Graphene plasmons in nanostructured environments

TOBIAS WENGER

Department of Microtechnology and Nanoscience (MC2)
Chalmers University of Technology
Gothenburg, Sweden 2017
Graphene plasmons in nanostructured environments

Tobias Wenger
Graphene plasmons in nanostructured environments
TOBIAS WENGER
Applied Quantum Physics Laboratory
Department of Microtechnology and Nanoscience
Chalmers University of Technology

Abstract

This thesis explores the combination of electromagnetism with electrons in graphene. Graphene is a one atom thick layer of carbon atoms which contains electrons that exhibit rather special properties in terms of their conduction abilities. The combination of electromagnetism with electrons gives rise to new solutions to the governing equations — solutions with characteristics not quite like normal electromagnetic radiation and not quite like electrons. These solutions, which also exist in normal conductors, are hybrids between electromagnetism and matter, and are usually referred to as plasmons. This thesis is a theoretical study of plasmons in graphene.

Graphene plasmons are investigated by calculating the conductivity of graphene, starting from a Hamiltonian describing the low-energy graphene electrons. The conductivity is calculated using linear response theory in terms of Green’s functions. In order to probe graphene plasmons with electromagnetic radiation, we consider a subwavelength dielectric grating as a coupling structure. We develop the necessary theory to calculate electromagnetic scattering from the combined system of graphene and grating. The techniques considered are a scattering matrix method and a finite element method. Both are applied to compute the scattering coefficients, which contain information about the graphene plasmons and are also obtainable in experiments. We use these techniques to study various aspects of graphene plasmons, such as their nonlocal properties and the effects of impurities. In addition, we examine the response of graphene plasmons to changes in the surrounding environment and apply this for sensing purposes. Finally, we show that graphene plasmons can be controlled using a DC current in the graphene sample.

Keywords: Graphene plasmons, nonlocal response, linear response Green’s functions, subwavelength electromagnetic scattering
List of publications

This thesis is based on the work contained in the following publications, appended in the end of the thesis:

I. **Optical signatures of nonlocal plasmons in graphene**
   Tobias Wenger, Giovanni Viola, Mikael Fogelström, Philippe Tassin, and Jari Kinaret

II. **High-sensitivity plasmonic refractive index sensing using graphene**
    Tobias Wenger, Giovanni Viola, Jari Kinaret, Mikael Fogelström, and Philippe Tassin
    2D Materials **4**, 025103 (2017)

III. **Graphene plasmons in the presence of adatoms**
    Giovanni Viola, Tobias Wenger, Jari Kinaret, and Mikael Fogelström

IV. **Graphene plasmons: impurities and nonlocal effects**
    Giovanni Viola, Tobias Wenger, Jari Kinaret, and Mikael Fogelström
    Manuscript to be submitted

V. **Current-controlled light scattering and asymmetric plasmon propagation in graphene**
    Tobias Wenger, Giovanni Viola, Jari Kinaret, Mikael Fogelström, and Philippe Tassin
    Manuscript submitted to Physical Review B
Specification of my contributions to the appended publications

I. The idea for this publication was developed during discussions among all authors. I performed the calculations and wrote the manuscript with comments from my co-authors.

II. I developed the idea, did the calculations, and wrote the manuscript. Input from my co-authors was continuous during this process.

III. I participated in the discussions and conception of the project as well as in the analysis of the results. I performed the calculations for the plasmon dispersions and co-authored the manuscript.

IV. For this project, I helped formulate the project and how to investigate the plasmons. I also participated in the analysis of the results and helped in the writing of the manuscript.

V. I developed the main idea and performed the calculations. I then analyzed the results together with my co-authors. I wrote the manuscript with comments from my co-authors.
Acknowledgments

First, I want to take this opportunity to thank Mikael Fogelström and Jari Kinaret for taking me in as a PhD student all those years ago. They managed to guide and tutor me over the years and I would like to think I know a great deal more than I used to. Thank you for this interesting time. I would also like to thank Philippe Tassin for his valuable supervision, guidance, and in-depth discussions.

My deepest gratitude to Giovanni Viola who always had time to discuss and share his vast knowledge of physics and related issues. The discussions we had were always engaging, interesting, and always took longer than we intended. Also, I thank Giovanni for his moral support in times when my spirit was low.

I feel also the need to thank the group in which this thesis work has been done, Applied Quantum Physics (AQP) at MC2, Chalmers. This naturally includes all members, past and present, that I have interacted with over the years. The coffee breaks and discussions about everything from physics to international tax law is perhaps the thing that I will miss the most.

I also wish to extend my thanks to all my friends for understanding me not being able to join so many activities lately. Also, thanks for all the times we have done something besides work that has allowed me to think about something else once in a while.

A special thanks to my family for being supportive of this whole endeavor and for understanding my less frequent calls and increased absent-mindedness these last months.

Sofia, thank you for your continuous support and encouragement.


Tobias Wenger
Göteborg, October, 2017

To all my teachers.
Contents

1 Introduction ........................................... 1
   1.1 Plasmonics ...................................... 2
   1.2 Graphene plasmonics ......................... 3
   1.3 Scope of the thesis - limitations .......... 5
   1.4 Thesis outline ................................ 6

2 Graphene ............................................. 7
   2.1 Electron band structure of graphene ....... 7
   2.1.1 Electron distribution and gating ....... 12
   2.2 Graphene production methods — different types of graphene 14
   2.3 Impurities and scattering in graphene .... 16

3 Green’s Functions and Linear Response Theory ....... 19
   3.1 Green’s functions ............................. 19
   3.1.1 Mathematical definition .................. 20
   3.1.2 Equilibrium Green’s function for graphene .... 21
   3.1.3 Non-equilibrium Green’s function formalism .... 22
   3.2 Linear response in Keldysh formalism ....... 27
   3.3 Graphene conductivity in the clean limit ...... 31
   3.4 Mermin relaxation time approximation ....... 33
   3.5 Green’s function self energy ............... 35
   3.5.1 Adatoms coupled to graphene ............. 37
   3.5.2 Dilute distribution of impurity sites — a self-consistent treatment .......... 38

4 The Maxwell Equations ............................... 41
Appendix D Rayleigh-Ritz Variational Method for the Poisson Equation

Appendix E Longitudinal (TM) Plasmon Dispersion Equation for Conducting Sheets

E.1 Longitudinal plasmon dispersion equation

E.2 Non-retarded limit and expansion in small plasmon losses
Chapter 1

Introduction

This is a thesis about plasmons in graphene. Plasmons are collective charge oscillations with an associated electromagnetic field that can occur in conductors [1]. Most of this thesis is focused on specific details of graphene plasmons, but to understand why graphene plasmons are of interest, it is interesting to first consider the larger perspective.

Electromagnetic radiation plays very important roles in all our lives; it provides energy from the sun, without which life on earth would be difficult at best, and it lets us see our surroundings. Visible light is nothing but a small part of the electromagnetic spectrum, with ultraviolet radiation and X rays being the same type of radiation but higher in frequency, and infrared radiation lower in frequency. Indeed, many more parts of the spectrum can be named but we refrain from listing them all. Suffice it to say that this plethora of different frequencies of electromagnetic radiation have numerous technological applications that affect our daily lives. The radio, X-ray imaging, lasers, screens of different kinds, and microwave ovens are all examples of technologies that rely on generating and receiving electromagnetic radiation. In particular, data transmission is one area where electromagnetic radiation has had a large impact. It allows us the send vast amounts of data in a wireless manner or over optical fibers to connect our computers. It also allows us to communicate with each other here on earth via radio and cell phones as well as to communicate with satellites, rovers, and astronauts far out in space.

Electromagnetic radiation is a good way to transmit information; it moves as fast as possible - with the speed of light. In addition, since electromagnetic radiation comes in different frequencies, it is possible to employ various clever data encoding schemes to transmit a lot of data in short amounts of time. Another important aspect of electromagnetic radiation is that we can use it to gather information about our environment, i.e., we can use electromagnetic radiation as a window into nature and
use it to study various phenomena. This ranges from the very large scale where we can detect radiation from distant quasars [2] to the very small scale where we can use visible light to image details down to 100 nm in size [3]. Thus we conclude that electromagnetic radiation is important, both scientifically and technologically, and new ways of controlling, generating, and detecting electromagnetic radiation is of significant interest for future applications.

When it comes to processing of information, electrons are favored over radiation. One of the most important inventions of the 20th century, the transistor, enables electronic computers to function the way they do. This has enabled electronic computers to completely permeate our civilization. Since plasmons are hybrid excitations, they have properties that are similar to both electromagnetic radiation and electrons. Perhaps it is possible to use this duality to extract the “best” properties from both worlds [4, 5].

1.1 Plasmonics

The study of electrons is called electronics; similarly, the field that studies plasmons is called plasmonics. The roots of the field of plasmonics can be traced back to the beginning of the 20th century. However, examples of the use of what is now known as plasmonic effects can actually be traced back much further than that. Well known examples are, for instance, the staining of church windows during medieval times and already the romans used plasmonic effects to color cups. However, these early uses of plasmonics where lacking in their understanding of the underlying physical effects. For such an understanding to be possible, a theory of electromagnetism was needed.

The unification of electricity and magnetism was made by Maxwell in 1865 [6], paving the road for investigation of electromagnetism in a wide range of settings. In 1902, Wood investigated the reflection properties of metallic gratings on a surface and observed unexplained features in their spectra [7], this is usually taken to be the first observation of plasmons in a scientific setting. Langmuir and Tonks experimented with gaseous plasmas in the 1920’s and found waves in the plasma [8], these are now known as Langmuir waves. In 1956, Pines realized there was a connection between Langmuir waves and the electron energy losses in the materials he was investigating. He published his work on energy losses in materials due to excitation of collective oscillations which he called plasmons [9]. Ritchie was around the same time investigating electron energy losses in thin films and discovered that plasmons could exist on or near the material surface.
1.2 Graphene plasmonics

and thus the notion of surface plasmons was born. In 1968, some 60 years after the experiment, Ritchie was able to explain the strange reflection behavior that Wood had found in terms of surface plasmons being excited at the gratings [10].

The fact that plasmons have a smaller wavelength than electromagnetic radiation at the same frequency, leads to localization of energy in the plasmon modes [11]. This causes large field enhancements which in turn leads to enhancement of interaction strengths, e.g., Raman enhancement [11, 12] and increased coupling to dipole emitters [13]. For instance, strong coupling to single molecules has been experimentally demonstrated at room temperature [14]. Furthermore, the ability to localize electromagnetic fields to scales that are subwavelength could enable miniaturized photonic circuits [5, 15], biosensors with high sensitivities [16, 17], as well as microscopy with subwavelength resolution [18].

However, the wave localization comes at a price. Large localization usually leads to large losses, meaning that plasmons become highly damped. A trade-off exists between large localization and long propagation lengths [1]. A common material with relatively small plasmon losses is silver; the surface plasmons can propagate up to 1 mm under the right conditions [15]. However, in this regime the plasmon wavelength is only a few percent smaller than the free-space wavelength [19].

1.2 Graphene plasmonics

Graphene plasmonics is a rather new field of research that merges the exciting properties of plasmons with those of graphene. The first theoretical investigations of graphene plasmons were performed in Refs. [20, 21], where the authors obtained the same $\omega \propto \sqrt{k}$ as normal 2DEG plasmons [22], but a scaling in density as $\omega \propto n^{1/4}$ instead of $\omega \propto \sqrt{n}$ as the 2DEG plasmons. The density dependence of the plasmon frequency in graphene (and 2DEG’s) can be utilized to tune the plasmon frequency by changing the electron density. In graphene, the electron density can be tuned by using a back gate to which a voltage is applied, thus forcing charges in or out of the graphene sheet [23]. The tunability of the plasmon energy of graphene is one of the main advantages of using graphene as a plasmonic material instead of using a more conventional material such as silver. Another difference between graphene plasmons and plasmons in metals is the amount of wave localization provided.

The fact that plasmons exhibit wave localization means that the wave-
length at a specific frequency is smaller for the plasmon than for free-space light. This is a common feature for plasmons, not only graphene plasmons, but the mismatch is larger in graphene plasmons than for conventional plasmons due to the larger localization. On flat surfaces, this means that incident light does not couple to the plasmons [1]. There are several methods to overcome this mismatch and allow incident light to couple to the plasmons. Chapter 5 contains a more thorough treatment of this subject and here we simply note that this mismatch can be overcome and graphene plasmons, as well as metal plasmons, can be probed using electromagnetic radiation. However, this requires special coupling structures.

The early theory for graphene plasmons treated clean graphene at zero temperature [20, 21]. This is sufficient for some purposes, such as understanding the frequency-wavelength relationship at long wavelength. However, it is not sufficient for understanding the propagation distance of graphene plasmons. In the clean limit at zero temperature, the prediction is that graphene plasmons propagate infinitely far, something clearly not observed in experiments [24, 25]. These experiments show that the graphene plasmons experience a significant damping and to model this, extensions of the previously mentioned idealized case are necessary. Impurity scattering effects on graphene plasmons were investigated in Ref. [26] and they found that for the amount of scattering that was estimated from experiments, the graphene plasmons would still exhibit large localization but with substantial losses as a result. The losses are substantial, but when considering also that the localization factor is large, the amount of loss is small. A generalization to finite temperature response in graphene was calculated in Ref. [27] and was also treated in Ref. [28], where finite temperature effects on the plasmon dispersion and the plasmon damping were also considered.

In graphene, as it turns out, the wavelength of electrons at the Fermi energy, called the Fermi wavelength, can be rather large. It can be several tens of nanometers for realistically doped graphene samples. This is comparable to the plasmon wavelength for certain frequencies and, as a result, the wave nature of electrons can affect the wave nature of the plasmons. This is usually called nonlocal effects, see chapter 4 for more on this subject. Nonlocal effects can also be important in metals. However, the Fermi wavelength of metals is almost two orders of magnitude smaller than in graphene. In silver and gold for example, the Fermi wavelength is approximately $5 \, \text{Å}$ [29] and so, a very small plasmon wavelength - or nanopatterning of the metal - is needed to resolve this. Nonlocal effects have been shown to limit the field enhancement resulting from plasmonic effects in
various metallic structures that are patterned on length scales comparable to 5 Å. Examples are spherical nanoparticles [30], half-cylinder grooves [31], and conical nanotips [32]. Nonlocal effects can also cause blue-shifts of plasmon resonances [30, 31] when compared with the local theory (which neglects nonlocal effects). Nonlocality has also been discussed in the context of making strong coupling with small dipole emitters more difficult [33].

By now, graphene plasmons have emerged as a promising platform at terahertz to mid-infrared frequencies [34] for a wide range of applications [35]. For instance, graphene plasmons have been used to enhance photodetectors [36, 37], to perform label-free sensing of biomolecules [38, 39], and for modulation of infrared radiation [40]. Furthermore, in the near future, stacking of various 2D materials is expected to become an additional factor to consider in the context of optoelectronic devices [35].

1.3 Scope of the thesis - limitations

This is a thesis focused on theory and as such, we are interested in constructing models of reality that can be used to explain and understand phenomena around us. All models are imperfect, in that approximations are made and effects that are expected to be less important are not included in the model. This is neatly captured in a quote by the statistician George Box who wrote: “Essentially, all models are wrong. But some are useful.” [41] p. 424. The meaning being that all models contain imperfections, but that might well be sufficient to obtain substantial information and understanding of the system in question.

In this thesis we have neglected the effect of phonons. This is done in order to keep focus on nonlocal effects and other sources of plasmon damping. A more detailed discussion about this can be found in section 2.3. Furthermore, we have neglected electron-electron interactions which have been shown to renormalize the Fermi velocity for small doping levels [42]. We consider doping levels larger than where these effects are appreciable.

What is included in this thesis is a treatment of finite temperatures, nonlocal effects, as well as impurities in different approximations. We shall see that this allows us to obtain a model that captures essential features of graphene plasmons and lets us calculate observable properties that can be compared with experiments.
1.4 Thesis outline

The rest of the thesis is divided into four chapters as well as five appendices. Chapter 2 introduces graphene and treats a tight-binding Hamiltonian to describe electrons moving in the carbon lattice. The chapter also contains the approximation of low energies which gives rise to the concept of Dirac electrons. In chapter 3, Green’s functions are introduced and utilized to formulate a linear response theory to calculate the graphene conductivity. This treatment is rather general and, in particular, the use of Keldysh Green’s functions makes the formalism valid also for systems that are initially out of equilibrium. The chapter also contains a treatment of how to introduce additions to the bare graphene Hamiltonian and how this can be handled in the Green’s function formalism by introducing the concept of self energies.

Chapter 4 treats Maxwell’s equations, energy propagation, nonlocality in response functions, and electromagnetic scattering. Furthermore, this chapter deals with solution methods for electromagnetic scattering in nanostructured environments necessary to probe graphene plasmons. Chapter 5 is devoted to graphene plasmons and their dispersion relation, propagation, and damping as well as how to couple radiation to graphene plasmons. In addition, the chapter includes a section about refractive index sensing using graphene plasmons and ends with a comparison between typical properties of graphene plasmons and metal plasmons.

The appendices consist of calculations that were deemed to lengthy to be put in the main text, but at the same time valuable enough to include in the thesis. Appendix A contains a derivation of how to arrive to the momentum-space Hamiltonian from the real-space tight-binding Hamiltonian for graphene. Appendix B treats how to compute the bare graphene Green’s function. Appendix C contains calculations for electromagnetic fields inside a grating region and D contains explicit calculations of a finite element method solution. Finally, appendix E is a derivation of the plasmon dispersion equation for thin conducting sheets.
Chapter 2

Graphene

In this chapter we calculate the band structure of graphene and, doing so, we explain the emergence of the so-called Dirac approximation that is used throughout the rest of the thesis. The Dirac approximation — as any approximation — has its limitations, but its greatest benefit is that it allows for analytic calculations to be carried out. The Dirac approximation is valid for low energies, which is the region of interest for graphene plasmonics, and thus it is sufficient for our purposes. In this chapter, we treat the basic electronic properties such as wave functions and electron dispersion, as well as electron distributions and doping level. We briefly discuss different graphene production methods and their benefits and drawbacks. The chapter ends with a treatment of impurities in graphene and the related concept of electron mobility — often used to quantify graphene sample quality.

2.1 Electron band structure of graphene

As mentioned in the introduction, graphene is a monolayer of carbon atoms arranged in what is usually called a honeycomb lattice or a hexagonal lattice, see Fig. 2.1. Carbon atoms have six electrons and two of these fill up the inner electronic orbital and stay inert, while four electrons reside in the outer orbital. Of these, three electrons hybridize and create bonds with the neighboring carbon atoms, called $\sigma$-bonds and the remaining electron occupies the so-called $\pi$-band. This band can hold two electrons per atom, and since there is only one electron per carbon atom, the result is a half filled $\pi$-band [43]. These $\pi$-electrons are the ones that are of importance for the electronic properties of graphene, but also the structure of the graphene lattice is important. There are two inequivalent atoms in the lattice structure of graphene, usually called the A and B atoms, and they build up the underlying triangular A and B lattices, see Fig. 2.1.
The lattice vectors, shown in Fig. 2.1, are

\[ \vec{a}_1 = \frac{a\sqrt{3}}{2} (\sqrt{3}, 1), \quad (2.1) \]
\[ \vec{a}_2 = \frac{a\sqrt{3}}{2} (\sqrt{3}, -1), \quad (2.2) \]

where \( a = 1.42 \, \text{Å} \), and the length of \( \vec{a}_1 \) and \( \vec{a}_2 \) is \( a\sqrt{3} \approx 2.46 \, \text{Å} \). The nearest-neighbor vectors, also shown in Fig. 2.1, are

\[ \vec{\delta}_1 = \frac{a}{2} (1, \sqrt{3}), \quad (2.3) \]
\[ \vec{\delta}_2 = \frac{a}{2} (1, -\sqrt{3}), \quad (2.4) \]
\[ \vec{\delta}_3 = a (-1, 0). \quad (2.5) \]

A minimal model for the \( \pi \)–orbital electrons is given by the tight-binding Hamiltonian [43]

\[ H = -t \sum_{\langle i,j \rangle, s} \left( a_{i,s}^{\dagger} b_{j,s} + h.c. \right), \quad (2.6) \]

where \( \langle .. \rangle \) denotes nearest-neighbor summation, \( s = (\uparrow, \downarrow) \) is a sum over spins, and \( t = 2.8 \, \text{eV} \) is the nearest-neighbor hopping amplitude. In
Eq. (2.6), $a_{i,s}$ ($a_{i,s}^\dagger$) is the annihilation (creation) operator of electrons at site $i$ and spin $s$ on the A sublattice and $b_{i,s}$ and $b_{i,s}^\dagger$ are defined similarly for the B sublattice. We now write the operators as a Fourier transform

$$a_{i,s} = \frac{1}{\sqrt{N}} \sum_{\vec{k}} e^{-i\vec{k} \cdot \vec{R}_i} a_s(k),$$

where $N$ is the number of unit cells, $\vec{R}_i$ is the position of site $i$, and the sum of $\vec{k}$ is over the Brillouin zone. The Hamiltonian in Eq. 2.6 can then be written as

$$H = \sum_{\vec{k},s} \begin{pmatrix} a_{\vec{k},s}^\dagger & b_{\vec{k},s}^\dagger \end{pmatrix} \begin{pmatrix} 0 & -t\phi_k \\ -t\phi_k^* & 0 \end{pmatrix} \begin{pmatrix} a_{\vec{k},s} \\ b_{\vec{k},s} \end{pmatrix},$$

where

$$\phi_k = \left( e^{-i\delta_1 \vec{k}} + e^{-i\delta_2 \vec{k}} + e^{-i\delta_3 \vec{k}} \right),$$

see Appendix A for details of the calculation. The eigenvalues of the Hamiltonian (2.8), i.e. the electron bands, are

$$E_k = \pm t|\phi_k|$$

where - (+) corresponds to the valence (conduction) band. Due to the half filling discussed earlier, the valence band is completely filled and the conduction band is completely empty for charge-neutral graphene. By explicitly inserting the expression for the $\delta$’s from Eqs. (2.3)-(2.5), the band structure given by Eq. (2.10) can be written in a somewhat more transparent way as [4, 43]

$$E_k = \pm t\sqrt{3 + 2 \cos(\sqrt{3}k_ya) + 4 \cos(\sqrt{3}k_ya/2) \cos(3k_xa/2)}.$$  

This result was first obtained by Wallace [44] who studied single-layer graphite as early as 1947. The dispersion in Eq. (2.11) is shown in Fig. 2.2 where we clearly see the points where the two bands touch. These points are called the Dirac points and there are six of them in the first Brillouin zone. Out of these six, only two are independent (i.e., cannot be connected by the reciprocal lattice vectors) and linearizing the dispersion around these points, we obtain the dispersion [4, 43]

$$E_k = \pm \frac{3ta}{2} |\vec{k}| = \pm v_F |\vec{k}|,$$

where $\vec{k}$ now is measured from the Dirac point we are linearizing around and we have defined the Fermi velocity $v_F = 3ta/2 \approx 10^6$ m/s.
Perhaps the most striking property that arises is that the low energy electrons behave as massless Dirac particles. The linear dispersion relation resembles the photon dispersion, but with the slope, i.e. the propagation speed, equal to $v_F$ instead of $c$ for photons. The speed of the Dirac electrons, $v_F$, is approximately $v_F \approx c/300 \ (10^6 \text{ m/s})$.

Above, we simply linearized the dispersion relation to obtain the linear dispersion around the Dirac points. However, it is also possible to linearize the Hamiltonian and solve for the eigenenergies and eigenvectors. There are only two inequivalent points in the full dispersion to linearize around; we call these $\vec{K}$ and $\vec{K}'$. From these two points it is possible to arrive at the other points by translating with the reciprocal lattice vectors, thus making the other points equivalent to $K$ and $K'$. The linearized Hamiltonians that one obtains from the linearization procedure are \[45\]

\[
\hat{H}_{\vec{K}}(\vec{k}) = v_F \begin{pmatrix} 0 & k_x - ik_y \\ k_x + ik_y & 0 \end{pmatrix} = v_F \vec{\sigma} \cdot \vec{k}, \tag{2.13}
\]

\[
\hat{H}_{\vec{K}'}(\vec{k}) = v_F \begin{pmatrix} 0 & k_x + ik_y \\ k_x - ik_y & 0 \end{pmatrix} = v_F (\vec{\sigma} \cdot \vec{k})^T, \tag{2.14}
\]

where $\vec{\sigma} = (\sigma_x, \sigma_y)$, i.e., a vector of Pauli matrices. Now, the Hamiltonian $\hat{H}_{\vec{K}}$ acts on spinors for the $\vec{K}$ valley and $\hat{H}_{\vec{K}'}$, on the spinors for the $\vec{K}'$
2.1 Electron band structure of graphene

valley, i.e.,

\[
\hat{H}_{\vec{k}} \begin{pmatrix} \psi_{KA}(\vec{k}) \\ \psi_{KB}(\vec{k}) \end{pmatrix} = E_\vec{k} \begin{pmatrix} \psi_{KA}(\vec{k}) \\ \psi_{KB}(\vec{k}) \end{pmatrix},
\]

(2.15)

\[
\hat{H}_{\vec{k}'} \begin{pmatrix} \psi_{K'A}(\vec{k}) \\ \psi_{K'B}(\vec{k}) \end{pmatrix} = E_\vec{k} \begin{pmatrix} \psi_{K'A}(\vec{k}) \\ \psi_{K'B}(\vec{k}) \end{pmatrix},
\]

(2.16)

where the \( \psi \) subscripts denote the valley and A or B the sub-lattice that the wave function belongs to. It may sometimes be convenient to put all this together in one matrix equation construct and this can be done by putting the Hamiltonians in block-diagonal form

\[
\begin{pmatrix} \hat{H}_{\vec{k}} & 0 \\ 0 & \hat{H}_{\vec{k}'} \end{pmatrix} \Psi(\vec{k}) = E_\vec{k} \Psi(\vec{k}),
\]

(2.17)

where

\[
\Psi(\vec{k}) = \begin{pmatrix} \psi_{K,A}(\vec{k}) \\ \psi_{K,B}(\vec{k}) \\ \psi_{K',A}(\vec{k}) \\ \psi_{K',B}(\vec{k}) \end{pmatrix}.
\]

(2.18)

In the interest of high degree of symmetry in the equations this is sometimes recast in a different form where one rearranges the \( K' \) spinor so that it has the same Hamiltonian as the \( K \) spinor:

\[
\begin{pmatrix} \hat{H}_{\vec{k}} & 0 \\ 0 & \hat{H}_{\vec{k}} \end{pmatrix} \Psi_R(\vec{k}) = v_F \left( \tau_0 \otimes \vec{\sigma} \cdot \vec{k} \right) \Psi_R(\vec{k}) = E_\vec{k} \Psi_R(\vec{k}),
\]

(2.19)

where

\[
\Psi_R(\vec{k}) = \begin{pmatrix} \psi_{K,A}(\vec{k}) \\ \psi_{K,B}(\vec{k}) \\ \psi_{K',B}(\vec{k}) \\ -\psi_{K',A}(\vec{k}) \end{pmatrix},
\]

(2.20)

denotes the rearranged 4-component vector and \( \tau_0 \) is a unit matrix in valley indices.

Using the Hamiltonians (2.13) and (2.14), the momentum space eigen-spinors can be calculated [43]:

\[
\psi_K(\vec{k}) = \begin{pmatrix} \psi_{KA} \\ \psi_{KB} \end{pmatrix} = \frac{1}{\sqrt{2}} \left( e^{-i\theta_\vec{k}/2} \pm e^{i\theta_\vec{k}/2} \right),
\]

(2.21)

\[
\psi_{K'}(\vec{k}) = \begin{pmatrix} \psi_{K'A} \\ \psi_{K'B} \end{pmatrix} = \frac{1}{\sqrt{2}} \left( e^{i\theta_\vec{k}/2} \pm e^{-i\theta_\vec{k}/2} \right),
\]

(2.22)
where $\theta_k = \arctan(k_x/k_y)$ and $+ (-)$ means conduction (valence) band, and corresponds to the sign in Eq. (2.12). The eigenvalues of the Hamiltonians (2.13) and (2.14) reproduce the linear electron dispersion in Eq. (2.12).

The low-energy picture that has now emerged, described by the linear energy spectrum in Eq. (2.12) and the eigenspinors in Eqs. (2.21) and (2.22), can be visualized by the Dirac cone. This is depicted in Fig. 2.3 which shows the two linear energy bands, the conduction band (positive energy) and the valence band (negative energy), and the Fermi energy denotes the highest occupied state at zero temperature. As already mentioned, neutral graphene is filled to the Dirac point, i.e., to the vertex of the cone in Fig. 2.3. It can be noted that the A, B sublattices are now encoded in the vectorial structure of the eigenspinors. It is often convenient to work with only one of the eigenspinors when performing calculations of physical observables. Then the degeneracy of $g_s g_v = 4$ from spin and valley degeneracies need to be inserted in the final result.

### 2.1.1 Electron distribution and gating

As depicted in Fig. 2.3, the graphene band structure contains electrons that are distributed in energy. The charge carriers in graphene are fermions, and as such they obey the Fermi distribution

$$n_F(E_k) = \frac{1}{1 + e^{(E_k - \mu)/(k_B T)}},$$

(2.23)
where \( \mu \) is the chemical potential, and \( E_k = \lambda v_F k \) is the energy of an electron at momentum \( k \) determined by Eq. 2.12. \( \lambda = \pm \) signifies the band, with + being the conduction band and − being the valence band. The Fermi distribution describes the distribution of electrons in equilibrium, defined by two parameters \( \mu \) and \( T \). The zero-temperature chemical potential is usually called the Fermi energy and the chemical potential is fixed by enforcing particle conservation when the temperature is increased from zero.

The carrier density — or, equivalently, the chemical potential — can be tuned in graphene using a metal gate as was first shown in Ref. [23]. Pristine, undoped graphene is charge neutral and has a Fermi energy at the Dirac point, i.e. at zero energy in Fig. 2.3. The Fermi energy is related to the charge carrier density in graphene as [46]

\[
E_F = v_F \sqrt{\pi n},
\]

(2.24)

or, equivalently,

\[
n = \frac{E_F^2}{v_F^2 \pi}.
\]

(2.25)

The Fermi energy can lie in the conduction band (electron doped) or in the valence band (hole doped). By changing the gate voltage between graphene and a metal gate, charges can be pushed into the graphene sample or extracted from it. Since these electrons (or holes) must occupy a state in the Fermi distribution, this raises or lowers the Fermi energy. The charge carrier density in graphene can be tuned significantly using a metal gate in close proximity to the graphene sheet. This effect is called the electric field effect and is present in more conventional conductors as well. However, the difference between these conductors and graphene is that the number of charge carriers in graphene is very small compared to other conductors, thus making small changes in the density more pronounced. Actually, pristine graphene has zero free charge carriers (at zero temperature) as the Fermi energy in this case lies at the Dirac point that separates the conduction and valence band. As an example of the large difference in number of charge carriers, we compare a \( \mu \text{m}^2 \) of graphene with a typical carrier density of \( 10^{12} \text{ cm}^{-2} \) and a \( \mu \text{m}^3 \) of silver with a carrier density of \( 5.86 \times 10^{22} \text{ cm}^{-3} \) [29]. The resulting number of charge carriers are \( 10^4 \) for the patch of graphene, and \( 5.86 \times 10^{10} \) for the silver cube. The difference of number of charge carriers is then more then a factor \( 10^6 \), illustrating the point above.

There are also other ways, besides using electrostatic gating, of changing the carrier density of graphene. Examples are chemical doping with foreign
Graphene atoms and molecules [47] as well as photo-excitation of charge carriers [48]. Chemical doping can occur via surface transfer of electrons to/from dopants on the graphene surface or via substitution of carbon atoms in the graphene lattice with other elements [47].

So far we discussed electrons in equilibrium, which are fully described by the temperature and the chemical potential which sets the number of electrons. In general, systems can be in a complicated non-equilibrium state. A special non-equilibrium state can be achieved by letting a time independent force act on the system for a long time until the system no longer evolves in time. This is called a stationary state and it is a close non-equilibrium analogue of an equilibrium state, with the similarity being that neither system evolves in time. Under the action of an electric field bias across graphene and some electron scattering, needed to provide resistance for the electron movement, the stationary distribution will take the form [49]

\[ f(E_k) = \frac{1}{1 + e^{(E_k - \beta \cdot \mathbf{k} - \mu)/(k_B T)}} , \]

(2.26)

where \( E_k \) is the standard graphene electron dispersion, \( \mu \) is the chemical potential, and \( \beta \) is the drift velocity. The difference between this stationary electron distribution and the equilibrium Fermi distribution is the appearance of the drift velocity. This shifts the electron distribution in a certain direction in \( k \)-space, so that more electrons move in a certain direction than in the opposite direction, i.e., a net DC current is flowing. This is only possible if a force is constantly forcing the electrons to move and the circuit is closed so that no charge build-up is occurring. This can be achieved by connecting two contacts and applying a DC bias across the graphene. How large the drift velocity is, i.e., how far from the equilibrium distribution the system evolves depends on the size of the voltage bias, the resistance of the sample, and subsequently on how large the current is. It can be noted that pushing a large DC current through graphene will tend to heat the graphene sample due to resistive heating. This can in principle be modeled since the temperature is already included in the distribution (2.26). In practice, however, the temperature is usually considered fixed and the heating effect is considered to be small, see e.g. Ref. [50]. This can be achieved with a good heat-sink attached to the system.

2.2 Graphene production methods — different types of graphene

Industrial use of graphene in large scale production requires low-cost production methods that are capable of producing good-quality graphene
samples [51]. There is also a need to develop laboratory-scale methods for good-quality graphene for use in prototype devices and fundamental studies. Graphene production presents a challenge and it was even believed for a long time that isolated graphene flakes were unstable due to thermal fluctuations [52] and, therefore, impossible to produce. From this standpoint, the successful isolation of graphene in Ref. [23] came as a surprise. Nowadays, graphene can be produced using a variety of methods [51] and a few select ones are mechanical exfoliation, chemical vapor deposition, epitaxial growth on silicon carbide, and liquid phase exfoliation. These methods have their benefits and drawbacks which we briefly discuss below.

