be termed 'false collectivity'. On the other

hand, we would not expect this to be the case for plasmonics. A plasmon is a collective oscillation of all the valence electrons, it would not be expected to be able to represent it as a single ph pair. To test this we introduce the inverse participation ratio, also known as the collectivity (Plasser et al., 2014)

$$\nu = \frac{1}{\sum_{i=1}^{N_{occ}} \lambda_i^4}. \tag{K.7}$$

This number, which can take values between 1 and N_{occ}, is calculated for each excitation and represents the number of ph pairs that contribute. The collectivity is a measure of the average rank of the transition density matrix (Luzanov and Zhikol, 2010), this gives the number of linearly independent columns or rows of a matrix. In the absence of electronic correlation, each excitation would correspond to a single ph pair and $\nu = 1$. In a correlated system, in general, an excitation will be a linear combination of ph pairs and $\nu > 1$. If all ph pairs need to be included to describe the excitation then $\nu = N_{occ}$, ν thus offers a convenient measure of electronic correlation. It has been used for instance to study collective charge-transfer excitons (Mukamel et al., 1997). Another potential measure of correlation effects between the ph pairs is provided by the recently introduced idea of measuring entanglement entropy (Plasser, 2016).

To extend the above discussion to cases with hp-ph coupling[3], such as time-dependent Hartree-Fock and Casida-TDDFT, involves some considerable effort; it is not straightforward to define the transition density matrix in this case (Etienne, 2015). We consider three alternative approaches:

[3] i.e. including de-excitation.

- Use the TDA, this involves throwing away the $\boldsymbol{\mathcal{Y}}$ eigenvector from the Casida super-matrix equation.
- The transition density matrix can be identified with either $\boldsymbol{\mathcal{Z}}_+$ or $\boldsymbol{\mathcal{Z}}_-$, see appendix I for details.
- A statistical mixture of the eigenvectors $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Y}}$, which was derived by (Luzanov and Zhikol, 2010).

The latter option is the most rigorous and gives

$$\nu_{xy} = \nu_x^{\sum_{ia} \mathcal{X}_{ia}^2} \nu_y^{\sum_{ia} \mathcal{Y}_{ia}^2}, \tag{K.8}$$

where ν_x and ν_y are the collectivity indices found from taking the transition density matrix equal to $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Y}}$ respectively. It was found that the collectivity calculated within the TDA did not compare well with other methods for the cases explored in this thesis, indicating that de-excitation is important. This is further confirmed by excitation spectra obtained within the TDA, they shows qualitative agreement with the Casida result but the TDA modes show a considerable blueshift, which is largest for the L-mode. Fortunately, the other two techniques agree well, for the

results in the main text we used the method provided by (Luzanov and Zhikol, 2010). A comparison of these two methods is shown in figure K.1 for the single Na_{10} chain. We observed that the statistical mixture method tended to give collectivities larger than the value derived from either $\mathcal{Z}_+$ or $\mathcal{Z}_-$.

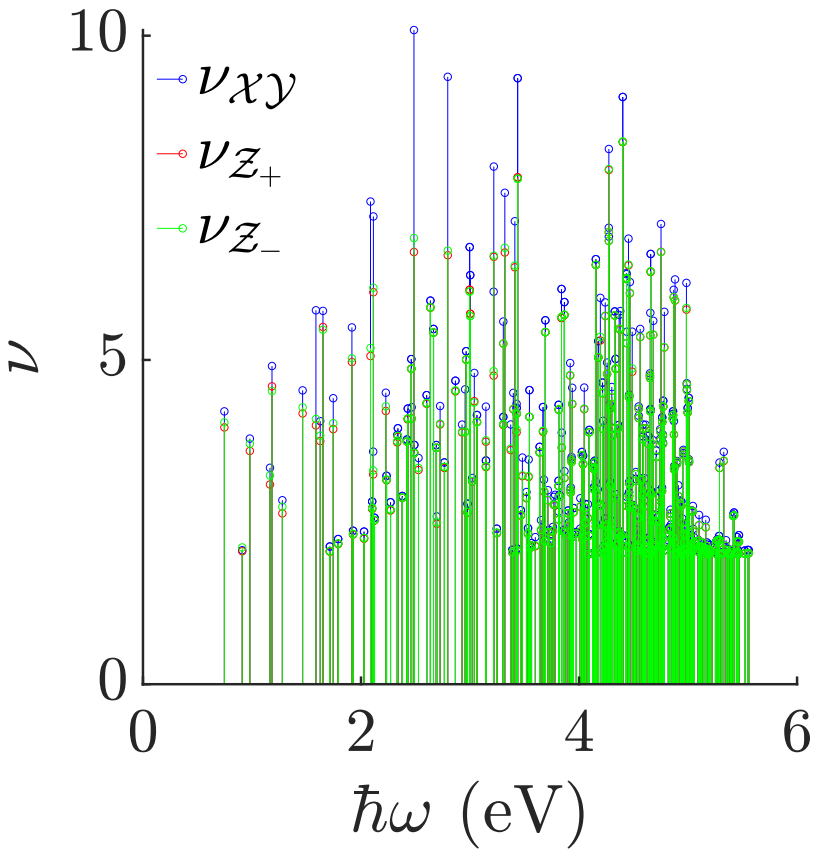


Figure K.1: A comparison of the three different ways to calculate the collectivity τ for each Casida excitation of the Na_{10} chain. $\tau_{\mathcal{XY}}$ is the statistical mixture method by (Luzanov and Zhikol, 2010), and $\tau_{\mathcal{Z}_+}$ and $\tau_{\mathcal{Z}_-}$ are the collectivities from taking $\mathcal{Z}_+$ and $\mathcal{Z}_+$ as the transition density matrix respectively. See appendix I for more details.

L Calculating The Field Enhancement From A TDDFT Calculation

The key quantity to calculating the FE is the induced density, from which the induced potential can be found

$$\begin{aligned} \nabla^2 \left(\Delta\phi(\mathbf{r},\omega)\right) &= e\frac{\Delta n(\mathbf{r},\omega)}{\epsilon_0} \\ \therefore \Delta\phi(\mathbf{r},\omega) &= \frac{-e}{4\pi\epsilon_0}\int d^3r' \, \frac{\Delta n(\mathbf{r}',\omega)}{|\mathbf{r}-\mathbf{r}'|}, \end{aligned} \tag{L.1}$$

and subsequently the induced electric field from the gradient

$$\Delta\mathbf{E}(\mathbf{r},\omega) = -\nabla\left(\Delta\phi(\mathbf{r},\omega)\right), \tag{L.2}$$

and finally the FE

$$FE(\mathbf{r},\omega) = \frac{|\mathbf{E}_{ext}(\mathbf{r},\omega) + \Delta\mathbf{E}(\mathbf{r},\omega)|}{|\mathbf{E}_{ext}(\mathbf{r},\omega)|}. \tag{L.3}$$

To calculate the FE we write the density matrix as[1] (Casida, 1995)

$$\Delta\rho_{ia} = \sum_{jb}\sqrt{\omega_{ai}}\frac{1}{\omega^2 - \mathcal{C}_{ia,jb}}\sqrt{\omega_{bj}}\Delta v_{jb}(\mathbf{r},\omega). \tag{L.4}$$

[1] In this appendix we will take the zero-temperature approximation and can ignore the Fermi functions, taking their effect into account via the usual summation convention.

We will rewrite this expression using the useful *spectral expansion*[2]

$$\frac{1}{\omega^2\mathbf{1} - \boldsymbol{\mathcal{C}}} = \sum_I \frac{1}{\Omega_I^2 - \omega^2}\boldsymbol{\mathcal{Z}}\boldsymbol{\mathcal{Z}}^\dagger. \tag{L.5}$$

[2] A tacit assumption that $\boldsymbol{\mathcal{C}}$ does not have a frequency dependence is used, this simplifies the normalisation of $\boldsymbol{\mathcal{Z}}$, see (Casida, 1995).

This gives

$$\Delta\rho_{ia}(\omega) = -\sum_I \frac{1}{\Omega_I^2 - \omega^2}\sum_{jb}\sqrt{\omega_{ai}}\mathcal{Z}^I_{ia}(\mathcal{Z}^I_{jb})^\dagger\sqrt{\omega_{bj}}\Delta v_{jb}(\mathbf{r},\omega). \tag{L.6}$$

The induced density can therefore be written as[3]

[3] For simplicity assuming real orbitals.

$$\begin{aligned} \Delta n(\mathbf{r},\omega) &= \sum_{ia}\psi_i(\mathbf{r})\psi_a(\mathbf{r})\Delta\rho_{ia}(\omega) \\ \therefore \Delta n(\mathbf{r},\omega) &= -\sum_I \frac{1}{\Omega_I^2 - \omega^2}\sum_{ia,jb}\psi_i(\mathbf{r})\psi_a(\mathbf{r})\sqrt{\omega_{ai}}\mathcal{Z}^I_{ia}(\mathcal{Z}^I_{jb})^\dagger\sqrt{\omega_{bj}}\Delta v_{jb}(\mathbf{r},\omega). \end{aligned} \tag{L.7}$$

Assuming the external perturbation wavelength is much bigger that the system under study, we can write in the dipole limit

$\Delta v_{jb}(\mathbf{r},\omega) = \mathbf{E}_{ext}\cdot\boldsymbol{\mu}_{ia}$. Also the finite lifetime term η is added by noting

$$\frac{1}{\Omega_I^2-\omega^2} = \lim_{\eta\to+0} -\frac{1}{2\Omega_I}\left(\frac{1}{\omega-\Omega_I+i\eta}-\frac{1}{\omega+\Omega_I+i\eta}\right), \tag{L.8}$$

and by writing as a combination of Lorentzians it is simple to identify that η is twice the linewidth γ. This completes the expression for the induced density

$$\Delta n(\mathbf{r},\omega) = \sum_I \frac{1}{2\Omega_I}\left(\frac{1}{\omega-\Omega_I+i\gamma/2}-\frac{1}{\omega+\Omega_I+i\gamma/2}\right)\times \sum_{ia,jb}\psi_i(\mathbf{r})\psi_a(\mathbf{r})\sqrt{\omega_{ai}}\mathcal{Z}^I_{ia}(\mathcal{Z}^I_{jb})^\dagger\sqrt{\omega_{bj}}\mathbf{E}_{ext}\cdot\boldsymbol{\mu}_{ia}, \tag{L.9}$$

but it can be rewritten in a more compact (and efficient) form by reordering the summation and identifying the transition density and transition dipole moment

$$\Delta n(\mathbf{r},\omega) = \sum_I \frac{1}{2}\left(\frac{1}{\omega-\Omega_I+i\gamma/2}-\frac{1}{\omega+\Omega_I+i\gamma/2}\right)\times \left(\sum_{ia}\frac{\psi_i(\mathbf{r})\psi_a(\mathbf{r})\sqrt{\omega_{ai}}\mathcal{Z}^I_{ia}}{\sqrt{\omega_I}}\right)\left(\sum_{jb}\frac{(\mathcal{Z}^I_{jb})^\dagger\sqrt{\omega_{bj}}\mathbf{E}_{ext}\cdot\boldsymbol{\mu}_{ia}}{\sqrt{\Omega_I}}\right) \tag{L.10}$$

giving the final full expression

$$\boxed{\Delta n(\mathbf{r},\omega) = \sum_I \frac{1}{2}\left(\frac{1}{\omega-\Omega_I+i\gamma/2}-\frac{1}{\omega+\Omega_I+i\gamma/2}\right)\Delta n^I(\mathbf{r})\boldsymbol{\mu}^I\cdot\mathbf{E}_{ext}.} \tag{L.11}$$

This agrees with the equation derived by (Rossi, 2013).

This equation takes into account contributions from *all* excitations, but we can achieve a considerable simplification if we assume we are at a frequency close to an excitation $\omega \approx \Omega_I$. Ignoring off-resonant terms and Taylor expanding gives

$$\boxed{\Delta n(\mathbf{r},\omega\approx\Omega_I) \approx \frac{1}{i\gamma}\Delta n^I(\mathbf{r})\boldsymbol{\mu}^I\cdot\mathbf{E}_{ext}.} \tag{L.12}$$

This equation has also been derived previously by Cocchi *et al* directly from the interacting density-density response function (Cocchi et al., 2012).