A small note on graphene quality is in order before we treat graphene production methods. The quality of graphene is often quantified in terms of the electron mobility. High mobility means a high-quality sample. The mobility is limited by electron scattering which may come from different sources. We leave a more detailed treatment of the mobility and impurity scattering to section 2.3. Here, it is enough to know that one of the main difficulties in graphene production is to avoid introducing anything to the graphene that causes electron scattering, as this will reduce the mobility and hence the graphene quality.

Mechanical exfoliation — sometimes referred to as the scotch-tape method — is the original production method used by Novoselov et al in Ref. [23], which resulted in their Nobel prize in 2010. This method starts with a thin flake of graphite, which is exfoliated repeatedly using a sticky tape, yielding high quality samples of graphene. However, because of its labor intensive nature and small graphene sample sizes, this method is not suitable for large-scale production. Due to the high mobilities, it remains the method of choice for fundamental studies and prototype devices [51]. Using this method, room temperature mobilities of 140,000 cm²/(Vs) [53] and low-temperature mobilities of 10⁶ cm²/(Vs) [42, 53] have been achieved.

Chemical Vapor Deposition (CVD) uses gaseous carbon to grow graphene on a substrate, usually copper although other substrates such as nickel or platinum are possible. This method is often seen as the most viable for large scale production as this can be performed over large areas of copper simultaneously. Using CVD, it is possible to produce graphene sheets as large as 30 inch (76 cm) across the diagonal [54]. The growth of graphene is a self-limiting process which stops the growth after the copper film is covered with graphene [51]. A drawback of the CVD method is that graphene ends up on top of copper, which is a conductor, something not very favorable for investigating the properties of graphene. To get graphene on
an insulator, many ways of either etching away the copper or transferring the graphene off the copper to some other substrate have been developed. However, this usually damages the graphene, especially when wet transfer methods are used. Using a dry transfer technique, mobilities up to $3 \times 10^6$ cm$^2$/Vs at low temperatures have recently been achieved when transferring onto hBN substrates \cite{55}, thus rivaling graphene produced using mechanical exfoliation.

Silicon carbide (SiC) is a layered material containing both silicon and carbon atoms. In a process called epitaxial growth, heating SiC to high temperatures facilitates evaporation of the Si atoms from the material and leaves graphene behind \cite{51}. This is not a self-limiting process and the result may be that different number of graphene layers are mixed together in different areas. If graphene is grown on the Si(0001)-face, it creates a graphene layer with many of the carbon atoms creating a bond to the underlying SiC, this is usually referred to as the buffer layer \cite{51}. By using hydrogen intercalation, this layer can be decoupled from the SiC to create a quasi-free-standing graphene layer on top of SiC. Using epitaxial growth on SiC, it is possible to produce cm-size graphene, but one of the main drawbacks of this method is the high price due to the high cost of the SiC wafers. Room-temperature mobilities ranging up to 10.000-30.000 cm$^2$/Vs have been achieved using epitaxial growth on SiC \cite{51}.

It is also possible to exfoliate graphite in a liquid environment, where ultrasound can be used to separate the graphene layers \cite{51}. This is called liquid phase exfoliation and is scalable and cheap, as it has no need for potentially expensive substrates. The downside with the method is its small flake sizes and limited graphene quality, this may limit its usefulness for electronic and plasmonic purposes. However, this method can for instance be used to produce conducting inks containing graphene, which enables graphene-based printable electronics \cite{56}. Furthermore, this method can be used to produce few-layer materials where one or several layers is a different material that has been intercalated into the graphite. This can be used to create materials with new exciting properties that can be different from those of graphene \cite{51}.

### 2.3 Impurities and scattering in graphene

In section 2.1, the graphene lattice was considered to be perfect and the equations were solved considering no edges and a completely translationally invariant system. However, we discussed in section 2.2 that depending on production method, there are various imperfections in the graphene lat-
Impurities and scattering in graphene

2.3 Impurities and scattering in graphene

tice as well as impurities attached to the surface of graphene. A complete
in-depth investigation of different imperfections and impurities is beyond
the scope of this thesis. We focus on general implications of imperfections
for experiments and what information can be gathered from them.

Phonons exist both in graphene itself as well as in any substrate that
may be used as a support for graphene. This has been shown to create a
hybridization between the plasmons and the phonon modes, close to the
bare phonon frequencies [57]. Considering that we often work below the
intrinsic phonon mode at approximately 50 THz [57], we assume that we
can neglect phonons in this thesis. The phonons can to some degree of
approximation be thought of as being included in the electron relaxation
time we include in our model. This approach was taken in Ref. [26], where
the authors also constructed an energy dependent relaxation time to more
accurately model the phonon effect.

Apart from phonons, the graphene lattice can contain charged impuri-
ties which seem to be a dominant source of electron scattering [58, 59],
limiting the DC conductivity. The lattice can also contain vacancies which
also affect the electron transport in graphene [46]. In addition, grain
boundaries can exist where different domains are separated by extended
lattice mismatches. It was found in Ref. [60] that midgap states may occur
in vacancies and along grain boundaries, and this can have similar effects
on the transport properties as charged impurity scattering. Furthermore,
the presence of grain boundaries have theoretically been found to destroy
the quantum hall effect in graphene, depending on the orientation of the
grain boundaries [61]. In addition, the presence of imperfections has been
shown to give rise to interesting behavior for electronic thermal transport
in graphene. In particular, the Wiedemann-Franz law has been shown to
be violated in the presence of randomly distributed impurities [62] as well
as in the presence of midgap states [60].

A common way to quantify the quality of graphene samples is to mea-
sure the mobility. The mobility is related to the DC conductivity as [23]
\[ \sigma = \frac{en\mu_m}{m}, \]  
where \( e \) is the electron charge, \( n \) is the carrier density, and we denote the
mobility with \( \mu_m \) to avoid confusion with the chemical potential (com-
monly denoted with \( \mu \)). Measuring the mobility effectively includes all
scattering processes discussed above. The mobility is also related to the
drift velocity as [29]
\[ v_d = \mu_mE, \]  
where \( v_d \) is the electron (or hole) drift velocity, i.e., the average velocity
of the charge carriers and $E$ is the applied electric field. The mobility is usually given in units of cm$^2$/(Vs) and it varies between different graphene samples and production methods as we saw in section 2.2. Mobility is often inferred by measuring the DC resistivity (conductivity), and by also knowing the carrier density it is possible to extract the carrier mobility using Eq. (2.27). The mobility is limited by all the different scattering sources discussed above. To improve the quality of graphene samples, it is of interest to know what the impurities are in order to remove their effect. In the original graphene article by Novoselov et al, Ref. [23], mobilities up to $10^4$ cm$^2$/(Vs) were found with essentially no temperature dependence, indicating that the mobility was impurity limited rather than phonon limited. Since the publication of this seminal article, much effort has been devoted to producing larger samples with higher mobilities, i.e., with less charged impurities. In more recent works, the mobility is increased by one order of magnitude, and at room temperature is found to be close to $10^5$ cm$^2$/(Vs) [53, 63, 64]. This can be achieved by cleaning the graphene in-situ by sending a current through it which “burns” the impurities away and/or by encapsulating the graphene with hexagonal boron nitride (hBN) [65]. The latter is a recent development which simultaneously achieves several things; hBN is very flat which leads to the graphene laying flat on top [65], it has a lattice constant close to graphene so the graphene lattice is unstrained [65], and it can encapsulate the graphene to protect from particles in the environment [66]. It can be noted that the room-temperature mobility has been theoretically shown to be limited to roughly $10^5$ cm$^2$/(Vs) due to phononic scattering [67].

By matching the conductivity in Eq. (2.27) with a simple Drude expression for the graphene conductivity, it is possible to relate the mobility and the DC relaxation time [26, 68]:

$$\tau = \frac{\mu m E_F}{ev_F^2}.$$  
(2.29)

This can be used to estimate relaxation times for the electrons from measurements of graphene mobility. Using data from experiments, these relaxation times range from 1 ps down to 20 fs [69], depending on the quality of graphene. As can be expected, the quality of graphene has implications for the plasmons in graphene. High-quality graphene, i.e., a long relaxation time, is advantageous for plasmonics; we will study this further in chapter 5.
Chapter 3

Green’s Functions and Linear Response Theory

In this chapter we introduce Green’s functions as a tool to calculate observable properties of graphene. We review basic concepts and then describe a linear response theory in terms of Green’s functions. We apply this to calculate the nonlocal graphene conductivity — the conductivity is central for the optical as well as the plasmonic properties investigated later in this thesis. We evaluate the conductivity and investigate the differences between the nonlocal and the local conductivity.

Up to this point, graphene has been treated as a perfect lattice where all damping pathways — giving rise to the real part of the conductivity — arise from the excitation of electron-hole pairs. As discussed in chapter 2, graphene always contains a multitude of impurities and imperfections and a perfect-lattice model cannot capture all the physics of real graphene samples. To address this limitation, we discuss models for including impurities in graphene. First we consider a commonly used method due to Mermin, where electron relaxation is introduced in a phenomenological way. Next, we take advantage of the Green’s function formalism and the ability to include impurities in a single-particle self energy term for the Green’s function.

3.1 Green’s functions

Green’s functions are used in many branches of physics and one of their main advantages in quantum mechanics is the rather straightforward way in which observables can be calculated. Observable quantities are of course an important part of condensed matter physics as it allows us to test theories with experiments.
3.1.1 Mathematical definition

Although we will focus on Green’s functions in quantum mechanics, it is instructive to first understand how the Green’s function can express the solution to a differential equation. Consider the differential equation

\[ L \phi(x) = f(x), \quad (3.1) \]

where \( L \) is a linear differential operator and \( f \) is some forcing function (source term). To solve this equation we must find a \( \phi \) such that it is fulfilled. The Green’s function allows us to do this for any function \( f \).

Consider the following equation:

\[ L G(x, x') = \delta(x - x'), \quad (3.2) \]

where \( G \) is the Green’s function of \( L \). By now writing an ansatz of the solution to Eq. (3.1) as

\[ \phi(x) = \int dx' G(x, x') f(x'), \quad (3.3) \]

we can insert this into the left hand side of Eq. (3.1) to obtain

\[ L \phi(x) = L \int dx' G(x, x') f(x') = \int dx' L G(x, x') f(x'). \quad (3.4) \]

Using the definition of the Green’s function, Eq. (3.2), we find

\[ L \phi(x) = \int dx' \delta(x - x') f(x') = f(x), \quad (3.5) \]

which means that the ansatz \( \phi(x) \) given by the Green’s function construction in Eq. (3.3) satisfies the original equation, Eq. (3.1). Note that we made no assumption on the function \( f(x) \), so the solution works for any function \( f \). This shows the power and generality of the Green’s function method. However, we did not actually find the Green’s function that satisfies Eq. (3.2) and this can in general be quite difficult, but once it is found for some differential operator \( L \), the solution for any source term can be written down easily.

The step to Green’s functions in quantum mechanics is to take Schrödinger’s equation, i.e.,

\[ (E - \hat{H}(x)) \psi(x) = 0, \quad (3.6) \]

as the operator \( L \). In Eq. (3.6), \( \hat{H}(x) \) is the Hamiltonian operator. We can then define the Green’s function as [70]

\[ (E - \hat{H}(x)) G(x, x') = \delta(x - x'). \quad (3.7) \]
This can be written in a representation independent form as [70]

\[ G^{r/a}(E) = \lim_{\eta \to 0} \left((E \pm i\eta)1 - \hat{H}\right)^{-1}, \]  

(3.8)

where + (−) gives the retarded (advanced) Green’s function and 1 is an identity operator of appropriate size.

### 3.1.2 Equilibrium Green’s function for graphene

By use of Eq. (3.8), the Dirac Hamiltonian in section 2.1 can be inverted and shown to have a Green’s function on the form

\[ G_0(\vec{k}, \varepsilon) = \frac{1}{2} \sum_{\lambda = \pm} \frac{1}{\varepsilon - \lambda E_k} \begin{pmatrix}
 1 \\
 1
\end{pmatrix}, \]  

(3.9)

where \( E_k = v_F k, \) \( k = |\vec{k}|, \) and \( \phi_k = \text{arg}(k_x + ik_y), \) see Appendix B for details of the calculation. The energy should be interpreted as \( \varepsilon \to \varepsilon + i\eta \) \( (\varepsilon \to \varepsilon - i\eta) \) where \( \eta \) is an infinitesimally small positive number for a correct definition of the retarded (advanced) Green’s function. The matrix structure of the Green’s function in Eq. (3.9) arises due to the underlying A,B sublattice in graphene.

For future reference we define

\[ h^r(\vec{k}, \varepsilon) = \frac{1}{2} \sum_{\lambda = \pm} \frac{1}{\varepsilon - \lambda E_k - \Sigma^r(\varepsilon)}, \]  

(3.10)

\[ f^r(\vec{k}, \varepsilon) = \frac{1}{2} \sum_{\lambda = \pm} \frac{\lambda}{\varepsilon - \lambda E_k - \Sigma^r(\varepsilon)}, \]  

(3.11)

and similarly for \( h^a \) and \( f^a. \) In writing Eqs. (3.10) and (3.11) we have also introduced the (retarded) self energy \( \Sigma^r(\varepsilon). \) We will discuss this further in section 3.1.3. Here, we simply mention that we regain the bare Green’s function in Eq. (3.9) by setting the retarded self energy \( \Sigma^r(\varepsilon) = -i\eta, \) and for the advanced self energy \( \Sigma^a(\varepsilon) = i\eta. \) This now lets us write the equilibrium Green’s function as

\[ G^x_0(\vec{k}, \varepsilon) = \begin{pmatrix}
 h^x(\vec{k}, \varepsilon) & f^x(\vec{k}, \varepsilon)e^{-i\phi_k} \\
 f^x(\vec{k}, \varepsilon)e^{i\phi_k} & h^x(\vec{k}, \varepsilon)
\end{pmatrix}, \]  

(3.12)

where \( x \in \{r, a\}. \) The definitions in Eqs. (3.10) and (3.11) can also be used for any self energies and we will use this later in section 3.2.

The equilibrium Green’s function in Eq. (3.9) can for instance be used
to compute the density of states for graphene electrons using the formula

$$\rho(\varepsilon) = -\frac{g_sg_v}{\pi} \text{Im} \left[ \text{Tr} \int \frac{d^2k}{(2\pi)^2} G_0(\vec{k}, \varepsilon) \right], \quad (3.13)$$

where Tr denotes trace and $g_s = g_v = 2$ are the spin and valley degeneracy factors. The off-diagonal Green’s function components vanish when performing the angular integral and only the radial integral remains:

$$\rho(\varepsilon) = -\frac{g_sg_v}{2\pi^2} \text{Im} \left[ \sum_{\lambda} \int_0^{k_c} \frac{dk}{\varepsilon - \lambda v_F k} \right]$$

$$= \frac{g_sg_v}{2\pi^2 v_F^2} \text{Im} \left[ \ln (\varepsilon - E_c) + \ln (\varepsilon + E_c) - 2 \ln(\varepsilon) \right], \quad (3.14)$$

and

$$\rho(\varepsilon) = \frac{g_sg_v}{2\pi^2 v_F^2} \left( \pi - \begin{cases} 2\text{Im} [\ln(\varepsilon)], & \varepsilon > 0 \\ 2\text{Im} [i\pi + \ln(-\varepsilon)], & \varepsilon < 0 \end{cases} \right)$$

$$= \frac{g_sg_v}{2\pi^2 v_F^2} |\varepsilon|, \quad (3.15)$$

where $k_c$ is a momentum cut-off and $E_c = v_F k_c$ is the energy cut-off. This energy cut-off is always much larger than $\varepsilon$ and thus the first logarithm in Eq. (3.15) is always negative and contributes $i\pi$ to the expression. The second term has a real positive argument and can be dropped and the last term can be written as

$$\rho(\varepsilon) = \frac{g_sg_v}{2\pi^2 v_F^2} |\varepsilon|, \quad (3.16)$$

which is the final answer for the bare density of states. Graphene has, in addition to linear dispersion for the charge carriers, a linear density of states. This linear density of states is symmetric with respect to the Dirac point ($\varepsilon = 0$).

### 3.1.3 Non-equilibrium Green’s function formalism

In this section we introduce a non-equilibrium Green’s function formalism, originally due to Kadanoff and Baym [71] and Keldysh [72]. We shall refer to this as the Keldysh Green’s function formalism or simply as the Keldysh formalism. We will only consider fermionic Green’s functions and we introduce general concepts needed in the rest of the thesis, rather than focus on detailed derivations. Such details can be found in numerous publications [70, 71, 72, 73, 74].

Non-equilibrium Green’s functions are far more general than equilibrium Green’s function in that — as the name suggests — they also allow
us to describe systems out of equilibrium. Many situations of interest in condensed matter physics, and elsewhere, involve the description of transport of energy, charge or some other quantity. Such systems are not in equilibrium and as such they fall outside the descriptive range of equilibrium Green’s functions. The standard way to handle this is to develop linear response theory in which the transport properties are described by using only the equilibrium properties of the system. Such descriptions can then be handled by linear response theory. The Keldysh formalism allows us to go beyond such response in powers of the perturbation. In this thesis, however, we develop a linear response theory, see section 3.2, in terms of the Keldysh formalism. Our reason for doing so is that the description of the linear response becomes rather general and many different physical aspects can be included in the modeling of the physical system. Much of the formalism can also be used to go beyond the linear response theory in a rather straightforward manner.

To define Green’s functions it is sometimes useful to introduce field operators $\psi^\dagger(rt)$ ($\psi(rt)$), which are the creation (annihilation) operators for the field at point $r$ and time $t$. We study fermionic field operators which obey the (equal time) anticommutation relations

$$\{\psi(rt), \psi^\dagger(r't')\} = \delta(r - r'), \quad (3.18)$$

where $\{ , \}$ denotes the anticommutator. The time-ordered (also called causal) Green’s function can be written in terms of these field operators as [70]

$$G(rt, r't') = -i\frac{\langle\psi_0|T[\psi(rt)\psi^\dagger(r't')]|\psi_0\rangle}{\langle\psi_0|\psi_0\rangle}, \quad (3.19)$$

where $T$ denotes the time-ordering operator. It explicitly puts the function with the largest time argument to the left in the expression. In the above definition of the Green’s function we use the notation that $rt$ denotes one argument which contains all spatial coordinates as well as the time coordinate. Thus, the Green’s function has two arguments as in section 3.1.1.

The time-ordered Green’s function turns out to allow for a well-defined perturbation expansion in equilibrium for perturbations to an unperturbed Hamiltonian. For non-equilibrium systems, this turns out to not be true in general and the insight of Kadanoff and Baym as well as Keldysh was that this problem can be handled by introducing more Green’s functions defined along a non-trivial contour in complex time coordinates. Time-ordering along this contour gives rise to new Green’s functions which expand the Green’s function space.
The Keldysh contour is depicted in Fig. 3.1 and it runs from minus infinity to infinity and back again. This ensures that we only make reference to the initial state \( t = -\infty \) and not the unknown final state \( t = \infty \) in the formalism. In equilibrium, the final state can safely be assumed to be the same as the initial state, but this is no longer true when the system is out of equilibrium. Ordering along this contour now means that on the upper \((+)\) branch, time ordering is performed using \( T \), and on the lower \((-)\) branch, operators are anti-time ordered, i.e., the opposite of \( T \).

There are now four possibilities for where the two time arguments lie on this contour and we summarize the resulting Green’s functions below:

1. \( t \in +, t' \in + \): **time-ordered Green’s function**
   \[
   G^{++}(rt, r't') = -i\langle \psi_0 | T [\psi(rt)\psi^\dagger(r't')] | \psi_0 \rangle.
   \]
   This is the same Green’s function as we defined for equilibrium.

2. \( t \in +, t' \in - \): **lesser Green’s function**
   \[
   G^{+-}(rt, r't') = i\langle \psi_0 | \psi^\dagger(r't')\psi(rt) | \psi_0 \rangle.
   \]
   All times in the + branch are earlier on the contour than any time on the - branch. We can immediately put the operators in the correct order and no time ordering operator is needed.

3. \( t \in -, t' \in + \): **greater Green’s function**
   \[
   G^{-+}(rt, r't') = -i\langle \psi_0 | \psi(rt)\psi^\dagger(r't') | \psi_0 \rangle.
   \]
   Same argumentation as the for the lesser Green’s function. The only difference is that we put the operators in the opposite order and pick up a minus sign due to the fermionic anticommutation of the field operators.

4. \( t \in -, t' \in - \): **anti time-ordered Green’s function**
   \[
   G^{--}(rt, r't') = -i\langle \psi_0 | \bar{T} [\psi(rt)\psi^\dagger(r't')] | \psi_0 \rangle.
   \]
   Here we have defined \( \bar{T} \) to do the opposite of \( T \), i.e., it puts the largest time to the right in the expression. This ensures that the resulting Green’s function has its operators ordered along the contour.
In what follow we will not explicitly write out the arguments of the Green’s functions but it is to be understood that when nothing else is noted they are real space Green’s functions with two arguments, \(rt\) and \(r't'\), as above. The four Green’s functions can be neatly put together in a 2-by-2 matrix form
\[
\tilde{G} = \begin{pmatrix} G^{++} & G^{+-} \\ G^{-+} & G^{--} \end{pmatrix},
\]
(3.20)
where the tilde on \(\tilde{G}\) denotes Keldysh space. It is important to keep in mind that the Keldysh matrix structure is completely unrelated to the A-B sublattice matrix structure of the equilibrium Green’s function in Eq. (3.9), although they appear similar in form.

From Eq. (3.20) it seems like there are four Green’s functions to solve for but upon inspection of the definitions 1-4 above, one can show that the functions are not all independent. For instance
\[
G^{++} + G^{--} = G^{+-} + G^{-+},
\]
(3.21)
and furthermore one can find the relationships
\[
G^r = G^{++} - G^{+-} = G^{-+} - G^{--},
\]
(3.22)
\[
G^a = G^{++} - G^{-+} = G^{+-} - G^{--},
\]
(3.23)
where \(r\) denotes retarded, \(a\) denotes advanced Green’s function. By also defining the Keldysh Green’s function as
\[
G^K = G^{++} + G^{--} = G^{+-} + G^{-+},
\]
(3.24)
we may rewrite the Keldysh space expression in Eq. (3.20) from the contour \((+-)\) into retarded, advanced, and Keldysh functions, using the unitary transformation
\[
\tilde{G} = U\tilde{G}U^{-1} = \begin{pmatrix} 0 & G^a \\ G^r & G^K \end{pmatrix},
\]
(3.25)
\[
U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix},
\]
(3.26)
where the caron on \(\tilde{G}\) denotes the so-called triangular representation of Keldysh space. The functions \(G^r\) and \(G^a\) are almost always related to each other [70], meaning there are only two independent functions to be determined, \(G^r\) and \(G^K\). In equilibrium, this reduces to one unknown function as expected, since the Keldysh Green’s function is then determined by the retarded and advanced components as [70]
\[
G^K_{eq}(p, \varepsilon) = f(\varepsilon) \left( G^r_{eq}(p, \varepsilon) - G^a_{eq}(p, \varepsilon) \right),
\]
(3.27)
where \( f(\varepsilon) = \tanh(\beta(\varepsilon - \mu)/2) = (1 - 2n_F(\varepsilon)) \), \( \beta = 1/(k_B T) \), and \( n_F(\varepsilon) \) is the Fermi function, Eq. (2.23).

Using the triangular representation (or equivalently the contour representation), it can be shown that analogous to the equilibrium case, it is possible to formulate a well defined perturbation expansion of the Green’s function and arrive at the Dyson equation [73]

\[
\hat{G} = \hat{g} + \hat{g} \circ \hat{\Sigma} \circ \hat{G},
\]

(3.29)

where \( \circ \) denotes summation/integration over common internal variables and \( \hat{g} = \begin{pmatrix} 0 & g^a_r \\ g^r K & g^K \end{pmatrix} \) is the Green’s function in absence of the perturbation which gives rise to the self energy term. It is important to note that in the triangular representation, \( \hat{\Sigma} \) is written

\[
\hat{\Sigma} = \begin{pmatrix} \Sigma^K & \Sigma^r \\ \Sigma^a & 0 \end{pmatrix}.
\]

(3.30)

Eq. (3.29) can be inverted to obtain

\[
(\hat{\varepsilon} - \hat{H}_0 - \hat{\Sigma}) \circ \hat{G} = \hat{\delta},
\]

(3.31)

where \( \hat{\delta} \) is a delta function in the remaining variables. The real-space full Dyson equation, Eq. (3.29), can also be Fourier transformed and assuming translational invariance of the system, we arrive at an expression for the full Green’s function in reciprocal space:

\[
\tilde{G}(p, \varepsilon) = \left( \tilde{G}_0(p, \varepsilon)^{-1} - \tilde{\Sigma}(p, \varepsilon) \right)^{-1} = (\hat{\varepsilon} - \hat{H}_0 - \hat{\Sigma}(p, \varepsilon))^{-1}.
\]

(3.32)

Writing out the equations in Eq. (3.29) explicitly yields

\[
G^{r/a} = g^{r/a} + g^{r/a} \circ \Sigma^{r/a} \circ G^{r/a},
\]

(3.33)

\[
G^K = g^K + g^K \circ \Sigma^K \circ G^K + g^r \circ \Sigma^K \circ G^a,
\]

(3.34)

where it can be seen that the non-equilibrium Keldysh Green’s function depends explicitly on the retarded/advanced Green’s function. The equation for the retarded/advanced component is identical to the equilibrium equation and its solution is needed in order to determine the Keldysh Green’s function which contains the non-equilibrium information. In Keldysh formalism, expectation values of observables are computed as

\[
\langle \hat{O}(xt) \rangle = \lim_{x't' \rightarrow xt^+} \frac{i}{2} \text{Tr} \left[ \hat{O}(xt) G^K(xt, x't') \right].
\]

(3.35)

This makes it clear that the Keldysh Green’s function provides a convenient way to obtain knowledge of a system since \( \hat{O}(xt) \) can be any observable. Examples of relevance in this thesis are the current and the density operators. Eq. (3.35) is valid for systems in equilibrium as well as for non-equilibrium systems.
3.2 Linear response in Keldysh formalism

Linear response theory is an important technique with which we may investigate non-equilibrium properties of quantum mechanical systems by considering them to be infinitesimally close to equilibrium. The main simplification arises from the fact that non-equilibrium properties in linear response actually only depend on the equilibrium or unperturbed quantities. For example, we will see that the electrical conductivity depends on the unperturbed Green’s function, albeit in a non-trivial way.

We will treat linear response within the Keldysh formalism, which is strictly speaking not necessary as was discussed in section 3.1.3. Many linear response treatments, see e.g., [75, 76, 77], are performed with equilibrium Green’s functions exactly due to the fact that the linear response depends on the unperturbed Green’s function which is taken to be in equilibrium. The general formalism presented here is valid also for systems where the unperturbed state is taken to be a non-equilibrium state. The explicit results we present in the end are, however, valid for systems in equilibrium only. We will make it clear during the derivation when we depart from the general case. We will see that the algebra involved in using Keldysh Green’s functions for the linear response is not very complicated due to the 2-by-2 matrix structure.

The starting point for a linear response calculation is to consider a total Hamiltonian on the form

\[ \hat{H} = \hat{H}_0 + \delta \hat{H}, \quad (3.36) \]

where \( \hat{H}_0 \) has a known (Keldysh) Green’s function \( \hat{G}_0 \) and \( \delta \hat{H} \) is a small perturbation written in Keldysh space. We write an ansatz for the Green’s function of the total Hamiltonian as \( \hat{G} = \hat{G}_0 + \delta \hat{G} \), i.e., the full Green’s function is only perturbed slightly from the unperturbed one. Inserting this ansatz into the Dyson equation, Eq. (3.31), we obtain

\[ (\tilde{\epsilon} - \hat{H}_0 - \delta \hat{H}) \circ (\hat{G}_0 + \delta \hat{G}) = \delta \quad (3.37) \]

and by using that by definition \( (\tilde{\epsilon} - \hat{H}_0) \circ \hat{G}_0 = \delta \), we obtain

\[ (\tilde{\epsilon} - \hat{H}_0 - \delta \hat{H}) \circ \delta \hat{G} = \delta \hat{H} \circ \hat{G}_0. \quad (3.38) \]

We now observe that the term \( \delta \hat{H} \circ \delta \hat{G} \) is second order in the perturbation and we drop it since we seek the linear response. For sufficiently small perturbations we expect this approximation to be good, indeed most materials exhibit a linear response to electromagnetic fields [78]. Dropping the second order term allows us to rewrite Eq. (3.38) into an equation...
determining the perturbation to the Green’s function as
\[
\delta \tilde{G} = \tilde{G}_0 \circ \delta \tilde{H} \circ \tilde{G}_0.
\] (3.39)

As alluded to above, we see that the perturbation to the Green’s function, \(\delta \tilde{G}\), indeed depends only on the unperturbed Green’s function together with the perturbation itself. We may now insert this into the formula for the expectation value, Eq. (3.35), and we obtain
\[
\langle \hat{O}(xt) \rangle = \lim_{x't' \to xt^+} \frac{i}{2} \text{Tr} \left[ \hat{O}(xt) \left( G^K_0( xt, x't') + \delta G^K( xt, x't') \right) \right]
\] (3.40)
and from this we define the unperturbed expectation value
\[
\langle \hat{O}(xt) \rangle_0 = \lim_{x't' \to xt^+} \frac{i}{2} \text{Tr} \left[ \hat{O}(xt)G^K_0( xt, x't') \right].
\] (3.41)
We may then write the deviation of the expectation value from its unperturbed value as
\[
\delta \hat{O}(xt) = \langle \hat{O}(xt) \rangle - \langle \hat{O}(xt) \rangle_0 = \lim_{x't' \to xt^+} \frac{i}{2} \text{Tr} \left[ \hat{O}(xt)\delta G^K( xt, x't') \right].
\] (3.42)

Eqs. (3.39) and (3.42) express the linear response of any observable due to a perturbation in terms of Keldysh Green’s functions. To proceed, we must specify the perturbation \(\delta \tilde{H}\) and also the observable \(\hat{O}\) we are interested in calculating. We must of course also specify the unperturbed Hamiltonian but in our case this is naturally the graphene Hamiltonian and the unperturbed Green’s function is given by Eq. (3.9).

We will derive an expression for the polarizability (charge-charge correlation function) which is defined as the charge density response to a scalar potential response \([75]\). This determines all quantities needed to apply the expression for the expectation value in Eq. (3.42). The observable we consider is the density that is written in A-B sublattice space as \(\hat{\rho} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\) and the perturbation to the Hamiltonian written on the Keldysh path is given by \(\delta \tilde{H} = \begin{pmatrix} \delta V & 0 \\ 0 & -\delta V \end{pmatrix}\). We remember that we used a rotation given by Eq. (3.27), and by transforming the perturbation using \(X \rightarrow UXU^{-1}\) we find that the perturbation in the triangular representation is written like
\[
\delta \tilde{H} = \begin{pmatrix} 0 & \delta V \\ \delta V & 0 \end{pmatrix}.
\] (3.43)
The perturbation has the form
\[
\delta V = \begin{pmatrix} \delta V_0 & 0 \\ 0 & \delta V_0 \end{pmatrix} e^{i(q \cdot x - \omega t)}.
\] (3.44)
where $\delta V_0$ is the (scalar) amplitude of the perturbation and the matrix structure is in A-B sublattice space (so this matrix should not be rotated in Keldysh space). It should be noted that by using the \( \circ \) operator we implicitly mean all real space functions to be functions of two real space arguments. The local perturbation in Eq. (3.44) has only one spatial argument, but we can easily remedy this by adding a $\delta(x - y)$ which makes the local nature of the perturbation explicit.

The perturbation given by Eqs. (3.43) and (3.44) is inserted into Eq. (3.39) to find the explicit expression for the Green’s function perturbation:

$$
\delta \hat{G} = \left( \begin{array}{cc} 0 & G^a_0 \\ G^r_0 & G^K_0 \end{array} \right) \circ \left[ \begin{array}{c} \delta V \\ 0 \end{array} \right] \circ \left( \begin{array}{cc} 0 & G^a_0 \\ G^r_0 & G^K_0 \end{array} \right)
$$

$$
= \left( \begin{array}{cc} 0 & G^K_0 \circ \delta V \circ G^a_0 \\ G^r_0 \circ \delta V \circ G^r_0 & G^K_0 \circ \delta V \circ G^a_0 + G^K_0 \circ \delta V \circ G^r_0 \end{array} \right).
$$

We conclude that the Keldysh part of the perturbation to the Green’s function is

$$
\delta G^K = G^K_0 \circ \delta V \circ G^a_0 + G^K_0 \circ \delta V \circ G^r_0.
$$

We are interested in this Green’s function since this allows us to calculate the average of the observable using Eq. (3.42). We must remember that \( \circ \) means summation over all internal variables, meaning that there are implicit integrations in the above expression for the Green’s function perturbation. To see this, we explicitly write out the first term with all its arguments and integrations explicitly

$$
G^K_0 \circ \delta V \circ G^a_0 = \int dy dy' G^K_0(x, y) \delta V(y) \delta(y - y') G^a_0(y', x')
$$

$$
= \int dy \ G^K_0(x, y) \delta V(y) G^a_0(y, x'),
$$

where the coordinates represent both spatial coordinate and time, and the delta function dependence comes from the local nature of the perturbation.

Now, we are interested in the Fourier transform of the expectation value and we take the Fourier transform of Eq. (3.42). By assuming that the Green’s functions exhibit translational invariance, we can arrive at an expression for the Fourier transform of the expectation value of the density perturbation as

$$
\delta n(q, \omega) = \frac{ig_s g_v}{2} \int \frac{d^2 p}{(2\pi)^2} \frac{d\varepsilon}{2\pi} \text{Tr} \left[ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \left( G^K_0(p, \varepsilon) G^a_0(p, \varepsilon) + G^r_0(p, \varepsilon) G^K_0(p, \varepsilon) \right) \delta V_0 \right],
$$

$$
+ G^r_0(p, \varepsilon) G^K_0(p, \varepsilon) \delta V_0 \right),
$$

(3.49)
where we have introduced the degeneracy factor $g_s g_v$, the unit matrix is the density operator in A-B sublattice space, $p_- = p - q$, $\varepsilon_- = \varepsilon - \omega$, and $\delta n$ denotes the expectation value of the deviation from the equilibrium density. From Eq. (3.49), we can extract a very general expression for the polarizability

$$\Pi(q, \omega) = \frac{ig_s g_v}{2} \int \frac{d^2 p}{(2\pi)^2} \frac{d\varepsilon}{2\pi} \left[ \Tr \left( G^K_0 (p, \varepsilon) G^a_0 (p_-, \varepsilon_-) + G^r_0 (p, \varepsilon) G^K_0 (p_-, \varepsilon_-) \right) \right],$$

(3.50)

which follows from the definition $\delta n(q, \omega) = \Pi(q, \omega) \delta V_0 [75]$.

To proceed, we now make explicit that the unperturbed Green’s function, $G_0$, is in equilibrium and we make use of the rewriting of the Green’s function defined in Eqs. (3.10)-(3.12). This can be inserted into Eq. (3.50), and performing the matrix multiplications we obtain

$$\Pi(q, \omega) = ig_s g_v \int \frac{d^2 p}{(2\pi)^2} \frac{d\varepsilon}{2\pi} \left( h^r h^K_+ + h^K h^a_- + \cos(\phi_p - \phi_p_-) (f^r f^K_+ + f^K f^a_-) \right),$$

(3.51)

where we used the notation that subscript “−” means that the function is evaluated at the argument $(p_-, \varepsilon_-)$ and no subscript means the function is evaluated at $(p, \varepsilon)$. Eq. (3.51) is valid for graphene Green’s functions with arbitrary self energies.

We now take the clean limit, i.e., we use the self energy $\Sigma^r(\varepsilon) = -i\eta^+$, $(\Sigma^a(\varepsilon) = i\eta^+)$ and we find that

$$\int d\varepsilon \left( \tanh(\varepsilon \pm \beta/2) h^r h^a_\pm - \tanh(\varepsilon \beta/2) h^a h^a_\pm \right) = 0,$$

(3.52)

$$\int d\varepsilon \left( \tanh(\varepsilon \pm \beta/2) f^r f^a_\pm - \tanh(\varepsilon \beta/2) f^a f^a_\pm \right) = 0,$$

(3.53)

which allows us to obtain

$$\Pi(q, \omega) = ig_s g_v \int \frac{d^2 p}{(2\pi)^2} \frac{d\varepsilon}{2\pi} \left[ (h^r h^a_+ + \cos(\phi_p - \phi_p_-) f^r f^a_+) \right.$$

$$\times \left( \tanh(\varepsilon \beta/2) - \tanh(\varepsilon \beta/2) \right) \right],$$

(3.54)

where we have used Eq. (3.28) to eliminate the Keldysh component. Performing the energy integral by the method of residues, we obtain contributions from the poles of $h$ and $f$, and the contributions from the poles of
the tanh terms cancel. By rewriting tanh in terms of the Fermi function and inserting the definitions of $h$ and $f$, we obtain the final expression

$$\Pi(q, \omega) = g_s g_v \lim_{\eta^+ \to 0^+} \int \frac{d^2 p}{(2\pi)^2} \sum_{\lambda, \lambda'} \frac{n_F(\lambda' E_{p-}) - n_F(\lambda E_p)}{\omega + \lambda' E_{p-} - \lambda E_p + i\eta^+} F_{\lambda\lambda'}(p),$$

(3.55)

$$F_{\lambda\lambda'}(p) = \frac{1 + \lambda \lambda' \cos(\phi_p - \phi_{p-})}{2}.$$  

(3.56)

Eq. (3.55) is a well-known expression for the clean-case polarizability and serves as the starting point for many linear response calculations. The graphene polarizability at zero temperature was first calculated by Wunsch et al [21] and by Hwang and Das Sarma [20]. This was later generalized to finite temperatures by Ramezanali et al [27]. In all these calculations, Eq. (3.55) was taken as a starting point and the calculations are valid in the clean limit.

It might seem that in our derivation here we took a complicated detour to re-derive a known expression, but in fact we accomplished a lot more. For the clean limit, it is indeed the case that starting from Eq. (3.55) would suffice, but during our derivation we arrived at general expressions for the polarizability that have a wider range of applicability. First of all, Eq. (3.50) is valid for non-equilibrium systems, and by taking this as the starting point we can compute the polarizability for any system with a known Hamiltonian. Furthermore, Eq. (3.51) is valid for graphene with impurities that can be captured by a Green’s function self energy. Thus, by taking this equation as our starting point we can calculate the linear response coefficient for graphene under rather general conditions. We only need to be able to compute an expression for the self energy of the problem. We leave such considerations for section 3.5 and in the following two sections we deal with the expression arrived at in the clean limit, i.e., Eq. (3.55).

### 3.3 Graphene conductivity in the clean limit

The longitudinal conductivity of graphene can be expressed in terms of the polarizability (charge-charge correlation function) [79] that was obtained in the previous section, or the current susceptibility (current-current correlation function) [75, 79]. Due to gauge invariance, these two are equivalent and the longitudinal conductivity is uniquely defined [77], and one may use the expression one finds most convenient. It should be noted that the transverse response is more general as it can handle both longitudinal
and transverse perturbations. Since we are interested in the longitudinal response, we use the polarizability to obtain the conductivity. This circumvents the complication of the anomalous commutator term needed to properly handle the current-current response in graphene [80, 81].

By definition, the polarizability \( \Pi(q, \omega) \) is related to the density response as [75]

\[
n(q, \omega) = \Pi(q, \omega)V(q, \omega),
\]

(3.57)

where \( V(q, \omega) \) is the perturbing potential. Now, the potential is related to the electric field by \( E = -\nabla V \), which in reciprocal space becomes \( E(q, \omega) = -iqV(q, \omega) \). Substituting this into Eq. (3.57), we obtain

\[
n(q, \omega) = \frac{i}{q}\Pi(q, \omega)E(q, \omega).
\]

(3.58)

The continuity equation in reciprocal space

\[
\omega n(q, \omega) = q j(q, \omega),
\]

(3.59)

can now be used to write Eq. (3.58) in terms of the current response as

\[
j(q, \omega) = \frac{i\omega}{q^2}\Pi(q, \omega)E(q, \omega).
\]

(3.60)

In Eq. (3.59), we have assumed that \( \vec{q} \parallel \vec{j} \), which is only valid in the longitudinal case. Thus, Eq. (3.60) is only valid for longitudinal perturbations. We recognize that the quantity relating the electric field with the current is nothing but the conductivity. From Eq. (3.60), we identify the longitudinal conductivity to be

\[
\sigma(q, \omega) = \frac{i\omega}{q^2}\Pi(q, \omega),
\]

(3.61)

which is valid for a general \( \Pi(q, \omega) \).

Now, the clean graphene conductivity is obtained by taking Eq. (3.61) with \( \Pi(q, \omega) \) given by Eq. (3.55). The full answer for the nonlocal conductivity is rather lengthy and we refrain from writing it down in full detail and refer to Refs. [20, 21, 27] for the full expressions. For zero temperature, the result can be written down in a closed form involving only known functions, while in the finite temperature result there remains an integral to be performed numerically. The local conductivity \( (q \to 0) \) at finite
temperature is \([82, 83]\)

\[
\sigma(\omega) = \sigma_1(\omega) + i\sigma_2(\omega),
\]

\[
\sigma_1(\omega) = \frac{e^2}{4} G(\omega/2),
\]

\[
\sigma_2(\omega) = \frac{e^2}{4} \left( \frac{8T}{\omega\pi} \ln \left[ 2 \cosh \left( \frac{\mu}{2T} \right) \right] 
+ \frac{4\omega}{\pi} \int_0^\infty \frac{G(x) - G(\omega/2)}{\omega^2 - 4x^2} dx \right),
\]

\[
G(x) = \frac{\sinh(x/T)}{\cosh(\mu/T) + \cosh(x/T)}, \quad (3.62)
\]

where \(\mu\) is the chemical potential and \(T\) is the temperature.

Figures showing the conductivity are provided in the next section, where we also compare the clean conductivity considered in the present section with the conductivity obtained when adding a relaxation time.

### 3.4 Mermin relaxation time approximation

The graphene conductivity obtained by using the polarizability in Eq. (3.55) was arrived at by assuming pristine graphene and includes thermal and nonlocal effects. One of the main drawbacks of the pristine conductivity is that it fails to reproduce a finite real DC conductivity and the well known real Drude peak for low frequencies. This peak at low frequencies is well known from metals and is experimentally found to exist also for graphene [84]. This means that graphene contains phonons or impurities of some kind, which are so far not included in our model for graphene electrons. The impurities could be carbon-atom dislocations or substitutions, grain boundaries in the graphene sheet, or particles that are stuck on the graphene surface. Quite possibly a combination of these effects is what creates the electron scattering which gives rise to the finite DC conductivity, see section 2.3 for a discussion.

One way to model electron relaxation is provided by including a phenomenological electron relaxation time. One such model, due to Mermin [85], dates back to 1970 and is widely used in graphene plasmonics to include plasmon losses, see e.g. Refs. [26, 68, 86, 87]. The Mermin relaxation time approximation uses the pristine response function which is calculated in the absence of the electron relaxation and provides a prescription for how to insert the relaxation time without changing the number of particles in the system. This would happen by using the naive substitution
The key realization in Mermin’s method is that the electrons, which are subject to the external electric field as well as relaxation by some scattering, are relaxing not to the global Fermi distribution but to a local equilibrium. This local equilibrium can have a spatial dependence as an effect of a spatially varying perturbation. For a detailed treatment of the Mermin relaxation time approximation, we refer to Refs. [85, 88]. In what follows we only give the final result and consider its implications.

The result from Mermin’s method is a polarizability of the form [85, 88]

\[
\Pi(q, \omega) = \frac{(1 + i \Gamma \frac{\omega}{\omega}) \Pi^0(q, \omega + i \Gamma)}{1 + i \Gamma \frac{\omega}{\omega} \Pi^0(q, \omega + i \Gamma)}, \tag{3.63}
\]

where \(\Pi^0\) refers to the clean-limit polarizability in Eq. (3.55). The Mermin polarizability can then be used in Eq. (3.61) to obtain the conductivity in the presence of the electron relaxation. It should be emphasized that the main benefit with Mermin’s method is its simplicity; it uses only the clean-limit result evaluated at complex energies in order to find the result which includes relaxation. However, the Mermin relaxation time approximation is phenomenological and the relaxation time is added by hand without any microscopic mechanism for the relaxation.

Fig. 3.2 shows the local conductivity, \(q \rightarrow 0\), for various values of the inverse relaxation time \(\Gamma\). The local conductivity for \(\Gamma = 0\), i.e., the clean case, is given by the expression in Eq. (3.62). It can be seen that increasing \(\Gamma\) substantially changes the conductivity and we note especially the appearance of the Drude peak for low frequencies.
Fig. 3.3 shows the nonlocal conductivity for $q/k_F = 0.3$ and otherwise identical parameters as the local conductivity in Fig. 3.2. A comparison of Figs. 3.2 and 3.3 is provided in Fig. 3.4 and shows that the conductivity changes rather substantially as a function of $q$. For this reason, it could be important to consider the nonlocality when dealing with perturbations of sufficiently small wavelengths. We will return to this in chapter 5.

3.5 Green’s function self energy

In the previous section we discussed an approach to include electron relaxation in a phenomenological way. In this section, we discuss a microscopic approach to add losses to the graphene electrons by including a scattering term in the Hamiltonian. Such a scattering term can be modeled using t-matrix formalism, see e.g., Refs. [73, 75, 89]. The t-matrix formalism is a method that treats impurity scattering with a random distribution of impurity sites. An average on impurity positions can be taken and this allows for evaluation of a self-energy term resulting from the impurities. We treat the case of a dilute density of impurity sites, dilute in the sense that the fraction of lattice sites containing impurities $n_{\text{imp}} \ll 1$. This simplifies the problem of finding the self energy since interference terms between different impurity sites can be neglected.

Using the t-matrix approach, we may set up a perturbation theory for the scattering term addition to the bare graphene Hamiltonian along the lines of Eq. (3.32). The scattering term contributes to the self energy that is to be included in the pristine graphene Green’s function in Eq. (3.9). Calculating the conductivity from the Green’s function with a self energy can be done using Eq. (3.51) and performing the integrals. Some care
must be taken when evaluating the poles of this Green’s function and we refer to Papers III and IV for a detailed treatment.

Using the t-matrix approach, the retarded component of the momentum-space Dyson equation, Eq. (3.32), in the case of graphene with pointlike S-wave scatterers, becomes

$$G^r(p,\varepsilon) = \frac{1}{2} \sum_{\lambda=\pm} \frac{1}{\varepsilon - \lambda E_p - \Sigma^r(\varepsilon)} \left( \frac{1}{\lambda e^{i\phi_p}} \lambda e^{-i\phi_p} \right).$$  \hspace{1cm} (3.64)

This is nothing but the bare graphene Green’s function together with a correction given by the self energy. The self energy — which in general depends also on $p$ — depends only on $\varepsilon$ due to the pointlike and S-wave nature of the scatterers. The addition of the scattering term to the bare graphene Hamiltonian is included in the self energy, and the task is to evaluate or calculate the self energy for any particular model of electron scattering. Below we consider two such models — one model for interaction with a distribution of adatoms coupled to the graphene lattice by tunneling, and one self-consistent model for evaluating the Green’s function together with the self energy for a dilute density of randomly distributed impurity sites.

In principle, the presence of a self energy can alter the linear response formulation that we considered in section 3.2. The perturbation applied to the system, can alter not only the Green’s function as we considered, but with a self energy there can also be an induced perturbation to lin-
ear order of the self energy. This is called vertex corrections and this is typically a small correction since it is proportional to the perturbation as well as the self energy itself. The self energy is already a correction to the bare Hamiltonian and is typically small, making the vertex correction suppressed. For the cases we consider below, i.e., within a t-matrix formulation of pointlike S-wave scatterers, it turns out that the vertex correction are identically zero.

3.5.1 Adatoms coupled to graphene

Atoms in close proximity to the graphene lattice — we will refer to them as adatoms — can cause the graphene electrons to tunnel into the adatoms, provided there is an empty electron state for the graphene electron to tunnel into. The electron can then tunnel back into the lattice and, in the limit of a dilute density, there are no direct tunneling events between the adatoms themselves. The Hamiltonian we consider for this reads

\[ \hat{H} = \sum_l \epsilon_0 \hat{d}^\dagger_l \hat{d}_l + \sum_{p,\lambda} \left( \lambda E_p \hat{c}^\dagger_{p\lambda} \hat{c}_{p\lambda} + \sum_l (t_{l,p,\lambda} \hat{d}^\dagger_l \hat{c}_{p\lambda} + \text{h.c.}) \right) \],

(3.65)

where \( \hat{d}_l \) (\( \hat{d}^\dagger_l \)) annihilates (creates) an electron on the adatom at position \( x_l \) and \( \hat{c}_{p,\lambda} \) (\( \hat{c}^\dagger_{p,\lambda} \)) annihilates (creates) an electron in the graphene lattice with momentum \( p \) in band \( \lambda \). Furthermore, \( p \) is the momentum, \( t_{l,p,\lambda} = \lambda \bar{t} e^{ip \cdot x_l} \), \( \bar{t} \) is the bare adatom-carbon atom coupling, \( \epsilon_0 \) is the energy level of the adatom, \( \lambda = \pm \) is the graphene band index and \( x_l \) is the position of the \( l \)-th impurity.

In the dilute limit with no interference effect, the individual adatoms can then be treated as an average density and the impurity-averaged self energy can be shown to take the form

\[ \Sigma^r(\epsilon) = \frac{|t|^2}{\epsilon + i\delta - \epsilon_0}, \]

(3.66)

\[ |t|^2 = n_{\text{imp}} |\bar{t}|^2, \]

(3.67)

where \( n_{\text{imp}} \) is the fraction of impurities per lattice site, and \( \delta \) is the bare inverse lifetime of the adatom electron states. The values of \( \epsilon_0, \delta, \) and \( \bar{t} \) are phenomenological parameters and external input from either first-principle calculations or experiments are needed to fix them. The adatom density \( n_{\text{imp}} \) is a free parameter that may vary between different graphene samples and this can be seen to affect the effective tunneling amplitude \( t \).

The appearance of the self energy term changes the poles of the Green’s function, which is apparent from Eq. (3.64). The self energy in Eq. (3.66)
inserted in the Green’s function, Eq. (3.64), yields the poles
\[ \varepsilon^\pm = \frac{\lambda E_p + \varepsilon_0 - i \delta \pm \sqrt{(\lambda E_p - \varepsilon_0 + i \delta)^2 + 4|t|^2}}{2}. \] (3.68)

Remember that we used the pole structure of the Green’s functions to arrive at the expression in Eq. (3.55) for \( \Pi(q, \omega) \), which gave us the conductivity. Thus, with the changed pole structure in the case of the adatom self energy, this equation is no longer valid and the starting point needs to be Eq. (3.51). We leave out this calculation here and refer to Paper III for the details.

### 3.5.2 Dilute distribution of impurity sites — a self-consistent treatment

To treat impurity scattering, we consider a distribution of random scattering sites with identical strength \( V_{\text{imp}} \). We consider a low density \( n_{\text{imp}} \) of such sites, meaning that only a small fraction of lattice sites contain a scatterer. Furthermore, we consider only s-wave scattering, i.e., the scattering is non-directional. These approximations are standard in t-matrix formulations of impurity scattering [73, 89].

The Hamiltonian for the scattering can be written as
\[ \hat{H} = \hat{H}_0 + \sum_m \hat{V}_m, \] (3.69)

where \( \hat{V}_m \) are the scattering potentials at some (random) sites. Using an averaging procedure over the random sites, it is possible to arrive at equations for the Green’s function (Dyson’s equation) as well as an equation for the self energy. The self energy depends on the impurity scattering as well as the full Green’s function itself, thus giving two self-consistent equations to solve. In the calculation, we neglect interference terms between the different impurity sites with the argument that they are far apart and should not influence each other. Thus the assumption of a dilute density of scatterers is important. We also note that the impurities are treated as fixed and “rigid” objects that do not absorb any energy or momentum in the scattering process.

Following the steps in the preceding discussion gives the Dyson equation together with an equation for the impurity averaged self energy:
\[ G^r(p, \varepsilon) = \frac{1}{2} \sum_{\lambda=\pm} \frac{1}{\varepsilon - \lambda E_p - \Sigma^r(\varepsilon)} \left( \begin{array}{c} \lambda e^{i \phi_p} \\ 1 \end{array} \right), \] (3.70)
\[ \Sigma^r(\varepsilon) = \frac{n_{\text{imp}} V_{\text{imp}}}{1 - V_{\text{imp}} \sum_p G(p, \varepsilon)}, \] (3.71)
and these need to be solved self-consistently. The polarizability, and hence the conductivity, can again be calculated using Eq. (3.51). The Green’s function obtained by numerically solving Eqs. (3.70) and (3.71) has poles that must also be found numerically. To this end, the numerical self energy obtained from the solution can be fitted to the form of the analytic self energy in Eq. (3.66), and the approximate poles can be found analytically. With these poles as initial guesses, a numerical root finder routine can be employed to find the exact poles of the numerical Green’s function. With the knowledge of the poles, it is possible to evaluate the conductivity, using Eq. (3.51), and proceed to investigate the plasmon properties. For details, we refer to Paper IV.
We discuss the Maxwell equations and some of their most important properties and related quantities. We pay special attention to the coupling between the electromagnetic field and matter, as described by the dielectric function and the conductivity. These quantities will play important roles later in the thesis when describing the plasmonic properties of materials. The work presented in this thesis is focused around the ability of graphene plasmons to affect electromagnetic radiation scattered from interfaces supporting plasmons. Before studying this directly, we here investigate electromagnetic propagation and scattering from a single sheet of graphene, and also scattering in more complex dielectric environments. Such environments are necessary for coupling electromagnetic radiation to the plasmons, but they also provide added difficulties in solving the scattering problem. At the end of the chapter, we treat in some detail two different methods for solving the scattering problem in such environments: an analytical model and a numerical model. Both methods have been used throughout the work leading to this thesis.

4.1 The Maxwell equations and light propagation

The basic equations that govern the interaction between electromagnetism and matter are the Maxwell equations. These equations describe classical electromagnetic fields of any frequency and wavelength under most imaginable conditions. Thus, they govern the behavior of visible light, radio waves, ultraviolet (UV) radiation, as well as infrared light and X rays. It is clear from this list that the study of Maxwell’s equations encompasses many different fields of technological as well as scientific relevance.
The macroscopic Maxwell equations [78]

\[ \nabla \cdot \vec{D} = \rho, \quad (4.1) \]
\[ \nabla \times \vec{H} - \frac{\partial \vec{D}}{\partial t} = \vec{j}, \quad (4.2) \]
\[ \nabla \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0, \quad (4.3) \]
\[ \nabla \cdot \vec{B} = 0, \quad (4.4) \]

contain four fields (\( \vec{E}, \vec{D}, \vec{H}, \vec{B} \)). These are called the electric field, displacement field, magnetic field strength and magnetic induction field respectively. In vacuum, we have the well known relationships \( \vec{D} = \epsilon_0 \vec{E} \) and \( \vec{B} = \mu_0 \vec{H} \), where \( \epsilon_0 \) and \( \mu_0 \) are the vacuum values of the permittivity and permeability, respectively, with \( \sqrt{\frac{1}{\epsilon_0 \mu_0}} = c \). In addition, \( \rho \) is the charge density and \( \vec{j} \) is the current density.

In a more general situation, there might be a medium present that behaves differently from vacuum. With “medium” we mean a material with charges that respond to external electric fields. Basically any material fits this description, but here we consider solid materials such as metals or semiconductors. The charges in the medium respond when an external field is applied and this can alter the relationships between the electromagnetic fields. In general [78],

\[ \vec{D} = \epsilon_0 \vec{E} + \vec{P}, \quad (4.5) \]
\[ \vec{B} = \mu_0 \vec{H} + \vec{M}, \quad (4.6) \]

where \( \vec{P} \) and \( \vec{M} \) are the polarization field and magnetization field respectively. This means that we add the polarization field to the vacuum value of the displacement field and likewise the magnetization to the vacuum value of the magnetic induction. The naming convention for the \( \vec{B} \) and \( \vec{H} \) fields changes between different sources in the literature, we will refer to both fields as the magnetic field.

In what follows, we will mainly be interested in the \( \vec{D} \) and \( \vec{E} \) fields since they are governing the ability of the medium to respond to electric fields. Furthermore, many materials exhibit a very weak magnetization and from now on we assume that the magnetization \( \vec{M} \) is zero and thus \( \vec{B} = \mu_0 \vec{H} \), i.e., the same as for vacuum. We will also make a very important simplification in the treatment of electric fields surrounding materials. We will only consider the linear response of the material, meaning that the response, i.e., \( \vec{P} \), is proportional to the external field \( \vec{E} \). There are different
ways to justify this, but the simplest justification is that this behavior is what most systems exhibit for accessible field strengths. At very high electric fields this assumption may break down, but we will not consider such cases. Consequently, we may now write the polarization as \( \vec{P} = \epsilon_0 \chi \vec{E} \), i.e.,

\[
\vec{D} = \epsilon_0 \vec{E} + \vec{P} = \epsilon_0 \vec{E}(1 + \chi),
\]

(4.7)

where \( \chi \) is the (electric) susceptibility of the medium. We may now define a new electric permittivity (compare with the vacuum electric permittivity) as \( \tilde{\epsilon} = \epsilon_0 (1 + \chi) \) and the dielectric function (relative permittivity) as

\[
\epsilon = \frac{\tilde{\epsilon}}{\epsilon_0} = 1 + \chi,
\]

(4.8)

giving us

\[
\vec{D} = \tilde{\epsilon} \vec{E}.
\]

(4.9)

The vacuum permittivity has been altered by the susceptible medium, thus giving rise to a different displacement field. From Eqs. (4.8) and (4.9), it is clear that a susceptibility of zero reproduces the vacuum value for the \( \vec{D} \) field.

4.1.1 Nonlocal electromagnetic properties

It is important to point out that the displacement field and the electric current can be nonlocal in space and time [1, 78]:

\[
\vec{D}(\vec{x}, t) = \int d\vec{x}' dt' \epsilon(\vec{x} - \vec{x}', t - t') \vec{E}(\vec{x}', t'),
\]

(4.10)

\[
\vec{j}(\vec{x}, t) = \int d\vec{x}' dt' \sigma(\vec{x} - \vec{x}', t - t') \vec{E}(\vec{x}', t').
\]

(4.11)

By considering the system to have been under continuous illumination for a very long time, the electric field, the dielectric function (\( \epsilon \)), and the conductivity (\( \sigma \)) can be written as Fourier transforms. By also applying a Fourier transform on the displacement field and the current we obtain

\[
\vec{D}(\vec{k}, \omega) = \epsilon(\vec{k}, \omega) \vec{E}(\vec{k}, \omega),
\]

(4.12)

\[
\vec{j}(\vec{k}, \omega) = \sigma(\vec{k}, \omega) \vec{E}(\vec{k}, \omega),
\]

(4.13)

since the Fourier transform has the property that it turns a convolution into a product. This shows that the equations linking different quantities are simpler in reciprocal space and that the response (i.e., the current and the displacement field) to an electric field with frequency \( \omega \) and modulation \( \vec{k} \), follows the same frequency and modulation. The quantities \( \epsilon \) and \( \sigma \) are
material parameters that are either calculated or measured and can then be included in the electromagnetic modeling. In general, \( \sigma \) and \( \epsilon \) can be second-order tensors, the fact that we only consider scalars reflects that we consider an isotropic medium.

From the Maxwell equations, it is possible to relate the bulk conductivity and permittivity as [1]

\[
\epsilon(\vec{k}, \omega) = 1 + \frac{i\sigma(\vec{k}, \omega)}{\epsilon_0 \omega}.
\] (4.14)

To properly model these material properties it is thus important to consider the frequency and momentum behavior. Modeling the \( \omega \)-dependence poses no significant problem as the values of material parameters can be measured rather easily and simple models for the frequency behavior of material parameters exist [78]. Such models are usually based on Drude-type behavior for the conductivity [90], and an oscillator-type model for the dielectric function [78]. The importance of the \( \vec{k} \) dependence of the dielectric constant and the conductivity is determined by the magnitude \( |\vec{k}| \) of the incident field, compared with the inverse of a typical length of the system. In metals such as gold and silver, both the lattice constant and the Fermi wavelength are on the order of a few Å [91], so for wavelengths much larger than this there is usually no need to include nonlocal effects in calculations. The exception is if the patterning of the material is on the order of a few nanometers, making it comparable to the Fermi wavelength. Nonlocal effects may then be of importance since the system is probing length scales where the quantum effects of electrons become noticeable.

Progress in nanofabrication has enabled the investigation of small metal nanoparticles — with dimensions close to the Fermi wavelength — and their electromagnetic response. It has been shown experimentally that such nanoparticles exhibit features not readily explained by a local response theory [92, 93, 94, 95]. This has created activity in modeling these nonlocal properties stretching back several decades in time. Early works focused on describing these size dependent effects by introducing a phenomenological size-dependent damping term [96, 97, 98]. More recently, semiclassical hydrodynamic models have been proposed, which include pressure effects [99] as well as charge diffusion effects [100]. In addition, electron tunneling between metallic structures in very close proximity can affect the optical response [101].

Effects of nonlocality have also been demonstrated for graphene structures [102]. To model graphene, we are aided by the fact that a rather
simple model of electrons can explain the observed low-energy band structure and this allows us to compute analytically the nonlocal conductivity from a microscopic picture, see chapter 3. This enables us to investigate the nonlocal properties of graphene and in particular the implications for graphene plasmons in rather straightforward terms.

4.1.2 Electromagnetic energy and propagation

An important concept in Maxwell’s equations is the energy carried by the electromagnetic field. The nature of its propagation is important for understanding and interpreting electromagnetic scattering from various objects. In this thesis we treat the electromagnetic field as a classical field and one question that arises is; how does the electromagnetic field propagate? The well-known answer is that it propagates as waves, which can be seen by forming a wave equation by combining Eqs. (4.2) and (4.3). It is then possible to arrive at wave equations for the electric and the magnetic fields.

For linear and non-dispersive media, the energy density contained in the electromagnetic field is [78]

\[ u = \frac{1}{2} \left( \vec{E} \cdot \vec{D} + \vec{H} \cdot \vec{B} \right) \]  

(4.15)

and can be written as

\[ u = \frac{1}{2} \left( \epsilon \vec{E}^2 + \mu \vec{H}^2 \right). \]  

(4.16)

The rate of change of the energy density is then

\[ \frac{\partial u}{\partial t} = \vec{E} \cdot \frac{\partial \vec{D}}{\partial t} + \vec{H} \cdot \frac{\partial \vec{B}}{\partial t}, \]  

(4.17)

and using Eqs. (4.2) and (4.3) as well as the vector identity

\[ \nabla \cdot \left( \vec{E} \times \vec{H} \right) = \vec{H} \cdot \left( \nabla \times \vec{E} \right) - \vec{E} \cdot \left( \nabla \times \vec{H} \right), \]  

(4.18)

Eq. (4.17) can be written as

\[ \frac{\partial u}{\partial t} + \nabla \cdot \left( \vec{E} \times \vec{H} \right) = -\vec{j} \cdot \vec{E}. \]  

(4.19)

This can be rewritten into

\[ \frac{\partial u}{\partial t} + \nabla \cdot \vec{S} = -\vec{j} \cdot \vec{E}, \]  

(4.20)
where we have defined the (instantaneous) Poynting vector
\[ \mathbf{S} = \mathbf{E} \times \mathbf{H}. \] (4.21)

From the structure of the Poynting vector, we can see that the energy propagation is always perpendicular to the electric and magnetic field. Eq. (4.21) is usually referred to as Poynting’s theorem and expresses energy conservation. To better illustrate this, we integrate Eq. (4.20) over some arbitrary volume \( V \), and use Gauss’ theorem (the divergence theorem) to obtain
\[
\int_V \frac{\partial u}{\partial t} \, dx^3 = - \int_V \mathbf{j} \cdot \mathbf{E} \, dx^3 - \int_{\partial V} \hat{n} \cdot \mathbf{S} \, dx^2, \tag{4.22}
\]
where \( \partial V \) denotes the boundary of \( V \), and \( \hat{n} \) is a unit vector perpendicular to the surface \( \partial V \) pointing outwards. This tells us that the rate of change of the energy inside the volume \( V \), is equal to the total rate of work performed inside the volume and the amount of energy flow through the boundary \( \partial V \) of the volume.

We are often interested in the energy flow of monochromatic plane waves
\[
\mathbf{E} = \mathbf{E}_0 e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}, \quad \mathbf{B} = \frac{1}{\omega} \mathbf{k} \times \mathbf{E}_0 e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}, \tag{4.23-4.24}
\]
where Eq. (4.3) has been used to relate the electric and magnetic fields. Note that we are now working with complex-valued notation for the fields, so we can calculate the time-averaged Poynting vector as
\[
\langle \mathbf{S} \rangle = \frac{1}{2} \text{Re} \left[ \mathbf{E} \times \mathbf{H}^* \right], \tag{4.25}
\]
which expresses energy conservation. Using Eq. (4.25) together with the plane wave fields in Eqs. (4.23) and (4.24) we obtain
\[
\langle \mathbf{S} \rangle = \hat{k} \frac{1}{2\eta} \left| \mathbf{E}_0 \right|^2, \tag{4.26}
\]
where \( \hat{k} \) is a unit vector pointing in the direction of the momentum \( \mathbf{k} \) and \( \eta = \sqrt{\frac{\mu}{\epsilon}} \) is the wave impedance (in vacuum this has the value \( \eta_0 \approx 376.7 \Omega \)).

Eq. (4.26) can be utilized to compute the energy flow across boundaries between different media. The energy flow depends on the square of the electric field amplitude as well as the permittivity and permeability of the medium in question. One way to investigate the amount of reflected and transmitted energy is thus to match plane waves across the boundary of
4.2 Electromagnetic scattering from a conducting interface

To study electromagnetic scattering from a conducting interface, such as a graphene sheet, we need the boundary conditions for electromagnetic fields at the interface. The boundary conditions for the electromagnetic fields are

\[
\begin{align*}
(D_2 - D_1) \cdot \vec{n} &= \rho_s, \\
(B_2 - B_1) \cdot \vec{n} &= 0, \\
\vec{n} \times (E_2 - E_1) &= 0, \\
\vec{n} \times (H_2 - H_1) &= j_s,
\end{align*}
\]

where subscript 1 (2) means the field evaluated above (below) the interface, \(\vec{n}\) is a vector with unit length perpendicular to the interface pointing from 1 to 2, and \(\rho_s\) and \(j_s\) is the surface charge and surface current, respectively. The charge and the current are understood to only exist on the boundary between the two media.

To study electromagnetic scattering from a conducting interface, the energy flow can be analyzed as described in the previous section. The normal-incidence results for reflectance and transmittance are

\[
\begin{align*}
\text{reflected energy, } R &= \left| \frac{\varepsilon_2 + \sigma(\omega) - \sqrt{\varepsilon_1 \varepsilon_2}}{\varepsilon_2 + \sigma(\omega) + \sqrt{\varepsilon_1 \varepsilon_2}} \right|^2, \\
\text{transmitted energy, } T &= \frac{4\varepsilon_2 \sqrt{\varepsilon_1 \varepsilon_2}}{\left| \varepsilon_2 + \sigma(\omega) + \sqrt{\varepsilon_1 \varepsilon_2} \right|^2},
\end{align*}
\]

from which we can obtain the absorbance from energy conservation as

\[
A = 1 - R - T = \frac{4\sqrt{\varepsilon_1 \varepsilon_2} \sigma_1(\omega)}{\left| \varepsilon_2 + \sigma(\omega) + \sqrt{\varepsilon_1 \varepsilon_2} \right|^2},
\]

where \(\sigma(\omega) = \sigma_1(\omega) + i\sigma_2(\omega)\). The fact that the scattering coefficients depend on the conductivity of the sheet makes the scattering coefficients frequency dependent. Normal-incidence light scattering from a planar
interface with no spatial variations only depends on the local conductivity, i.e., the results do not depend on momentum $k$.