M Classical Fit For The Sodium Chain

The classical fit to the Na_{10} chain discussed in chapter 5 was obtained using a prolate ellipsoid model, see equations 1.38 and 1.40. Following Yan and Gao, we define $R_1 = \Delta + (N-1)\frac{d}{2}$ and $R_2 = R_3 = \Delta$, where $d = 3.08$ Å is the bond length used in the simulation and Δ is a measure of the electron spill-out at the chain edges, which is used as a fitting parameter (Yan and Gao, 2008). Yan and Gao used a fixed bulk value for the plasma frequency in the Drude dielectric function, effectively adding another fitting parameter. Instead we define a chain size-dependent plasma frequency, this seems more appropriate given the small size of the systems

$$\begin{aligned}\omega_P^2 = \frac{ne^2}{\epsilon_0 m} = \frac{Ne^2}{V\epsilon_0 m} = \frac{3Ne^2}{4\pi R_1 R_2 R_3 \epsilon_0 m} \\ = \frac{3Ne^2}{4\pi(\Delta + (N-1)d/2)\Delta^2 \epsilon_0 m},\end{aligned} \tag{M.1}$$

meaning we need only fit Δ. For large numbers of electrons, $N \to \infty$, the plasma frequency becomes independent of N as would be expected. Once Δ has been fitted, the depolarization factor L_1 can be calculated, see equation 1.40. L_2 can also be calculated via the relation: $L_1 + 2L_2 = 1$. From the depolarization factors the mode energies can immediately be found. Applying this procedure, the results shown in figure 5.6 are obtained, with a value of $\Delta = 1.45d$ for the fitting parameter.

If instead we take ω_P to be fixed at the bulk value[1], then a best fit of $\Delta = 1.13d$ is found. The absorption cross section and the results of the fit for the L-mode energy are shown in figure M.1(a) and (b) respectively. Another potential method is to consider the ratio of the T-mode and L-mode energies, $\frac{\omega_T}{\omega_L}$, which is convenient as the plasma frequency does not need to be given an explicit value. It is simple to show that

$$\frac{\omega_T}{\omega_L} = \sqrt{\frac{1-L_1}{2L_1}}, \tag{M.2}$$

which can be compared to the TDDFT results. The fitting parameter is obtained for the largest chain and is found to be $\Delta = 0.79d$, the results are shown in figure M.1(c). In figure M.1(d) a cross sectional view of the ground state density, along a short axis and calculated

[1] We choose $\omega_P = 3.83$ eV as used by (Yan and Gao, 2008).

by DFT, is shown along with the radii for the fitting parameter Δ from the three different methods discussed.

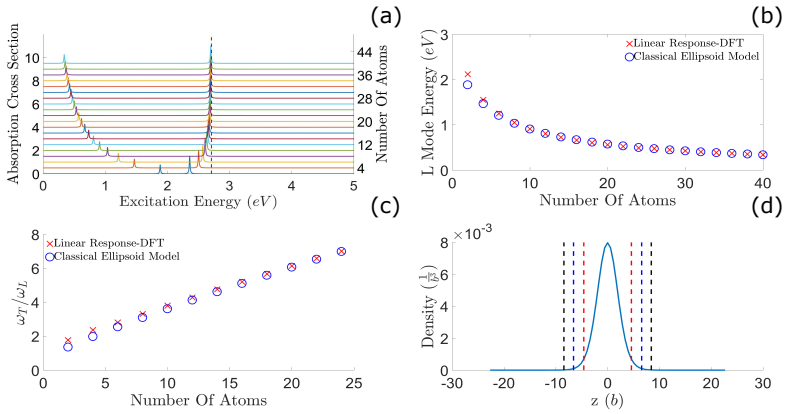


Figure M.1: (a) Normalized absorption cross section calculated using the classical prolate ellipsoid model for Na chains ranging from 2 to 40 atoms. The dashed blue line indicates the transverse plasmon frequency using $\omega_P = 3.83$ eV. (b) Fit for the longitudinal plasmon energy using a bulk plasma frequency, $\omega_P = 3.83$ eV, with $\Delta = 1.13d$. (c) Ratio of the T-mode and L-mode energies calculated from the Casida and the classical ellipsoid method. (d) DFT calculated ground state density of a Na_{10} chain for $x = y = 0$, with radii corresponding to fitting parameters $\Delta = 1.45d,\ 1.13d$ and $0.79d$ indicated by the black, blue and red dashed lines respectively.

Acronyms

ALDA adiabatic local density approximation. 79, 81, 82, 120

BDP bonding dipole plasmon. 60, 61, 67, 132–134

BEM boundary element method. 39, 40, 46–49, 55, 70, 103, 105, 106, 114

CIS configuration interaction singles. 82, 170, 175

CTP charge transfer plasmon. 60, 61, 132–134

DFPT density-functional perturbation theory. 87, 121, 128

DFT density-functional theory. 11, 21, 25, 34, 46, 58, 60, 65, 69, 72, 74–78, 82, 85, 107, 119, 122, 123, 182

EELS electron energy loss spectroscopy. 27, 54, 60

eh electron-hole. 11, 26, 40, 79, 81, 82, 84, 119, 120, 126, 132

EM electromagnetic. 9, 16–18, 22, 26, 27, 29, 32, 34–36, 38, 41–43, 48, 52, 55, 58, 65, 128, 135, 141, 143, 146, 156

FDTD finite difference time domain. 20, 21, 46–48, 137

FE field enhancement. 10, 12, 35, 36, 44, 54, 55, 58, 60, 62, 71, 82, 83, 90, 94, 96, 100–103, 105–117, 119–121, 127–130, 132–135, 137–139, 152, 179

GMM generalised multiparticle Mie. 46, 49, 105, 114

GNOR generalized nonlocal optical response. 67, 68

HM hydrodynamic model. 11, 60, 64–70, 106, 107, 117, 162

HOMO highest occupied molecular orbital. 61, 74, 124, 126, 133

hp hole-particle. 82, 175, 177

IR infrared. 12, 21, 29, 30, 35, 52, 53, 56, 57, 89, 90, 100, 110, 114, 120, 137

KS Kohn-Sham. 73–75, 77–84, 122–125, 131, 134, 165–167, 169, 170, 175

L-mode longitudinal mode. 39, 48, 49, 123–129, 131–133, 177, 181, 182

LAM local analogue model. 69, 70, 106

LDA local density approximation. 25, 66, 68, 74–76, 79, 85, 106, 118, 131

LO longitudinal optic. 55, 56, 93, 98

LSP localised surface plasmon. 32–36, 38–41, 43–45, 55, 58, 64, 68, 69, 90, 101, 106, 108, 110, 113, 121, 149, 159

LSPhP localised surface phonon polariton. 89, 90, 93, 98, 110, 113

LUMO lowest occupied molecular orbital. 61, 124, 126, 133

ph particle-hole. 82, 175–177

QED quantum electrodynamics. 91, 135

RPA random phase approximation. 26, 27, 66, 71, 81, 82, 85, 92, 169

SERS surface-enhanced Raman spectroscopy. 12, 43, 62, 115, 120

SP surface plasmon. 25, 27–32, 34, 52, 60, 62, 145–147, 150

SPhP surface phonon polariton. 13, 55–57

SPP surface plasmon polariton. 11, 28–32, 35, 41, 52–54, 89, 90, 93–96, 98, 108, 143–147

T-mode transverse mode. 39, 48, 49, 123, 125, 127, 132, 181, 182

TC-mode transverse central mode. 123–131

TDA Tamm-Dancoff approximation. 82, 83, 175, 177

TDDFT time-dependent density-functional theory. 11, 65, 70–72, 74, 76–80, 82–86, 106–108, 117, 119, 120, 127, 135, 169, 170, 177, 181

TE-mode transverse end mode. 123, 125–127, 131

TM transverse magnetic. 29, 53, 54, 89, 90, 143

TO transverse optic. 55, 56, 95

UV ultraviolet. 17, 19, 22, 120, 137

VSH vector spherical harmonic. 49, 153, 155–157

XC exchange-correlation. 66, 68, 72–75, 78–83, 118, 122, 123, 161, 167, 169

Bibliography

Alessandro Alabastri, Xiao Yang, Alejandro Manjavacas, Henry O Everitt, and Peter Nordlander. Extraordinary light-induced local angular momentum near metallic nanoparticles. *ACS nano*, 10(4):4835–4846, 2016.

Philip B Allen. Single particle versus collective electronic excitations. In *From Quantum Mechanics to Technology*, pages 125–141. Springer, 1996.

E Altewischer, MP Van Exter, and JP Woerdman. Plasmon-assisted transmission of entangled photons. *Nature*, 418(6895):304, 2002.

Andrea Alù and Nader Engheta. Optical nanotransmission lines: synthesis of planar left-handed metamaterials in the infrared and visible regimes. *JOSA B*, 23(3):571–583, 2006.

Andrea Alu, Mário G Silveirinha, Alessandro Salandrino, and Nader Engheta. Epsilon-near-zero metamaterials and electromagnetic sources: Tailoring the radiation phase pattern. *Physical review B*, 75(15):155410, 2007.

Marcos M Alvarez, Joseph T Khoury, T Gregory Schaaff, Marat N Shafigullin, Igor Vezmar, and Robert L Whetten. Optical absorption spectra of nanocrystal gold molecules. *The Journal of Physical Chemistry B*, 101(19):3706–3712, 1997.

AT Amos and GG Hall. Single determinant wave functions. *Proc. R. Soc. Lond. A*, 263 (1315):483–493, 1961.

M Anderegg, B Feuerbacher, and B Fitton. Optically excited longitudinal plasmons in potassium. *Physical Review Letters*, 27(23):1565, 1971.

Philip Warren Anderson. *Basic notions of condensed matter physics*. Benjamin-Cummings, 1984.

Xavier Andrade, David Strubbe, Umberto De Giovannini, Ask Hjorth Larsen, Micael JT Oliveira, Joseba Alberdi-Rodriguez, Alejandro Varas, Iris Theophilou, Nicole Helbig, Matthieu J Verstraete, et al. Real-space grids and the octopus code as tools for the development of new simulation approaches for electronic systems. *Physical Chemistry Chemical Physics*, 17(47):31371–31396, 2015.

Jeffrey N Anker, W Paige Hall, Olga Lyandres, Nilam C Shah, Jing Zhao, and Richard P Van Duyne. Biosensing with plasmonic nanosensors. *Nature materials*, 7(6): 442–453, 2008.

George Arfken. *Mathematical Methods for Physicists*. Academic Press, Inc., 1985.

NW Ashcroft and ND Mermin. *Solid State Physics*. Cengage Learning, 1976.

Harry A Atwater and Albert Polman. Plasmonics for improved photovoltaic devices. *Nature materials*, 9(3):205–213, 2010.

Baptiste Auguié, José Lorenzo Alonso-Gómez, Andrés Guerrero-Martínez, and Luis M Liz-Marzán. Fingers crossed: Optical activity of a chiral dimer of plasmonic nanorods. *The Journal of Physical Chemistry Letters*, 2(8):846–851, 2011.

Guillaume Baffou and Romain Quidant. Thermo-plasmonics: using metallic nanostructures as nano-sources of heat. *Laser & Photonics Reviews*, 7(2):171–187, 2013.

Reuben M Bakker, Dmitry Permyakov, Ye Feng Yu, Dmitry Markovich, Ramón Paniagua-Domínguez, Leonard Gonzaga, Anton Samusev, Yuri Kivshar, Boris Luk'yanchuk, and Arseniy I Kuznetsov. Magnetic and electric hotspots with silicon nanodimers. *Nano Letters*, 15(3):2137–2142, 2015.

R v Baltz. Plasmons and surface plasmons in bulk metals, metallic clusters, and metallic heterostructures. In *Spectroscopy and Dynamics of Collective Excitations in Solids*, pages 303–338. Springer, 1997.

M Barbry, P Koval, F Marchesin, R Esteban, AG Borisov, J Aizpurua, and D Sánchez-Portal. Atomistic near-field nanoplasmonics: reaching atomic-scale resolution in nanooptics. *Nano letters*, 15(5):3410–3419, 2015.

Ingrid D Barcelos, Alisson R Cadore, Leonardo C Campos, Angelo Malachias, K Watanabe, T Taniguchi, Francisco CB Maia, Raul Freitas, and Christoph Deneke. Graphene/h-bn plasmon–phonon coupling and plasmon delocalization observed by infrared nano-spectroscopy. *Nanoscale*, 7(27):11620–11625, 2015.

William Barnes. Quantum physics: Survival of the entangled. *Nature*, 418(6895): 281–282, 2002.

William L Barnes, Alain Dereux, and Thomas W Ebbesen. Surface plasmon subwavelength optics. *Nature*, 424(6950):824–830, 2003.

Stefano Baroni, Stefano De Gironcoli, Andrea Dal Corso, and Paolo Giannozzi. Phonons and related crystal properties from density-functional perturbation theory. *Reviews of Modern Physics*, 73(2):515, 2001.

Rüdiger Bauernschmitt and Reinhart Ahlrichs. Treatment of electronic excitations within the adiabatic approximation of time dependent density functional theory. *Chemical Physics Letters*, 256(4-5):454–464, 1996.