Fig. 4.1 shows the scattering coefficients, Eqs. (4.31)-(4.33) for a graphene sheet in vacuum using the local conductivity from Eq. (3.62), with $E_F = 0.2$ eV for different temperatures. There is a transition in all scattering coefficients at $\hbar \omega = 2E_F$ which represents the onset of interband transitions. For zero temperature this is a sharp transition which is smeared out due to thermal broadening of the Fermi distribution for non-zero temperatures. For high frequencies compared with the Fermi energy, the absorption approaches the value 2.3% which is characteristic for graphene [104].

It is worth noticing that the scattering results in Fig. 4.1, although containing information about the graphene sample, contains no signature of graphene plasmons. The reason for this is that electromagnetic radiation does not couple to plasmons unless the environment contains a coupling structure. In the next section, we treat solution methods for the scattering problem in such structures, but we postpone the treatment of graphene plasmons and their properties to chapter 5.

4.3 Solution methods for complex light scattering problems

In the previous section we discussed the rather simple solution to the electromagnetic scattering problem for a planar and homogeneous sheet. In this case, the conductivity included in the scattering coefficients depended only on the frequency and not the momentum, making the local conductivity sufficient for this problem. As alluded to above, it is impossible to probe plasmons in such a configuration [1]. The key issue is to bridge the
Figure 4.2: A graphene sheet with a grating. The red arrows represent electromagnetic radiation used to probe the plasmons.

wavelength mismatch between plasmons and the incident electromagnetic radiation. This will be discussed in more detail in chapter 5; here we only note that this is a problem that needs to be overcome in order to study plasmons. One way to achieve this is to introduce periodicity in either graphene or the surrounding environment. In such environments it can be important to consider the nonlocal behavior of the conductivity.

The route taken in this thesis is to pattern the dielectric environment surrounding the graphene. The environment is patterned into an array that is periodic in one dimension and translationally invariant in the other, see Fig. 4.2. This is usually called a dielectric grating and the periodicity is used to facilitate the coupling. Simply put, the inverse of the grating periodicity, $k = \frac{2\pi}{d}$, is supplied to the momentum of the incoming light and when this matches the plasmon momentum, coupling between the incident radiation and the graphene plasmon is possible.

The solution procedure for such infinitely periodic structures is more complicated than for the scattering problem of the homogeneous sheet that was treated in section 4.2. The solution procedure needs to treat the periodic structure as well as the conducting sheet and include the possibility of evanescent modes inside the periodic structure. In the following, we discuss two different methods to solve this scattering problem.
In section 4.3.1 we discuss the scattering matrix (S-matrix) method, which is a very general method to treat scattering problems in different branches of physics. One advantage with this method is the possibility to carry out much of the calculation using matrix algebra with comparatively small matrices, thus reducing computation times. Another advantage is the flexibility in dividing the problem domain into different regions with separate scattering matrices and the possibility to combine them into one scattering matrix in the end.

Section 4.3.2 contains a treatment of the Finite Element Method (FEM) with focus on the basic concepts, since there exists numerous commercial softwares that automatically perform these steps. The main difference between FEM and the S-matrix approach is that while FEM is a real space solver, the S-matrix is a momentum space solver that works in reciprocal space. One implication of this is that FEM often requires heavy numerical calculations while the S-matrix method in principle only needs final evaluation of matrix multiplications to yield the final answer. A benefit of using FEM in general, is its ability to handle arbitrary geometries without much added difficulty. The treatment of FEM in section 4.3.2 includes a discussion of the Floquet boundary condition, an important concept for periodic structures when the incident radiation contains a phase shift between the periodic boundaries.

### 4.3.1 Scattering matrix method

In this section, we cover the basic concepts of the S-matrix method, for a more detailed treatment we refer to Appendix B in Paper I and the article from which the method is adapted, Ref. [105]. Note that in our treatment here, we restrict the discussion to normal incidence.

The periodic scattering problem is depicted in Fig. 4.3 which shows the unit cell and the division of the problem into separate regions. The starting point of the S-matrix method is to use the periodicity and write the electric and magnetic fields as Fourier series of the form

\[
A(x) = \sum_n A_n e^{i k_n x},
\]

(4.34)

where \(k_n = 2\pi n / d\), \(d\) is the length of the unit cell, and \(n\) is an integer. The dielectric function and the sheet conductivity are also written as Fourier
4.3 Solution methods for complex light scattering problems 51

Region 1
Grating region
Region 2
Region 3

Figure 4.3: The division of regions in space for the S-matrix method. Region 1 is a semi-infinite half space, the grating region consists of the periodic dielectric function in Eq. (4.35), region 2 is an infinitely thin air film, and region 3 is another semi-infinite half space. Regions 2 and 3 are separated by graphene, i.e., by the graphene sheet conductivity and thus the boundary condition for the electromagnetic fields is different from the other (non-conducting) boundaries between the different regions.

**Eqs. (4.27)-(4.30)**

\[
\epsilon(x) = \frac{\epsilon_1 + \epsilon_2}{2} - \sum_{n \in \text{odd}} \epsilon_n e^{ik_n},
\]

\[
\sigma(x, \omega) = 
\]

where \(\epsilon_1 (\epsilon_2)\) is the dielectric constant in the shaded (non-shaded) grating region, \(\epsilon_n = (\epsilon_2 - \epsilon_1)/\pi j \sin(\pi j/2)\) are the Fourier coefficients of the dielectric function, \(\sigma(k_n, \omega)\) is the nonlocal (local) graphene conductivity treated in chapter 3. This makes it possible to reformulate the scattering problem in terms of the unknown coefficients \(A_n\) and recast the problem into a system of linear equations that can be efficiently handled using matrix algebra. By enforcing boundary conditions between different regions [Eqs. (4.27)-(4.30)], it is possible to write down matrix relationships between the Fourier coefficients in all regions. We first divide the problem into different scattering regions, see Fig. 4.3, and we write the electric and magnetic fields in free space, Regions 1, 2, and 3, as Fourier sums:

\[
E^\pm_x = \sum_j \Omega_j^\pm e^{i k_j x} e^{i \beta_j z}, \quad (4.37)
\]

\[
H^\pm_y = \sum_j \Lambda_j^\pm e^{i k_j x} e^{i \beta_j z}, \quad (4.38)
\]

where each region has its own set of expansion coefficients and \(k_j^2 + \beta_j^2 =\)
\( \omega^2 \). This means that if the in-plane momentum \( k_j \) of a particular mode \( j \) is larger than \( \omega \), by necessity \( \beta_j \) is imaginary and the mode is evanescent (decaying) in the \( z \) direction (perpendicular to the graphene sheet). Note that the field configuration in Eqs. (4.37) and (4.38) is valid for the electric field in the \( x \) direction, the magnetic field in the \( y \) direction and normal incidence only. The S-matrix method itself is more general and could be adapted to a more general case, but we here consider only this simplified configuration. From the Maxwell equations, it follows that there is a relationship between the field amplitudes of the electric field \( \Omega_j \) and the magnetic field \( \Lambda_j \) of the form

\[
\Lambda_j^\pm = \pm \frac{1}{\sqrt{1 - k_j^2/\omega^2}} \Omega_j^\pm, \tag{4.39}
\]

which allows us to write the fields in free space in terms of only the electric field amplitudes \( \Omega_j^\pm \). We can now write the electric and magnetic field amplitudes in a matrix structure as

\[
\begin{pmatrix}
\vec{E}_m(x, z) \\
\vec{H}_m(x, z)
\end{pmatrix} = \begin{pmatrix}
S_0 & S_0 \\
T_0 & -T_0
\end{pmatrix}
\begin{pmatrix}
\vec{\Omega}_m^+(x, z) \\
\vec{\Omega}_m^-(x, z)
\end{pmatrix}, \tag{4.40}
\]

where \( S_0 \) is a unit matrix, \( T_0 \) is a matrix with \( \frac{1}{\sqrt{1-k_j^2/\omega^2}} \) on the diagonal, and we have introduced the notation

\[
\vec{\Omega}_m^\pm(x, z) = (\ldots \Omega_{-m}^\pm e^{i(k_{-m}x \pm \beta_{-m}z)}, \ldots, \Omega_0^\pm e^{\pm i\omega z}, \ldots \Omega_{m}^\pm e^{i(k_{m}x \pm \beta_{m}z)}, \ldots)^T, \tag{4.41}
\]

and \( \vec{E}_m(x, z) \) (\( \vec{H}_m(x, z) \)) contains all the space dependent field amplitudes of the electric (magnetic) field. The full electromagnetic fields in free space can be obtained by summing all the components of these vectors.

The above discussion gives a flavor of the calculation at hand. These fields can be used to match field amplitudes around the graphene sheet to find the graphene S matrix. A more complicated calculation is necessary to write down the fields inside the grating region and to find the S matrix for the grating. More details about these S-matrices can be found in appendix C. The two S-matrices are given in Eqs. (C.9) and (C.13), and they can be combined to give a final S-matrix from region 1 to region 3. This can be accomplished by rewriting the matrices in terms of transfer matrices, reshuffling the terms and multiplying the T-matrices, and finally reshuffling the answer to S-matrix form again. Following this approach, the final S-matrix from region 1 to region 3 is

\[
S_{\text{tot}} = \begin{pmatrix}
t_2 q_1 t_1 & r_2 + t_2 r_1 q_2 t_2 \\
r_1 + t_1 r_2 q_1 t_1 & t_1 q_2 t_2
\end{pmatrix}, \tag{4.42}
\]
4.3 Solution methods for complex light scattering problems

where we have defined $r_{1/2}$ and $t_{1/2}$ as the components of the S-matrices that can be found in Eqs. (C.9) and (C.13) of appendix C:

$$S_{\text{grating}} = \begin{pmatrix} t_1 & r_1 \\ r_1 & t_1 \end{pmatrix},$$  \hspace{1cm} (4.43)

$$S_{\text{graphene}} = \begin{pmatrix} t_2 & r_2 \\ r_2 & t_2 \end{pmatrix},$$  \hspace{1cm} (4.44)

and

$$q_1 = (S_0 - r_1 r_2)^{-1},$$  \hspace{1cm} (4.45)

$$q_2 = (S_0 - r_2 r_1)^{-1}. \hspace{1cm} (4.46)$$

Eq. (4.42) contains the final result of the S-matrix method and is a 2-by-2 block matrix where every block is an $N \times N$ matrix, with $N$ being the number of amplitudes in the Fourier expansions. The number of modes used in the expansion sets an effective cut-off in momentum that cannot be too large, or the matrices involved become too large to handle. In practice, $N$ is found by checking for convergence of the returned solutions as $N$ is increased. The reflection and transmission amplitudes can be read from the final S-matrix by identifying the propagating mode. It should be noted that in general a periodic structure such as the one we consider may have many propagating modes. However, this is only the case if the wavelength of the incident light is shorter than the period of the structure and this is never the case in the structures we consider. We can thus safely ignore the possibility of grating diffraction in our subwavelength structures and only the zero mode is propagating. It is important to note that the evanescent modes inside the grating play an important role in determining the amplitude of the propagating mode.

4.3.2 Finite element method

The Finite Element Method (FEM) is a widely used method for solving partial differential equations and the fields of application include structural mechanics, fluid dynamics, and electromagnetics. The method is usually attributed to Courant [106] who used it in 1943 to study vibrations. The main idea in FEM is to subdivide the problem domain into subdomains (elements) over which basis functions are defined over single elements. These basis functions may be simple and they contain coefficients that are determined by solving a system of equations that is constructed for the particular problem being solved. The solution obtained is in general not exact, and the quality of the solution depends on how well the basis
functions can represent the exact solution as well as the amount of elements used in the subdivision of the original domain. In FEM, the original partial differential equation is transformed into a system of equations for finding the basis function coefficients. For time-independent problems, the resulting system of equations are coupled algebraic equations and time dependent problems result in a system of coupled ordinary differential equations. The main accomplishment of FEM is the reformulation of the original partial differential equation into a system of equations which has a more straightforward solution method.

In this thesis, we deal with Maxwell’s equations in the frequency domain, where time is explicitly Fourier transformed, leaving us with a problem which is effectively time-independent. We thus focus on the time-independent case, which yields a system of coupled algebraic equations to solve. Depending on the problem, the equations can be linear and such a system of coupled linear equations can be represented by a matrix containing the linear coefficients and a vector containing the unknowns. Many efficient ways of solving such matrix-represented equations exist and can be employed to obtain an approximate solution to the original problem. Many different FEM softwares exist, the FEM analysis presented in this thesis has been performed with COMSOL [107].

The typical problem to be solved by FEM is the boundary value problem formulated on some domain $\Omega$ which has a boundary $\Gamma$, i.e., $\Gamma = \partial \Omega$. This can be written as a differential equation on $\Omega$ as

$$\mathcal{L}\phi = f,$$

(4.47)

where $\mathcal{L}$ is a differential operator. For Eq.(4.47) to have a unique solution it has to be complemented with some boundary condition on $\Gamma$. A FEM solver contains the following basic steps:

1. Discretize the domain into elements;
2. Choose basis functions defined on the elements;
3. Create the system of equations using either
   (a) the Rayleigh-Ritz variational method or
   (b) the Galerkin method;
4. Solve the system of equations.

These steps can best be illustrated by an explicit example and even though the simulations performed in the rest of the thesis are on two-dimensional domains, we will here consider the one-dimensional case. We do so for
simply as higher-dimensional cases become increasingly cluttered due to the complicated node numbering schemes which tends to hide the basic structure of the method.

We consider two infinite parallel plates which are kept at different potentials, separated by a distance \( L \) in the \( x \) direction. The medium between the plates is assumed to have a constant charge distribution and the voltage in the system can be described the Poisson equation with a source:

\[
-\frac{\partial^2}{\partial x^2} V(x) = \frac{\rho(x)}{\epsilon}, \tag{4.48}
\]

\[
\rho(x) = \rho_0, \tag{4.49}
\]

on the domain \( \Omega : x \in [0, L] \), subject to the boundary condition \( V(0) = 0 \) \( V \), \( V(L) = 1 \) \( V \) (Dirichlet boundary conditions). For simplicity we set \( L = 1 \) \( m \), \( \epsilon = 1 \) \( F/m \), and \( \rho_0 = 1 \) \( C/m^3 \) and the exact solution is

\[
V(x) = \frac{(3x - x^2)}{2} \quad [V], \tag{4.50}
\]

which can be found by direct integration and use of the boundary conditions to determine the integration constants. Of course, since we can find the exact solution to this problem there is no need to use FEM to find an approximate solution. However, the strength of FEM is its ability to find approximate solutions to essentially any problem and the present problem serves as a demonstrational example where we can compare the approximate solution with the exact solution. We will now see how to find the approximate FEM solution to this problem by going step by step through the list above. In the first step, we divide the domain into three elements (subdomains), thereby creating four nodes, see Fig. 4.4. In the second step we choose the basis functions to be linear functions within the subdomains of the form

\[
\tilde{V}(x) = V_i \frac{x_{i+1} - x}{x_{i+1} - x_i} + V_{i+1} \frac{x - x_i}{x_{i+1} - x_i}, \tag{4.51}
\]

for \( x_i < x < x_{i+1} \), \( i = 1, 2, 3 \), and \( V_i \) are four unknown coefficients to be found. The boundary condition immediately gives that \( V_1 = 0 \) and \( V_4 = 1 \) and the two remaining coefficients remain to be calculated. This is accomplished in the remaining two steps where we first construct a system
of equations and then solve them. As seen in the list above, there is a choice to be made between the Rayleigh-Ritz method and the Galerkin method. We choose the Rayleigh-Ritz method, which is a variational method and might seem more familiar to physicists. Appendix D describes this method briefly and applies it to determine the coefficients in Eq. 4.51. The solution obtained is

\[ V_1 = 0, \quad V_2 = \frac{4}{9}, \quad V_3 = \frac{7}{9}, \quad V_4 = 1, \quad (4.52) \]

where \( V_1 \) and \( V_4 \) were known already from the boundary conditions.

The solution obtained in appendix D is shown in Fig. 4.5 together with the exact solution and the two can be seen to be in good agreement. By increasing the number of elements, the FEM solution becomes closer to the exact solution at the expense of a larger system of equations to solve. At some point it becomes cumbersome to perform the algebra explicitly like we did above and the resulting system of equations can be conveniently recast into matrix form for which effective solution methods can be used to obtain the unknown coefficients. In the present example it is obviously not necessary to use the FEM solver since we know the exact solution. However, for many problems it may be hard or impossible to find the exact solution, especially in higher dimensions and complicated domain geometries. This is where FEM shows its true strength: it can give approximate solutions for general problems in general geometries.
Floquet boundary condition for periodic structures

We now turn to a special type of boundary condition needed for simulations of non-normal incidence on periodic grating structures. The simulation domains that have been treated in this thesis are periodic grating structures which, although infinite in extension, can be simulated on a single unit cell by taking advantage of the discrete translational symmetry. This is in contrast to the simplified example above where we terminated the domain using Dirichlet boundary conditions. The problem we are interested in is electromagnetic scattering on a periodic structure, i.e., the reflection and transmission from the structure.

Fig. 4.6 shows a schematic view of the system under consideration. The middle unit cell is repeated indefinitely to the left and right, making the system extension infinite. The key to simulating this system is realizing that the solution is also periodic with the same periodicity as the underlying physical system, as well as a possible phase due to misalignment between the periodicity and the source. Mathematically this can be expressed as [108]

\[ u(x + mL) = u(x)e^{-im\phi}, \]  

(4.53)

where \( L \) is the width of the unit cell, \( m \) is an integer, and \( \phi \) is the phase shift between two adjacent unit cells. Eq. (4.53) implies that the total function \( u(x) \) is a product of a periodic function \( u_p(x) \) and a phase factor as

\[ u(x) = u_p(x)e^{-i\phi x/L}, \]  

(4.54)

and can be used as a boundary condition of the form

\[ u(0) = u(L)e^{-i\phi}, \]  

(4.55)

which can be used in simulations to investigate periodic structures. This is known as a Floquet boundary condition and it is a generalization of a periodic boundary condition. Setting \( \phi = 0 \) in Eq. (4.55) reduces to the periodic condition.

We will now investigate some of the implications of using the Floquet boundary condition. Since the function \( u_p(x) \) is periodic, it has a Fourier series expansion of the form

\[ u_p(x) = \sum_{m=-\infty}^{\infty} A_m e^{ik_m x}, \]  

(4.56)

where \( A_m \) are Fourier expansion coefficients, \( k_m = 2\pi m/L \), and \( L \) is the width of the unit cell, see Fig. 4.6. Combining Eqs. (4.54) and (4.56) we
obtain

\[ u(x) = \sum_{-\infty}^{\infty} A_m e^{i(k_m - \phi/L)x}, \]  

(4.57)

which is known as Floquet’s theorem. In condensed matter physics, this is known as Bloch’s theorem and is used to treat band structures in crystals that exhibit periodicity [29], much like the situation we are considering here. To illustrate the use of Floquet’s theorem, we consider the electromagnetic scattering depicted in Fig. 4.6, where radiation is incident from above on the infinitely periodic structure. The radiation may be incident with an angle and can be written as

\[ u^0(x,y) = A e^{i(k_0^0x + k_0^0y)}, \]  

(4.58)

where \( k_0 = \sqrt{(k_0^0 x)^2 + (k_0^0 y)^2} \) is the propagation constant and \( A \) is a constant amplitude. The incident radiation is periodic in space with a phase shift between adjacent unit cells in the \( x \)-direction of \( k_0^0 L \). The solution for the field above the structure can be written as a superposition of the incident field and the scattered field which, according to Floquet’s theorem, can be written as the product of a periodic function and a phase factor as

\[ u(x,y) = u^0(x,y) + u_p(x,y)e^{ik_0^0x}. \]  

(4.59)

By Fourier expanding the periodic part of the scattered field this can be
written as \[108\]

\[ u(x, y) = u^0(x, y) + \sum_{m=-\infty}^{\infty} u^{sc}_{p,m} e^{ik_{m,x}x + ik_{y,m}y} e^{ik_0^0 x}, \] (4.60)

where \( \kappa_m = 2\pi m/L \) defines the Fourier modes. The total wave vector in the \( x \)-direction is now

\[ k_{x,m} = \kappa_m + k_0^0, \] (4.61)

and the wave vector in the \( y \)-direction can be determined to be

\[ k_{y,m} = \begin{cases} \sqrt{(k_0^0)^2 - (k_{x,m})^2}, & \text{for } (k_0^0)^2 \geq (k_{x,m})^2, \\ i\sqrt{(k_{x,m})^2 - (k_0^0)^2}, & \text{for } (k_0^0)^2 \leq (k_{x,m})^2. \end{cases} \] (4.62)

From Eq. (4.61), it is clear that the periodic structure can add momentum in the periodic direction and from Eq. (4.62) we see that not all of the modes can be propagating. This depends on the incidence angle, the wavelength of the incident light, and the periodicity of the grating.

The directions of the reflected propagating waves are \( \varphi_m = \cos^{-1}(-k_{x,m}/k_0^0) \), which using Eq. (4.61) can be rewritten into

\[ \varphi_m = \cos^{-1}\left( -\frac{k_0^0 + \kappa_m}{k_0^0} \right) = \pi - \cos^{-1}\left( \cos(\varphi_0) + \frac{m\lambda_0}{L} \right). \] (4.63)

The angles in the above equations are defined from the \( x \) axis, so normal incidence as depicted in Fig. 4.6 is incident with an angle \( \varphi_0 = \pi/2 \).

The main insight we get from Eq. (4.63) is that the scattered field from the periodic structure consists of at least one propagating mode \( (m = 0) \) and several evanescent modes that decay away from the structure. For normal incidence, we can also see that in order for the grating to induce scattered propagating modes beyond the fundamental mode, the incident wavelength \( \lambda_0 \) has to be smaller than \( L \). Gratings where multiple propagating modes are created are called diffraction gratings and can for instance be used to separate colors from wide spectrum light. The gratings studied in this thesis are deep-subwavelength, and thus \( \lambda_0 > L \) and no propagating modes beyond the fundamental mode are created. This means that the transmitted field will propagate in the same direction as the the incident field and the reflected mode will be specularly reflected, i.e., at an angle \( \pi - \varphi_0 \).

This concludes our discussion of the Floquet boundary condition that was used in COMSOL to perform simulations for non-normal angles of incidence. In the S-matrix method, we only studied normal incidence and thus we only considered periodic boundary conditions. However, the
method itself is general enough to be able to handle also non-normal incidence by using the Floquet boundary condition described above.
Chapter 5

Graphene Plasmons

In this chapter, we introduce graphene plasmons and investigate their electromagnetic field configuration and propagation properties. We also discuss their subwavelength nature and the implications of this for probing graphene plasmons with electromagnetic radiation. In addition, the chapter contains a discussion of an application for graphene plasmons in sensing and discusses sensing figures of merit and electromagnetic field localization provided by the graphene plasmons. This chapter uses results and theory from the preceding chapters and refers to the relevant chapters where possible.

We end this chapter with a discussion and comparison between graphene and metals as plasmonic materials. Due to the rather different geometry of graphene and bulk metals, the resulting plasmons are rather different in nature and the governing equations are quite different. Of course, metals can be patterned and much attention has been paid in the literature to metal nanoparticles, where electromagnetic energy can be confined to very small regions. We list some advantages and disadvantages of plasmons in the different paradigms.

5.1 Graphene plasmon dispersion relation

Plasmons are collective oscillations in the electron density of metals, which have an associated electromagnetic field [1]. Since graphene contains free charge carriers, it also supports plasmons. One way to investigate graphene plasmons is to consider a graphene sheet embedded between two dielectrics. Graphene is described by its conductivity and is included in the electromagnetic problem as a boundary condition. This explains why we paid so much attention to the conductivity in chapter 3. Electromagnetic boundary conditions were also treated in chapter 4. An investiga-
tion of solutions for bound modes at the graphene interface described by the graphene conductivity gives the determining equation for longitudinal plasmons [26]:

\[ 1 + \frac{iq \sigma(q, \omega)}{\omega(\varepsilon_1 + \varepsilon_2)} = 0, \]  

(5.1)

where we have assumed that \( q \gg \omega \), see appendix E for details of the derivation. In the derivation, we have assumed a decaying mode profile in the \( z \) direction (perpendicular to the graphene plane), an electric field in the \( x-z \) directions and a magnetic field in the \( y \) direction. The momentum \( q \) is taken to be in the \( x \) direction, this direction is transverse to the magnetic field, so this mode is sometimes called a Transverse-Magnetic (TM) mode. It is also called a longitudinal mode where longitudinal means that the corresponding oscillation occurs in the direction of propagation. Sound waves are another example of longitudinal waves.

Graphene also supports another type of confined mode; a Transverse Electric (TE) mode [109]. Such bound modes cannot exist at metal interfaces [1] and the fact that they can exist in graphene is due to the unusual behavior of the charge carriers in graphene [109]. In this thesis we focus on the TM plasmons, as they exhibit a larger confinement than the TE modes which are only weakly confined [4].

The longitudinal (TM) plasmon dispersion can be obtained analytically in the zero-temperature and long-wavelength limit, i.e., \( q \to 0 \). By considering the (clean) local conductivity, Eq. (3.62), in the low energy limit, \( \omega \ll E_F \), we obtain \( \sigma(\omega) = ie^2E_F/(\pi \omega) \). Inserting this in the nonretarded dispersion equation, Eq. (5.1), we obtain the long-wavelength plasmon dispersion

\[ \hbar \omega = \sqrt{\frac{4\alpha \hbar c E_F q}{(\varepsilon_1 + \varepsilon_2)}}, \]  

(5.2)

where \( \alpha = e^2/(4\pi c \varepsilon_0 \hbar) \approx 1/137 \) and we have briefly reinserted \( \hbar \) and \( c \) for clarity. The plasmon energy depends on the square root of the wave vector, which is in stark contrast to the linear dispersion of photons and, as a consequence, plasmons with different energies propagate with different velocities. It also means that the plasmon group velocity and the phase velocity never coincide, but are related by \( v_g = v_{ph}/2 \).

The fact that the plasmon energy depends on the Fermi energy makes it clear that it is possible to affect graphene plasmons by changing the doping. As discussed in chapter 2, the doping level in graphene can be controlled electrostatically and through this the plasmon energy can be controlled. Typically, graphene plasmons are in the terahertz to mid-
infrared regime [4]. From the dispersion relation it is also evident that the plasmon energy depends on the dielectric environment surrounding the graphene sheet. From this we may conclude that there is a possibility to utilize graphene plasmons to sense the local environment. We investigate this further in section 5.4.

The $\sqrt{q}$-behavior obtained in Eq. (5.2) is also obtained for plasmons in 2DEGs, in fact it is a consequence of the two-dimensionality [4]. By using that $E_F = \sqrt{\pi n}$ we can see that the graphene plasmon energy depends on $n^{1/4}$, in contrast to the 2DEG dependence of $n^{1/2}$. As a side note, it is amusing to note that water waves obey the same dispersion relation as long-wavelength graphene plasmons. Deep-water waves obey the dispersion relation $\omega = \sqrt{gq}$ [110], where $g \approx 10 \text{ m/s}^2$ is the gravitational constant on earth (the corresponding “acceleration” for graphene plasmons is on the order of $10^{21} \text{ m/s}^2$).

We will now use the plasmon dispersion and show that graphene plasmons indeed facilitate a strong electromagnetic field localization. By comparing the wavelength of free-space electromagnetic radiation with the plasmon wavelength, we can get an expression for the wavelength of both modes. From the dispersion relations we obtain the wavelength ($\lambda = 2\pi/q$) for a given frequency as

\begin{equation}
\lambda_0 = \frac{2\pi c}{\omega},
\end{equation}

\begin{equation}
\lambda_{pl} = \frac{4\pi \alpha c E_F}{\hbar \omega^2},
\end{equation}

where $\lambda_0$ is the free-space wavelength of electromagnetic radiation, $\lambda_{pl}$ is the plasmon wavelength, and we have taken the dielectric environment to be vacuum, i.e., $\varepsilon_1 = \varepsilon_2 = 1$. From this we can easily compute the ratio between the free-space wavelength and the plasmon wavelength as

\begin{equation}
\frac{\lambda_0}{\lambda_{pl}} = \frac{\hbar \omega}{2\alpha E_F},
\end{equation}

which is plotted as the red dashed line in the right panel of Fig. 5.1. The ratio expressed in Eq. (5.5) is valid in the small-frequency ($\hbar \omega \ll E_F$) limit and should strictly speaking not be extended beyond $\hbar \omega = E_F$. However, solving the dispersion relation using the nonlocal conductivity numerically, we can obtain the same ratio but valid for all frequencies. This is shown as the blue solid line in Fig. 5.1 and for small frequencies it agrees with the low-energy expansion expression obtained above. Both the local and nonlocal ratio in the right panel of Fig. 5.1 tend to zero as $\omega \to 0$, which signals that plasmon wavelength becomes much larger.
than the wavelength of electromagnetic radiation in this limit. This is an artifact of the non-retarded approximation \((q \gg \omega)\) which was made to arrive at Eq. (5.1) and it is only relevant in the very small frequency limit. Should the fully retarded equation be used, the ratio would tend to unity in the small frequency limit. For energies larger than \(0.1 E_F\), i.e., the energies relevant for graphene plasmonics, this correction has no effect on the results and can safely be neglected.

From Fig. 5.1, one might conclude that graphene plasmons facilitate a localization up to a factor of 300 of the incident radiation. However, this is arrived at by assuming no losses (clean graphene at \(T = 0\ K\)) and it turns out that under more realistic conditions, plasmons with large confinement are associated with larger losses. We turn to an investigation of plasmon losses in graphene in the next section.

### 5.2 Graphene plasmon propagation and damping

The results for the dispersion in the previous section are only valid in the clean graphene case and at zero temperature. For nonzero temperatures and/or damping mechanisms added, the graphene conductivity acquires a real part in addition to its imaginary part. The plasmon dispersion equation, Eq. (5.1), then becomes two equations, the real part and the imaginary part, and cannot be solved in general with only one variable. To accommodate this, we let \(q = q_1 + iq_2\) be a complex variable and we
obtain $q_1$ and $q_2$ by solving both the real and imaginary parts of the dispersion equation. Since the conductivity is a rather complicated function of $q$, finding the solution could in general be quite hard for the coupled equations. To simplify the solution procedure we realize that we are mostly interested in weakly damped plasmons, i.e., $q_2/q_1 \ll 1$, and we may expand the dispersion equation in this small parameter and keeping only the lowest-order term we obtain

$$1 = \frac{q_1 \sigma_2(q_1, \omega)}{\omega(\varepsilon_1 + \varepsilon_2)}, \quad (5.6)$$

$$\frac{q_2}{q_1} = \frac{\sigma_1(q_1, \omega)}{\partial_{q_1}(q_1 \sigma_2(q_1, \omega))}. \quad (5.7)$$

We note that Eq. (5.6) is the same equation as we would obtain by taking Eq. (5.1) directly and assuming no losses, i.e., neglecting $\sigma_1(q, \omega)$. By solving Eq. (5.6), we may use the solution to evaluate the expression in Eq. (5.7) to obtain the value of $q_2$, i.e., the losses. The value of $q_2$ can be seen to be proportional to $\sigma_1$, i.e., the real part of the conductivity. In effect, a completely imaginary conductivity yields zero plasmon losses.

An investigation of the real part of the clean graphene-conductivity shows that there are no losses at zero temperature in a certain region in $q$-$\omega$ space. This is shown in the top left panel of Fig. 5.2, where the real part conductivity is shown in a large region of $q$-$\omega$ space. There is a triangle where $\sigma_1(q, \omega) = 0$ and the graphene plasmons experience no losses in this region. The regions where the conductivity becomes non-zero are associated with interband and intraband excitation of electron-hole pairs. This is the only loss channel that exists in clean graphene. Inside the triangle, these transitions cannot take place due to Pauli blocking.

The top right panel of Fig. 5.2 shows the real conductivity for graphene at $T = 300$ K for $\Gamma = 0.012$ eV($\tau = 50$ fs). The $\Gamma$ is added using the Mermin relaxation time approximation, see section 3.4 The previously lossless triangle now has a non-zero conductivity due to the electron relaxation as well as non-zero temperature. The temperature effect arises due to thermal smearing of the electron distribution which makes the Pauli blocking not perfect close to the triangle edges. The electron relaxation time affects the entire region inside the triangle.

The plasmon wavelength is given by $\lambda_{pl} = 2\pi/q_1$ and the propagation distance is given by $L_p = 1/(2q_2)$ [4]. This definition takes the rather strict view that the plasmon is considered decayed when the intensity of the plasmon has dropped to $e^{-1} \approx 0.37$ of its initial value.

A convenient way to quantify the propagation of plasmons is to calculate
Figure 5.2: Real part of the graphene conductivity, $\sigma(q, \omega)$ at $T = 0$ and $\Gamma = 0$ (top left panel) and at $T = 300$ K $\Gamma = 0.012$ eV (top right panel). The relaxation time is included using the Mermin relaxation time approximation. The bottom panel shows a cut of the real part at $q/k_F = 0.2$ of the top panels (shown as vertical lines in the top panels). All results are shown for $E_F = 0.2$ eV. The white triangle in the top left panel shows the region where the real conductivity is identically zero for clean graphene at $T = 0$ K. This triangle is highlighted with black dashed lines that are also superposed on the top right panel where the real part is no longer zero. The non-zero real part will translate into non-zero plasmon losses, see Eq. (5.7).
the propagation distance in units of the plasmon wavelength, i.e. $L_p/\lambda_p$. By studying the definitions of $q_1$ and $q_2$ above, this ratio can be expressed as

$$\frac{L_p}{\lambda_p} = \frac{q_1}{4\pi q_2}, \quad (5.8)$$

i.e., the inverse of the small expansion parameter $q_2/q_1$ used above. From the ratio in Eq. (5.8), we see that (as expected) undamped plasmons have an infinite propagation length. Below, we investigate the propagation distances, as quantified by the ratio in Eq. (5.8), for realistic relaxation times. It should be noted that this is done here using the Mermin relaxation time approximation.

Fig. 5.3 shows the propagation length for different relaxation times and at room temperature. The left panel shows the propagation length in units of the plasmon wavelength and the right panel shows the propagation length units of nanometers (log-scale). The typical number of oscillations obtained using the Mermin relaxation time approximation agrees qualitatively with the number of plasmon oscillations found in experiments [24, 66]. However, this is only indicative since the relaxation time approximation is a phenomenological model.

The left panel of Fig. 5.3 shows that $q_1/q_2$ has a maximum around $\hbar\omega = E_F$ and experiences a sharp drop for larger frequencies. This drop is due to interband transitions that rapidly damps out the plasmon mode. We conclude that the maximum frequency of plasmons under realistic conditions are only slightly larger than $E_F$, larger frequencies are completely damped. From Fig. 5.1 we see that this limits the field localization to around $\alpha^{-1} \approx 137 [4]$. This is still a large localization factor, but not as
large as 300 that was achieved when losses were neglected in section 5.1.

5.3 Coupling electromagnetic radiation to graphene plasmons

It was shown in Fig. 5.1 that the plasmon wavelength can be much smaller than the electromagnetic free-space wavelength at the same frequency. As a consequence, there is a large wavelength mismatch between electromagnetic radiation and plasmons of the same frequency, making excitation of plasmons with radiation impossible on flat surfaces [1]. This is true also for plasmons in metals and several ways to overcome this problem have been developed over the years [1]. Examples are prism coupling [111, 112], grating coupling [10], and near-field coupling [113]. Plasmons can also be probed using other methods, such as electron energy loss spectroscopy [114, 115]. From now on we focus only on electromagnetic coupling to plasmons.

For graphene plasmons, patterning the graphene into microribbon arrays [57, 116] was an early method used for coupling. The microribbon array acts as a grating and supplies the coupling between incident radiation and graphene plasmons. The plasmons can then be probed by illuminating the array with electromagnetic radiation and monitoring the transmission or reflection. Using microribbon arrays, the authors in Ref. [116] were able to demonstrate the gate-tunability of graphene plasmons.

Another route to couple to graphene plasmons is by using nanotip illumination [24, 25, 117]. This creates an evanescent field surrounding the nanotip, placed in close proximity to the graphene sample, through which electromagnetic radiation couples to the plasmons. By now, this method has been used in many works to study various aspects of graphene plasmons. Examples are coupling to localized states [118], studies of graphene edge plasmons [119, 120], nonlocal effects [102], and plasmon-induced photocurrent [121]. In addition, illumination of metal nanoantennas, deposited on top of graphene, with a laser can be used to excite plasmons [122].

Subwavelength patterning of the environment surrounding the graphene sample has also been used to couple electromagnetic radiation to graphene plasmons. Periodic arrays of metal patterning [40, 123] as well as dielectric gratings [124, 125] have been investigated. In this thesis, we focus on dielectric gratings in proximity to a graphene sheet in order to facilitate the necessary coupling between the incident radiation and the graphene plasmons. A grating together with a graphene sheet is depicted in Fig. 4.2.
and methods for solving the electromagnetic scattering problem in such geometries were treated in section 4.3.

5.4 Refractive index sensing using graphene plasmons

One application of plasmons is for label-free sensing of the local environment [17], which can be used for drug discovery, food safety, and environmental monitoring [16]. Many biomolecules are active in the mid-infrared which makes graphene plasmons ideal for sensing such molecules. Graphene plasmons are found to be in this part of the spectrum and as we have also seen, they are tunable by means of external gating. This can be used for taking spectroscopic fingerprints of biomolecules, something already demonstrated by Rodrigo et al [38]. Recently, also double layer graphene has been used to further enhance the sensing capabilities of graphene plasmons [126].

Plasmonic sensors work by probing some aspect of plasmons in the sensor, such as the plasmon resonance frequency, and monitor this as the environment changes [16, 17]. The change in the environment can then be inferred by reading out the change in the plasmonic signal. The benefit of using plasmons for this is that they can dramatically increase the electromagnetic field strength in the region surrounding the sensor, which increases the response to small changes in the environment [1]. However, there are limitations that arise from plasmon losses, which degrade the sensitivity and need to be considered. In the next section, we address the question: “how good is sensing using graphene plasmons and how to quantify this?”

5.4.1 Figures of merit for refractive index sensing

To investigate sensing performance, it is convenient to introduce a Figure of Merit (FOM) that captures aspect(s) of the sensing scheme. We will consider the response of graphene plasmons to changes in the dielectric environment. The sensor setup we consider consists of a graphene sheet together with a dielectric grating to provide coupling between electromagnetic radiation and the plasmons. This setup can be probed by incident radiation and the transmission and reflection can be measured. The transmission has a dip as the frequency sweeps across the sensor and the position of this resonance constitutes the sensing signal. Now, this dip has a finite width due to the fact that the plasmon experiences losses as well as being coupled to radiation. The width of the peak can mask small
shifts in the plasmon frequency and in practice limits the ability to detect small variations in the dielectric environment, i.e., it limits the sensitivity of the proposed sensing scheme.

For the sensing scheme described above, a commonly used FOM is the frequency shift of the plasmon resonance due to a small refractive index change, divided by the resonance width \[127\]:

\[
FOM_{\text{bulk}} = \frac{m}{\Gamma_p},
\]

\[
m = \frac{\partial \omega_p}{\partial n},
\]

where \(\omega_p\) is the plasmon frequency, \(\Gamma_p\) is the width of the plasmon resonance, and \(n\) denotes refractive index. The quantity \(m\) expresses the change in resonance frequency per refractive index unit. We refer to the FOM in Eq. (5.9) as the bulk FOM for reasons that will become clear below.

The bulk FOM lacks any notion of thickness of the medium experiencing the refractive index change. This is expected to be a good measure if the entire field localized by the plasmon is covered with the medium to be sensed. However, this might not be the case when thin layers of biomolecules are investigated. To address this, a thickness dependent FOM can be introduced \[128, 129\]

\[
FOM = \frac{m}{\Gamma_p},
\]

\[
\frac{\partial \omega_p}{\partial n} = m \left(1 - e^{-2t/L_d}\right),
\]

which is nothing but a redefinition of the quantity \(m\). The quantity \(L_d\) is the decay length of the plasmons and quantifies how far from the graphene sheet the fields of the plasmon extend. In the case where the thickness \(t \gg L_d\), this reduces to the bulk FOM defined above. By introducing the thickness in the FOM, it is possible to study the difference between having the sensor completely and partially immersed in the dielectric medium to be sensed. The sensing is more effective the more the plasmon’s electromagnetic field is covered, but since graphene plasmons are confined close to the graphene sheet, already thin layers of sensing medium completely cover the graphene plasmon. With the thickness added to the modeling of the plasmon frequency shift, it becomes possible to also use graphene plasmons to monitor thicknesses of thin layers of biomolecules with a known refractive index.

For an investigation of refractive index sensing using graphene plasmons
quantified by the above FOMs, we refer to Paper II.

5.5 Comparison between graphene plasmons and plasmons in metals

This section is devoted to a comparison between graphene plasmons and plasmons in metals. Metal plasmonics is a vast subject and this section does not give an overview of the entire field. Rather, focus is on aspects of the physical differences — and similarities — of plasmons in the two cases. In metal plasmonics, there are several possible geometries to be compared with. The two cases that are treated here are metal-dielectric interface plasmons and plasmons in small metal nanoparticles.

The similarity between plasmons in metals and plasmons in graphene, is that they both arise as solutions to Maxwell’s equations around the conductor in question. The presence of conduction electrons modifies the electromagnetic field solutions, giving rise to the hybrid nature of plasmons. A flat metal surface is the metal geometry that most resembles the graphene plasmons that have been considered in this thesis. These plasmon modes can propagate along the interface between the metal and the dielectric and we treat this case below in section 5.5.1.

It turns out that interface plasmons in metal-dielectric interfaces are quite different from graphene plasmons. Perhaps the most striking difference is that metal interface plasmons exhibit much less localization, i.e., the wavelength of the interface plasmons is close to the electromagnetic wavelength for the same frequency. This is in large contrast to graphene plasmons where the localization can be on the order of 100. To study plasmons with more comparable localization properties, section 5.5.2 treats spherical nanoparticles which support localized (non-propagating) plasmon modes which exhibit large field localization comparable with graphene plasmons. Section 5.5.3 contains a summary of the metal plasmon properties and a comparison with graphene plasmons.

5.5.1 Metal-dielectric interface plasmons

By performing a similar calculation as was done for the graphene plasmon dispersion relation in appendix E, it is possible to arrive at the corresponding plasmon dispersion equation for the metal-dielectric interface TM plasmons [1]:

\[ q = \frac{\omega}{c} \sqrt{\frac{\varepsilon_1 \varepsilon_2}{\varepsilon_1 + \varepsilon_2}}. \]  

(5.13)
The material parameters of importance are the dielectric constants of the two materials above and below the interface. Setting the dielectric to be air ($\varepsilon_1 = 1$) and taking the metal to be silver we get the results shown in Fig. 5.4. The experimental data for the dielectric function of silver is from Ref. [130], obtained at [131]. It turns out that TE plasmons cannot exist at metal interfaces [1], in contrast with their existence at a graphene interface [109].

From the solution to Eq. (5.13), it is easy to compute some derived properties such as the propagation length in units of the plasmon wavelength, defined in Eq. (5.8), as well as the wave localization provided by the plasmons in silver. These quantities are shown in Fig. 5.5; these modes have properties that are quite distinct from graphene plasmons. At low energies, the metal interface plasmons propagate very far, on the order of $10^4$ plasmon wavelengths as shown in the top panel. In actual distances this can be up to millimeters, this is indeed found in experiments on low-energy plasmons in silver [15]. Such a long propagation length is possible due to the poor confinement at these frequencies. The bottom panel shows the localization factor and at low frequencies, the plasmon wavelength approaches the free space electromagnetic wavelength. At these frequencies the plasmon acquires a more light-like nature and is essentially a grazing incidence light beam [1].
5.5 Comparison between graphene plasmons and plasmons in metals

Figure 5.5: Calculated properties for plasmon in a silver-air interface obtained from the solution shown in Fig. 5.4. Top: Propagation length for surface plasmons in silver. The propagation length (in units of the plasmon wavelength) for low-energy surface plasmons can be seen to approach $10^4$. Bottom: Localization factor for silver plasmons. The most confined surface plasmons in silver are localized by a factor 1.8 at the expense of becoming essentially completely damped (see top panel for the same frequency). This damping occurs due to the onset of interband transitions which is also visible in Fig. 5.4.
5.5.2 Localized surface plasmons in metal nanoparticles

The study of localized resonances in metal nanoparticles was studied a long time ago by Mie [132], who essentially solved the electromagnetic scattering problem for metal particles. The problem with this solution is that it is expressed as infinite sums which are difficult to gain insight from. So even though Mie theory has a vast predictive power in the sense that any scattering can be calculated on a computer that evaluates the sum to arbitrary precision, it is of interest to develop simpler models of perhaps limited validity that can allow for more insight due to its simpler structure.

One such simplification arises in the limit of a small sphere of radius \( a \), where \( a \ll \lambda \), i.e., the radius of the sphere is much smaller than the incident wavelength; this limit is often referred to as Rayleigh scattering. The problem can then be treated electrostatically and the solution for the electric field is [1]

\[
\begin{align*}
\vec{E} &= \begin{cases} 
\frac{3\varepsilon_2}{\varepsilon(\omega)+2\varepsilon_2} \vec{E}_0, & r < a, \\
\vec{E}_0 + \frac{3\vec{n} (\vec{n} \cdot \vec{p}) - \vec{p}}{4\pi\varepsilon_0 \varepsilon_2} \frac{1}{r^3}, & r > a,
\end{cases} \\
\vec{p} &= 4\pi\varepsilon_0 \varepsilon_2 a^3 \frac{\varepsilon(\omega) - \varepsilon_2}{\varepsilon(\omega) + 2\varepsilon_2} \vec{E}_0,
\end{align*}
\]

where \( \varepsilon(\omega) \) is the dielectric function for the metal in the sphere, \( \varepsilon_2 \) is the dielectric constant of the surrounding medium, \( \vec{E}_0 \) is the incident electric field, and \( \vec{p} \) is the dipole moment of the sphere. Note the appearance of the factor \( (\varepsilon(\omega) + 2\varepsilon_2)^{-1} \) which gives rise to an enhancement of the fields when \( \text{Re}\{\varepsilon(\omega)\} = -2\varepsilon_2 \). This is called the Fröhlich condition and the mode associated with this is called the dipole surface plasmon [1]. A general term for resonances in metal nanoparticles is Localized Surface Plasmon Resonances (LSPRs). For a small silver sphere in air (\( \varepsilon_2 = 1 \)), the dipole resonance condition occurs at \( \hbar \omega = 3.5 \text{ eV} \) (\( \lambda = 354 \text{ nm} \)), using data from [130].

Fig. 5.6 shows the electric field strength, Eq. (5.14), in units of the incident field strength (in the static approximation) for a small silver sphere on resonance at \( \lambda = 354 \text{ nm} \). The incident field is polarized in the \( z \) direction and the field is enhanced by the presence of the small metal sphere and reaches an enhancement of 15 on the top and bottom of the sphere. The magnitude of the field enhancement is set by the imaginary part of the dielectric function, which provides the limiting value for \( (\varepsilon(\omega) + 2\varepsilon_2)^{-1} \) when the resonance condition is met. A (hypothetical) lossless dielectric material would give an infinite field enhancement exactly on resonance.
5.5 Comparison between graphene plasmons and plasmons in metals

The conclusion is that also metal nanoparticles have the ability to confine light, similar to graphene plasmons.

The above treatment was the simplest case of scattering on a nanoparticle. In the more general case where the size of the sphere approaches the incident wavelength there are more modes appearing than just the dipole mode. A treatment of such modes requires more sophisticated approximations or use of full Mie theory.

Finally, we mention that in terms of enhancement of the electromagnetic field there is another method that can be used and that is the so called lightning-rod effect [12]. This is the enhancement of electromagnetic fields around sharp tips and edges of metal structures. This can be used to further enhance the field around a metal sphere by placing it close to a sharp corner of another metal object. One such implementation is called a bow-tie nanonantenna where gold bow ties can enhance the field more than 30-fold [133].
### Table 5.1: Comparison between properties of graphene plasmons, metal interface plasmons, and localized surface plasmon resonances (LSPRs).

<table>
<thead>
<tr>
<th></th>
<th>Propagation</th>
<th>Confinement</th>
<th>Tunable</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graphene plasmons</td>
<td>short</td>
<td>strong</td>
<td>yes</td>
<td>mid-infrared</td>
</tr>
<tr>
<td>Interface plasmons</td>
<td>long</td>
<td>weak</td>
<td>no</td>
<td>near infrared - visible</td>
</tr>
<tr>
<td>LSPR</td>
<td>-</td>
<td>medium</td>
<td>no</td>
<td>visible - UV</td>
</tr>
</tbody>
</table>

#### 5.5.3 Summary of graphene versus metals for plasmonics

Table 5.1 provides a summary of the plasmon properties that were treated earlier in this section. This summary highlights the fact that graphene plasmons are different from plasmons in metals in that they exhibit a very strong confinement and as a result, they do not propagate very far. In addition, the frequency range for graphene plasmons is in the mid-infrared, whereas metal plasmonics is usually done at larger frequencies.

Metal plasmons can provide either very long propagation length (interface plasmons) or a rather large field enhancement (LSPR), but not both at once. Perhaps one of the most important advantages of graphene plasmons is that they are tunable. This opens up possibilities for tunable photonic devices, in the mid-infrared, that can be very compact since the graphene plasmon is very small in wavelength.
Chapter 6

Overview of Appended Publications

This thesis is based on the work presented in Papers I-V, appended at the end of the thesis. This chapter provides a brief description of the papers as well as a summary of the main results.

6.1 Paper I

In this publication, we investigated the nonlocal properties of graphene plasmons and compared and contrasted them with graphene plasmons in the local approximation. To this end, we calculated the graphene conductivity, using linear response theory, for both the nonlocal and local cases and solved the plasmon dispersion equation for them both. Doing so, we computed the plasmon dispersion relations and we found, as expected, that the local approximation is rather accurate for small \( q/k_F \), while for larger \( q/k_F \) the more accurate full nonlocal dispersion starts showing significant deviations from the local one. We showed that this difference occurs over a large range of temperatures, from zero temperature up to \( T/T_F = 1 \), and that larger temperatures tend to make the local approximation agree slightly better with the nonlocal dispersion.

In addition, we studied the optical response of graphene plasmons in a nanostructured dielectric environment (subwavelength grating). The focus was on comparing the optical responses for graphene plasmons in the local and nonlocal models. The main finding was that below \( T/T_F = 0.15 \), the reflection, transmission, and absorption as well as the width of the resonance features differed significantly between the two cases. Above \( T/T_F = 0.15 \), the two cases gave rather similar results and the main difference was then in the frequency shift from the difference in the plasmon dispersion. It should be noted that \( T/T_F = 0.15 \) can easily extend several hundred kelvin above room temperature for realistic doping levels. This
means that nonlocal effects can be of importance to correctly capture the optical properties of graphene plasmons at room temperature.

6.2 Paper II

In this paper, we applied graphene plasmons for sensing purposes as well as quantified the sensing ability of the graphene plasmons using standard figures of merit (FOMs). In our calculations we included the effects of temperature, nonlocality, and also a scattering time that affects the electrons as they propagate in the lattice. The scattering time was included phenomenologically in a number conserving Mermin relaxation time.

Our calculations showed that for thick layers of sensing material close to the sensor, the FOM was 11. This is a rather good number although larger FoMs have been achieved in the literature. However, the main strength of our sensor comes from the extreme confinement of the graphene plasmons, which creates a high sensitivity very close — on the order of 10 nm — to the graphene surface. Together with the tunability of the graphene plasmons, this constitutes an attractive platform for taking spectroscopic fingerprints of vibrational modes in biomolecules. Thus allowing for selective sensing of very thin layers of biomolecules using our proposed sensor setup.

6.3 Paper III

Here, we studied a low density of adatoms coupled to the graphene lattice by tunneling. The tunnel coupling as well as the energy level of the adatoms were chosen to match those of hydrogen atoms and we extracted these parameters from DFT studies. Using this model, we calculated the conductivity within linear response and observed that the conductivity could change significantly depending on the adatom density as well as the position of the adatom energy level with respect to the Fermi energy.

Furthermore, we used the conductivity to solve the plasmon dispersion equation and investigated the graphene plasmon properties. In particular we studied the plasmon dipsersion, the plasmon propagation length, and the plasmon damping induced by the adatoms. We found that the plasmon dispersion could be changed significantly even for a small density of adatoms and we observed a level splitting between the plasmons and the adatom energy level. The level splitting produced two distinct plasmon branches, one high energy branch and one low energy branch. The high
energy branch was damped by the presence of the adatoms and obtained a smaller propagation distance while the low energy branch was much less affected.

The adatom influence on the graphene plasmons occur for rather small densities of adatoms and this effect can therefore be used for sensing purposes. We showed in the paper how the plasmon width and frequency is affected by the impurities. For sensing, a grating setup similar to the ones in Papers I and II could be used in order to couple to the plasmons and read out these properties.

### 6.4 Paper IV

In Paper IV, we investigated impurities in graphene in a self-consistent model which fully included nonlocal effects. The calculations showed that an impurity band emerges in the density of states and the energy position of this band depends on the impurity density and strength of the impurities. These are free parameters in our model and we investigated the results for different parameter values, finding that the impurity band always exists and can result in a substantial real part in the conductivity below $2E_F$. We also found that the plasmons are very sensitive to the presence of the impurity band, especially when the impurity band has the same energy as the plasmons. Since the position of the impurity band depends on the strength of the impurities as well as the density of impurities, this implies that the damping of graphene plasmons can depend strongly on the type of impurities in graphene.

Furthermore, we compared our nonlocal self-consistent impurity model with the commonly used Mermin relaxation time model, finding rather large differences. This highlights that there may be a need for more complex models of impurities, such as the one considered in this paper, than what is normally used in the literature. However, further studies, both experimental and theoretical, are needed to further quantify the impurities normally occurring in graphene in order to deduce the proper model and parameters for impurity-induced plasmon damping.

### 6.5 Paper V

This publication contains an investigation of graphene plasmons in non-equilibrium graphene. We considered graphene to be biased with a DC voltage to produce an in-plane current in the graphene sample. We studied
the resulting effects on the graphene plasmons by calculating the graphene conductivity in the non-equilibrium graphene. We found asymmetries in the plasmon dispersion and plasmons propagating parallel (anti-parallel) with the current obtained an enhanced (decreased) propagation length. We also discussed how this effect can be measured in electromagnetic scattering experiments in a dielectric grating environment. The ability to control the plasmon propagation using a DC current could potentially be used to create miniaturized photonic devices such as optical modulators.

Moreover, we showed how the increased propagation length leads to narrower linewidths in the electromagnetic scattering signals. This can potentially lead to enhanced sensitivity for graphene plasmon-based sensors like the one considered in Paper II, since narrow linewidths increase the sensing figure of merit.
Chapter 7

Summary

In this thesis we treated plasmons in graphene starting from a microscopic Hamiltonian for the low-energy electrons in graphene. This Hamiltonian has been used pervasively in the literature and provides a solid foundation in which calculations are manageable and are not completely performed in large computer simulations. From the microscopic Hamiltonian, we derived a macroscopic property, the graphene conductivity, which can be used as input to Maxwell’s equations to investigate electromagnetic modes. To perform the linear response calculation, we used a general Keldysh Green’s function formalism which in addition to being useful in equilibrium, is also able to handle perturbations of non-equilibrium states. The price to pay is the addition of new Green’s functions that need to be handled and this is done using a 2-by-2 matrix structure in Keldysh space.

When the linear response in terms of Keldysh Green’s functions was done, we had arrived at general expressions for the polarizability of graphene in terms of Green’s functions and self energies. The self energies are one of the strengths of the Green’s functions method; it is possible to include additions to the bare graphene Hamiltonian and formulate a perturbation series in terms of a self energy. The self energy can be calculated for various additions to the bare graphene Hamiltonian (impurities etc.) and can then simply be included in the calculation of the graphene conductivity. It can be noted that in the clean-graphene case, we arrived — as we must — to the same expression for the polarizability as previous authors.

Armed with the conductivity, we were in a position to investigate the plasmon mode surrounding graphene sheets. These modes turn out to be deeply subwavelength and special care must be taken to interact with plasmons using electromagnetic radiation. In anticipation of the required nanopatterning, we spent some time treating Maxwell’s equation in subwavelength patterned environments and, in particular, we studied two different solution methods to the electromagnetic scattering problem in
such cases; the S-matrix method and FEM analysis.

Finally, we arrived at a treatment of graphene plasmons and started out by studying the graphene dispersion equation and a treatment of losses in graphene plasmons. Losses turn out to have a considerable effect on the graphene plasmons and limits the field localization to values around $\alpha^{-1} \approx 137$ for realistic values of damping in graphene. In this thesis, we have studied graphene plasmons with nonlocal effects included and for frequencies comparable with $E_F$, we saw that the dispersion relation in the less accurate local approximation deviates considerably from the nonlocal dispersion relation. Also, we observed that the presence of the interband continuum limits the frequency at which plasmons can propagate and introduces a considerable damping. This is not present in the local approximation, which breaks down at frequencies close to $E_F$.

Plasmon losses are traditionally introduced using the Mermin relaxation time approximation (RTA), which is practical due to its simplicity. Its limitation, however, is that it introduces a phenomenological relaxation time without any microscopic mechanism. The results we display in this thesis are obtained using the RTA, but the methods we have developed, i.e., Green’s functions with self energies can be used to also investigate microscopic scattering mechanisms. This was studied in Papers III and IV where adatoms and random scattering sites are investigated in terms of their effect on the graphene plasmons. The conclusion from these papers is that the microscopic details of the scattering that occurs matters for the losses induced in the plasmons. Further studies are needed, both theoretical and experimental, in order to find the best model for damping of graphene plasmons. This may well turn out to be different for different graphene samples due to differences in production method, environment etc.

In the appended publications we have also investigated nonlocal effects on graphene plasmons and their signature in light scattering experiments. Nonlocal effects can be important to consider when the plasmon wavelength is comparable with the Fermi wavelength. We also studied plasmon losses and their effects on the light scattering signals and applied this for sensing purposes. In addition, we studied the possibility of controlling graphene plasmons using DC currents in the graphene sheet.
Appendices
Appendix A

Graphene Hamiltonian and Electron Bands

In this appendix we derive the momentum space Hamiltonian in Eq. (2.8) starting from the real space tight-binding Hamiltonian in Eq. (2.6).

For convenience, we reprint the real-space tight-binding Hamiltonian in Eq. (2.6), which is the starting point of the calculation. The Hamiltonian reads

\[
H = -t \sum_{\langle i,j \rangle, s} \left( a_{i,s}^\dagger b_{j,s} + h.c. \right) \tag{A.1}
\]

and \langle .. \rangle denotes nearest-neighbor summation, \( s = (\uparrow, \downarrow) \) is a sum over spins, and \( t = 2.8 \) eV is the nearest-neighbor hopping amplitude. We also reprint Eqs. (2.3)-(2.5), the nearest-neighbor vectors that connect an atom in the graphene lattice with its nearest neighbors:

\[
\vec{\delta}_1 = \frac{a}{2} \left( 1, \sqrt{3} \right), \tag{A.2}
\]

\[
\vec{\delta}_2 = \frac{a}{2} \left( 1, -\sqrt{3} \right), \tag{A.3}
\]

\[
\vec{\delta}_3 = a (-1, 0). \tag{A.4}
\]

We now insert a Fourier transform of the real-space operators, i.e., \( a_{i,s} = \frac{1}{\sqrt{N}} \sum_k e^{-i\vec{k} \cdot \vec{R}_i} a_s(k) \) (and similarly for \( b_{i,s} \)) into Eq. (A.1) and obtain

\[
H = -\frac{t}{N} \sum_{\langle i,j \rangle, s} \sum_k \sum_{k'} \left( e^{i(\vec{k} - \vec{k}') \cdot \vec{R}_j} a_{i,s}^\dagger(k) b_{s}(k') + h.c. \right). \tag{A.5}
\]

Considering the lattice site \( i \), for site \( j \) to be a nearest neighbor they necessarily obey the relationship \( \vec{R}_j = \vec{R}_i + \vec{\delta}_m \) where \( m = (1, 2, 3) \), so we
can replace the nearest-neighbor sum with a sum over all unit cells, \( i \), and
all three nearest neighbors \( m \). Making these replacement to the sums in
the Hamiltonian, we get

\[
H = -\frac{t}{N} \sum_{i,m,s} \sum_{\vec{k}} \sum_{\vec{k}'} \left( e^{i\vec{R}_i \cdot (\vec{k} - \vec{k}')} e^{-i\vec{\delta}_m \cdot \vec{k}'} a^\dagger_s(k) b_s(k') + H.c. \right). \tag{A.6}
\]

Now, we use the fact that the sum on \( i \) becomes \( N\delta_{k,k'} \) and then use this
Kronecker delta to remove the sum on \( k' \) to obtain

\[
H = -t \sum_{\vec{k},s} \left[ \left( e^{-i\vec{\delta}_1 \cdot \vec{k}} + e^{-i\vec{\delta}_2 \cdot \vec{k}} + e^{-i\vec{\delta}_3 \cdot \vec{k}} \right) a^\dagger_s(k) b_s(k) + h.c. \right], \tag{A.7}
\]

where we have written out the sum on \( m \) explicitly. This can be written
in matrix form as

\[
H = \begin{pmatrix} a^\dagger_s(k) & b^\dagger_s(k) \end{pmatrix} \begin{pmatrix} 0 & -t\phi_k \\ -t\phi_k^* & 0 \end{pmatrix} \begin{pmatrix} a_s(k) \\ b_s(k) \end{pmatrix}, \tag{A.8}
\]

where we have defined

\[
\phi_k = \left( e^{-i\vec{\delta}_1 \cdot \vec{k}} + e^{-i\vec{\delta}_2 \cdot \vec{k}} + e^{-i\vec{\delta}_3 \cdot \vec{k}} \right). \tag{A.9}
\]

Eq. (A.8) is exactly Eq. (2.8).
Appendix B

Graphene Green’s Function

In this appendix we calculate the Green’s function for graphene, i.e., Eq. (3.9) in the main text.

The definition of the Green’s function can be written \[ G(\varepsilon) = (\varepsilon \mathbf{1} - H)^{-1}, \]
where \( \mathbf{1} \) denotes a unit matrix of suitable size. By having knowledge of the eigenvalues and eigenvectors of the Hamiltonian, one may construct the Green’s function satisfying Eq. (B.1) by writing the Green’s function [70]:

\[ G(\varepsilon) = \sum_{\lambda} \frac{|\psi_\lambda \rangle \langle \psi_\lambda|}{\varepsilon - E_\lambda}, \]

where \( |\psi_\lambda \rangle \) are the eigenvectors, \( E_\lambda \) are the eigenvalues and the sum on \( \lambda \) runs over all eigenvalues.

We retype now the one-valley graphene Hamiltonian from Eq. (2.13) for practical purposes:

\[ \hat{H}_K(\vec{k}) = v_F \begin{pmatrix} 0 & k_x - i k_y \\ k_x + i k_y & 0 \end{pmatrix}. \]

By performing a simple rewrite of this Hamiltonian we get

\[ \hat{H}_K(\vec{k}) = v_F \begin{pmatrix} 0 & ke^{i \phi_k} \\ ke^{-i \phi_k} & 0 \end{pmatrix}, \]

where \( k = \sqrt{k_x^2 + k_y^2} \) and \( \phi_k = \text{arg}(k_x + ik_y) \). The eigenvalues and eigenvectors of the one-valley Hamiltonian are

\[ E_\lambda = \lambda v_F k \]
\[ |\psi_\lambda \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} \lambda e^{i \phi_k} \\ 1 \end{pmatrix}, \]
where $\lambda = \pm$ denotes the valence band ($-$) and conduction band ($+$) respectively. Inserting the eigenvalues and eigenvectors in Eq. (B.2) we obtain

$$G(\vec{k}, \varepsilon) = \frac{1}{2} \sum_{\lambda = \pm} \frac{1}{\varepsilon - \lambda v_F k} \begin{pmatrix} 1 & \lambda e^{-i\phi_k} \\ \lambda e^{i\phi_k} & 1 \end{pmatrix}$$

which is exactly the Green’s function in Eq. (3.9) of the main text.
Appendix C

S-matrix Method Calculation

Here, we perform the most technical parts of the S-matrix calculation as well as write down the S-matrices for the grating region and the graphene sheet. The notation is introduced in section 4.3.1.

To write the waves in the grating region as propagating waves, we first find the eigenmodes that give us the propagating modes and decaying modes inside this region. After some matrix algebra involving Maxwell’s equations, we arrive at the eigenproblem for the field inside the grating region:

\[ \partial_z^2 \tilde{E}_l = P_{l,m} \tilde{E}_m, \]  

where the vector arrow denotes a vector of field amplitudes and

\[ P_{l,m} = T_{1,l,n} T_{n,m}^T, \]

\[ T_{1,l,n} = i \omega \delta_{n,l} - \frac{i}{\omega} \varepsilon_{n-1} k_n k_l, \]

\[ T_{2,l,n} = i \omega \varepsilon_{n-l}. \]

The solution to (C.1) can be recast in terms of the eigenmodes \( \tilde{E}_m \) as

\[ \partial_z^2 \tilde{E}_l = S_a^{-1} P S_a \tilde{E}_m, \]

where \( S_a \) is a matrix with the eigenvectors of \( P_{l,m} \) as its columns and \( D \) is a matrix with the eigenvalues of \( P_{l,m} \) on its diagonal. The relationship between the actual electric field and the eigenmodes is \( \tilde{E}_l = S_a \tilde{E}_m \) and the total eigenmode is written as

\[ \tilde{E}_m(z) = \tilde{E}_m^+ e^{i \gamma_m z} + \tilde{E}_m^- e^{-i \gamma_m z} \]

where we have defined \( \gamma_m \) as the diagonal elements of the matrix \( \gamma \) which satisfies

\[ \gamma^2 = -D. \]
By defining yet another matrix $T_a = i T_1^{-1} S_a \gamma$, we can write down the actual electric and magnetic fields inside the grating region as
\[
\begin{pmatrix}
\vec{E}_m(z) \\
\vec{H}_m(z)
\end{pmatrix}
= \begin{pmatrix} S_a & S_a \\ T_a & -T_a \end{pmatrix}
\begin{pmatrix}
\vec{E}_m^+(z) \\
\vec{E}_m^-(z)
\end{pmatrix},
\] (C.8)
where $\vec{E}_m(z)$ and $\vec{H}_m(z)$ are vectors containing the field amplitudes for the total fields in the grating region.