Felix Benz, Christos Tserkezis, Lars O Herrmann, Bart De Nijs, Alan Sanders, Daniel O Sigle, Laurynas Pukenas, Stephen D Evans, Javier Aizpurua, and Jeremy J Baumberg. Nanooptics of molecular-shunted plasmonic nanojunctions. *Nano letters*, 15(1):669–674, 2014.

A Bergara, JB Neaton, and NW Ashcroft. Ferromagnetic instabilities in atomically thin lithium and sodium wires. *International journal of quantum chemistry*, 91(2): 239–244, 2003.

David J Bergman and Mark I Stockman. Surface plasmon amplification by stimulated emission of radiation: quantum generation of coherent surface plasmons in nanosystems. *Physical review letters*, 90(2):027402, 2003.

Stephan Bernadotte, Ferdinand Evers, and Christoph R Jacob. Plasmons in molecules. *The Journal of Physical Chemistry C*, 117(4):1863–1878, 2013.

Sébastien Bidault, F. Javier García de Abajo, and Albert Polman. Plasmon-based nanolenses assembled on a well-defined dna template. *Journal of the American Chemical Society*, 130(9):2750–2751, 2008.

Felix Bloch. Bremsvermögen von atomen mit mehreren elektronen. *Zeitschrift für Physik A Hadrons and Nuclei*, 81(5):363–376, 1933.

A Boardman. Electromagnetic surface modes. hydrodynamic theory of plasmon-polaritons on plane surfaces, 1982.

Craig F Bohren and Donald R Huffman. *Absorption and scattering of light by small particles*. John Wiley & Sons, 2008.

Alexandra Boltasseva and Harry A Atwater. Low-loss plasmonic metamaterials. *Science*, 331(6015):290–291, 2011.

Tim J Booth, Peter Blake, Rahul R Nair, Da Jiang, Ernie W Hill, Ursel Bangert, Andrew Bleloch, Mhairi Gass, Kostya S Novoselov, Mikhail I Katsnelson, et al. Macroscopic graphene membranes and their extraordinary stiffness. *Nano letters*, 8 (8):2442–2446, 2008.

Matthias Brack. The physics of simple metal clusters: self-consistent jellium model and semiclassical approaches. *Reviews of modern physics*, 65(3):677, 1993.

Victor W Brar, Min Seok Jang, Michelle Sherrott, Josue J Lopez, and Harry A Atwater. Highly confined tunable mid-infrared plasmonics in graphene nanoresonators. *Nano letters*, 13(6):2541–2547, 2013.

Victor W Brar, Min Seok Jang, Michelle Sherrott, Seyoon Kim, Josue J Lopez, Laura B Kim, Mansoo Choi, and Harry Atwater. Hybrid surface-phonon-plasmon polariton modes in graphene/monolayer h-bn heterostructures. *Nano letters*, 14(7): 3876–3880, 2014.

Ricardo A Broglia, Gianluca Colò, Giovanni Onida, and H Eduardo Roman. *Solid state physics of finite systems: metal clusters, fullerenes, atomic wires*. Springer Science & Business Media, 2013.

Garnett W Bryant, F Javier García de Abajo, and Javier Aizpurua. Mapping the plasmon resonances of metallic nanoantennas. *Nano letters*, 8(2):631–636, 2008.

Luca Bursi, Arrigo Calzolari, Stefano Corni, and Elisa Molinari. Light-induced field enhancement in nanoscale systems from first-principles: the case of polyacenes. *ACS Photonics*, 1(10):1049–1058, 2014.

Luca Bursi, Arrigo Calzolari, Stefano Corni, and Elisa Molinari. Quantifying the plasmonic character of optical excitations in nanostructures. *ACS Photonics*, 3(4): 520–525, 2016.

Joshua D Caldwell, Orest J Glembocki, Yan Francescato, Nicholas Sharac, Vincenzo Giannini, Francisco J Bezares, James P Long, Jeffrey C Owrutsky, Igor Vurgaftman, Joseph G Tischler, et al. Low-loss, extreme subdiffraction photon confinement via silicon carbide localized surface phonon polariton resonators. *Nano letters*, 13(8): 3690–3697, 2013.

Joshua D Caldwell, Lucas Lindsay, Vincenzo Giannini, Igor Vurgaftman, Thomas L Reinecke, Stefan A Maier, and Orest J Glembocki. Low-loss, infrared and terahertz nanophotonics using surface phonon polaritons. *Nanophotonics*, 4(1):44–68, 2015.

Joshua D Caldwell, Igor Vurgaftman, Joseph G Tischler, Orest J Glembocki, Jeffrey C Owrutsky, and Thomas L Reinecke. Atomic-scale photonic hybrids for mid-infrared and terahertz nanophotonics. *Nature nanotechnology*, 11(1):9, 2016.

Javier Cambiasso. *Light-matter interactions in lossy and lossless media*. PhD thesis, 2017.

Hu Cang, Anna Labno, Changgui Lu, Xiaobo Yin, Ming Liu, Christopher Gladden, Yongmin Liu, and Xiang Zhang. Probing the electromagnetic field of a 15-nanometre hotspot by single molecule imaging. *Nature*, 469(7330):385–388, 2011.

Mark E Casida. Time-dependent density functional response theory for molecules. In *Recent Advances In Density Functional Methods: (Part I)*, pages 155–192. World Scientific, 1995.

Andres Castellanos-Gomez. Why all the fuss about 2d semiconductors? *Nature Photonics*, 10(4):202, 2016.

Alberto Castro, Miguel AL Marques, Julio A Alonso, and Angel Rubio. Optical properties of nanostructures from time-dependent density functional theory. *Journal of Computational and Theoretical Nanoscience*, 1(3):231–255, 2004.

Alberto Castro, Heiko Appel, Micael Oliveira, Carlo A Rozzi, Xavier Andrade, Florian Lorenzen, Miguel AL Marques, EKU Gross, and Angel Rubio. Octopus: a tool for the application of time-dependent density functional theory. *physica status solidi (b)*, 243(11):2465–2488, 2006.

Hoon Cha, Jun Hee Yoon, and Sangwoon Yoon. Probing quantum plasmon coupling using gold nanoparticle dimers with tunable interparticle distances down to the subnanometer range. *ACS nano*, 8(8):8554–8563, 2014.

Chi-Fan Chen, Cheol-Hwan Park, Bryan W Boudouris, Jason Horng, Baisong Geng, Caglar Girit, Alex Zettl, Michael F Crommie, Rachel A Segalman, Steven G Louie, et al. Controlling inelastic light scattering quantum pathways in graphene. *Nature*, 471(7340):617, 2011.

Jianing Chen, Michela Badioli, Pablo Alonso-González, Sukosin Thongrattanasiri, Florian Huth, Johann Osmond, Marko Spasenović, Alba Centeno, Amaia Pesquera, Philippe Godignon, et al. Optical nano-imaging of gate-tunable graphene plasmons. *Nature*, 487(7405):77, 2012.

Junsheng Chen, Alexey Krasavin, Pavel Ginzburg, Anatoly V Zayats, Tõnu Pullerits, and Khadga Jung Karki. Evidence of high-order nonlinearities in supercontinuum white-light generation from a gold nanofilm. *ACS Photonics*, 5(5): 1927–1932, 2018.

Yiguo Chen, Yan Francescato, Joshua D Caldwell, Vincenzo Giannini, Tobias WW Maß, Orest J Glembocki, Francisco J Bezares, Thomas Taubner, Richard Kasica, Minghui Hong, et al. Spectral tuning of localized surface phonon polariton resonators for low-loss mid-ir applications. *Acs Photonics*, 1(8):718–724, 2014.

Rohit Chikkaraddy, Bart de Nijs, Felix Benz, Steven J Barrow, Oren A Scherman, Edina Rosta, Angela Demetriadou, Peter Fox, Ortwin Hess, and Jeremy J Baumberg. Single-molecule strong coupling at room temperature in plasmonic nanocavities. *Nature*, 535(7610):127–130, 2016.

CR Chris Wang, Stuart Pollack, Douglas Cameron, and Manfred M Kappes. Optical absorption spectroscopy of sodium clusters as measured by collinear molecular beam photodepletion. *The Journal of Chemical Physics*, 93(6):3787–3801, 1990.

Ming-Wen Chu, Viktor Myroshnychenko, Cheng Hsuan Chen, Jing-Pei Deng, Chung-Yuan Mou, and F Javier García de Abajo. Probing bright and dark surface-plasmon modes in individual and coupled noble metal nanoparticles using an electron beam. *Nano letters*, 9(1):399–404, 2008.

C Ciracì, RT Hill, JJ Mock, Y Urzhumov, AI Fernández-Domínguez, SA Maier, JB Pendry, A Chilkoti, and DR Smith. Probing the ultimate limits of plasmonic enhancement. *Science*, 337(6098):1072–1074, 2012.

Caterina Cocchi, Deborah Prezzi, Alice Ruini, Enrico Benassi, Marilia J Caldas, Stefano Corni, and Elisa Molinari. Optical excitations and field enhancement in short graphene nanoribbons. *The journal of physical chemistry letters*, 3(7):924–929, 2012.

Tyler L Cocker, Devin Baillie, Miles Buruma, Lyubov V Titova, Richard D Sydora, Frank Marsiglio, and Frank A Hegmann. Microscopic origin of the drude-smith model. *Physical Review B*, 96(20):205439, 2017.

Maria Laura Coluccio, Francesco Gentile, Gobind Das, Annalisa Nicastri, Angela Mena Perri, Patrizio Candeloro, Gerardo Perozziello, Remo Proietti Zaccaria, Juan Sebastian Totero Gongora, Salma Alrasheed, Andrea Fratalocchi, Tania Limongi, Giovanni Cuda, and Enzo Di Fabrizio. Detection of single amino acid mutation in human breast cancer by disordered plasmonic self-similar chain. *Science advances*, 1(8):e1500487, 2015.

G Contreras, AK Sood, and M Cardona. Raman scattering by intervalley carrier-density fluctuations in n-type si: Intervalley and intravalley mechanisms. *Physical Review B*, 32(2):924, 1985.

Emiliano Cortés, Wei Xie, Javier Cambiasso, Adam S Jermyn, Ravishankar Sundararaman, Prineha Narang, Sebastian Schlücker, and Stefan A Maier. Plasmonic hot electron transport drives nano-localized chemistry. *Nature Communications*, 8, 2017.

Joel D Cox and F Javier García De Abajo. Electrically tunable nonlinear plasmonics in graphene nanoislands. *Nature communications*, 5:5725, 2014.

Yao Cui, Adam Lauchner, Alejandro Manjavacas, F Javier García de Abajo, Naomi J Halas, and Peter Nordlander. Molecular plasmon–phonon coupling. *Nano letters*, 16 (10):6390–6395, 2016.

Jianhua Dai, Frantisek Čajko, Igor Tsukerman, and Mark I Stockman. Electrodynamic effects in plasmonic nanolenses. *Physical Review B*, 77(11):115419, 2008.

S Dai, Z Fei, Q Ma, AS Rodin, M Wagner, AS McLeod, MK Liu, W Gannett, W Regan, K Watanabe, et al. Tunable phonon polaritons in atomically thin van der waals crystals of boron nitride. *Science*, 343(6175):1125–1129, 2014.

S Dai, Q Ma, MK Liu, T Andersen, Z Fei, MD Goldflam, M Wagner, K Watanabe, T Taniguchi, M Thiemens, et al. Graphene on hexagonal boron nitride as a tunable hyperbolic metamaterial. *Nature nanotechnology*, 10(8):nnano–2015, 2015.

Gobind Das, Salma Alrasheed, Maria Laura Coluccio, Francesco Gentile, Annalisa Nicastri, Patrizio Candeloro, Giovanni Cuda, Gerardo Perozziello, and Enzo Di Fabrizio. Few molecule sers detection using nanolens based plasmonic nanostructure: application to point mutation detection. *RSC Adv.*, 6:107916–107923, 2016.

Basab B Dasgupta. Surface plasmon dispersion for very small metallic spheres: A quantum mechanical formulation. *Zeitschrift für Physik B Condensed Matter*, 27(1): 75–79, 1977.

Christin David and F Javier García de Abajo. Spatial nonlocality in the optical response of metal nanoparticles. *The Journal of Physical Chemistry C*, 115(40): 19470–19475, 2011.

TJ Davis and DE Gómez. Colloquium: An algebraic model of localized surface plasmons and their interactions. *Reviews of Modern Physics*, 89(1):011003, 2017.

FJ García De Abajo and A Howie. Retarded field calculation of electron energy loss in inhomogeneous dielectrics. *Physical Review B*, 65(11):115418, 2002.

Walt A De Heer. The physics of simple metal clusters: experimental aspects and simple models. *Reviews of Modern Physics*, 65(3):611, 1993.

Walt A de Heer, Kathy Selby, Vitaly Kresin, Jun Masui, Michael Vollmer, A Chatelain, and WD Knight. Collective dipole oscillations in small sodium clusters. *Physical review letters*, 59(16):1805, 1987.