We are now in a position to match the field amplitudes across the boundaries between the regions in Fig. 4.3. The graphene conductivity enters as a discontinuity of the parallel component of the magnetic field, see Eq. (4.30), and note that $j_s = \sigma E_x$. First, we consider the wave matching from region 1 to region 2, i.e., across the grating region, giving us a scattering matrix
\[
S_{\text{grating}} = \begin{pmatrix}
Q e^{i \gamma h} (p - m p^{-1} m) & Q (e^{i \gamma h} m p^{-1} e^{i \gamma h} p - m) \\
Q (e^{i \gamma h} m p^{-1} e^{i \gamma h} p - m) & Q e^{i \gamma h} (p - m p^{-1} m)
\end{pmatrix},
\] (C.9)
with
\[
Q = (p - e^{i \gamma h} m p^{-1} e^{i \gamma h} m)^{-1}, \quad (C.10)
\]
\[
p = S_a^{-1} + T_a^{-1} T_0, \quad (C.11)
\]
\[
m = S_a^{-1} - T_a^{-1} T_0. \quad (C.12)
\]
The $h$-dependence comes from the propagation of fields inside the grating region and the rather complicated matrix expressions arise from the non-trivial matching of wave amplitudes into and out of the grating region. Second, we consider the S-matrix across the graphene sheet, which gives us
\[
S_{\text{graphene}} = \begin{pmatrix}
2M & -M T_0^{-1} \sigma \\
-M T_0^{-1} \sigma & 2M
\end{pmatrix},
\] (C.13)
\[
M = (2S_0 + T_0^{-1} \sigma)^{-1}. \quad (C.14)
\]
where $\sigma$ is a diagonal matrix with $\sigma(k_j, \omega)$ on the diagonal. For more detail on how to find the expressions in Eqs. (C.9) and (C.13), we refer to appendix B of Paper I.
Appendix D

Rayleigh-Ritz Variational Method for the Poisson Equation

In this appendix, we discuss the Rayleigh-Ritz variational method for the Poisson equation given by Eq. (4.48). In this method, we must first construct a functional whose minimum corresponds to the differential equation in question, in our case Eq. (4.48). The functional for our problem is given by [108]

\[ F(\tilde{V}) = \int_{0}^{L} dx \left[ \left( \frac{1}{2} \frac{d\tilde{V}}{dx} \right)^{2} - \rho \tilde{V} \right], \tag{D.1} \]

which can be verified by performing a variation \( \delta V \) on the functional around the solution \( V \). Such a variation needs to be zero for an arbitrary variation that goes to zero at the end points. In a first step, a variation of Eq. (D.1) leads to

\[ \int_{0}^{L} dx \left( \frac{dV}{dx} \frac{d\delta V}{dx} - \rho \delta V \right) = 0, \tag{D.2} \]

and integrating the first term by parts we obtain

\[ \left. \frac{dV}{dx} \delta V \right|_{x=0}^{x=L} - \int_{0}^{L} dx \delta V \left( \frac{d^{2}V}{dx^{2}} + \rho \right) = 0. \tag{D.3} \]

The first term is zero since the variation \( \delta V \) is zero at the boundary and since the equality should hold for small and arbitrary variations, we conclude that the expression in the parenthesis must be zero. Comparing the expression in the parenthesis with Eq. (4.48), we see that they are identical and we thus conclude that \( V \) must satisfy this differential equation.

We can now use the functional in Eq. (D.1) to determine the unknown coefficients in the solution ansatz, Eq. (4.51). We substitute the ansatz
into Eq. (D.1) and obtain

\[ F = \sum_{i=1}^{3} \left[ \frac{1}{2} \int_{x_i}^{x_{i+1}} dx \left( \frac{V_{i+1} - V_i}{x_{i+1} - x_i} \right)^2 \right. \]

\[ - \left. \int_{x_i}^{x_{i+1}} dx \left( V_i \frac{x + x_{i+1} - x - x_i}{x_{i+1} - x_i} + V_{i+1} \frac{x - x_i}{x_{i+1} - x_i} \right) \right]. \]

(Equation D.4)

(Equation D.5)

Evaluating the integrals, performing the sum, and simplifying we obtain

\[ F = 3 \left( V_2^2 + V_3^2 + \frac{1}{2} - V_2V_3 - V_3 \right) - \frac{1}{3} \left( V_2 + V_3 + \frac{1}{2} \right). \]

(Equation D.6)

To minimize the functional, we now take the derivatives with respect to \( V_2 \) and \( V_3 \) and equate the expressions to zero. This gives the equations:

\[ \frac{\partial F}{\partial V_2} = 6V_2 - 3V_3 - \frac{1}{3} = 0, \]

(Equation D.7)

\[ \frac{\partial F}{\partial V_3} = -3V_2 + 6V_3 - \frac{10}{3} = 0, \]

(Equation D.8)

and solving these gives us the remaining coefficients.

The approximate FEM solution is now determined by the basis function, Eq. (4.51), together with the coefficients

\[ V_1 = 0, \quad V_2 = \frac{4}{9}, \quad V_3 = \frac{7}{9}, \quad V_4 = 1. \]

(Equation D.9)

Fig. 4.5 shows this solution together with the exact solution and they can be seen to be in good agreement.
Appendix E

Longitudinal (TM) Plasmon Dispersion Equation for Conducting Sheets

In this appendix we derive the longitudinal (TM) plasmon dispersion equation around conducting sheets, i.e., Eq. (5.1) in the main text. The sheet can in general be any sheet that is thin enough (thin compared with the plasmon wavelength).

E.1 Longitudinal plasmon dispersion equation

We start by considering a conducting sheet in the $z = 0$ plane, i.e., the sheet spans the $x$ and $y$ direction and the $z$ direction is perpendicular to the conducting sheet. We will refer to the half-plane $z > 0$ as $I$ and the half-plane $z < 0$ as $II$, they are assumed to have dielectric constants $\varepsilon_1$ and $\varepsilon_2$, respectively.

We are interested in electromagnetic field configurations that are confined around the sheet, i.e., the fields should have a $e^{i(qx-\beta|z|-i\omega t)}$ behaviour, where $\beta = \sqrt{q^2 - \omega^2\varepsilon_{1,2}}$. We now need to make use of the boundary conditions for electromagnetic fields at a conducting boundary. The ones needed for our purposes are [78]

$$D_z^I - D_z^{II} = \rho, \quad (E.1)$$  
$$E_x^I - E_x^{II} = 0, \quad (E.2)$$  
$$H_y^{II} - H_y^I = j = \sigma(q, \omega)E_x^{I/II}, \quad (E.3)$$

where $\rho$ is the sheet charge density, $j$ is the sheet current, and the superscripts $I$ and $II$ denote the fields just above and below the conducting
sheet. In the last equality in (E.3), we have used Ohm’s law to express the current in terms of the electric field and the conductivity. The boundary conditions (E.1)-(E.3) tells us that the discontinuity in the perpendicular component of the displacement field is set by the charge density, the in-plane component of the electric field is continuous, and the in-plane component of the magnetic field has a discontinuity which is set by the current density.

Two of Maxwells equations from chapter 4 with no currents or charges (the current and charge mentioned previously exist only at the graphene interface) are

\[ \nabla \times \vec{E} = i\omega \vec{H} \]  
\[ \nabla \times \vec{H} = -i\omega \vec{E} \varepsilon_{1,2}. \]  

Writing these equations in their components for the field configuration for TM plasmons \((E_x, E_z, H_y)\) we obtain

\[
\begin{align*}
\partial_z E_x - \partial_x E_z &= i\omega H_y \quad (E.6) \\
\partial_x H_y &= -\omega E_z \varepsilon_{1,2} \quad (E.7) \\
\partial_z H_y &= i\omega E_x \varepsilon_{1,2}. \quad (E.8)
\end{align*}
\]

Using (E.7) we can express \(E_z\) as

\[ E_z = \frac{-qH_y}{\omega \varepsilon_{1,2}}. \]  

(E.9)

Inserting this in (E.6), we get

\[ \partial_z E_x = i \left( \omega - \frac{q^2}{\omega \varepsilon_{1,2}} \right) H_y. \]  

(E.10)

Remembering the \(e^{-\beta|z|}\)-behavior of the confined modes we have to be careful evaluating the \(z\)-derivative, and we get

\[ -\text{sign}(z)\beta E_{x}^{I/II} = \frac{i}{\omega \varepsilon_{1,2}} \left( \varepsilon_{1,2} \omega^2 - q^2 \right) H_{y}^{I/II}, \]  

(E.11)

where \(z > 0\) for region I and \(z < 0\) for region II. From this we write \(H_y\) as

\[ H_{y}^{I/II} = \text{sign}(z) \frac{i \sqrt{\varepsilon_{1,2}}}{\sqrt{\frac{q^2}{\omega^2 \varepsilon_{1,2}} - 1}} E_{x}^{I/II}, \]  

(E.12)

where we have used that \(\beta = \sqrt{q^2 - \omega^2 \varepsilon_{1,2}}\). We now insert this expression
in the boundary condition, (E.3), and rearrange the terms to obtain

\[ E_0 \left( \sqrt{\frac{\varepsilon_1}{\omega^2 \varepsilon_1}} + \sqrt{\frac{\varepsilon_2}{\omega^2 \varepsilon_2}} + i \sigma(q, \omega) \right) = 0, \]  

(E.13)

where we have relabeled \( E_x^I = E_x^II = E_0 \). To fulfill this equality, either the electric field \( E_0 \) is zero, or the terms within the parentheses is vanishing. By taking the terms inside the parentheses and setting them to zero, we obtain

\[ \frac{\sqrt{\varepsilon_1}}{\sqrt{\frac{q^2}{\omega^2 \varepsilon_1} - 1}} + \frac{\sqrt{\varepsilon_2}}{\sqrt{\frac{q^2}{\omega^2 \varepsilon_2} - 1}} + i \sigma(q, \omega) = 0 \]  

(E.14)

which is the longitudinal (TM) plasmon dispersion equation for thin conducting sheets. By using the sheet conductivity for graphene for \( \sigma(q, \omega) \), we obtain the plasmon dispersion equation for graphene. We call the mode transverse magnetic (TM) since the magnetic field (\( \hat{y} H_y \)) is transverse to the direction of propagation (\( \hat{x} q \)).

### E.2 Non-retarded limit and expansion in small plasmon losses

The plasmon dispersion equation in Eq. (E.14) is in many cases unnecessarily cumbersome to work with, (partly) due to the presence of the square-root factors in the expression. Since graphene plasmons often fulfill the condition \( q \gg \omega/c \), it is possible to expand the square roots and keep only the lowest order term. Using a Taylor expansion of the form

\[ \frac{1}{\sqrt{x^2 - 1}} = \frac{1}{x} + \frac{1}{2x^3} + ... \]  

(E.15)

and keeping only the first term, Eq. (E.14) can be written

\[ 1 + \frac{iq \sigma(q, \omega)}{\omega(\varepsilon_1 + \varepsilon_2)} = 0, \]  

(E.16)

which has a simpler form than the original equation and is often used in practice. Eq. (E.16) is often referred to as the non-retarded dispersion equation and one should remember that it is only valid in the regime \( q \gg \omega/c \). However, this is usually the case for graphene plasmons and the limitation is in practice only important when considering the limiting behavior for \( q \to 0 \).

One thing to notice with Eq. (E.16) is that it is a complex equation, which for complex \( \sigma(q, \omega) \) needs a complex wave vector to solve. The
complex wave vector, \( q = q_1 + iq_2 \), is introduced and the wavelength of the mode is given by \( \lambda = \frac{2\pi}{q_1} \) and the plasmon losses are given by \( q_2 \). Now, for small plasmon losses, \( \frac{q_2}{q_1} \ll 1 \), we can expand Eq. (E.16) to lowest order in \( q_2 \) and we get two equations

\[
1 = \frac{q_1 \sigma_2(q_1, \omega)}{\omega(\varepsilon_1 + \varepsilon_2)} \quad \text{(E.17)}
\]

\[
\frac{q_2}{q_1} = \frac{\sigma_1(q_1, \omega)}{\partial q_1(q_1 \sigma_2(q_1, \omega)))} \quad \text{(E.18)}
\]

where \( \sigma(q_1, \omega) = \sigma(q_1, \omega) + i\sigma_2(q_1, \omega) \). Inspecting Eqs. (E.17) and (E.18), we observe that the first one is the same equation as Eq. (E.16) under the assumption of no losses and it does not contain \( q_2 \). Solving this equation then lets us evaluate \( q_2 \) using the second equation; this provides information about the plasmon losses. Note that this allows us to also make a consistency check on the validity of the assumption of small losses, or \( q_2/q_1 \ll 1 \).

Having knowledge of both \( q_1 \) and \( q_2 \), (wavelength and losses), it is possible to determine the propagation length of the plasmons. One ambiguity that arises is that since the mode is exponentially decaying, we need to make a decision about what when to say that the plasmon “has decayed” and is no longer propagating. Following Ref. [4], we take the definition that the plasmon has decayed when the intensity of the oscillation has dropped by a factor \( e^{-1} \approx 0.37 \). This gives us that the propagation length \( L_p \) in units of the plasmon wavelength \( \lambda_p \) is [4]

\[
\frac{L_p}{\lambda_p} = \frac{q_1}{4\pi q_2} \quad \text{(E.19)}
\]
Bibliography


[57] H. Yan, T. Low, W. Zhu, Y. Wu, M. Freitag, X. Li, F. Guinea, P. Avouris, and F. Xia, “Damping pathways of mid-infrared plasmons


Optical signatures of nonlocal plasmons in graphene

Tobias Wenger, Giovanni Viola, and Mikael Fogelström
Department of Microtechnology and Nanoscience (MC2), Chalmers University of Technology, SE-412 96 Göteborg, Sweden

Philippe Tassin and Jari Kinaret
Department of Physics, Chalmers University of Technology, SE-412 96 Göteborg, Sweden

(Received 15 July 2016; published 14 November 2016)

We theoretically investigate under which conditions nonlocal plasmon response in monolayer graphene can be detected. To this purpose, we study optical scattering off graphene plasmon resonances coupled using a subwavelength dielectric grating. We compute the graphene conductivity using the random phase approximation (RPA) obtaining a nonlocal conductivity, and we calculate the optical scattering of the graphene-grating structure. We then compare this with the scattering amplitudes obtained if graphene is modeled by the local RPA conductivity commonly used in the literature. We find that the graphene plasmon wavelength calculated from the local model may deviate up to 20% from the more accurate nonlocal model in the small-wavelength (large-\( q \)) regime. We also find substantial differences in the scattering amplitudes obtained from the two models. However, these differences in response are pronounced only for small grating periods and low temperatures compared to the Fermi temperature.

DOI: 10.1103/PhysRevB.94.205419

I. INTRODUCTION

Monolayer graphene has attracted much attention due to its remarkable electronic and optical properties [1–5]. For instance, monolayer graphene has broadband absorption of 2.3% [6], which is quite substantial since graphene is only one atom thick. The doping level in monolayer graphene is also tunable by applying external gating [7], and it exhibits an optical response ranging from terahertz to optical frequencies [8]. The exciting properties of graphene arise from a combination of its two-dimensional nature and its hexagonal lattice structure. Together these properties make the low-energy electrons obey an effective massless Dirac equation [9], which also has consequences for the collective plasmon excitations in graphene.

Plasmons in graphene have been known for quite some time [10–13] and exhibit strong confinement of the electromagnetic fields [14]. The plasmon wavelength is much smaller than the free-space wavelength of light, for instance making it possible to achieve subwavelength resolution microscopy [15], and facilitates strong light-matter interaction [16]. Furthermore, the strong field confinement of graphene plasmons has recently been used for ultra-sensitive detection of molecules [17,18]. Other proposed applications of graphene plasmons include modulators, filters, polarizers, and photodetectors [19,20].

However, due to their small wavelength it is challenging to interact with and to detect graphene plasmons, and many different schemes have been proposed. Examples include introducing metal antennas on top of the graphene surface [21], patterning the graphene into microribbon arrays [22–24] or microdisk arrays [25,26], using total internal reflection [27], introducing a periodic spatial modulation of the graphene conductivity [28,29], and placing a nanotip close to the graphene surface [30–35]. Another approach is to pattern the substrate into a grating [36] which has been experimentally demonstrated in Refs. [37,38].

In this paper, we theoretically investigate the optical scattering of a system consisting of a subwavelength dielectric grating and a doped monolayer graphene sheet, as shown in Fig. 1. The scattering amplitudes are computed using a scattering matrix method [39], and the graphene enters our electromagnetic problem as a conducting boundary condition. We calculate the graphene conductivity using the random phase approximation (RPA), yielding a nonlocal conductivity \( \sigma(q,\omega) \). A local expression can be obtained by taking the limit \( \sigma(q \rightarrow 0,\omega) \)—this is usually called the local RPA result in literature. The combined system of graphene together with a subwavelength dielectric grating has been treated previously [40–43] using the local RPA. The local RPA result is expected to correctly describe long-wavelength plasmons (where \( q \) is small), but since much interest in plasmons arises from their small wavelengths (where \( q \) is large), it is important to also investigate nonlocal effects.

The conductivity of graphene has been the subject of much research lately and in particular the effects of disorder [44,45], phonons [14,46], and electron-electron interaction [44,47] have been investigated using various approaches. Plasmon-phonon hybridization has also been experimentally investigated in Refs. [48,49]. We assume that we are far from resonance with any phonon in our system, and we also assume that our samples are clean enough to neglect impurities. We also neglect electron-electron interaction effects in our treatment. Our focus will be on quantifying the nonlocal effects (nonzero \( q \)) by comparing with the local RPA. Previous studies of nonlocal effects in graphene include Refs. [50–52]. It was found that nonlocal effects influence the plasmon dispersion and also the plasmon width in both nanoribbons and nanodisks. We also investigate the temperature dependence of our results.

The paper is organized as follows: In Sec. II we calculate the plasmon dispersion and quantify the intrinsic plasmon width. Section III contains our calculated results for the reflectance, transmittance, and absorbance in the combined system of graphene and subwavelength dielectric grating for one specific grating. In Sec. IV we investigate the optical response for various grating periodicities and temperatures.
II. LONGITUDINAL SURFACE PLASMON MODES

In order to study the plasmons, it is convenient to calculate the plasmon dispersion, i.e., the relationship between energy and momentum of the plasmon mode. This has previously been studied at zero temperature in Refs. [12,13] and at finite temperature in Refs. [14,53,54]. Having a conductor in between two dielectrics, an equation for modes confined to the conductor can be derived from Maxwell’s equations [14]:

\[
\begin{align*}
\frac{\epsilon_1}{\sqrt{q^2 - \frac{\omega^2}{c^2}}} + \frac{\epsilon_i}{\sqrt{q^2 - \frac{\omega^2}{c^2}}} + \frac{i q}{\omega \epsilon_0} = 0, \tag{1}
\end{align*}
\]

where \( q \) is the in-plane wave vector, \( \sigma(q,\omega) \) is the sheet conductivity of graphene, and \( \epsilon_1 \) and \( \epsilon_i \) are the relative dielectric constants above and below the graphene sheet. Solving the real part of Eq. (1) for \( \omega \) as a function of \( q \), we obtain the plasmon dispersion. The nonlocal conductivity of graphene is computed using linear response theory (details are given in Appendix A).

Another convenient way to investigate intrinsic plasmon properties is to calculate the spectral function of density fluctuations [55,56]

\[
S(q,\omega) = -\frac{1}{v_q} \operatorname{Im} \left[ \frac{1}{\epsilon(q,\omega)} \right] = \frac{1}{v_q} \frac{\epsilon_2(q,\omega)}{\epsilon_1(q,\omega)^2 + \epsilon_2(q,\omega)^2}, \tag{2}
\]

where \( \epsilon(q,\omega) = \epsilon_1(q,\omega) + i \epsilon_2(q,\omega) \). The RPA expression for the dielectric function is [55,56]

\[
\epsilon(q,\omega) = 1 - v_q \Pi(q,\omega), \tag{3}
\]

where \( v_q = \frac{e^2}{\sqrt{\sigma_0(\epsilon_1 + \epsilon_i)}} \) and \( \Pi(q,\omega) \) is the polarizability. For a definition of the polarizability see Appendix A.

In order to relate equations (1) and (2), we rewrite Eq. (1) using the so-called nonretarded approximation, i.e., \( q \gg \omega/c \). Since we are interested in the plasmon behavior of strongly localized plasmons this is a valid approximation. Equation (1) then becomes

\[
1 + \frac{i q}{\epsilon(\epsilon_1 + \epsilon_i)\omega} \sigma(q,\omega) = 0 \tag{4}
\]

and from Appendix A we have

\[
\sigma(q,\omega) = \frac{i e^2 \omega}{q^2} \Pi(q,\omega). \tag{5}
\]

Inserting this in Eq. (4), we get

\[
1 - v_q \Pi(q,\omega) = 0 \tag{6}
\]

which by definition is equivalent to

\[
\epsilon(q,\omega) = 0. \tag{7}
\]

This equation is often used to determine the plasmon dispersion in the literature. Here we see it emerging as a nonretarded approximation to Eq. (1).

The width of the surface plasmons can be estimated by substituting the definition of the dielectric function into Eq. (2) for the spectral function, yielding

\[
S(q,\omega) = \frac{-\Pi_2(q,\omega)}{(1 - v_q \Pi_1(q,\omega))^2 + v_q^2 \Pi_2(q,\omega)}, \tag{8}
\]

where \( \Pi_1 (\Pi_2) \) is the real (imaginary) part of \( \Pi \). Close to the plasmon frequency \( \omega_p \), \( \epsilon_1(q,\omega) = 1 - v_q \Pi_1(q,\omega) \) may be expanded as \( \epsilon_1(q,\omega) \approx -v_q(\omega - \omega_p)\partial_{\omega} \Pi_1(q,\omega)|_{\omega=\omega_p} \), and inserting this into the spectral function we obtain

\[
S(q,\omega) = I_0(q,\omega) \frac{\gamma'(q,\omega)^2}{(\omega - \omega_p)^2 + \gamma'(q,\omega)^2}, \tag{9}
\]

where we have defined

\[
\gamma(q,\omega) = \frac{\Pi_2(q,\omega)}{\partial_{\omega} \Pi_1(q,\omega)|_{\omega=\omega_p}}, \tag{10}
\]

\[
I_0(q,\omega) = -\frac{1}{v_q \Pi_1(q,\omega)}, \tag{11}
\]

Close to the plasmon frequency, this resembles a Lorentzian with height \( I_0(q,\omega) \) and half width at half maximum (HWHM) \( \gamma(q,\omega) \). This is strictly only true if \( I_0(q,\omega) \) and \( \gamma(q,\omega) \) are constant close to the plasmon resonance.

Figure 2(a) shows the plasmon dispersion for four different temperatures obtained by solving the real part of Eq. (1) using \( \sigma(q,\omega) \) [57]. Temperature effects on the plasmon dispersion are modest at small \( T / T_F \) but shift the dispersion curve significantly at temperatures \( T / T_F \approx 1 \). We clearly observe a nonmonotonic behavior for the dispersion as a function of the temperature, which was previously discussed in Ref. [58].

In Fig. 2(a), we wish to clarify the white triangle. This area is defined by the real part of \( \sigma(q,\omega) \) [or imaginary part of \( \Pi(q,\omega) \)] being identically zero at zero temperature due to Pauli blocking of interband transitions. The spectral function then becomes a delta function, as the width of the Lorentzian goes to zero, and the plasmon mode is an infinitely sharp energy state. However, this is only true in the absence of impurities and at zero temperature. Adding impurities and/or going to nonzero temperatures \( \Pi_2(q,\omega) \) nonzero, resulting in a nonzero plasmon width. Below we investigate how the plasmon width is affected by nonzero temperatures.

Figure 2(b) shows the intrinsic plasmon width obtained from Eq. (10) for the dispersions in Fig. 2(a). We see that the zero-temperature plasmon width is indeed zero inside the white triangle in Fig. 2(a). The zero-temperature plasmon obtains a nonzero width for \( q / k_F \gtrsim 0.74 \), which is where the zero-temperature dispersion crosses from the white triangle to the gray shaded area in Fig. 2(a). The temperature can be seen to affect the high energy plasmons (large \( q / k_F \)) more than the low energy plasmons (small \( q / k_F \)). However, when the
temperature is of the order of the Fermi temperature, the low energy plasmons are affected too.

Figure 3(a) shows a comparison between plasmon dispersions obtained from the nonlocal RPA model and the local RPA model. We calculate the dispersion for zero temperature and then compute the plasmon wavelengths as a function of the energy plasmons. Having obtained the plasmon dispersions, we can determine the interband transitions. The white triangle is defined by the ratio $T/T_F = 0.0$, $T/T_F = 0.1$ (green dot-dashed line), $T/T_F = 0.3$ (orange dashed line), and $T/T_F = 1.0$ (red dotted line).

The intrinsic plasmon width of the dispersions in (a) is obtained by solving the real part of Eq. (10). In both figures 2(b) and 3(b), the ratio $T/T_F = 0.0$ (blue solid line), $T/T_F = 0.1$ (green dot-dashed line), $T/T_F = 0.3$ (orange dashed line), and $T/T_F = 1.0$ (red dotted line).

In this section, we calculate the reflectance, transmittance, and absorbance from the graphene-grating structure shown in Fig. 1. The dielectric grating has a dielectric constant of $\epsilon_r = 3$, and we take the dielectric rods to have the same width, $d/2$, as height (aspect ratio of 1). The dielectric rods are placed periodically along the $x$ axis with a distance of $d/2$ between them. The length of the effective unit cell of the periodicity is then $d$. In this section, we use $d \approx 80$ nm, which for these parameters corresponds to $q/k_F = 0.5$. Our results indicate that the aspect ratio of the dielectric rods does not play a significant role for the scattering properties. We restrict our treatment to longitudinal electric fields, meaning that the electric field is aligned as in Fig. 1, along the direction of periodicity, and the magnetic field is thus always along the grating rods. This restriction of longitudinal fields is quite natural as we want to investigate the response to longitudinal graphene plasmons. For simplicity we also restrict our treatment to normal angles of incidence.

We use the scattering matrix method explained in Appendix B, and we calculate the reflectance $R$, transmittance $T$, and absorbance $A$ defined as

\[
R = |r|^2 \quad (12)
\]
\[
T = |t|^2 \quad (13)
\]
\[
A = 1 - R - T, \quad (14)
\]

where $r$ and $t$ are the reflection and transmission amplitudes calculated in Appendix B, and the equation for the absorbance comes from energy conservation. We set the Fermi energy of the graphene sheet to $E_F = 0.1$ eV, corresponding to $n \approx 0.8 \times 10^{12}$/cm$^2$. As was shown earlier in Fig. 2(a), the temperature effect on the plasmon dispersion is determined by the ratio $T/T_F$. In this section, we investigate three different temperatures, $T/T_F = 0$, $T/T_F = 0.05$, and $T/T_F = 0.1$, which correspond to $T = 0$ K, $T = 58$ K, and $T = 116$ K, for the chosen Fermi energy. In order to compare the optical properties between $\sigma(q, \omega)$ (nonlocal RPA) and $\sigma(q, \omega)$ (local RPA), we do all calculations with identical parameters for both cases. The results of our calculations are shown in Fig. 4.

Figure 4 shows the reflectance, transmittance, and absorbance of our structure as a function of frequency. The top row shows the results for the local RPA, and the bottom row shows the results for the nonlocal RPA. The insets in Figs. 4(c) and 4(f) show the relevant plasmon dispersion together with a
FIG. 3. (a) Dispersion relations illustrating the differences between using nonlocal RPA (solid lines) and local RPA (dashed lines). The blue lines are obtained for $T/T_F = 0.0$ and the red lines for $T/T_F = 1.0$. (b) The differences between the surface plasmon wavelengths obtained from the dispersions in (a). The blue line is the relative difference between the results at $T/T_F = 0.0$, and the red line is the relative differences between the result at $T/T_F = 1.0$. (c) Relative difference in the plasmon width obtained from Eq. (10) using nonlocal RPA and local RPA. The colors represent the same colors as in Fig. 2, i.e., $T/T_F = 0.0$ (blue solid line), $T/T_F = 0.1$ (green dot-dashed line), $T/T_F = 0.3$ (orange dashed line), and $T/T_F = 1.0$ (red dotted line). For $q/k_F \lesssim 0.74$, both the zero-temperature width from local RPA and nonlocal RPA are zero and the relative difference is undefined. This is the reason for the blue curve ending abruptly.

dashed line indicating the grating-induced momentum. In the top row of Fig. 4 (local RPA results), we observe clear plasmon resonances and the peak positions are in good agreement with the calculated plasmon dispersions shown in the inset of Fig. 4(c). The plasmon response is visible in reflectance, transmittance, and absorbance with the exception of the zero-temperature absorbance which is zero everywhere. This is because the local RPA conductivity at zero temperature has a vanishing real part for all energies below $2E_F$ (remember that we have no impurities or phonons in our model). Physically the vanishing real part is due to the interband transitions being Pauli blocked at zero temperature. However, for nonzero

FIG. 4. Top row: results using $\sigma(\omega)$. Bottom row: results using $\sigma(q,\omega)$. In all figures $T = 0$ K (blue solid lines), $T = 58$ K (orange dashed lines), and $T = 116$ K (red dotted lines). (a),(d) Reflectance. (b),(e) Transmittance. (c),(f) Absorbance with an inset showing relevant plasmon dispersions together with a dashed line indicating the grating induced momentum $q = 2\pi/d$. Note that the frequency axes are different in the top row and bottom row.

205419-4
temperatures, we see that the interband transitions are allowed due to thermal smearing of the Fermi functions, and this leads to nonzero absorbance. We observe that the plasmon frequency shifts towards lower frequencies when the temperature is increased. Increasing temperatures also lead to an increased broadening of the plasmon peaks together with a decrease of the reflectance, transmittance, and absorbance peaks/dips. Even in the rather small $T/T_F = 0.1$ results (red dotted lines), the reflectance and transmittance peaks/dips have become substantially less pronounced compared with the zero-temperature results.

The bottom row of Fig. 4 shows the calculated optical properties obtained using the nonlocal RPA conductivity. Similarly to the local RPA calculation, we observe resonance peaks in both reflectance, transmittance, and absorbance. We also observe a frequency shift towards higher frequencies compared to local RPA results for the same parameters. (Note that the frequency axes are different in the top and bottom rows of Fig. 4.) We also see that the plasmon dispersion shifts to lower frequencies when raising the temperature. This corresponds well with the dispersions shown in the inset of Fig. 4(f). Comparing the top and bottom rows, we see that using nonlocal RPA predicts smaller peaks/dips for nonzero temperatures compared to local RPA. This is clearly visible by comparing for instance the reflectance shown in the orange dashed lines in Figs. 4(a) and 4(d). The difference in the reflectance for these two cases clearly illustrates the importance of taking into account the momentum dependence of the conductivity, at least in this particular case. The absorbance we obtain in the structure is rather large on resonance with the plasmon, up to 50%, and in contrast to the local RPA result the zero temperature absorbance is nonzero. The appearance of nonzero absorbance can be understood by considering the $q$ dependence of the nonlocal RPA conductivity. Due to the grating structure, any given frequency couples to many $q$ vectors, allowing the plasmon (also at zero temperature) to couple to the electron-hole continuum which gives rise to a nonzero real part of the conductivity. In our model this is encoded in the Fourier series expansion of the conductivity, similar to the way we understood the nonzero absorbance at zero temperature in the local RPA results. The sharp features in the RPA is also responsible for the overall lowering of the reflectance peaks and transmittance dips in the RPA results compared to the local RPA results.

We wish to once again point out that we have neglected impurities and phonons in our model, and we have instead focused on the momentum dependence and the temperature effects. Of course, since the temperature effect depends on the ratio $T/T_F$, an obvious way to reduce the temperature effect is to go towards larger doping levels. We also wish to point out that the $q$ dependence introduces an additional lowering of the reflectance peaks and transmittance dips and in addition broadens the resonances compared to the local RPA. Additional broadening is an effect one would normally associate with impurities and/or electron-electron interactions, and care should be taken when analyzing scattering results using the local RPA with regard to these effects.

IV. ANALYSIS OF THE PLASMON RESONANCES FOR VARIOUS GRATING PERIODICITIES

In order to further investigate the plasmon resonance signatures and their temperature dependence, we consider several different grating periodicities. For each grating periodicity, giving rise to $q = 2\pi/d$ where $d$ is the grating periodicity, we find the peak reflectance, transmittance, and absorbance given by the plasmon resonance. We also calculate the width of the plasmon resonance peaks. We calculate these quantities using both the local RPA and nonlocal RPA conductivity. This allows us to quantify the behavior of the plasmon resonances as a function of grating periodicity, and by comparing the local RPA with the nonlocal RPA results, we can quantify the difference between them. We calculate these results for the temperatures $T/T_F = 0.0$, $T/T_F = 0.05$, $T/T_F = 0.1$, and $T/T_F = 0.15$. We point out that the results in Sec. III were obtained using $q/k_F = 0.5$. The results shown in Fig. 5 are produced by solving the scattering problem as in Sec. III for a multitude of grating periodicities, and for each periodicity we find the peak reflectance, transmittance, and absorbance. Also the width of the transmittance peaks was extracted and is shown in Figs. 6(a) and 6(b).

The top row in Fig. 5 shows the peak value of the reflectance, transmittance, and absorbance for various grating periodicities obtained with the local RPA conductivity. At zero temperature the reflectance always peaks to unity and the transmittance decreases to zero, so the absorbance must be zero. The local RPA predicts that the reflectance peak becomes smaller for larger temperatures—this is in agreement with the behavior we observed in Sec. III. Also, the transmittance dip becomes smaller as the temperature increases and we see that, for any temperature, there is a particular grating period for which the absorbance is 50%. At a temperature of $T/T_F = 0.15$ only gratings with small $q/k_F$ (large grating distance $d$) will show scattering signatures from the plasmon resonances.

The bottom row in Fig. 5 shows the peak value of the reflectance, transmittance, and absorbance for various grating periodicities obtained with the nonlocal RPA conductivity. We see that at zero temperature (blue solid lines) there is a nontrivial sharp behavior of the peak heights as a function of $q/k_F$. This is in stark contrast to the smooth behavior at zero temperature in the local RPA results. The sharp features can be understood by again considering the $q$ dependence of the conductivity, similar to the way we understood the nonzero absorbance at zero temperature in Sec. III. For a given frequency the light is coupled through the grating to many different momenta, and at zero temperature the conductivity has sharp features that we are probing. These sharp features are quickly smeared out when going to nonzero temperatures; this can be seen in the green dot-dashed lines in Figs. 5(d) and 5(f) where the sharp features only remain for small $q/k_F$. In the larger temperature results the sharp features have completely vanished and we approach the local RPA results. It is worth noting that there is a rather significant absorbance in the system due to the interaction with the single particle continuum, also at small temperatures. This is the case for both the local RPA and the nonlocal RPA results. However, the location and form of this peak as a function of grating periodicity and temperature is quite different between the two models. The absorbance has
FIG. 5. Scattering coefficients along the plasmon dispersion. Top row: results using $\sigma(\omega)$. Bottom row: results using $\sigma(q,\omega)$. In all figures $T/T_F = 0$ (blue solid lines), $T/T_F = 0.05$ (green dot-dashed lines), and $T/T_F = 0.1$ (orange dashed lines) and $T/T_F = 0.15$ (red dotted lines). (a),(d) Reflectance peak height. (b),(e) Transmittance peak height. (c),(f) Absorbance peak height.

FIG. 6. (a) Width (FWHM) of the transmittance peaks used in Fig. 5(b), i.e., using local RPA. (b) Width (FWHM) of the transmittance peaks used in Fig. 5(e), i.e., using RPA. In both figures $T/T_F = 0$ (blue solid lines), $T/T_F = 0.05$ (green dot-dashed lines), $T/T_F = 0.1$ (orange dashed lines), and $T/T_F = 0.15$ (red dotted lines).

V. CONCLUSIONS

In conclusion, we have theoretically investigated the scattering properties of monolayer graphene on a subwavelength dielectric grating, and we find that such a structure, if properly tuned, can exhibit very large reflectance and transmittance signatures when the external light is on resonance with the graphene plasmon. We have compared intrinsic plasmon
properties as well as scattering results obtained by using the local RPA conductivity $\sigma(q \rightarrow 0, \omega)$ with results obtained by using the nonlocal RPA conductivity $\sigma(q, \omega)$. We find that small grating periodicities (large $q/k_F$) have the largest discrepancies in the optical properties obtained from the two results, indicating that in this regime the local RPA is no longer valid. A less obvious result we find is that the discrepancies between local RPA and nonlocal RPA are largest at small temperatures, but as the temperature is increased the discrepancies become smaller, and around $T/T_F = 0.15$ the differences have almost completely vanished. For large temperatures, $T/T_F > 0.15$, the local RPA conductivity is a good approximation that correctly captures the scattering properties of graphene on a subwavelength dielectric grating.

Furthermore, we find that the optical scattering amplitudes are heavily degraded by temperature effects, which makes grating experiments at room temperature challenging. We anticipate that the grating will induce a perturbation at multiple wavevectors and that the calculated wavelength of the plasmons may differ up to 20%. For example, for a grating periodicity $d \approx 80$ nm and doping level $E_F = 0.1$ eV or $n \approx 0.8 \times 10^{12}$ cm$^{-2}$, the difference in the calculated plasmon frequency obtained from the two models is around 2 THz. In an experiment this blueshift may be misinterpreted if the experimental results are compared with the local RPA dispersion, but it is in fact a nonlocal effect.

For some applications such as solar cells, a large absorbance is beneficial and plasmonic enhancements may help improve on current technologies [60]. We point out that graphene on a subwavelength dielectric grating structure absorbs up to 50% of the incident radiation close to resonance with the plasmon and that in some cases this absorbance approaches the theoretical maximum [25].

Our results could be used to make a refractive index sensor, and our narrow plasmon resonances should make it possible to detect very small refractive index changes in the dielectric environment surrounding the graphene and subwavelength dielectric grating. This refractive index sensor would be tunable since the plasmon resonance frequency is tunable by electrostatic gating.

ACKNOWLEDGMENTS

The authors wish to thank the Knut and Alice Wallenberg foundation for financial support. We also acknowledge fruitful discussions with Samuel Lara Avila, Timur Shegai, Hans He, Sergey Kubatkin, and Tomas Löfwander.

APPENDIX A: LINEAR RESPONSE THEORY AND CONDUCTIVITY

In this paper we treat graphene within linear response theory; the unperturbed graphene Hamiltonian is [9]

$$\hat{H}_0 = v_F \sigma \cdot \vec{k} = \begin{pmatrix} 0 & v_F(k_x - i k_y) \\ v_F(k_x + i k_y) & 0 \end{pmatrix}$$  \hspace{1cm} (A1)

with the graphene sheet confined to the $xy$ plane and $v_F$ being the Fermi velocity of graphene. We use $\hbar = 1$ in the paper. We assume that intervalley scattering is absent so both graphene valleys are independent of each other. Similarly we assume spin up and down to be independent of each other; hence spin and valley degrees of freedom are both degenerate and contribute a factor 2 each in the final answer.

The real space Dyson equation is

$$[\delta - \hat{H}_0] \circ \hat{G}_0 = \delta,$$  \hspace{1cm} (A2)

where $\circ$ means integration/summation over all internal variables. Solving Dyson’s equation for graphene we obtain the unperturbed Green’s function for one valley [1,61]

$$\hat{G}_{0}(\vec{k}, \epsilon) = \frac{1}{2} \sum_{\lambda = \pm 1} \left( \frac{1}{\lambda \epsilon + \lambda v_F k} \right) \left( 1 - i \epsilon \delta \right),$$  \hspace{1cm} (A3)

where $k = |\vec{k}|$ and $\delta_0 = \text{arg}(k_x + i k_y)$.

We now perturb the graphene Hamiltonian with an external longitudinal electric field. For a longitudinal electric field we can write the electric field in terms of a potential as $\vec{E} = i \vec{q} \phi$. We write the perturbation Hamiltonian as (notice that this perturbation is spin and valley independent)

$$\delta \hat{H}(x) = -e \sum_n \phi_n(x) e^{-i \omega t + i q x} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$= \sum_n \frac{i e}{q_n} E_n(x) e^{-i \omega t + i q x} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$  \hspace{1cm} (A4)

where we sum over many perturbation wavelengths since we anticipate that the grating will induce a perturbation at multiple $q_n$. The Dyson equation with the perturbation becomes

$$[\hat{\delta} - \hat{H}_0 - \delta \hat{H}] \circ \hat{G} = \delta.$$  \hspace{1cm} (A5)

In linear response theory the full Green’s function is linearly perturbed by the perturbation so we write the full Green’s function as $\hat{G} = \hat{G}_0 + \delta \hat{G}$. Inserting this ansatz in Eq. (A5), to zeroth and first order in the perturbation we get

$$\langle \delta - \hat{H}_0 \rangle \circ \hat{G}_0 = \delta$$  \hspace{1cm} (A6)

$$\langle \delta - \hat{H}_0 \rangle \circ \delta \hat{G} - \delta \hat{H} \circ \hat{G}_0 = 0,$$  \hspace{1cm} (A7)

where the higher order term in the perturbation is omitted since we are doing linear response theory. From these equations we immediately see that to linear order the perturbation to the Green’s function can be written

$$\delta \hat{G}(x, x') = \hat{G}_0 \circ \delta \hat{H} \circ \hat{G}_0$$  \hspace{1cm} (A8)

which means that to linear order in the perturbation the correction to the Green’s function is determined by the unperturbed Green’s function together with the perturbation. For fermions we can write the density response from the perturbation as [62]

$$\langle \delta \hat{\rho}(x, t) \rangle = -i \text{Tr} [\delta \hat{G}(x, x^+)]$$  \hspace{1cm} (A9)
and evaluating the equal argument Green’s function perturbation starting from Eq. (A8) we obtain

$$
\delta G(x,t,x',t') = \sum_n \sum_{k,e} \int d\vec{k} \frac{1}{i\omega_m} \frac{i\epsilon G_0(\vec{k} + \vec{q}_n, i\epsilon_m + i\omega_m)}{\vec{q}_n} e^{i\vec{q}_n \cdot \vec{x} - i\omega t},
$$

(A10)

where $\vec{q}_n = q_n \vec{k}$. Inserting this expression into Eq. (A9), performing the Matsubara summation over $i\epsilon_m$ [55] and doing the trace in sublattice space we obtain

$$
\langle \delta \hat{h}(x,t) \rangle = \sum_n \langle \delta \hat{n}_n(x,t) \rangle
$$

(A11)

$$
\langle \delta \hat{n}_n(x,t) \rangle = i g_s g_v \sum_{\lambda,k} \int d\vec{k} \frac{f_{\lambda,k} - f_{\lambda,k+q_n}}{i\omega_m - \lambda' v_F |\vec{k} + \vec{q}_n| + \lambda v_F k}
$$

$$
\times \frac{i}{2} (1 + \lambda \lambda' \cos(\phi_{k+q_n} - \phi_k)) \frac{\epsilon G_n}{\vec{q}_n} e^{i\vec{q}_n \cdot \vec{x} - i\omega t},
$$

(A12)

where $f_{\lambda,k}$ is the Fermi function and we have inserted $g_s$ and $g_v$ for spin and valley degeneracy. The continuity equation

$$
e \partial_x \langle n_\alpha(x,t) \rangle = - \nabla \cdot \langle j_\alpha(x,t) \rangle$$

(A13)

$$
\Rightarrow e \omega q_n \langle n_\alpha(q_n,\omega) \rangle = \langle j_\alpha(q_n,\omega) \rangle
$$

(A14)

gives us the current as

$$
\langle j(q,\omega) \rangle = \sum_n \langle j_n(q,\omega) \rangle
$$

(A15)

$$
\langle j_n(q,\omega) \rangle = \sigma(q,\omega) \delta(q - q_n) E_n
$$

(A16)

$$
\sigma(q,\omega) = \frac{i g_s g_v e \omega^2}{q^2}
$$

$$
\times \sum_{\lambda,k} \int d\vec{k} \frac{f_{\lambda,k} - f_{\lambda,k+q_n}}{i\omega_m - \lambda' v_F |\vec{k} + \vec{q}_n| + \lambda v_F k}
$$

$$
\times \frac{i}{2} (1 + \lambda \lambda' \cos(\phi_{k+q_n} - \phi_k)).
$$

(A17)

These equations tell us that for each mode $q_n$ of the perturbation there is an associated current component at the same wave number, and the total current is obtained by adding all these contributions. Performing analytic continuation on the conductivity $i\omega_m \rightarrow \omega + i\eta$ and defining the polarizability we obtain the final answer

$$
\sigma(q,\omega) = \frac{i e \omega^2}{q^2} \Pi(q,\omega)
$$

(A18)

$$
\Pi(q,\omega) = \lim_{\eta \rightarrow 0^+} \frac{g_s g_v}{2} \int d\vec{k} \sum_{\lambda,k} \frac{f_{\lambda,k} - f_{\lambda,k+\vec{k}'}}{i\omega + i\eta - \lambda' v_F |\vec{k}' + \vec{q}_n| + \lambda v_F k}
$$

$$
\times (1 + \lambda \lambda' \cos \phi_{\lambda,k+\vec{k}'}).
$$

(A19)

where $\vec{k}' = \vec{k} + \vec{q}_n$, $\Pi(q,\omega)$ is the polarizability (density-density correlation function) and $\phi_{\lambda,k} = \phi_{x+q} - \phi_k$.

The only thing left to do is to compute the polarizability integral in Eq. (A19). This was computed in Refs. [12,13] for zero temperature, and this was later generalized to finite temperatures in Ref. [58]. In the paper we refer to the result in equations (A18) and (A19) as the nonlocal RPA result or simply the RPA result.

In this paper we also use the conductivity at zero wave vector where the answer simplifies to [63,64]

$$
\sigma(\omega) = \Re [\sigma(\omega)] + i \Im [\sigma(\omega)]
$$

(A20)

$$
\Re [\sigma(\omega)] = \frac{e^2}{4} G(\omega/2)
$$

(A21)

$$
\Im [\sigma(\omega)] = \frac{e^2}{4} \frac{8T}{\omega \pi} \ln \left[ 2 \cosh \left( \frac{\mu}{2T} \right) \right]
$$

$$
+ \frac{4\omega}{\pi} \int_0^\infty \frac{G(x) - G(\omega/2)}{\omega^2 - 4\epsilon^2} dx
$$

(A22)

$$
G(x) = \frac{\sinh(x/T)}{\cosh(\mu/T) + \cosh(x/T)}
$$

(A23)

where $\mu$ is the chemical potential and $T$ is the temperature. In the paper we refer to equations (A20)–(A23) as the local RPA result or the local RPA conductivity.

**APPENDIX B: SCATTERING MATRIX METHOD**

We use a scattering matrix method [39] to investigate the optical properties of the subwavelength dielectric grating-graphene structure. This is a convenient way to solve electromagnetic scattering problems because it allows the problem to be subdivided into smaller pieces, if necessary, and then recombined to obtain the solution to the full problem. In this Appendix we use units where $\mu_0 = \epsilon_0 = c = \hbar = 1$ and $e^2 = 4\pi \alpha$, where $\alpha \approx 1/137$ is the fine structure constant.

Our structure consists of several distinct regions, see Fig. 7. The upper region, region 1, is a semi-infinite half space from which the EM radiation is incident; underneath is the periodic

![FIG. 7. Figure showing the different regions considered in the scattering problem. Region 2 is the thin air film, and the shaded areas represent the periodic dielectric function.](image-url)
subwavelength dielectric grating (subwavelength dielectric grating). Region 2 is an infinitely thin air film, and region 3 is a semi-infinite half space that fills the bottom half space into which the transmitted radiation is propagating. The graphene is here considered as an infinitely thin conducting sheet underneath the subwavelength dielectric grating and appears as a modified boundary condition when matching the waves between the thin air film and the semi-infinite half space.

The scattering matrix method is based on Fourier decomposing the electric and magnetic fields, and by matching the boundary conditions we relate the Fourier amplitudes between the different regions. The periodicity is set by the grating region; here we take the dielectric function to be a periodic step function between the values $\varepsilon_r = 1$ and $\varepsilon_r = 3$ in the $x$ direction, with total periodicity set by the distance $d$. Putting the dielectric step of width $d/2$ in the middle the Fourier series expansion of the dielectric function is

$$\varepsilon(x) = \frac{\varepsilon_0 + \varepsilon_1}{2} - \sum_{j=\text{odd}} a_j e^{ik_j x}$$

$$a_j = \frac{\varepsilon_1 - \varepsilon_0}{\pi j} \sin \left( \frac{\pi j}{2} \right)$$

$$k_j = \frac{2\pi j}{d}.$$ (B1)

We restrict our treatment to normal incidence with the incident electric field parallel to the periodic axis and we call it longitudinal. In our setup this makes the incident electric field in the $x$ direction.

In order to find expressions for free space fields we need two of Maxwell’s equations, namely

$$\nabla \times \vec{E} + \partial_t \vec{B} = 0$$

$$\nabla \times \vec{H} - \partial_t \vec{D} = 0,$$ (B4)

where the space and time dependence of the fields is not written explicitly. By applying a rotation to one of the above equations and then inserting the other in the resulting equation we can obtain the dispersion relation for a Maxwell field as

$$\omega^2 = \beta^2 + q^2,$$ (B5)

where $q$ is now in the $x, y$ direction which spans the graphene plane, and $\beta$ is in the $z$ direction perpendicular to the graphene plane. Notice that a mode is evanescent in the $z$ direction if $\omega^2 < q^2$ since this forces $\beta$ to be imaginary.

Using equations (B4) and (B5) with harmonic time dependence we obtain (remember that $\mu_0 = \varepsilon_0 = 1$ so that $\vec{E} = \vec{D}$ and $\vec{H} = \vec{B}$)

$$\nabla \times \vec{E} = i\omega \vec{H}$$

$$\nabla \times \vec{H} = -i\omega \vec{E}.$$ (B6)

Writing these equations out in their components (for $E_x$, $E_z$, and $H_y$) we get

$$\partial_t E_x - \partial_z E_z = i\omega H_y$$

$$\partial_t H_y = -i\omega E_z$$

$$\partial_t H_z = i\omega E_x.$$ (B7)

and using the second equation to eliminate $E_z$ in the first one we get

$$i\omega H_y + \frac{i}{\omega} \partial_z^2 H_y = \partial_z E_x.$$ (B8)

We now write out the electric and magnetic fields as Fourier sums

$$E_y = \sum_j \Omega_j^+ e^{ik_j x} e^{-i\beta_j z}$$

$$H_y = \sum_j \Omega_j^- e^{ik_j x} e^{+i\beta_j z},$$ (B9)

where $+$ ($-$) denotes waves propagating in the positive (negative) $z$ direction, and from Eq. (B6) $\beta_j^2 = \omega^2 - k_j^2$. (B10)

Now we plug the above Fourier expansions into Eq. (B12) and multiply by $e^{-ik_j x}$ and integrate over $x$ to project out equations for the amplitudes obtaining

$$\left( \omega - \frac{k_j^2}{\omega} \right) \Omega_j^+ = \pm \beta_j \Omega_j^-.$$ (B11)

Now, inverting the prefactor on the left and remembering that $\beta_j$ is defined by Eq. (B15) we obtain a relationship between the magnetic field amplitudes $\Omega_j$ and electric field amplitudes $\Omega_j^+$ as

$$\Omega_j^+ = \pm \frac{1}{\sqrt{1 - k_j^2/\omega^2}} \Omega_j^-.$$ (B12)

This relationship means that we can fully determine the magnetic field if we know the electric field and vice versa. We choose to work with the electric field amplitudes $\Omega_j$’s as our fundamental objects. We write

$$\Omega_j^+ = \pm T_{0,j} \Omega_j^-,$$ (B13)

where

$$T_{0,j} = \frac{1}{\sqrt{1 - k_j^2/\omega^2}}.$$ (B14)

We now introduce $\vec{E}_m = (\ldots E_{-n} \ldots E_0 \ldots E_{-n} \ldots)^T$ which is a vector containing all electric field amplitudes and similarly $\vec{H}_m = (\ldots H_{-n} \ldots H_0 \ldots H_{-n} \ldots)^T$ for the magnetic field. In the same manner we write the unknown Fourier expansion coefficients $\Omega_m^+ = (\ldots \Omega_n^+ \ldots \Omega_0^+ \ldots \Omega_{-n}^+ \ldots)^T$, and we may then write the matrix relationship

$$\left( \begin{array}{c} \vec{E}_m \\ \vec{H}_m \end{array} \right) = \left( \begin{array}{cc} S_0 & S_0 \\ T_0 & -T_0 \end{array} \right) \left( \begin{array}{c} \Omega_m^+ \\ \Omega_m^- \end{array} \right).$$ (B15)
where \( T_0 \) is a diagonal matrix with \( T_{0,j} \) on the diagonal and \( S_0 \) is the unit matrix. Reintroducing the spatial dependence in Eq. (B20) we get the expression for the fields in free space as

\[
\begin{pmatrix}
E_m(x,z) \\
H_m(x,z)
\end{pmatrix} = \begin{pmatrix}
S_0 & S_0 \\
T_0 & -T_0
\end{pmatrix} \begin{pmatrix}
\tilde{\Omega}_m^{\pm}(x,z) \\
\tilde{\Omega}_m^{\pm}(x,z)
\end{pmatrix},
\]

where

\[
\tilde{\Omega}_m^{\pm}(x,z) = (\ldots \Omega_{-m}^{\pm} e^{i(k_{-m} x \pm \beta_{-m} z)} \ldots \Omega_{+m}^{\pm} e^{i(k_{+m} x \pm \beta_{+m} z)} \ldots)^T.
\]

We point out that the electric field in equations (B20) and (B21) is the component along the \( x \) axis and the magnetic field is the field along the \( y \) axis.

### 2. Wave matching at the graphene interface

In order to compute the \( S \) matrix for the graphene sheet we must wave match across a conducting interface with the boundary conditions [65]

\[
\begin{align*}
\tilde{H}_x(0^-) - \tilde{H}_x(0^+) &= j_x = \sigma \tilde{E}_x(0^+) \\
\tilde{E}_x(0^-) - \tilde{E}_x(0^+) &= 0,
\end{align*}
\]

where \( + \) denotes the fields above the graphene sheet and \( - \) the fields below. In terms of field amplitudes, see Fig. 7, these boundary conditions become

\[
\begin{align*}
\Omega_2^- + \Omega_2^+ &= \Omega_3^- \\
\Omega_2^- - \Omega_2^+ &= (S_0 + T_0^{-1} \sigma) \Omega_3^+.
\end{align*}
\]

By adding and subtracting these two equations we get

\[
\begin{align*}
2\Omega_2^- &= (2S_0 + T_0^{-1} \sigma) \Omega_3^- \\
2\Omega_2^+ &= -T_0^{-1} \sigma \Omega_3^-
\end{align*}
\]

from which we can solve

\[
\begin{align*}
\Omega_3^- &= (S_0 + \frac{1}{2} T_0^{-1} \sigma)^{-1} \Omega_2^- \\
\Omega_3^+ &= (S_0 + \frac{1}{2} T_0^{-1} \sigma)^{-1} T_0^{-1} \sigma \Omega_2^-.
\end{align*}
\]

Also, since the graphene scattering problem has inversion symmetry around \( z = 0 \), the reflection and transmission amplitudes from the other side are identical. This gives us the final \( S \) matrix for a graphene sheet as

\[
S_{\text{graphene}} = \begin{pmatrix}
\frac{2M}{2} & -M T_0^{-1} \sigma \\
-M T_0^{-1} \sigma & \frac{2M}{2}
\end{pmatrix}
\]

where \( \sigma \) is a diagonal matrix with \( \sigma(k_j, \omega) \) on the diagonal.

### 3. Fields in the grating

Following Ref. [39] we rewrite Maxwell’s equations inside the grating into an eigenproblem for \( E_x \). We use equations (B5) and (B4) (no currents inside the grating) with harmonic time dependence giving us

\[
\nabla \times \vec{E} = i\omega \vec{H}
\]

and the difference between this and the free space expression is the periodic dielectric function \( \epsilon(x) \). We only have the components \( E_x, E_z, \) and \( H_z \), and our system is translationally invariant in the \( y \) direction so all fields are \( y \) independent. Inserting this in the above equations gives us the following three equations

\[
\begin{align*}
\partial_z E_x - \partial_t E_x &= i\omega H_y \\
\partial_x E_y - \partial_t E_y &= -i\omega \epsilon(x) E_z \\
\partial_x H_y - \partial_t H_y &= -i\omega \epsilon(x) E_x.
\end{align*}
\]

Using the second of these we get an expression for \( E_z \) as

\[
E_z = \frac{i \epsilon^{-1}(x)}{\omega} \partial_x H_y,
\]

and plugging this back into the other two we obtain

\[
\begin{align*}
\partial_z E_x &= i\omega H_y + \frac{i}{\omega} \partial_x (\epsilon^{-1}(x) \partial_t H_y) \\
\partial_x H_y &= i\omega \epsilon(x) E_x.
\end{align*}
\]

We now Fourier expand all fields, including \( \epsilon^{-1}(x) \), giving us the following expressions

\[
\begin{align*}
\sum_j \partial_z E_j e^{ik_j x} &= i\omega \sum_j H_j e^{ik_j x} \\
-\frac{i}{\omega} \sum_{j,l} \epsilon^{-1}_l E_l(k_j + k_l) e^{i(k_j + k_l)x} \\
\sum_j \partial_x H_j e^{ik_j x} &= i\omega \sum_{j,l} \epsilon_j E_l e^{i(k_j + k_l)x}
\end{align*}
\]

where the fields are expanded with \( k_j \) given by Eq. (B3). In order to project out equations for the Fourier coefficients we apply \( \int dx \ e^{-i k_j x} \) to both equations giving us

\[
\begin{align*}
\partial_x E_n &= i\omega \sum_l \delta_{n,l} H_l - \frac{i}{\omega} \sum_l \epsilon^{-1}_{n-l} \delta_{n-l} k_l H_l \\
\partial_x H_n &= i\omega \sum_l \epsilon_{n,l} E_l,
\end{align*}
\]

which can be recast into matrix form as

\[
\begin{align*}
\partial_x \tilde{E}_n &= T_n^{-1} \tilde{H}_l \\
\partial_x \tilde{H}_n &= T_n^{-2} \tilde{E}_l
\end{align*}
\]

where as before \( \tilde{E}_n = (\ldots E_{-m} \ldots E_0 \ldots E_m \ldots)^T \) and \( \tilde{H}_n = (\ldots H_{-m} \ldots H_0 \ldots H_m \ldots)^T \) are column vectors containing the
field amplitudes and

\[ T_{1,m}^{l,n} = i \omega \delta_{b,l} - \frac{i}{\omega} \epsilon_{m-1} \epsilon_{n} k_{m} k_{l} \]  
(B47)

\[ T_{2,m}^{l,n} = i \omega \epsilon_{m-1} \]  
(B48)

are matrices. Combining these equations we obtain the eigenproblem for \( \hat{E}_l \) as

\[ \hat{a}_{\gamma} \hat{E}_l = T_{1,m}^{l,n} T_{2,m}^{l,n} \hat{E}_m. \]  
(B49)

Now, to solve this we introduce

\[ \tilde{E}_m = S_{0}^{-1} \hat{E}_l \iff \hat{E}_l = S_{0}^{-1} \tilde{E}_m. \]  
(B50)

where \( S_{0} \) is a matrix which has the eigenvectors of \( P \) as its columns. \( \tilde{E}_m \) is also a vector, but we have dropped the arrow for brevity. Multiplying Eq. (B49) with \( S_{0}^{-1} \) from the left and inserting a unity on the RHS we get

\[ \hat{a}_{\gamma} \hat{E}_l = S_{0}^{-1} P S_{0} \hat{E}_m, \]  
(B51)

where \( S_{0} \) denotes the unit matrix. According to Eq. (B50) this can be written in terms of \( \tilde{E}_m \) as

\[ \hat{a}_{\gamma} \hat{E}_l = S_{0}^{-1} P S_{0} \tilde{E}_m, \]  
(B52)

and the matrix \( D \) contains the eigenvalues of \( P \) on its diagonal [66]. For convenience we define a new diagonal matrix

\[ \gamma^2 = -D, \]

where the diagonal matrix \( \gamma \) has \( \gamma_m \) on its diagonal. We can now write the solution for \( \tilde{E}_m(z) \) as

\[ \tilde{E}_m(z) = \tilde{E}_m^{+} e^{i \gamma_m z} + \tilde{E}_m^{-} e^{-i \gamma_m z}, \]  
(B53)

and according to Eq. (B50) the proper electric field is

\[ \tilde{E}_m(z) = S_{0} \tilde{E}_m^{+} e^{i \gamma_m z} + \tilde{E}_m^{-} e^{-i \gamma_m z}, \]  
(B54)

and according to Eq. (B45) the magnetic field is

\[ \tilde{H}_m(z) = i \tilde{T}_m^{-1} S_{0} \gamma_m (\tilde{E}_m^{+} e^{i \gamma_m z} - \tilde{E}_m^{-} e^{-i \gamma_m z}) \]  
(B55)

from which we define

\[ \tilde{T}_m = i \tilde{T}_m^{-1} S_{0} \gamma. \]

We now define new \( z \) dependent vectors

\[ \tilde{E}_m^{+}(z) = (\cdots \tilde{E}_m^{+} e^{i \gamma_m z} \cdots) \tilde{E}_m^{+} e^{i \gamma_m z}, \]
\[ \tilde{E}_m^{-}(z) = (\cdots \tilde{E}_m^{-} e^{-i \gamma_m z} \cdots) \tilde{E}_m^{-} e^{-i \gamma_m z}, \]

and we can finally write down the matrix expression

\[ \begin{pmatrix} \tilde{E}_m(z) \\ \tilde{H}_m(z) \end{pmatrix} = \begin{pmatrix} S_{0} & \tilde{T}_m^{-1} \gamma_m \tilde{E}_m^{+}(z) \\ \tilde{T}_m \end{pmatrix} \begin{pmatrix} \tilde{E}_m^{+}(z) \\ \tilde{E}_m^{-}(z) \end{pmatrix} \]  
(B56)

which is the solution for the fields inside the grating.

4. Wave matching in the grating

To find the \( S \) matrix of the grating we start with the matching across the plane \( z = h \), see Fig. 7, which gives

\[ \begin{pmatrix} S_{0} & S_{0} \tilde{\Omega}_{1}^{+} \\ T_{0} & -T_{0} \end{pmatrix} \begin{pmatrix} \tilde{\Omega}_{1}^{+} \\ \tilde{\Omega}_{2}^{+} \end{pmatrix} = \begin{pmatrix} S_{0} & S_{0} \tilde{\Omega}_{1}^{-} \\ T_{0} & -T_{0} \end{pmatrix} \begin{pmatrix} \tilde{\Omega}_{1}^{-} \\ \tilde{\Omega}_{2}^{-} \end{pmatrix}, \]  
(B57)

and we also know how to propagate the field inside the grating from \( z = h \) to \( z = 0 \), namely

\[ \begin{pmatrix} \tilde{E}_{0}^{+} \\ \tilde{E}_{0}^{-} \end{pmatrix} = \begin{pmatrix} e^{i \gamma h} & 0 \\ 0 & e^{-i \gamma h} \end{pmatrix} \begin{pmatrix} \tilde{E}^{+}(0) \\ \tilde{E}^{-}(0) \end{pmatrix}. \]  
(B58)

The matching across \( z = 0 \) into the thin air film is done similarly

\[ \begin{pmatrix} S_{0} & S_{0} \tilde{\Omega}_{1}^{+} \\ T_{0} & -T_{0} \end{pmatrix} \begin{pmatrix} \tilde{\Omega}_{1}^{+} \\ \tilde{\Omega}_{2}^{+} \end{pmatrix} = \begin{pmatrix} S_{0} & S_{0} \tilde{\Omega}_{1}^{-} \\ T_{0} & -T_{0} \end{pmatrix} \begin{pmatrix} \tilde{\Omega}_{1}^{-} \\ \tilde{\Omega}_{2}^{-} \end{pmatrix}. \]  
(B59)

Now we insert Eq. (B58) into (B57) and in the resulting equation we plug the expression for \( (\tilde{E}^{+}(0), \tilde{E}^{-}(0))^{T} \) obtained from Eq. (B59). Then we make sure to remove the numerically unstable behavior from the resulting equation by multiplying so that we only have \( e^{i \gamma h} \) (and not \( e^{-i \gamma h} \)). The result is

\[ \begin{pmatrix} S_{0}^{-1} + T_{0}^{-1} T_{0} \\ e^{i \gamma h} (S_{0}^{-1} + T_{0}^{-1} T_{0}) \end{pmatrix} \begin{pmatrix} S_{0}^{-1} - T_{0}^{-1} T_{0} \\ e^{i \gamma h} (S_{0}^{-1} - T_{0}^{-1} T_{0}) \end{pmatrix} = \begin{pmatrix} \tilde{\Omega}_{1}^{+} \\ \tilde{\Omega}_{2}^{+} \end{pmatrix} \begin{pmatrix} \tilde{\Omega}_{1}^{-} \\ \tilde{\Omega}_{2}^{-} \end{pmatrix}. \]  
(B60)

and defining new variables we write this as

\[ \begin{pmatrix} p & m \\ e^{i \gamma h} m & e^{i \gamma h} p \end{pmatrix} \begin{pmatrix} \tilde{\Omega}_{1}^{+} \\ \tilde{\Omega}_{2}^{+} \end{pmatrix} = \begin{pmatrix} e^{i \gamma h} p & e^{i \gamma h} m \\ m & p \end{pmatrix} \begin{pmatrix} \tilde{\Omega}_{1}^{-} \\ \tilde{\Omega}_{2}^{-} \end{pmatrix}. \]  
(B61)

where

\[ p = S_{0}^{-1} + T_{0}^{-1} T_{0} \]  
(B62)

\[ m = S_{0}^{-1} - T_{0}^{-1} T_{0} \]  
(B63)

We can now rewrite this on \( S \)-matrix form

\[ \begin{pmatrix} \tilde{\Omega}_{1}^{+} \\ \tilde{\Omega}_{2}^{+} \end{pmatrix} = S_{\text{grating}} \begin{pmatrix} \tilde{\Omega}_{1}^{-} \\ \tilde{\Omega}_{2}^{-} \end{pmatrix}, \]  
(B64)

where

\[ S_{\text{grating}} = \begin{pmatrix} Q e^{i \gamma h} (p - mp^{-1} m) & Q e^{i \gamma h} m p^{-1} \left(e^{i \gamma h} m p - m\right) \\ Q (e^{i \gamma h} m p^{-1} e^{i \gamma h} m p - m) & Q e^{i \gamma h} (p - mp^{-1} m) \end{pmatrix} \]  
(B65)

with

\[ Q = (p - e^{i \gamma h} m p^{-1} e^{i \gamma h} m)^{-1}. \]  
(B66)

5. Combining two \( S \) matrices

Finally we now seek the total \( S \) matrix of both the grating structure and the graphene interface, i.e., the \( S \) matrix obtained
by combining equations (B31) and (B65). Writing

\[
S_{\text{grating}} = \begin{pmatrix} t_1 & r_1 \\ r_1 & t_1 \end{pmatrix} \quad (B67)
\]

\[
S_{\text{graphene}} = \begin{pmatrix} t_2 & r_2 \\ r_2 & t_2 \end{pmatrix} \quad (B68)
\]

and remembering how they are defined, see Fig. 7, we can write down the relationships

\[
\begin{pmatrix} \Omega^2_2 \\ \Omega^2_1 \end{pmatrix} = S_{\text{grating}} \begin{pmatrix} \Omega^2_1 \\ \Omega^2_2 \end{pmatrix} \quad (B69)
\]

\[
\begin{pmatrix} \Omega^2_3 \\ \Omega^2_2 \end{pmatrix} = S_{\text{graphene}} \begin{pmatrix} \Omega^2_2 \\ \Omega^2_3 \end{pmatrix} \quad (B70)
\]

and the \(S\)-matrix we seek is the total \(S\)-matrix that relates

\[
\begin{pmatrix} \Omega^2_3 \\ \Omega^2_1 \end{pmatrix} = S \begin{pmatrix} \Omega^2_1 \\ \Omega^2_3 \end{pmatrix} \quad (B71)
\]

which means we must eliminate \(\Omega^2_2\). Eliminating both \(67\) we obtain the total \(S\) matrix as

\[
S_{\text{tot}} = \begin{pmatrix} t_2q_2t_1 & r_2 + t_2r_1q_2t_2 \\ r_1 + t_1r_2q_1t_1 & t_1q_2t_2 \end{pmatrix} \quad (B72)
\]

with

\[
q_1 = (s_0 - r_1r_2)^{-1} \quad (B73)
\]

\[
q_2 = (s_0 - r_2r_1)^{-1}. \quad (B74)
\]

Equations (B72)–(B74) represent the final answer in our scattering matrix method. The answer is a block matrix where each block is \(N \times N\), \(N\) being the number of Fourier modes used in the expansion of the fields and the grating dielectric function. Since the \(S\) matrix relates incoming fields with outgoing fields it is now simple to multiply the \(S\) matrix with the incoming field amplitudes and read out the reflected and transmitted fields. We point out that in this paper we work in a regime such that there is only one propagating mode in the reflected field and one propagating mode in the transmitted field. All other modes are evanescent and as such they do not carry any energy away from the structure. To investigate the power flow in the direction perpendicular to the graphene, the \(z\) direction, it is thus only necessary to determine one field component amplitude for the reflected field and one for the transmitted field. This is the reflection amplitude \(r\) and the transmission amplitude \(t\) used in the paper to calculate the reflectance, transmittance, and absorbance.
It is true in general that $A^+ X A = Y$ where $Y$ contains the eigenvalues of $X$ on the diagonal if $A$ is constructed as the matrix that has the eigenvectors of $X$ as its columns.