Sandra de Vega and F Javier García de Abajo. Plasmon generation through electron tunneling in graphene. *ACS Photonics*, 4(9):2367–2375, 2017.

Giuliana Di Martino, Yannick Sonnefraud, Mark S Tame, Stéphane Kéna-Cohen, F Dieleman, ŞK Özdemir, MS Kim, and Stefan A Maier. Observation of quantum interference in the plasmonic hong-ou-mandel effect. *Physical Review Applied*, 1(3): 034004, 2014.

Baoquan Ding, Zhengtao Deng, Hao Yan, Stefano Cabrini, Ronald N Zuckermann, and Jeffrey Bokor. Gold nanoparticle self-similar chain structure organized by dna origami. *Journal of the American Chemical Society*, 132(10):3248–3249, 2010.

JA Dionne, LA Sweatlock, HA Atwater, and A Polman. Planar metal plasmon waveguides: frequency-dependent dispersion, propagation, localization, and loss beyond the free electron model. *Physical Review B*, 72(7):075405, 2005.

Greta Donati, David B Lingerfelt, Christine M Aikens, and Xiaosong Li. Molecular vibration induced plasmon decay. *The Journal of Physical Chemistry C*, 121(28): 15368–15374, 2017.

Bruce T Draine. The discrete-dipole approximation and its application to interstellar graphite grains. *The Astrophysical Journal*, 333:848–872, 1988.

Andreas Dreuw and Martin Head-Gordon. Single-reference ab initio methods for the calculation of excited states of large molecules. *Chemical reviews*, 105(11):4009–4037, 2005.

Luping Du, Dingyuan Tang, and Xiaocong Yuan. Edge-reflection phase directed plasmonic resonances on graphene nano-structures. *Optics Express*, 22(19):22689–22698, 2014.

Adam D Dunkelberger, Chase T Ellis, Daniel C Ratchford, Alexander J Giles, Mijin Kim, Chul Soo Kim, Bryan T Spann, Igor Vurgaftman, Joseph G Tischler, James P Long, et al. Active tuning of surface phonon polariton resonances via carrier photoinjection. *Nature Photonics*, 12(1):50, 2018.

Thomas W Ebbesen, H J Lezec, H F Ghaemi, Tineke Thio, and PA Wolff. Extraordinary optical transmission through sub-wavelength hole arrays. *Nature*, 391(6668): 667, 1998.

Eleftherios N Economou. *Green's functions in quantum physics*, volume 3. Springer, 1983.

H Ehrenreich and Morrel H Cohen. Self-consistent field approach to the many-electron problem. *Physical Review*, 115(4):786, 1959.

W Ekardt. Work function of small metal particles: Self-consistent spherical jellium-background model. *Physical Review B*, 29(4):1558, 1984a.

W Ekardt. Dynamical polarizability of small metal particles: self-consistent spherical jellium background model. *Physical review letters*, 52(21):1925, 1984b.

W Ekardt. Size-dependent photoabsorption and photoemission of small metal particles. *Physical Review B*, 31(10):6360, 1985.

Alexandros Emboras, Jens Niegemann, Ping Ma, Christian Haffner, Andreas Pedersen, Mathieu Luisier, Christian Hafner, Thomas Schimmel, and Juerg Leuthold. Atomic scale plasmonic switch. *Nano letters*, 16(1):709–714, 2015.

J e Enkovaara, Carsten Rostgaard, J Jørgen Mortensen, Jingzhe Chen, M Dułak, Lara Ferrighi, Jeppe Gavnholt, Christian Glinsvad, V Haikola, HA Hansen, et al. Electronic structure calculations with gpaw: a real-space implementation of the projector augmented-wave method. *Journal of Physics: Condensed Matter*, 22(25): 253202, 2010.

Ruben Esteban, Andrei G Borisov, Peter Nordlander, and Javier Aizpurua. Bridging quantum and classical plasmonics with a quantum-corrected model. *Nature communications*, 3:825, 2012.

Pablo G Etchegoin, EC Le Ru, and M Meyer. An analytic model for the optical properties of gold. *The Journal of chemical physics*, 125(16):164705, 2006.

Thibaud Etienne. Transition matrices and orbitals from reduced density matrix theory. *The Journal of chemical physics*, 142(24):244103, 2015.

LA Falkovsky. Optical properties of graphene. In *Journal of Physics: Conference Series*, volume 129, page 012004. IOP Publishing, 2008.

Yurui Fang and Xiaorui Tian. Resonant surface plasmons of a metal nanosphere can be considered in the way of propagating surface plasmons. *arXiv preprint arXiv:1412.2664*, 2014.

Jacob A Faucheaux, Alexandria LD Stanton, and Prashant K Jain. Plasmon resonances of semiconductor nanocrystals: physical principles and new opportunities. *The journal of physical chemistry letters*, 5(6):976–985, 2014.

Zhe Fei, Gregory O Andreev, Wenzhong Bao, Lingfeng M Zhang, Alexander S. McLeod, Chen Wang, Margaret K Stewart, Zeng Zhao, Gerardo Dominguez, Mark Thiemens, et al. Infrared nanoscopy of dirac plasmons at the graphene–sio2 interface. *Nano letters*, 11(11):4701–4705, 2011.

Zhe Fei, AS Rodin, GO Andreev, W Bao, AS McLeod, M Wagner, LM Zhang, Z Zhao, M Thiemens, G Dominguez, et al. Gate-tuning of graphene plasmons revealed by infrared nano-imaging. *Nature*, 487(7405):82, 2012.

Peter J Feibelman. Surface electromagnetic fields. *Progress in Surface Science*, 12(4): 287–407, 1982.

Alexander L Fetter and John Dirk Walecka. *Quantum theory of many-particle systems*. Courier Corporation, 2012.

Richard Phillips Feynman. Forces in molecules. *Physical Review*, 56(4):340, 1939.

RP Feynman. *Statistical Mechanics, A Set of Lectures, California, Institute of Technology*. WA Benjamin, Inc. Advanced Book Program Reading, Massachusets, 1972.

Jamie M Fitzgerald and Vincenzo Giannini. Perspective on molecular quantum plasmonic nanoantennas. *J. Opt*, 19(060401):060401, 2017.

Jamie M Fitzgerald and Vincenzo Giannini. Battling retardation and nonlocality: The hunt for the ultimate plasmonic cascade nanolens. *ACS Photonics*, 5(6):2459–2467, 2018.

Jamie M Fitzgerald, Prineha Narang, Richard V Craster, Stefan A Maier, and Vincenzo Giannini. Quantum plasmonics. *Proceedings of the IEEE*, 104(12):2307–2322, 2016.

Jamie M Fitzgerald, Sam Azadi, and Vincenzo Giannini. Quantum plasmonic nanoantennas. *Physical Review B*, 95(23):235414, 2017.

Benjamin Foerster, Anneli Joplin, Katharina Kaefer, Sirin Celiksoy, Stephan Link, and Carsten Sönnichsen. Chemical interface damping depends on electrons reaching the surface. *ACS nano*, 11(3):2886–2893, 2017.

Mark Fox. *Optical properties of solids*. OUP Oxford, 2002.

Yan Francescato. *New frequencies and geometries for plasmonics and metamaterials*. PhD thesis, Imperial College London, 2014.

Shiwu Gao and Zhe Yuan. Emergence of collective plasmon excitation in a confined one-dimensional electron gas. *Physical Review B*, 72(12):121406, 2005.

Vincenzo Giannini, Antonio I Fernández-Domínguez, Susannah C Heck, and Stefan A Maier. Plasmonic nanoantennas: fundamentals and their use in controlling the radiative properties of nanoemitters. *Chemical reviews*, 111(6):3888–3912, 2011.

Rebecca L Gieseking, Mark A Ratner, and George C Schatz. Quantum mechanical identification of quadrupolar plasmonic excited states in silver nanorods. *The Journal of Physical Chemistry A*, 120(46):9324–9329, 2016a.

Rebecca L Gieseking, Mark A Ratner, and George C Schatz. Semiempirical modeling of ag nanoclusters: New parameters for optical property studies enable determination of double excitation contributions to plasmonic excitation. *The Journal of Physical Chemistry A*, 120(26):4542–4549, 2016b.

Pavel Ginzburg and Anatoly V Zayats. Localized surface plasmon resonances in spatially dispersive nano-objects: phenomenological treatise. *ACS nano*, 7(5): 4334–4342, 2013.

Patricia Gisbert-Quilis, Matteo Masetti, Judit Morla-Folch, Jamie M Fitzgerald, Nicolas Pazos-Perez, Eduardo Garcia-Rico, Vincenzo Giannini, Ramon A Alvarez-Puebla, and Luca Guerrini. The structure of short and genomic dna at the interparticle junctions of cationic nanoparticles. *Advanced Materials Interfaces*, 4(17), 2017.

Gabriele Giuliani and Giovanni Vignale. *Quantum theory of the electron liquid*. Cambridge university press, 2005.

Morten N Gjerding, Mohnish Pandey, and Kristian S Thygesen. Band structure engineered layered metals for low-loss plasmonics. *Nature communications*, 8:15133, 2017.

Michael D Goldflam, Guang-Xin Ni, Kirk W Post, Zhe Fei, Yuting Yeo, Jun You Tan, Aleksandr S Rodin, Brian C Chapler, Barbaros özyilmaz, Antonio H Castro Neto, et al. Tuning and persistent switching of graphene plasmons on a ferroelectric substrate. *Nano letters*, 15(8):4859–4864, 2015.

Cláudia Gomes Silva, Raquel Juárez, Tiziana Marino, Raffaele Molinari, and Hermenegildo García. Influence of excitation wavelength (uv or visible light) on the photocatalytic activity of titania containing gold nanoparticles for the generation of hydrogen or oxygen from water. *Journal of the American Chemical Society*, 133(3): 595–602, 2010.

Paulo André Dias Gonçalves and Nuno MR Peres. *An introduction to graphene plasmonics*. World Scientific, 2016.

Chen Gong and Marina S Leite. Noble metal alloys for plasmonics. *ACS Photonics*, 3 (4):507–513, 2016.

LP Gor'Kov and GM Eliashberg. Minute metallic particles in an electromagnetic field. *Sov. Phys. JETP*, 21(940), 1965.

Dmitri K Gramotnev and Sergey I Bozhevolnyi. Plasmonics beyond the diffraction limit. *Nature photonics*, 4(2):83–91, 2010.

Jean-Jacques Greffet, Remi Carminati, Karl Joulain, Jean-Philippe Mulet, Stephane Mainguy, and Yong Chen. Coherent emission of light by thermal sources. *Nature*, 416(6876):61, 2002.

D. J. Griffiths. *Introduction to Electrodynamics*. Pearson Education, 2008.

Christoph Große, Alexander Kabakchiev, Theresa Lutz, Romain Froidevaux, Frank Schramm, Mario Ruben, Markus Etzkorn, Uta Schlickum, Klaus Kuhnke, and Klaus Kern. Dynamic control of plasmon generation by an individual quantum system. *Nano letters*, 14(10):5693–5697, 2014.

Christopher R Gubbin, Francesco Martini, Alberto Politi, Stefan A Maier, and Simone De Liberato. Strong and coherent coupling between localized and propagating phonon polaritons. *Physical review letters*, 116(24):246402, 2016.

Emilie B Guidez and Christine M Aikens. Quantum mechanical origin of the plasmon: from molecular systems to nanoparticles. *Nanoscale*, 6(20):11512–11527, 2014.

Hellmut Haberland, Bernd von Issendorff, Ji Yufeng, and Thomas Kolar. Transition to plasmonlike absorption in small hg clusters. *Physical review letters*, 69(22):3212, 1992.

Ghazal Hajisalem, Mohammedreza S Nezami, and Reuven Gordon. Probing the quantum tunneling limit of plasmonic enhancement by third harmonic generation. *Nano letters*, 14(11):6651–6654, 2014.

Hermann Haken and Hans Christoph Wolf. *Molecular physics and elements of quantum chemistry: introduction to experiments and theory*. Springer Science & Business Media, 2013.

Encai Hao and George C Schatz. Electromagnetic fields around silver nanoparticles and dimers. *The Journal of chemical physics*, 120(1):357–366, 2004.

Judith Harl. *The linear response function in density functional theory: Optical spectra and improved description of the electron correlation*. PhD thesis, 2008.

Christian Heck, Julia Prinz, André Dathe, Virginia Merk, Ondrej Stranik, Wolfgang Fritzsche, Janina Kneipp, and Ilko Bald. Gold nanolenses self-assembled by dna origami. *Acs Photonics*, 4(5):1123–1130, 2017.

Christian Heck, Yuya Kanehira, Janina Kneipp, and Ilko Bald. Placement of single proteins within the sers hot spots of self-assembled silver nanolenses. *Angewandte Chemie International Edition*, 57:7444–7447, 2018.

Reinier W Heeres, Leo P Kouwenhoven, and Valery Zwiller. Quantum interference in plasmonic circuits. *Nature nanotechnology*, 8(10):719–722, 2013.

R Hillenbrand, T Taubner, and F Keilmann. Phonon-enhanced light–matter interaction at the nanometre scale. *Nature*, 418(6894):159, 2002.

So Hirata and Martin Head-Gordon. Time-dependent density functional theory within the tamm–dancoff approximation. *Chemical Physics Letters*, 314(3-4):291–299, 1999.

Pierre Hohenberg and Walter Kohn. Inhomogeneous electron gas. *Physical review*, 136(3B):B864, 1964.

Ulrich Hohenester. Simulating electron energy loss spectroscopy with the mnpbem toolbox. *Computer Physics Communications*, 185(3):1177–1187, 2014.

Ulrich Hohenester and Andreas Trügler. Mnpbem–a matlab toolbox for the simulation of plasmonic nanoparticles. *Computer Physics Communications*, 183(2): 370–381, 2012.

Christiane Höppener, Zachary J Lapin, Palash Bharadwaj, and Lukas Novotny. Self-similar gold-nanoparticle antennas for a cascaded enhancement of the optical field. *Physical review letters*, 109(1):017402, 2012.

H Hövel, S Fritz, A Hilger, U Kreibig, and Michael Vollmer. Width of cluster plasmon resonances: bulk dielectric functions and chemical interface damping. *Physical Review B*, 48(24):18178, 1993.

Hao Hu, Jingjing Zhang, Stefan A. Maier, and Yu Luo. Enhancing third-harmonic generation with spatial nonlocality. *ACS Photonics*, 5(2):592–598, 2018.

Wen Chu Huang and Juh Tzeng Lue. Quantum size effect on the optical properties of small metallic particles. *Physical Review B*, 49(24):17279, 1994.

Yu-Hui Huang, Ken-Ming Lin, TC Leung, and Che Ting Chan. Plasmonlike resonances in atomic chains: A time-dependent density-functional theory study. *Physical Review B*, 90(7):075418, 2014.

Christian Huck, Jochen Vogt, Tomáš Neuman, Tadaaki Nagao, Rainer Hillenbrand, Javier Aizpurua, Annemarie Pucci, and Frank Neubrech. Strong coupling between phonon-polaritons and plasmonic nanorods. *Optics express*, 24(22):25528–25539, 2016.

Marinko Jablan, Marin Soljačić, and Hrvoje Buljan. Plasmons in graphene: fundamental properties and potential applications. *Proceedings of the IEEE*, 101(7): 1689–1704, 2013.

John David Jackson. *Classical electrodynamics*. John Wiley & Sons, 2007.

Zubin Jacob. Nanophotonics: hyperbolic phonon–polaritons. *Nature materials*, 13(12): 1081, 2014.

Saman Jahani and Zubin Jacob. All-dielectric metamaterials. *Nature nanotechnology*, 11(1):23, 2016.

Prashant K Jain, Susie Eustis, and Mostafa A El-Sayed. Plasmon coupling in nanorod assemblies: optical absorption, discrete dipole approximation simulation, and exciton-coupling model. *The Journal of Physical Chemistry B*, 110(37):18243–18253, 2006.

Mary B James and David J Griffiths. Why the speed of light is reduced in a transparent medium. *American journal of physics*, 60(4):309–313, 1992.

Lasse Jensen, Christine M Aikens, and George C Schatz. Electronic structure methods for studying surface-enhanced raman scattering. *Chemical Society Reviews*, 37(5):1061–1073, 2008.

Yajie Jiang, Supriya Pillai, and Martin A Green. Realistic silver optical constants for plasmonics. *Scientific reports*, 6:30605, 2016.

Peter B Johnson and R-W_ Christy. Optical constants of the noble metals. *Physical review B*, 6(12):4370, 1972.

Water Johnstone, George Stewart, T Hart, and Brian Culshaw. Surface plasmon polaritons in thin metal films and their role in fiber optic polarizing devices. *Journal of Lightwave Technology*, 8(4):538–544, 1990.

Marcus Jones, Chaiwat Engtrakul, Wyatt K Metzger, Randy J Ellingson, Arthur J Nozik, Michael J Heben, and Garry Rumbles. Analysis of photoluminescence from solubilized single-walled carbon nanotubes. *Physical Review B*, 71(11):115426, 2005.

Long Ju, Baisong Geng, Jason Horng, Caglar Girit, Michael Martin, Zhao Hao, Hans A Bechtel, Xiaogan Liang, Alex Zettl, Y Ron Shen, et al. Graphene plasmonics for tunable terahertz metamaterials. *Nature nanotechnology*, 6(10):630, 2011.

Heesun Jung, Hoon Cha, Daedu Lee, and Sangwoon Yoon. Bridging the nanogap with light: continuous tuning of plasmon coupling between gold nanoparticles. *ACS nano*, 9(12):12292–12300, 2015.

Radoslaw Jurga, Stefania D'Agostino, Fabio Della Sala, and Cristian Ciracì. Plasmonic nonlocal response effects on dipole decay dynamics in the weak-and strong-coupling regimes. *The Journal of Physical Chemistry C*, 121(40):22361–22368, 2017.

Martti Kauranen and Anatoly V Zayats. Nonlinear plasmonics. *Nature Photonics*, 6 (11):737–748, 2012.

LV Keldysh, Alexei A Maradudin, and DA Kirzhnitz. *The dielectric function of condensed systems*. Elsevier, 2012.

Richard M Kerber, Jamie M Fitzgerald, Doris E Reiter, Sang Soon Oh, and Ortwin Hess. Reading the orbital angular momentum of light using plasmonic nanoantennas. *ACS Photonics*, 4(4):891–896, 2017.

Galina Khitrova, HM Gibbs, M Kira, Stephan W Koch, and Axel Scherer. Vacuum rabi splitting in semiconductors. *Nature Physics*, 2(2):81, 2006.

Arman S Kirakosyan, Mark I Stockman, and Tigran V Shahbazyan. Surface plasmon lifetime in metal nanoshells. *Physical Review B*, 94(15):155429, 2016.

Charles Kittel. *Introduction to solid state physics*. John Wiley & Sons, 1966.

Dario Knebl, Anton Hörl, Andreas Trügler, Johannes Kern, Joachim R Krenn, Peter Puschnig, and Ulrich Hohenester. Gap plasmonics of silver nanocube dimers. *Physical Review B*, 93(8):081405, 2016.

Janina Kneipp, Xiangting Li, Margaret Sherwood, Ulrich Panne, Harald Kneipp, Mark I Stockman, and Katrin Kneipp. Gold nanolenses generated by laser ablation-efficient enhancing structure for surface enhanced raman scattering analytics and sensing. *Analytical chemistry*, 80(11):4247–4251, 2008.

WD Knight, Keith Clemenger, Walt A de Heer, Winston A Saunders, MY Chou, and Marvin L Cohen. Electronic shell structure and abundances of sodium clusters. *Physical review letters*, 52(24):2141, 1984.

RJ Koch, Th Seyller, and JA Schaefer. Strong phonon-plasmon coupled modes in the graphene/silicon carbide heterosystem. *Physical Review B*, 82(20):201413, 2010.

Walter Kohn and Lu Jeu Sham. Self-consistent equations including exchange and correlation effects. *Physical review*, 140(4A):A1133, 1965.

AV Korol and AV Solov'Yov. Comment on "photoexcitation of a volume plasmon in c 60 ions". *Physical review letters*, 98(17):179601, 2007.

Valeri N Kotov, Bruno Uchoa, Vitor M Pereira, F Guinea, and AH Castro Neto. Electron-electron interactions in graphene: Current status and perspectives. *Reviews of Modern Physics*, 84(3):1067, 2012.

VG Kravets, George Zoriniants, Christopher P Burrows, F Schedin, C Casiraghi, P Klar, AK Geim, William L Barnes, and AN Grigorenko. Cascaded optical field enhancement in composite plasmonic nanostructures. *Physical review letters*, 105(24): 246806, 2010.

U Kreibig. Electronic properties of small silver particles: the optical constants and their temperature dependence. *Journal of Physics F: Metal Physics*, 4(7):999, 1974.

U Kreibig. The transition cluster-solid state in small gold particles. *Solid State Communications*, 28(9):767–769, 1978.

Uwe Kreibig and Michael Vollmer. *Optical properties of metal clusters*, volume 25. Springer Science & Business Media, 2013.

Ryogo Kubo. Electronic properties of metallic fine particles. i. *Journal of the Physical Society of Japan*, 17(6):975–986, 1962.

Stephan Kuemmel and Matthias Brack. Quantum fluid dynamics from density-functional theory. *Physical Review A*, 64(2):022506, 2001.

PW Langhoff, ST Epstein, and M Karplus. Aspects of time-dependent perturbation theory. *Reviews of Modern Physics*, 44(3):602, 1972.

Adam Lauchner, Andrea E Schlather, Alejandro Manjavacas, Yao Cui, Michael J McClain, Grant J Stec, F Javier García de Abajo, Peter Nordlander, and Naomi J Halas. Molecular plasmonics. *Nano letters*, 15(9):6208–6214, 2015.

Carlo Maria Lazzarini, Levato Tadzio, Jamie M Fitzgerald, José A Sánchez-Gil, and Vincenzo Giannini. Linear ultrafast dynamics of plasmon and magnetic resonances in nanoparticles. *Physical Review B*, 96(23):235407, 2017.

Eric Le Ru and Pablo Etchegoin. *Principles of Surface-Enhanced Raman Spectroscopy: and related plasmonic effects*. Elsevier, 2008.

Sarah Lerch and Björn M Reinhard. Quantum plasmonics: Optical monitoring of dna-mediated charge transfer in plasmon rulers. *Advanced Materials*, 28(10): 2030–2036, 2016.

Jean Lermé, Hatim Baida, Christophe Bonnet, Michel Broyer, Emmanuel Cottancin, Aurélien Crut, Paolo Maioli, Natalia Del Fatti, Fabrice Vallée, and Michel Pellarin. Size dependence of the surface plasmon resonance damping in metal nanospheres. *The Journal of Physical Chemistry Letters*, 1(19):2922–2928, 2010.

Ke Li, Jamie M Fitzgerald, Xiaofei Xiao, Joshua D Caldwell, Cheng Zhang, Stefan A Maier, Xiaofeng Li, and Vincenzo Giannini. Graphene plasmon cavities made with silicon carbide. *ACS Omega*, 2(7):3640–3646, 2017.

Kuiru Li, Mark I Stockman, and David J Bergman. Self-similar chain of metal nanospheres as an efficient nanolens. *Physical review letters*, 91(22):227402, 2003.

Xiaoguang Li, Di Xiao, and Zhenyu Zhang. Landau damping of quantum plasmons in metal nanostructures. *New Journal of Physics*, 15(2):023011, 2013.

Yilei Li, Hugen Yan, Damon B Farmer, Xiang Meng, Wenjuan Zhu, Richard M Osgood, Tony F Heinz, and Phaedon Avouris. Graphene plasmon enhanced vibrational sensing of surface-adsorbed layers. *Nano letters*, 14(3):1573–1577, 2014.

Yonghui Li and CA Ullrich. Time-dependent transition density matrix. *Chemical Physics*, 391(1):157–163, 2011.

Yonghui Li and Carsten A Ullrich. The particle–hole map: A computational tool to visualize electronic excitations. *Journal of chemical theory and computation*, 11(12): 5838–5852, 2015.

Zhipeng Li, Zhilin Yang, and Hongxing Xu. Comment on "self-similar chain of metal nanospheres as an efficient nanolens". *Physical review letters*, 97(7):079701, 2006.

Shujing Liang, Chao Li, Chunlei Zhang, Yunsheng Chen, Liang Xu, Chenchen Bao, Xiaoyong Wang, et al. Cd44v6 monoclonal antibody-conjugated gold nanostars for targeted photoacoustic imaging and plasmonic photothermal therapy of gastric cancer stem-like cells. *Theranostics*, 5(9):970, 2015.

I Lindau and PO Nilsson. Experimental verification of optically excited longitudinal plasmons. *Physica Scripta*, 3(2):87, 1971.

Stephan Link and Mostafa A. El-Sayed. Spectral properties and relaxation dynamics of surface plasmon electronic oscillations in gold and silver nanodots and nanorods. *The Journal of Physical Chemistry B*, 103(40):8410–8426, 1999.

Fang Liu, Pingbing Ming, and Ju Li. Ab initio calculation of ideal strength and phonon instability of graphene under tension. *Physical Review B*, 76(6):064120, 2007.

Yu Liu and Roy F Willis. Plasmon-phonon strongly coupled mode in epitaxial graphene. *Physical Review B*, 81(8):081406, 2010.

Yu Liu, Roy F Willis, KV Emtsev, and Th Seyller. Plasmon dispersion and damping in electrically isolated two-dimensional charge sheets. *Physical Review B*, 78(20): 201403, 2008.

Julian A Lloyd, Soon Hock Ng, Amelia CY Liu, Ye Zhu, Wei Chao, Toon Coenen, Joanne Etheridge, Daniel E Gómez, and Udo Bach. Plasmonic nanolenses: Electrostatic self-assembly of hierarchical nanoparticle trimers and their response to optical and electron beam stimuli. *ACS nano*, 11(2):1604–1612, 2017.

Francis E Low. *Classical field theory: electromagnetism and gravitation*. John Wiley & Sons, 2008.

Tony Low, Andrey Chaves, Joshua D Caldwell, Anshuman Kumar, Nicholas X Fang, Phaedon Avouris, Tony F Heinz, Francisco Guinea, Luis Martin-Moreno, and Frank Koppens. Polaritons in layered two-dimensional materials. *Nature materials*, 16 (2):182, 2017.

Mark B Lundeberg, Yuanda Gao, Reza Asgari, Cheng Tan, Ben Van Duppen, Marta Autore, Pablo Alonso-González, Achim Woessner, Kenji Watanabe, Takashi Taniguchi, et al. Tuning quantum nonlocal effects in graphene plasmonics. *Science*, page eaan2735, 2017. ISSN 0036-8075.

Stig Lundqvist and Norman H March. *Theory of the inhomogeneous electron gas*. Springer Science & Business Media, 2013.

Yu Luo, AI Fernandez-Dominguez, Aeneas Wiener, Stefan A Maier, and JB Pendry. Surface plasmons and nonlocality: a simple model. *Physical review letters*, 111(9): 093901, 2013.

AA Lushnikov and AJ Simonov. Surface plasmons in small metal particles. *Zeitschrift für Physik A Hadrons and Nuclei*, 270(1):17–24, 1974.

Theresa Lutz, Christoph Große, Christian Dette, Alexander Kabakchiev, Frank Schramm, Mario Ruben, Rico Gutzler, Klaus Kuhnke, Uta Schlickum, and Klaus Kern. Molecular orbital gates for plasmon excitation. *Nano letters*, 13(6):2846–2850, 2013.

AV Luzanov and OA Zhikol. Electron invariants and excited state structural analysis for electronic transitions within cis, rpa, and tddft models. *International Journal of Quantum Chemistry*, 110(4):902–924, 2010.

Jie Ma, Zhi Wang, and Lin-Wang Wang. Interplay between plasmon and single-particle excitations in a metal nanocluster. *Nature communications*, 6:10107, 2015.

DW Mackowski and MI Mishchenko. A multiple sphere t-matrix fortran code for use on parallel computer clusters. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 112(13):2182–2192, 2011.

Stefan A Maier. Plasmonic field enhancement and sers in the effective mode volume picture. *Optics Express*, 14(5):1957–1964, 2006.

Stefan Alexander Maier. *Plasmonics: fundamentals and applications*. Springer Science & Business Media, 2007.

Alejandro Manjavacas, Federico Marchesin, Sukosin Thongrattanasiri, Peter Koval, Peter Nordlander, Daniel Sanchez-Portal, and F Javier García de Abajo. Tunable molecular plasmons in polycyclic aromatic hydrocarbons. *ACS nano*, 7(4):3635–3643, 2013.

Dana Codruta Marinica, Mario Zapata, Peter Nordlander, Andrey K Kazansky, Pedro M Echenique, Javier Aizpurua, and Andrei G Borisov. Active quantum plasmonics. *Science advances*, 1(11):e1501095, 2015.

DC Marinica, AK Kazansky, P Nordlander, J Aizpurua, and AG Borisov. Quantum plasmonics: nonlinear effects in the field enhancement of a plasmonic nanoparticle dimer. *Nano letters*, 12(3):1333–1339, 2012.

Richard L Martin. Natural transition orbitals. *The Journal of chemical physics*, 118(11): 4775–4777, 2003.

Richard M Martin. *Electronic structure: basic theory and practical methods*. Cambridge university press, 2004.

L Martin-Moreno, FJ Garcia-Vidal, HJ Lezec, KM Pellerin, Tineke Thio, JB Pendry, and TW Ebbesen. Theory of extraordinary optical transmission through subwavelength hole arrays. *Physical review letters*, 86(6):1114, 2001.

L Marton, J Arol Simpson, HA Fowler, and N Swanson. Plural scattering of 20-kev electrons in aluminum. *Physical Review*, 126(1):182, 1962.

Kathryn M Mayer and Jason H Hafner. Localized surface plasmon resonance sensors. *Chemical reviews*, 111(6):3828–3857, 2011.

CW McCurdy and LC Cusachs. Simplification of the rpa secular equation. *The Journal of Chemical Physics*, 55(4):1994–1995, 1971.

Kevin M McPeak, Sriharsha V Jayanti, Stephan JP Kress, Stefan Meyer, Stelio Iotti, Aurelio Rossinelli, and David J Norris. Plasmonic films can easily be better: Rules and recipes. *ACS photonics*, 2(3):326–333, 2015.

Andrew R Melnyk and Michael J Harrison. Theory of optical excitation of plasmons in metals. *Physical Review B*, 2(4):835, 1970.

Eugene Merzbacher. *Quantum mechanics*. Wiley International Edition, 1970.

Barbara J Messinger, K Ulrich Von Raben, Richard K Chang, and Peter W Barber. Local fields at the surface of noble-metal microspheres. *Physical Review B*, 24(2):649, 1981.

Sean Molesky, Christopher J Dewalt, and Zubin Jacob. High temperature epsilon-near-zero and epsilon-near-pole metamaterial emitters for thermophotovoltaics. *Optics express*, 21(101):A96–A110, 2013.

Rafael A Molina, Dietmar Weinmann, and Rodolfo A Jalabert. Oscillatory size dependence of the surface plasmon linewidth in metallic nanoparticles. *Physical Review B*, 65(15):155427, 2002.

R Carmina Monreal, Tomasz J Antosiewicz, and S Peter Apell. Competition between surface screening and size quantization for surface plasmons in nanoparticles. *New Journal of Physics*, 15(8):083044, 2013.

R Carmina Monreal, Tomasz J Antosiewicz, and S Peter Apell. Diffuse surface scattering in the plasmonic resonances of ultralow electron density nanospheres. *The journal of physical chemistry letters*, 6(10):1847–1853, 2015.

Alexander Moroz. Electron mean free path in a spherical shell geometry. *The Journal of Physical Chemistry C*, 112(29):10641–10652, 2008.

Alexander Moroz. Depolarization field of spheroidal particles. *JOSA B*, 26(3): 517–527, 2009.

Jens Jørgen Mortensen, Lars Bruno Hansen, and Karsten Wedel Jacobsen. Real-space grid implementation of the projector augmented wave method. *Physical Review B*, 71(3):035109, 2005.

N Asger Mortensen, Søren Raza, Martijn Wubs, Thomas Søndergaard, and Sergey I Bozhevolnyi. A generalized non-local optical response theory for plasmonic nanostructures. *Nature communications*, 5:3809, 2014.

Seth M Morton, Daniel W Silverstein, and Lasse Jensen. Theoretical studies of plasmonics using electronic structure methods. *Chemical reviews*, 111(6):3962–3994, 2011.

Shaul Mukamel, Sergei Tretiak, Thomas Wagersreiter, and Vladimir Chernyak. Electronic coherence and collective optical excitations of conjugated molecules. *Science*, 277(5327):781–787, 1997.

Colleen L Nehl, Nathaniel K Grady, Glenn P Goodrich, Felicia Tam, Naomi J Halas, and Jason H Hafner. Scattering spectra of single gold nanoshells. *Nano Letters*, 4(12): 2355–2359, 2004.

Guang-Xin Ni, Hong-Zhi Yang, Wei Ji, Seung-Jae Baeck, Chee-Tat Toh, Jong-Hyun Ahn, Vitor M Pereira, and Barbaros Özyilmaz. Tuning optical conductivity of large-scale cvd graphene by strain engineering. *Advanced Materials*, 26(7):1081–1086, 2014.

GX Ni, L Wang, MD Goldflam, M Wagner, Z Fei, AS McLeod, MK Liu, F Keilmann, B Özyilmaz, AH Castro Neto, et al. Ultrafast optical switching of infrared plasmon polaritons in high-mobility graphene. *Nature Photonics*, 10(4):244, 2016.

A Yu Nikitin, T Low, and L Martin-Moreno. Anomalous reflection phase of graphene plasmons and its influence on resonators. *Physical Review B*, 90(4):041407, 2014.

N Nilius, TM Wallis, and W Ho. Development of one-dimensional band structure in artificial gold chains. *Science*, 297(5588):1853–1856, 2002.

Mauro Nisoli, Piero Decleva, Francesca Calegari, Alicia Palacios, and Fernando Martín. Attosecond electron dynamics in molecules. *Chemical reviews*, 117(16): 10760–10825, 2017.

Fernando Nogueira, Alberto Castro, and Miguel AL Marques. A tutorial on density functional theory. In *A Primer in Density Functional Theory*, pages 218–256. Springer, 2003.

Peter Nordlander, C Oubre, E Prodan, K Li, and MI Stockman. Plasmon hybridization in nanoparticle dimers. *Nano letters*, 4(5):899–903, 2004.

Lukas Novotny. Effective wavelength scaling for optical antennas. *Physical Review Letters*, 98(26):266802, 2007.

Lukas Novotny. Strong coupling, energy splitting, and level crossings: A classical perspective. *American Journal of Physics*, 78(11):1199–1202, 2010.

Lukas Novotny and Bert Hecht. *Principles of nano-optics*. Cambridge university press, 2012.

Philippe Nozières and David Pines. *Theory of quantum liquids*. Hachette UK, 1999.

Sara Núñez Sánchez, Martin Lopez-Garcia, Mohamed M Murshidy, Asmaa Gamal Abdel-Hady, Mohamed Serry, Ali M Adawi, John G Rarity, Ruth Oulton, and William L Barnes. Excitonic optical tamm states: A step toward a full molecular–dielectric photonic integration. *ACS Photonics*, 3(5):743–748, 2016.

Ekmel Ozbay. Plasmonics: merging photonics and electronics at nanoscale dimensions. *science*, 311(5758):189–193, 2006.

Edward D Palik. *Handbook of optical constants of solids*, volume 3. Academic press, 1998.

Joel H Parks and Stephen A McDonald. Evolution of the collective-mode resonance in small adsorbed sodium clusters. *Physical review letters*, 62(19):2301, 1989.

Mishu Paul and P Balanarayan. Electronic rearrangement in molecular plasmons: An electron density and electrostatic potential-based study. *ChemPhysChem*, 2018.

G Pellegrini, G Mattei, V Bello, and P Mazzoldi. Interacting metal nanoparticles: Optical properties from nanoparticle dimers to core-satellite systems. *Materials Science and Engineering: C*, 27(5-8):1347–1350, 2007.

Giovanni Pellegrini, Michele Celebrano, Marco Finazzi, and Paolo Biagioni. Local field enhancement: comparing self-similar and dimer nanoantennas. *The Journal of Physical Chemistry C*, 120(45):26021–26024, 2016.

Matthew Pelton. Modified spontaneous emission in nanophotonic structures. *Nature Photonics*, 9(7):427–435, 2015.

Bo Peng, David B Lingerfelt, Feizhi Ding, Christine M Aikens, and Xiaosong Li. Real-time tddft studies of exciton decay and transfer in silver nanowire arrays. *The Journal of Physical Chemistry C*, 119(11):6421–6427, 2015.

Z Penzar and M Šunjić. Dynamical dipole and quadrupole response properties of small metallic particles: Random phase approximation and the modified spherical infinite barrier model. *Solid state communications*, 52(8):747–751, 1984.

John P Perdew and Stefan Kurth. Density functionals for non-relativistic coulomb systems in the new century. In *A primer in density functional theory*, pages 1–55. Springer, 2003.

John P Perdew and Yue Wang. Accurate and simple analytic representation of the electron-gas correlation energy. *Physical Review B*, 45(23):13244, 1992.

MGUJ Petersilka, UJ Gossmann, and EKU Gross. Excitation energies from time-dependent density-functional theory. *Physical Review Letters*, 76(8):1212, 1996.

Eric F Pettersen, Thomas D Goddard, Conrad C Huang, Gregory S Couch, Daniel M Greenblatt, Elaine C Meng, and Thomas E Ferrin. Ucsf chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry*, 25(13):1605–1612, 2004.

GiovanniMaria Piccini, Remco WA Havenith, Ria Broer, and Mauro Stener. Gold nanowires: a time-dependent density functional assessment of plasmonic behavior. *The Journal of Physical Chemistry C*, 117(33):17196–17204, 2013.

Anatoliy Pinchuk and Uwe Kreibig. Interface decay channel of particle surface plasmon resonance. *New Journal of Physics*, 5(1):151, 2003.

David Pines. *Elementary excitations in solids: lectures on phonons, electrons, and plasmons*, volume 5. WA Benjamin, 1964.

JM Pitarke, VM Silkin, EV Chulkov, and PM Echenique. Theory of surface plasmons and surface-plasmon polaritons. *Reports on progress in physics*, 70(1):1, 2006.

Felix Plasser. Entanglement entropy of electronic excitations. *The Journal of Chemical Physics*, 144(19):194107, 2016.

Felix Plasser and Hans Lischka. Analysis of excitonic and charge transfer interactions from quantum chemical calculations. *Journal of chemical theory and computation*, 8(8):2777–2789, 2012.

Felix Plasser, Michael Wormit, and Andreas Dreuw. New tools for the systematic analysis and visualization of electronic excitations. i. formalism. *The Journal of chemical physics*, 141(2):024106, 2014.

CJ Powell and JB Swan. Origin of the characteristic electron energy losses in aluminum. *Physical Review*, 115(4):869, 1959a.

CJ Powell and JB Swan. Origin of the characteristic electron energy losses in magnesium. *Physical Review*, 116(1):81, 1959b.

E Prodan and P Nordlander. Exchange and correlations effects in small metallic nanoshells. *Chemical physics letters*, 349(1-2):153–160, 2001.

Emil Prodan and Peter Nordlander. Electronic structure and polarizability of metallic nanoshells. *Chemical physics letters*, 352(3-4):140–146, 2002.

Marina Quijada, R Díez Muiño, Andrei G Borisov, JA Alonso, and Pedro M Echenique. Lifetime of electronic excitations in metal nanoparticles. *New Journal of Physics*, 12(5):053023, 2010.

Michael Quinten. Optical constants of gold and silver clusters in the spectral range between 1.5 ev and 4.5 ev. *Zeitschrift für Physik B Condensed Matter*, 101(2):211–217, 1996.

Michael Quinten. *Optical properties of nanoparticle systems: Mie and beyond*. John Wiley & Sons, 2010.

MG Raizen, RJ Thompson, RJ Brecha, HJ Kimble, and HJ Carmichael. Normal-mode splitting and linewidth averaging for two-state atoms in an optical cavity. *Physical Review Letters*, 63(3):240, 1989.

Aleksandar D Rakić, Aleksandra B Djurišić, Jovan M Elazar, and Marian L Majewski. Optical properties of metallic films for vertical-cavity optoelectronic devices. *Applied optics*, 37(22):5271–5283, 1998.

Søren Raza, Giuseppe Toscano, Antti-Pekka Jauho, Martijn Wubs, and N Asger Mortensen. Unusual resonances in nanoplasmonic structures due to nonlocal response. *Physical Review B*, 84(12):121412, 2011.

Søren Raza, Nicolas Stenger, Shima Kadkhodazadeh, Søren V Fischer, Natalie Kostesha, Antti-Pekka Jauho, Andrew Burrows, Martijn Wubs, and N Asger Mortensen. Blueshift of the surface plasmon resonance in silver nanoparticles studied with eels. *Nanophotonics*, 2(2):131–138, 2013a.

Søren Raza, Wei Yan, Nicolas Stenger, Martijn Wubs, and N Asger Mortensen. Blueshift of the surface plasmon resonance in silver nanoparticles: substrate effects. *Optics express*, 21(22):27344–27355, 2013b.

Søren Raza, Sergey I Bozhevolnyi, Martijn Wubs, and N Asger Mortensen. Nonlocal optical response in metallic nanostructures. *Journal of Physics: Condensed Matter*, 27 (18):183204, 2015.

Thomas Reiners, Christoph Ellert, Martin Schmidt, and Hellmut Haberland. Size dependence of the optical response of spherical sodium clusters. *Physical review letters*, 74(9):1558, 1995.

MJ Rice, WR Schneider, and S Strässler. Electronic polarizabilities of very small metallic particles and thin films. *Physical Review B*, 8(2):474, 1973.

Kenneth Franklin Riley, Michael Paul Hobson, and Stephen John Bence. *Mathematical methods for physics and engineering*. Cambridge university press, 2006.

RH Ritchie. Plasma losses by fast electrons in thin films. *Physical Review*, 106(5):874, 1957.

Nicholas Rivera, Ido Kaminer, Bo Zhen, John D Joannopoulos, and Marin Soljačić. Shrinking light to allow forbidden transitions on the atomic scale. *Science*, 353(6296): 263–269, 2016.

EJ Robbins, RE Leckenby, and P Willis. The ionization potentials of clustered sodium atoms. *Advances in Physics*, 16(64):739–744, 1967.

Charles W Robson, Kieran A Fraser, and Fabio Biancalana. Giant ultrafast kerr effect in superconductors. *Physical Review B*, 95(21):214504, 2017.

Tuomas Rossi. Simulating electric field enhancement in plasmonic nanomaterials. Master's thesis, Aalto University School Of Science, 2013.

Tuomas P Rossi, Asier Zugarramurdi, Martti J Puska, and Risto M Nieminen. Quantized evolution of the plasmonic response in a stretched nanorod. *Physical review letters*, 115(23):236804, 2015.

Tuomas P Rossi, Mikael Kuisma, Martti J Puska, Risto M Nieminen, and Paul Erhart. Kohn–sham decomposition in real-time time-dependent density-functional theory: An efficient tool for analyzing plasmonic excitations. *Journal of chemical theory and computation*, 13(10):4779–4790, 2017.

Erich Runge and Eberhard KU Gross. Density-functional theory for time-dependent systems. *Physical Review Letters*, 52(12):997, 1984.

R Ruppin. Optical properties of a plasma sphere. *Physical Review Letters*, 31(24):1434, 1973.

R Ruppin. Optical properties of small metal spheres. *Physical Review B*, 11(8):2871, 1975.

JRM Saavedra, Ana Asenjo-Garcia, and F Javier García de Abajo. Hot-electron dynamics and thermalization in small metallic nanoparticles. *ACS Photonics*, 3(9): 1637–1646, 2016.

Kevin J Savage, Matthew M Hawkeye, Rubén Esteban, Andrei G Borisov, Javier Aizpurua, and Jeremy J Baumberg. Revealing the quantum regime in tunnelling plasmonics. *Nature*, 491(7425):574, 2012.

Andrea E Schlather, Nicolas Large, Alexander S Urban, Peter Nordlander, and Naomi J Halas. Near-field mediated plexcitonic coupling and giant rabi splitting in individual metallic dimers. *Nano letters*, 13(7):3281–3286, 2013.

Jonathan A Scholl, Ai Leen Koh, and Jennifer A Dionne. Quantum plasmon resonances of individual metallic nanoparticles. *Nature*, 483(7390):421, 2012.

Jonathan A Scholl, Aitzol García-Etxarri, Ai Leen Koh, and Jennifer A Dionne. Observation of quantum tunneling between two plasmonic nanoparticles. *Nano letters*, 13(2):564–569, 2013.

Jon A Schuller, Thomas Taubner, and Mark L Brongersma. Optical antenna thermal emitters. *Nature Photonics*, 3(11):658, 2009.

SWJ Scully, ED Emmons, MF Gharaibeh, RA Phaneuf, ALD Kilcoyne, AS Schlachter, S Schippers, A Müller, HS Chakraborty, ME Madjet, et al. Photoexcitation of a volume plasmon in c_{60} ions. *Physical review letters*, 94(6):065503, 2005.

Luca Sementa, Andrea Marini, Giovanni Barcaro, Fabio R Negreiros, and Alessandro Fortunelli. Atomistic quantum plasmonics of gold nanowire arrays. *ACS Photonics*, 1(4):315–322, 2014.

David J Shelton, Igal Brener, James C Ginn, Michael B Sinclair, David W Peters, Kevin R Coffey, and Glenn D Boreman. Strong coupling between nanoscale metamaterials and phonons. *Nano letters*, 11(5):2104–2108, 2011.

Ranjan Singh and Nikolay Zheludev. Materials: Superconductor photonics. *Nature photonics*, 8(9):679, 2014.

Rajarshi Sinha-Roy, Pablo Garcia-Gonzalez, Hans-Christian Weissker, Frank Rabilloud, and Antonio I Fernandez-Dominguez. Classical and ab initio plasmonics meet at sub-nanometric noble metal rods. *ACS Photonics*, 4(6):1484–1493, 2017.

EA Stern and RA Ferrell. Surface plasma oscillations of a degenerate electron gas. *Physical Review*, 120(1):130, 1960.

Greg Sun, Jacob B Khurgin, and Alexander Bratkovsky. Coupled-mode theory of field enhancement in complex metal nanostructures. *Physical Review B*, 84(4):045415, 2011.

Attila Szabo and Neil S Ostlund. *Modern quantum chemistry: introduction to advanced electronic structure theory*. Courier Corporation, 2012.

Mark S Tame, KR McEnery, ŞK Özdemir, J Lee, SA Maier, and MS Kim. Quantum plasmonics. *Nature Physics*, 9(6):329–340, 2013.

Shu Fen Tan, Lin Wu, Joel KW Yang, Ping Bai, Michel Bosman, and Christian A Nijhuis. Quantum plasmon resonances controlled by molecular tunnel junctions. *Science*, 343(6178):1496–1499, 2014.

Philippe Tassin, Thomas Koschny, Maria Kafesaki, and Costas M Soukoulis. A comparison of graphene, superconductors and metals as conductors for metamaterials and plasmonics. *Nature Photonics*, 6(4):259, 2012.

Philippe Tassin, Thomas Koschny, and Costas M Soukoulis. Graphene for terahertz applications. *Science*, 341(6146):620–621, 2013.

Tatiana V Teperik, Peter Nordlander, Javier Aizpurua, and Andrei G Borisov. Robust subnanometric plasmon ruler by rescaling of the nonlocal optical response. *Physical review letters*, 110(26):263901, 2013.

Tatiana V Teperik, Andrey K Kazansky, and Andrei G Borisov. Electron tunneling through water layer in nanogaps probed by plasmon resonances. *Physical Review B*, 93(15):155431, 2016.

RJ Thompson, G Rempe, and HJ Kimble. Observation of normal-mode splitting for an atom in an optical cavity. *Physical Review Letters*, 68(8):1132, 1992.

Sukosin Thongrattanasiri, Frank HL Koppens, and F Javier García De Abajo. Complete optical absorption in periodically patterned graphene. *Physical review letters*, 108(4):047401, 2012.

KJ Tielrooij, L Orona, A Ferrier, M Badioli, G Navickaite, S Coop, S Nanot, B Kalinic, T Cesca, L Gaudreau, et al. Electrical control of optical emitter relaxation pathways enabled by graphene. *Nature Physics*, 11(3):281, 2015.

Josef Tiggesbäumker, Lars Köller, Karl-Heinz Meiwes-Broer, and Ansgar Liebsch. Blue shift of the mie plasma frequency in ag clusters and particles. *Physical Review A*, 48(3):R1749, 1993.

P Törmä and William L Barnes. Strong coupling between surface plasmon polaritons and emitters: a review. *Reports on Progress in Physics*, 78(1):013901, 2014.

Giuseppe Toscano, Jakob Straubel, Alexander Kwiatkowski, Carsten Rockstuhl, Ferdinand Evers, Hongxing Xu, N Asger Mortensen, and Martijn Wubs. Resonance shifts and spill-out effects in self-consistent hydrodynamic nanoplasmonics. *Nature communications*, 6:7132, 2015.

Emily Townsend and Garnett W Bryant. Plasmonic properties of metallic nanoparticles: The effects of size quantization. *Nano letters*, 12(1):429–434, 2011.

Emily Townsend and Garnett W Bryant. Which resonances in small metallic nanoparticles are plasmonic? *Journal of Optics*, 16(11):114022, 2014.

Norman Troullier and José Luís Martins. Efficient pseudopotentials for plane-wave calculations. *Physical review B*, 43(3):1993, 1991.

Andreas Trügler. *Optical properties of metallic nanoparticles*. Springer, 2011.

CA Ullrich and V Turkowski. Time-dependent density-functional theory for electronic excitations in materials: basics and perspectives. *arXiv preprint arXiv:0808.2021*, 2008.

Carsten A Ullrich. *Time-dependent density-functional theory: concepts and applications*. OUP Oxford, 2011.

Mattin Urbieta, Marc Barbry, Yao Zhang, Peter Koval, Daniel Sánchez-Portal, Nerea Zabala, and Javier Aizpurua. Atomic-scale lightning rod effect in plasmonic picocavities: A classical view to a quantum effect. *ACS nano*, 12(1):585–595, 2018.

Ashkan Vakil and Nader Engheta. Transformation optics using graphene. *Science*, 332(6035):1291–1294, 2011.

Hendrik C van de Hulst. *Light scattering by small particles*. Dover Publications, 1957.

Alejandro Varas, Pablo García-González, Johannes Feist, FJ García-Vidal, and Angel Rubio. Quantum plasmonics: from jellium models to ab initio calculations. *Nanophotonics*, 5(3):409–426, 2016.

Borislav Vasić, Goran Isić, and Radoš Gajić. Localized surface plasmon resonances in graphene ribbon arrays for sensing of dielectric environment at infrared frequencies. *Journal of Applied Physics*, 113(1):013110, 2013.

Igor Vasiliev, Serdar Öğüt, and James R Chelikowsky. First-principles density-functional calculations for optical spectra of clusters and nanocrystals. *Physical Review B*, 65(11):115416, 2002.

Michelle S Vezie, Sheridan Few, Iain Meager, Galatia Pieridou, Bernhard Dörling, Raja Shahid Ashraf, Alejandro R Goñi, Hugo Bronstein, Iain McCulloch, Sophia C Hayes, et al. Exploring the origin of high optical absorption in conjugated polymers. *Nature materials*, 15(7):746, 2016.

Alexandre Vial, Anne-Sophie Grimault, Demetrio Macías, Dominique Barchiesi, and Marc Lamy de La Chapelle. Improved analytical fit of gold dispersion: Application to the modeling of extinction spectra with a finite-difference time-domain method. *Physical Review B*, 71(8):085416, 2005.

Feng Wang and Y Ron Shen. General properties of local plasmons in metal nanostructures. *Physical review letters*, 97(20):206806, 2006.

L Wang, I Meric, PY Huang, Q Gao, Y Gao, H Tran, T Taniguchi, K Watanabe, LM Campos, DA Muller, et al. One-dimensional electrical contact to a two-dimensional material. *Science*, 342(6158):614–617, 2013.

Daniel R Ward, Falco Hüser, Fabian Pauly, Juan Carlos Cuevas, and Douglas Natelson. Optical rectification and field enhancement in a plasmonic nanogap. *Nature nanotechnology*, 5(10):732, 2010.

Jürgen Waxenegger, Andreas Trügler, and Ulrich Hohenester. Plasmonics simulations with the mnpbem toolbox: Consideration of substrates and layer structures. *Computer Physics Communications*, 193:138–150, 2015.

Maggie L Weber and Katherine A Willets. Nanoscale studies of plasmonic hot spots using super-resolution optical imaging. *MRS bulletin*, 37(8):745–751, 2012.

Wolfgang SM Werner, Kathrin Glantschnig, and Claudia Ambrosch-Draxl. Optical constants and inelastic electron-scattering data for 17 elemental metals. *Journal of Physical and Chemical Reference Data*, 38(4):1013–1092, 2009.

Peter R Wiecha, Leo-Jay Black, Yudong Wang, Vincent Paillard, Christian Girard, Otto L Muskens, and Arnaud Arbouet. Polarization conversion in plasmonic nanoantennas for metasurfaces using structural asymmetry and mode hybridization. *Scientific reports*, 7:40906, 2017.

Achim Woessner, Mark B Lundeberg, Yuanda Gao, Alessandro Principi, Pablo Alonso-González, Matteo Carrega, Kenji Watanabe, Takashi Taniguchi, Giovanni Vignale, Marco Polini, et al. Highly confined low-loss plasmons in graphene–boron nitride heterostructures. *Nature materials*, 14(4):421, 2015.

DM Wood and NW Ashcroft. Quantum size effects in the optical properties of small metallic particles. *Physical review B*, 25(10):6255, 1982.

Robert Williams Wood. Xlii. on a remarkable case of uneven distribution of light in a diffraction grating spectrum. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 4(21):396–402, 1902.

Lin Wu, Huigao Duan, Ping Bai, Michel Bosman, Joel KW Yang, and Erping Li. Fowler–nordheim tunneling induced charge transfer plasmons between nearly touching nanoparticles. *ACS nano*, 7(1):707–716, 2012.

Yanwen Wu, Chengdong Zhang, N Mohammadi Estakhri, Yang Zhao, Jisun Kim, Matt Zhang, Xing-Xiang Liu, Greg K Pribil, Andrea Alù, Chih-Kang Shih, et al. Intrinsic optical properties and enhanced plasmonic response of epitaxial silver. *Advanced Materials*, 26(35):6106–6110, 2014.

Martijn Wubs. Classification of scalar and dyadic nonlocal optical response models. *Optics express*, 23(24):31296–31312, 2015.

Chunlei Xia, Chunrong Yin, and Vitaly V Kresin. Photoabsorption by volume plasmons in metal nanoclusters. *Physical review letters*, 102(15):156802, 2009.

Fengnian Xia, Han Wang, Di Xiao, Madan Dubey, and Ashwin Ramasubramaniam. Two-dimensional material nanophotonics. *Nature Photonics*, 8(12):899, 2014.

Xiaofei Xiao, Xiaofeng Li, Joshua D Caldwell, Stefan A Maier, and Vincenzo Giannini. Theoretical analysis of graphene plasmon cavities. *Applied Materials Today*, 12:283–293, 2018.

Kazuhiro Yabana and GF Bertsch. Time-dependent local-density approximation in real time. *Physical Review B*, 54(7):4484, 1996.

Kazuhiro Yabana, T Sugiyama, Y Shinohara, T Otobe, and GF Bertsch. Time-dependent density functional theory for strong electromagnetic fields in crystalline solids. *Physical Review B*, 85(4):045134, 2012.

Maiku Yamaguchi and Katsuyuki Nobusada. Large hyperpolarizabilities of the second harmonic generation induced by nonuniform optical near fields. *The Journal of Physical Chemistry C*, 120(41):23748–23755, 2016.

Hugen Yan, Xuesong Li, Bhupesh Chandra, George Tulevski, Yanqing Wu, Marcus Freitag, Wenjuan Zhu, Phaedon Avouris, and Fengnian Xia. Tunable infrared plasmonic devices using graphene/insulator stacks. *Nature nanotechnology*, 7(5):330, 2012a.

Hugen Yan, Zhiqiang Li, Xuesong Li, Wenjuan Zhu, Phaedon Avouris, and Fengnian Xia. Infrared spectroscopy of tunable dirac terahertz magneto-plasmons in graphene. *Nano letters*, 12(7):3766–3771, 2012b.

Hugen Yan, Fengnian Xia, Zhiqiang Li, and Phaedon Avouris. Plasmonics of coupled graphene micro-structures. *New Journal of Physics*, 14(12):125001, 2012c.

Hugen Yan, Tony Low, Wenjuan Zhu, Yanqing Wu, Marcus Freitag, Xuesong Li, Francisco Guinea, Phaedon Avouris, and Fengnian Xia. Damping pathways of mid-infrared plasmons in graphene nanostructures. *Nature Photonics*, 7(5):394, 2013.

Jun Yan and Shiwu Gao. Plasmon resonances in linear atomic chains: Free-electron behavior and anisotropic screening of d electrons. *Physical Review B*, 78(23):235413, 2008.

Jun Yan, Zhe Yuan, and Shiwu Gao. End and central plasmon resonances in linear atomic chains. *Physical review letters*, 98(21):216602, 2007.

Lei Yan, Fangwei Wang, and Sheng Meng. Quantum mode selectivity of plasmon-induced water splitting on gold nanoparticles. *ACS nano*, 10(5):5452–5458, 2016.

Wei Yan. Hydrodynamic theory for quantum plasmonics: Linear-response dynamics of the inhomogeneous electron gas. *Physical Review B*, 91(11):115416, 2015.

Honghua U Yang, Jeffrey D'Archangel, Michael L Sundheimer, Eric Tucker, Glenn D Boreman, and Markus B Raschke. Optical dielectric function of silver. *Physical Review B*, 91(23):235137, 2015.

Huayan Yang, Yu Wang, Xi Chen, Xiaojing Zhao, Lin Gu, Huaqi Huang, Juanzhu Yan, Chaofa Xu, Gang Li, Junchao Wu, et al. Plasmonic twinned silver nanoparticles with molecular precision. *Nature communications*, 7:12809, 2016.

Xiaoxia Yang, Zhipei Sun, Tony Low, Hai Hu, Xiangdong Guo, F Javier García de Abajo, Phaedon Avouris, and Qing Dai. Nanomaterial-based plasmon-enhanced infrared spectroscopy. *Advanced Materials*, 30(20):1704896, 2018.

Yi Yang, Owen D Miller, Thomas Christensen, John D Joannopoulos, and Marin Soljacic. Low-loss plasmonic dielectric nanoresonators. *Nano letters*, 17(5):3238–3245, 2017.

Tomokazu Yasuike, Katsuyuki Nobusada, and Michitoshi Hayashi. Collectivity of plasmonic excitations in small sodium clusters with ring and linear structures. *Physical Review A*, 83(1):013201, 2011.

Jun Yin, Harish NS Krishnamoorthy, Giorgio Adamo, Alexander M Dubrovkin, Yidong Chong, Nikolay I Zheludev, and Cesare Soci. Plasmonics of topological insulators at optical frequencies. *NPG Asia Materials*, 9(8):e425, 2017.

Sd C Ying. Hydrodynamic response of inhomogeneous metallic systems. *Il Nuovo Cimento B (1971-1996)*, 23(1):270–281, 1974.

Tomoyuki Yoshie, Axel Scherer, J Hendrickson, G Khitrova, HM Gibbs, G Rupper, C Ell, OB Shchekin, and DG Deppe. Vacuum rabi splitting with a single quantum dot in a photonic crystal nanocavity. *Nature*, 432(7014):200, 2004.

Renwen Yu, Valerio Pruneri, and F Javier García de Abajo. Active modulation of visible light with graphene-loaded ultrathin metal plasmonic antennas. *Scientific reports*, 6:32144, 2016.

A Zangwill and Paul Soven. Resonant photoemission in barium and cerium. *Physical Review Letters*, 45(3):204, 1980.

Gülis Zengin, Martin Wersäll, Sara Nilsson, Tomasz J Antosiewicz, Mikael Käll, and Timur Shegai. Realizing strong light-matter interactions between single-nanoparticle plasmons and molecular excitons at ambient conditions. *Physical review letters*, 114(15):157401, 2015.

Pu Zhang, Johannes Feist, Angel Rubio, Pablo García-González, and FJ García-Vidal. Ab initio nanoplasmonics: The impact of atomic structure. *Physical Review B*, 90(16):161407, 2014.

Runmin Zhang, Luca Bursi, Joel D Cox, Yao Cui, Caroline M Krauter, Alessandro Alabastri, Alejandro Manjavacas, Arrigo Calzolari, Stefano Corni, Elisa Molinari, et al. How to identify plasmons from the optical response of nanostructures. *ACS nano*, 11(7):7321–7335, 2017.

Bo Zhao and Zhuomin M Zhang. Strong plasmonic coupling between graphene ribbon array and metal gratings. *ACS Photonics*, 2(11):1611–1618, 2015.

Jie Zheng, Philip R Nicovich, and Robert M Dickson. Highly fluorescent noble-metal quantum dots. *Annu. Rev. Phys. Chem.*, 58:409–431, 2007.

Wenqi Zhu and Kenneth B Crozier. Quantum mechanical limit to plasmonic enhancement as observed by surface-enhanced raman scattering. *Nature communications*, 5:5228, 2014.

Wenqi Zhu, Ruben Esteban, Andrei G. Borisov, Jeremy J. Baumberg, Peter Nordlander, Henri J. Lezec, Javier Aizpurua, and Kenneth B. Crozier. Quantum mechanical effects in plasmonic structures with subnanometre gaps. *Nature Communications*, 7: 11495, 2016.

Xiaolong Zhu, Weihua Wang, Wei Yan, Martin B Larsen, Peter Bøggild, Thomas Garm Pedersen, Sanshui Xiao, Jian Zi, and N Asger Mortensen. Plasmon–phonon coupling in large-area graphene dot and antidot arrays fabricated by nanosphere lithography. *Nano letters*, 14(5):2907–2913, 2014.

Rashid Zia, Jon A Schuller, Anu Chandran, and Mark L Brongersma. Plasmonics: the next chip-scale technology. *Materials today*, 9(7-8):20–27, 2006.

Jorge Zuloaga, Emil Prodan, and Peter Nordlander. Quantum description of the plasmon resonances of a nanoparticle dimer. *Nano letters*, 9(2):887–891, 2009.

Jorge Zuloaga, Emil Prodan, and Peter Nordlander. Quantum plasmonics: optical properties and tunability of metallic nanorods. *ACS nano*, 4(9):5269–5276, 2010.