Since all entities in the equation are matrices, the order plays no role in the algebraic expression obtained. The expressions are of course equivalent, however this might not be obvious at first glance. The answer we give represents the one the authors believe to be most compact.
High-sensitivity plasmonic refractive index sensing using graphene

Tobias Wenger\textsuperscript{1}, Giovanni Viola\textsuperscript{1}, Jari Kinaret\textsuperscript{2}, Mikael Fogelström\textsuperscript{1} and Philippe Tassin\textsuperscript{2}

\textsuperscript{1} Department of Microtechnology and Nanoscience (MC2), Chalmers University of Technology, SE-412 96 Göteborg, Sweden
\textsuperscript{2} Department of Physics, Chalmers University of Technology, SE-412 96 Göteborg, Sweden

E-mail: wenger@chalmers.se

Abstract

We theoretically demonstrate a high-sensitivity graphene-plasmon-based refractive index sensor working in the mid-infrared at room temperature. The bulk figure of merit of our sensor reaches values above 10, but the key aspect of our proposed plasmonic sensor is its surface sensitivity, which we examine in detail. We have used realistic values regarding doping level and electron relaxation time, which is the limiting factor for the sensor performance. Our results show quantitatively the high performance of graphene-plasmon-based refractive index sensors working in the mid-infrared.

1. Introduction

Surface plasmons are collective charge density oscillations on conducting surfaces, which have been used for sensing purposes over the last two decades [1, 2]. Recently, graphene has emerged as a new plasmonic material, active in the terahertz to mid-infrared part of the spectrum [3]. Graphene plasmons combine low loss and strong field confinement with a unique external tunability [4, 5]. This makes them attractive for applications such as modulators [3] and photodetectors [6], but also for chemical sensors and biosensors [7]. In contrast to other plasmonic materials, such as silver and gold, the ability to gate the graphene sample offers the possibility to create plasmonic devices that selectively probe for different molecules, as well as obtaining broadband spectroscopic fingerprints from biomolecules. Surface plasmon resonances help achieve label-free, high-throughput detection and screening of biomolecules for drug discovery, genomics, bioengineering, and environmental monitoring [1, 2] and such systems are expected to have a large impact in the future [8].

Previously, graphene nanoribbons [9–12] and graphene disks [13] have been studied for biosensing and chemical sensing purposes. The common feature in these works is the use of graphene plasmons and their field localizing properties to enhance the sensing signal. Li–Farmer–Rodrigo [9–11] studied vibrational modes of molecules and it was found that coupling them to graphene plasmons enhances the signal by a factor of 5 [9] and that a graphene-plasmon-based sensor can detect minute amounts of gas, down to levels of 50 zeptomol/\(\mu\)m\(^2\) [10]. In addition, [11] demonstrated a gate tunable sensor capable of extracting the permittivity of a molecular layer on the graphene sensor, thereby enabling selective sensing of biomolecules.

In this article, we study theoretically the graphene plasmon frequency shift as a tool to detect small changes in the refractive index surrounding the graphene sheet. To investigate the frequency shift of the plasmons, we use the optical reflectance and/or transmittance of graphene on a subwavelength dielectric grating (see figure 1). The grating is needed to overcome the large momentum mismatch between the incident light and the graphene plasmons. Previous studies have used a local random phase approximation (RPA) to treat the graphene conductivity. However, a nonlocal approach is known to better approximate the plasmon dispersion as well as incorporate losses in light scattering more correctly [14]. In this article we use a nonlocal graphene conductivity obtained from RPA [14–17]. We also quantify the sensing figure of merit (FoM) of graphene plasmons and examine its dependence on the thickness of the material to be sensed. This is important for sensing small amounts of biomolecules and chemicals on the sensor surface.

2. Graphene plasmons and damping

To use graphene plasmons, we first determine their frequency-wavelength relationship, i.e. the plasmon dispersion relation. This information is used to make the grating periodicity match a suitable plasmon frequency for the sensor setup. The graphene plasmon dispersion can be found by solving the non-retarded, \(q \gg \omega/c\), dispersion equation expanded to first order in \(qz\) [14, 17]:

\[
\epsilon_g(q, \omega) = \epsilon_g(\omega) - \omega^2 / \epsilon_g(\omega) c^2 - q^2 / \epsilon_g(\omega) c^2 \approx \omega^2 / \epsilon_g(\omega) c^2 - q^2 / \epsilon_g(\omega) c^2
\]
σω, making the grating periodicity
is the complex wave vector.
H, sometimes called the inverse
damping, i.e. the
q = q1 = q2, the
electron relaxation
creating the opti-
and the results in literature for the relaxation time
varies between tens of femtoseconds to above one
dielectric constant and vacuum. The resulting plasmon
Equations (1) and (2) include the lowest-order estimate
the graphene sheet and
Since the grating fills half of the volume below the
graphene, ε2 is taken as the average of the grating
dielectric constant and vacuum. The resulting plasmon
resonance frequencies from this equation are in good
agreement with our numerical calculations.
The ratio q1/q2, sometimes called the inverse
damping ratio, is a measure of how many oscillations the
plasmon makes before it is damped out completely.
For sensing purposes, this number should be as large as
possible since it indicates well defined plasmon
resonances with narrow line widths. Plasmons in
graphene are damped by several sources, such as
impurities, phonons, and finite temperature, which
adversely affect the performance of the sensor. In this
article, we include losses through the relaxation-time
approximation [17, 18], in which it is assumed that
the combined effect of all electron losses (phonons,
impurities, etc) contribute to a total electron relaxation
time τ. The total τ can be estimated from either DC
transport measurements or terahertz measurements,
and the results in literature for the relaxation time
varies between tens of femtoseconds to above one
picosecond [19]. For concreteness, we set τ = 1 ps,
which can be achieved by current annealing [20].
Figure 2 shows the inverse plasmon damping ratio
for several different scattering times versus q1/k0, the
inverse grating periodicity normalized by the Fermi
wave vector. All curves are calculated at room temper-
and for a carrier density of n0 = 8 · 10^{11} cm^{-2}
(E_F = 0.33 eV). A large inverse damping ratio indicates
that the plasmon is a well defined resonance. It is clear
from the figure that there is an optimum in terms of
damping that occurs around q1/k0 ≈ 0.15. This means
that for these specific parameters there is an optimal
grating periodicity to obtain low plasmon damping.
The optimum occurs as the temperature damps the
plasmons with large q1/k0 and the electron relaxation
damps the plasmons with small q1/k0, creating the opti-
um at an intermediate value of q1/k0.
There is also an intrinsic graphene phonon at an
energy of 0.2 eV, or approximately 50 THz. For larger
energies, the plasmon becomes severely damped due to
the phonon coupling [3, 17]. This additional damping
could in principle be added to the relaxation time, but
we choose instead to restrict ourselves to energies below
the phonon energy. We do this as we anticipate that long
relaxation times are essential for the performance of
the sensor.
Keeping to energies below 50 THz, and with
the knowledge of the optimal damping, we set
q1/k0 = 0.1, making the grating periodicity d = 130 nm
and the plasmon frequency approximately 40 THz
(λ ≈ 7.5 µm). The width of one grating is d/2 and the
empty space between the gratings also has a width d/2.
The grating height is set to be H = 130 nm and the
grating material has a dielectric constant ε3 = 3, i.e. the
grating is assumed to be lossless in the frequency range
we consider. These grating parameters remain fixed
throughout the article.

3. Refractive index sensing

Figure 1 shows the proposed sensing structure, where
graphene is on top of a subwavelength dielectric
grating. The substance layer on top of the graphene has
a thickness $t$, which is initially set to $t = H = 130$ nm.

To assess the potential of this graphene structure as a refractive index sensor, we calculate its optical scattering properties under normal incidence. This is calculated with a finite-element method solver (COMSOL), in which calculations are performed using one unit cell of the grating with periodic boundary conditions. The top and bottom of the computational domain have absorbing boundary conditions. Graphene enters into the solver as a conducting boundary condition which carries a current induced by the external electric field. The nonlocal real space graphene conductivity $\sigma_{xx}$ is computed by performing a discrete Fourier transform of the nonlocal momentum space graphene conductivity as $\sigma_{\omega} = \sum_{q} \sigma(q_{x}, \omega) e^{i q_{x} x}$ and $q_{x} = 2\pi j/d$. The graphene current is obtained by $j(x) = \int \sigma(x - x', \omega) E(x') dx'$, where the integration is over graphene in the unit cell. Note that it is the grating periodicity $d$ that enables the incoming light to overcome the large momentum mismatch and probe plasmons at large $q$. The polarization of the incident light is set to be parallel with the induced momentum from the grating periodicity, $q_{x}E$, in order to probe longitudinal plasmons. The nonlocal momentum space conductivity is calculated using linear response theory [14] and a number conserving relaxation-time approximation is used to include the damping [17, 18].

Figure 3 shows the calculated optical reflectance and transmittance from the graphene structure for different values of the refractive index of the substance on top of the graphene surface. It is clear that the scattering peaks in both reflectance and transmittance are sensitive to the refractive index change and the more the value of the refractive index increases the larger the wavelength becomes. Note the different scales on the y-axes in figure 3, the height of the reflectance peak is roughly 3% and the dip of the transmittance is approximately ten times larger. The results in figure 3 are calculated using a substance layer thickness $t = 130$ nm (see figure 1).

Figure 4 shows the magnitude of the $x$-component of the electric field on resonance (three unit cells are shown). The outlines of the grating can be seen below the graphene and the majority of the electric field is located outside the grating material.

To quantify the performance of the proposed refractive index sensor, we compute a sensing FoM. Note that the (free-space) wavelength shift of the resonance is linear in the refractive index change, i.e.

$$\Delta \lambda = m \Delta n,$$

where $m$ denotes the resonance wavelength shift per refractive index unit (RIU) change and has units of meters per RIU. $m$ can be obtained from equation (3)

$$m = \frac{\partial \Delta \lambda}{\partial n}.$$

A commonly used FoM is the so called bulk FoM [21, 22]

$$\text{FoM}_{\text{bulk}} = \frac{m}{\Gamma_{W}},$$

where $\Gamma_{W}$ denotes the full width at half maximum bandwidth of the resonance. Figure 5 shows the extracted peak positions from the transmittance resonances in figure 3 versus the refractive index of the substance layer. By fitting a line through the data points, the slope $\partial \Delta \lambda / \partial n = 2.5 \mu$m/RIU is obtained and the resonance widths can be extracted from figure 3. Substituting these values in equation (4), we obtain a bulk FoM of 10.7 for the sensor.

Comparing with literature, this is a rather large number, meaning that our graphene-plasmon based sensor exhibits a competitive bulk sensing performance. However, bulk FoMs are not always the relevant quantity for the actual merit of sensing applications...
as pointed out in [23, 24]. The reason is that the bulk FoM is calculated by changing the refractive index of a big volume surrounding the sensor. In biosensing, it is often more important to be sensitive to local refractive index changes close to the surface of the sensor. If the localization of the modes is poor, i.e. if the resonant modes used to probe the index change extend far away from the surface, the bulk FoM can overestimate the performance of the sensor, since it needs a thick layer of refractive index material in order to be as sensitive as the bulk FoM suggests. These considerations make it interesting to quantify the sensor performance for a varying thickness of the substance layer. To this end, the authors in [23, 24] propose a new way to quantify refractive index sensing using the formula

\[ \Delta \lambda = m \Delta n (1 - e^{-2t/L_d}) \]  

(5)

where \( t \) is the layer thickness of the substance layer, \( L_d \) is the decay length, and \( m \) is the sensitivity. This formula explicitly includes the layer thickness dependence of the spectral shifts. We point out that this is nothing but a redefinition of the sensitivity \( m \), the spectral shifts of the resonances, \( \Delta \lambda \), are of course the same. Note that in the limit of layer thickness \( t \to \infty \), we regain the bulk expression for the sensitivity \( m \), i.e. equation (3). Differentiating equation (5) with respect to \( n \) we obtain an equation for the spectral shift per RIU:

\[ \frac{\partial \lambda}{\partial n} = m(1 - e^{-2t/L_d}) \]  

(6)

meaning that the spectral shift per RIU is decreasing with decreasing substance layer thickness \( t \).

Figure 6 shows the transmittance resonance position as a function of refractive index for different substance layer thicknesses \( t \). As expected from equation (5), the shift gets smaller as the substance layer thickness is reduced, i.e. the slope of the fitting curves is reduced.

Figure 7 (dots) shows the spectral shifts per RIU obtained from the lines in figure 6. Figure 7 also shows a fit of this data to equation (6) (solid line), obtaining \( m = 2480 \text{ nm/RIU} \) and \( L_d = 21 \text{ nm} \). It is clear that the spectral shifts are reduced as the thickness \( t \) is reduced, the fit also allows us to extract information about the shifts for very small \( t \). Note the axis on the right hand side of figure 7, where the spectral shift is divided by the resonance width, thereby obtaining a thickness dependent FoM. For thicknesses above \( 2L_d \) (42 nm), the FoM reaches the bulk value and below \( L_d \) (21 nm) the FoM drops to zero. Note that even for thicknesses of a few nanometers the FoM is 2–3, meaning that the sensor is capable of sensing very thin substance layers on the sensor surface.

The authors in [23, 24] propose to use a second order derivative of the spectral shift described by equation (5), giving

\[ \frac{\partial^2 \lambda}{\partial n \partial \Delta n} = \frac{2m}{L_d} e^{-2t/L_d} \]  

(7)

which allows the limit \( t \to 0 \) to be taken to extract information about the surface sensitivity defined by the prefactor \( \frac{2m}{L_d} \). From the fit in figure 7, \( m \) and \( L_d \) are extracted to obtain a surface sensitivity prefactor of \( 2m/L_d = 237 \). This is a dimensionless number that can be used to compare different sensors. We note that in [23, 24], surface sensitivity prefactors below 10 are reported. This highlights the ability of graphene plasmons to strongly localize electromagnetic fields, creating a high sensitivity at the surface.

4. Discussion

We have demonstrated that graphene allows for refractive index sensors with better performance than traditional refractive index sensors. We wish to point out that graphene losses limit the sensing FoM. By increasing the graphene quality, it is possible to reach even higher FoMs than reported here.

Furthermore, temperature plays a role in limiting the FoM. We point out that the sensor operates at room temperature and that all calculations are performed at this temperature. Larger temperatures will reduce the FoM of the sensor slightly as it will increase the width of the plasmon resonance. However, temperature effects are exponentially suppressed in the Fermi distribution, so we do not expect a significant change for moderate
temperature increases. Conversely, reducing the temperature decreases the width of the plasmon resonance and thus increases the sensor performance. In addition, decreasing the temperature tends to increase the mobility of graphene, this can be understood by considering that lowering the temperature limits the phonon relaxation scattering. This leads to a longer electron relaxation time \( \tau \), which as previously stated is a main source of plasmon damping in our model. We find that by increasing to a more optimistic estimate for the relaxation time of \( \tau = 3 \) ps (throughout the article we have used \( \tau = 1 \) ps) we obtain a bulk FoM above 20. In the limit of pristine graphene (i.e. an infinite scattering time, but still at room temperature and nonlocal effects included), we obtain a bulk FoM of roughly 40, which provides a best-case scenario for the FoM, showing the potential of ultra-sensitive graphene-plasmon based refractive index sensors. Note that the temperature and the relaxation time both enter the equations as fractions of the Fermi energy. Thus, by increasing the Fermi energy and keeping the absolute temperature and relaxation time fixed, their effects on the sensitivity become smaller.

By incorporating a metal gate, the Fermi energy of graphene can be tuned and thus change the resonance frequency of the plasmon. Using the fact that both temperature and damping induced by electron relaxation enters the conductivity as fractions of the Fermi energy and that long-wavelength plasmons have \( \omega_p \propto \sqrt{E_p} \), we estimate that the sensor retains a bulk FoM above 10 for frequencies down to 38 THz (7.9 \( \mu \)m). As previously discussed, the working frequency of the sensor is constrained from above by an intrinsic graphene phonon at 50 THz not included in our model. We work well below this in our calculations, which are performed for approximately 40 THz, but it is not unreasonable to relax this restriction slightly and allow the working frequency to be as high as 42 THz (7.14 \( \mu \)m). This would give a working range of the sensor between 38–42 THz (7.9–7.14 \( \mu \)m), where the bulk FoM is above 10. A larger tuning range towards lower frequencies is possible by sacrificing the FoM.

The results presented in section 3 were all obtained using normal incidence on the graphene-grating structure. However, in experiments using, for example, a Gaussian beam, there will be a small spread of angles of incidence, which in principle could affect the sensing performance. Also, for practical reasons it could be useful to know the importance of alignment in experiments. To investigate this, we have performed calculations of the bulk FoM for angles of incidence up to 15\(^\circ\) from normal incidence. The results are shown in figure 8. The bulk FoM is essentially flat, meaning that for moderate angles of incidence the sensor performance remains high. We have also investigated the sensitivity to offsets from the \( x \)-direction in the linear polarization and we find that for an offset of a few degrees the impact is negligible. For polarization offsets up to 10\(^\circ\) the bulk FoM remains well above 10. We conclude that the sensor is quite robust to deviations in polarization angle and angle of incidence.

Even though the bulk FoM of our graphene-plasmon-based sensor is large, it is not the largest achieved in the literature [25, 26]. However, we emphasize that the main strength of our plasmon sensor is its ability to sense refractive index changes that occur very close to the surface. This is due to the strong electromagnetic field localization facilitated by the plasmons and it is the reason for the enhanced sensing properties of graphene-based devices in, for instance, [9–11, 13]. In this article, we have taken a more general view and calculated the quantitative performance of graphene-plasmon-based refractive-index sensing devices.

Using phase sensitive measurements, it is possible to enhance the FoM [8, 27]. However, this requires a more complicated setup to perform phase measurements of the reflected and/or transmitted light. We note that the phase of the reflected light in our calculations has a pronounced step as the frequency sweeps across the plasmon resonance and the shape of this step is quite robust against losses. Such steps in phase are the basis for the FoM enhancements in [8, 27]. This makes the prospect of phase sensitive measurements together with graphene plasmons very interesting for future investigation.
As a final comment, we emphasize that the square subwavelength dielectric grating was chosen for its simplicity. It is very possible, even likely, that further optimization of this geometry will lead to even better performance of graphene-plasmon-based sensors. An interesting possibility is to combine the dielectric structures that we have investigated with metallic nanoparticles used in traditional refractive index sensors.

5. Conclusions

We have theoretically demonstrated a graphene-plasmon based refractive index sensor working in the mid-infrared, exhibiting a large bulk FoM close to 11. We have thoroughly investigated the performance of the sensor by performing calculations for a reduced thickness of the substance layer. We show that even for a substance layer as thin as a few nanometers the sensor has a FoM of 2–3, making it possible to sense very small refractive index changes close to the sensor surface. This sensitivity comes from the ability of the graphene plasmons to localize the electromagnetic fields very close to the graphene surface. For distances above roughly 40 nm the sensor achieves its bulk FoM, meaning that the electromagnetic fields are well localized within this region. The performance of the sensor is limited by electron relaxation and finite temperature. Reducing either the temperature or the electron scattering, the performance of the sensor may be further improved. We believe that our quantitative findings support the high potential of graphene-plasmon-based refractive index sensors in the mid-infrared.

Acknowledgments

Financial support from the Knut and Alice Wallenberg Foundation (KAW), the Swedish Foundation for Strategic Research (SSF), and the Swedish Research Council (VR) is gratefully acknowledged. Part of the work was supported by VR Grants No. 2014-5275 and 2016-03603.

References

[29] Low T and Avouris P 2014 Graphene plasmonics for terahertz to mid-infrared applications ACS Nano 8 1086–101
Graphene plasmons in the presence of adatoms

Giovanni Viola\textsuperscript{1}, Tobias Wenger\textsuperscript{1}, Jari Kinaret\textsuperscript{2} and Mikael Fogelström\textsuperscript{1,3}

\textsuperscript{1} Department of Microtechnology and Nanoscience (MC2), Chalmers University of Technology, SE-412 96 Göteborg, Sweden
\textsuperscript{2} Department of Physics, Chalmers University of Technology, SE-412 96 Göteborg, Sweden
\textsuperscript{3} Author to whom any correspondence should be addressed.

E-mail: mikael.fogelstrom@chalmers.se

Keywords: graphene plasmons, nonlocal response, graphene impurities, transport in graphene, plasmonic sensing

Abstract

We theoretically investigate graphene plasmons in the presence of a low density of adatoms on the graphene surface. The adatoms can significantly modify the conductivity and plasmonic properties of graphene and may produce a level splitting with the plasmon mode, resulting in two plasmon branches. The high energy branch exhibits large losses and the low energy branch exhibits low losses. Our model may also be considered as a simple model for molecules on graphene and we show that graphene plasmons are sensitive to such changes in the environment. Our microscopic treatment of plasmons and adatoms shows the sensitivity of plasmons and highlights the potential of graphene plasmons for sensing purposes.

1. Introduction

Graphene has recently emerged as an attractive plasmonic material at terahertz and mid-infrared frequencies [1–3]. Among the benefits of graphene as a plasmonic material are its tunable optical properties [4], low losses [5], and large confinement of electromagnetic fields under the right conditions. Field localization by a factor of up to 200 has been predicted [6] which facilitates strong light–matter interactions [7]. Considering that large field localization leads to large plasmon losses, a more conservative estimate of the field localization is \(1/\alpha\), where \(\alpha \approx 1/137\) is the fine-structure constant [6, 8].

Recently, graphene plasmons have been studied using nanotips [9, 10], subwavelength gratings [11–13], metal nanoantennas [14, 15] and nanoribbons [16]. Possible applications include label-free molecular sensing [17, 18], photonic modulators [14, 19], as well as ultrafast photodetectors [20], showcasing the versatility of graphene as a plasmonic material.

However, the quality of graphene still limits many proposed applications [21], and high quality graphene devices are labor-intensive to fabricate. Even the cleanest graphene samples exhibit some momentum relaxation [22], and thus a theoretical analysis of various loss mechanisms is of much interest [23–28]. The high frequency relaxation is of fundamental importance since the plasmons of interest are in this regime. Graphene conductivity in a relaxation time approximation [29, 30] was investigated by Rana [31] and by Jablan \textit{et al} [5] who found substantial plasmon losses for realistic relaxation times. This was also found in the experiments by Chen \textit{et al} [9] and by Fei \textit{et al} [10].

Graphene has previously been considered for sensing purposes, see for instance [32, 33], and the large surface-to-volume ratio is one of the main advantages of graphene in this regard. Chemical sensing has been explored in the mid-infrared part of the spectrum where plasmons have been exploited to detect changes in refractive indices [34, 35] and vibrational states in biomolecules [18, 36]. These applications show much promise for plasmonic-based sensing in the future. However, also electronic transitions in atoms and molecules can couple to graphene plasmons but have so far not been analyzed in detail. Electronic properties of molecules and atoms that adsorb on the graphene surface have been studied extensively using various computational methods [37–42] and different substances have different coupling strengths and energy of the electronic levels. This variability between substances makes it possible to consider graphene plasmonic-based sensors which have the ability to selectively detect various compounds. We use these previous results to investigate adatom effects on the
plasmonic properties of graphene. Our approach is complementary to previous works on plasmonic-based sensing in graphene as we study the plasmon response to microscopic degrees of freedom, rather than to changes in the dielectric environment [35] or to vibrational modes [18, 36, 43].

In this article we develop and analyze a model for uncorrelated adatoms, coupled to the graphene surface by tunneling (see figure 1). Since the adatoms are not the only imperfections in graphene we also include an electron relaxation time which we include in a number conserving manner following the Mermin prescription [5, 29]. This relaxation time describes, phenomenologically, all sources of damping except the adatoms we are investigating. We explore the effects of adatoms on the single-particle properties of graphene as well as on the conductivity. We focus on the graphene surface plasmon mode and investigate its dispersion and the related losses. We find that plasmons close to resonance with the transition from the adatom energy level to above the Fermi energy become lossy. Furthermore, depending on the density of adatoms, their presence can split the plasmon mode into two separate branches, one low energy branch which experiences low losses and one high energy branch experiencing high degree of losses. We discuss how this can be used for ultra-sensitive sensing under the right conditions.

The article is organized as follows: in section 2 we treat the graphene plasmon dispersion and the graphene loss function. In section 3 we develop a manybody description of the system and derive an expression for the nonlocal longitudinal conductivity \( \sigma(q, \omega) \). Finally, in section 4 we analyze the effects of the adatoms on plasmons, in particular on the plasmon dispersion relation and damping. In appendix A we present a derivation of two central equations in the article, equations (3) and (4). Appendix B gives details of the microscopic model and in appendix C a simplified expression for the susceptibility tensor is derived, from which the conductivity is obtained. Throughout the article we use \( \hbar = c = 1 \).

2. Graphene plasmons

Longitudinal plasmons confined at a conducting interface between two dielectrics with relative dielectric constants \( \varepsilon_1 \) and \( \varepsilon_2 \) satisfy the dispersion relation [5, 6]

\[
\frac{\varepsilon_1}{\kappa_1(q, \omega)} + \frac{\varepsilon_2}{\kappa_2(q, \omega)} + \frac{i\sigma(q, \omega)}{\omega\varepsilon_0} = 0,
\]

where \( \kappa_{1,2}(q, \omega) = \sqrt{q^2 - \varepsilon_{1,2}\omega^2/c^2} \) and \( \varepsilon_0 \) is the vacuum dielectric constant. Here \( \sigma(q, \omega) \) is the longitudinal nonlocal conductivity of the graphene layer. The longitudinal conductivity describes the response to a longitudinal electric field meaning that \( q \) and \( E \) are parallel and \( |q| = q \). In the non-retarded limit, \( q \gg \sqrt{\varepsilon_{1,2}\omega/c} \), equation (1) reduces to the simpler expression

\[
(\varepsilon_1 + \varepsilon_2) + \frac{i\sigma(q, \omega)}{\varepsilon_0\omega} = 0.
\]

This is a good approximation as long as the wavelength of the mode is much smaller than the free space wavelength, which is the regime we are investigating.

In general equation (2) is a complex equation and for any given \( \omega \) it can be solved by allowing complex wave vectors, \( q = q_1 + iq_2 \). Physically this means that the corresponding oscillation of the mode is damped. For weak damping, we can expand equation (2) in small \( q_2/q_1 \), and separating the real and imaginary parts of the
conductivity, \( \sigma(q, \omega) = \sigma_l(q, \omega) + i \sigma_2(q, \omega) \), we obtain to lowest order in \( q \)

\[
\sigma_1(q, \omega) + \varepsilon_2(q) - \frac{q_1 \sigma_2(q, \omega)}{\omega \varepsilon_0} = 0,
\]

under the assumption that no losses are present in the substrate, i.e. \( \varepsilon_{1,2} \) are real. For a more detailed derivation of equations (3) and (4), see appendix A. Equation (3) is identical to equation (2) with the assumption of no losses, i.e. \( \sigma_l(q, \omega) = 0 \). The plasmons are given by equation (4) and in the second equality we have used equation (2) to express \( \sigma_2(q, \omega) \) on the plasmon mode. Equations (3) and (4) show explicitly that in the low-loss limit the graphene plasmon dispersion is determined by \( \sigma_2(q, \omega) \) while the plasmon losses are given by the ratio of \( \sigma_2(q, \omega) \) to the effective velocity \( v_{\text{eff}} = \frac{\omega \varepsilon_l(q, \omega)}{\varepsilon_0(q_1 + \varepsilon_2)} \) times \( \varepsilon_0(q_1 + \varepsilon_2) \). The effective velocity is given by the plasmon phase velocity \( \omega/q \) and a nonlocal correction given by the \( q_1 \) derivative of \( \sigma_2(q_1, \omega) \). The smaller the effective velocity, the higher the dissipative loss.

An alternative way to describe plasmons is to analyze the imaginary part of the current correlation function evaluated in the random-phase approximation. This is the spectral function of current fluctuations and describes where in \( q - \omega \) space it is possible to deposit energy. For this reason it is sometimes called the loss function and it is defined as [30, 44]

\[
S_{ij}(q, \omega) \equiv -\frac{1}{\nu_j} \text{Im} \left[ \frac{1}{\varepsilon(q, \omega)} \right] = \frac{\omega \sigma_l(q, \omega)}{1 + \frac{\omega^2 \nu_j(q, \omega)}{\varepsilon(\varepsilon_0(q_1 + \varepsilon_2))}}
\]

with \( \nu_j = \frac{\dot{e}_{ij}^{\nu}}{\varepsilon_0(q_1 + \varepsilon_2)} \). The loss function has peaks where equation (2) is satisfied and these peaks are interpreted as signatures of the plasmons [30].

The plasmon dispersion, i.e., solutions to equations (3) and (4) as well as the loss function are discussed in section 4. First, the conductivity of graphene with adatoms needs to be calculated.

3. Microscopic model for graphene conductivity

The system considered consists of a pristine infinite graphene sheet [45, 46] where momentum relaxation is added to model losses in graphene [5, 29]. We introduce the convention that clean graphene means graphene with the finite momentum relaxation. To the clean graphene, we add a dilute distribution of identical adatoms that are modeled as non-magnetic Fano-Anderson localized states [30] coupled to graphene by tunneling as sketched in figure 1. When the coverage is dilute, correlations between adatoms are unimportant, and each adatom can be considered independently. Here, dilute means \( n_{\text{imp}} \ll 1 \), where \( n_{\text{imp}} \) is the fraction of adatoms per lattice site. The total system can be described by a tight binding Hamiltonian that includes the tight binding Hamiltonian of graphene, the Hamiltonian of the adatoms, and the hopping between graphene and adatoms. Effective hopping Hamiltonians for different adatoms on graphene are obtained in [39–42], by use of DFT modeling of the composed system. In the following we assume a spin degenerate system, where the spin degree of freedom is included as a spin degeneracy factor \( g_s = 2 \). An individual adatom is situated on the graphene atom at site \( \mathbf{x} \) connected to a single graphene lattice site by a hopping parameter \( \tilde{t} \). The adatom has a single energy level at \( \varepsilon_0 \), measured relative to the charge neutrality point of clean graphene, and has an intrinsic lifetime \( \delta^{-1} \). These are phenomenological parameters that are inputs of the model and in this section we explore the general features of the model for various adatom parameters. In section 4 we obtain parameters for hydrogen from [40] to examine a simple adatom.

The Hamiltonian of the system with a single adatom is

\[
\hat{H} = \varepsilon_0 \hat{d}^\dagger \hat{d} + \sum_{\mathbf{i}} t_{\mathbf{i}} \hat{\psi}_{\mathbf{i}}^\dagger \hat{\psi}_{\mathbf{i}} + \tilde{t} \hat{d}^\dagger \hat{\psi}_{\mathbf{x}} + \text{h.c.,}
\]

where \( \hat{d} \) is the annihilation operator for the electron on the adatom and \( \hat{\psi}_{\mathbf{i}} \) is the annihilation operator of graphene electrons on site \( \mathbf{i} \). The sum in equation (6) is over all nearest neighbor sites of the graphene lattice and \( t_{\mathbf{i}} \) is the corresponding hopping parameter. The Hamiltonian is quadratic in the operators and can be diagonalized also for the case of a dilute density of adatoms, see appendix B for details.

In the regime of dilute density of adatoms and in the Dirac approximation, the Green’s function of the Hamiltonian is given by [47]

\[
G_{0}\mathbf{p}(\varepsilon) = \sum_{\lambda = \pm} \frac{\theta_{\hbar\lambda}(\mathbf{p}, \varepsilon)}{2} \begin{pmatrix} 1 \varepsilon + i\lambda \varepsilon \lambda e^{-i\hbar\lambda} \
\varepsilon + i\lambda \varepsilon \lambda e^{i\hbar\lambda} \end{pmatrix},
\]
due to the coupling to the adatoms. The change of
is the fraction of adatoms per lattice site. The effect of the adatoms is captured by
\(\epsilon^p = \epsilon - \lambda E_p - \Sigma_{\text{imp}}^R(\epsilon)\), where
\(\Sigma_{\text{imp}}^R(\epsilon) = \frac{|t|^2}{(\epsilon^+ - \epsilon_0 + i\delta)}\), \(\epsilon^+ = \epsilon + i0^+\),
\[\text{(7b)}\]
\[\text{(7c)}\]

where \(|t|^2 = n_{\text{imp}} |\hat{t}|^2\) and \(n_{\text{imp}}\) is the fraction of adatoms per lattice site. The effect of the adatoms is captured by the self-energy \(\Sigma_{\text{imp}}^R(\epsilon)\). The poles of the single-particle Green’s function define the single-particle states and in the presence of the adatoms the resulting single-particle bands are hybridized between the bare adatoms and bare graphene single-particle energy bands \(\lambda E_p\), where \(\lambda = \pm\) is the graphene band index. The poles of the retarded Green’s function and the renormalization factors are found to be
\[\epsilon^\pm(\mathbf{p}) = \frac{\lambda E_p - \epsilon_0 - i\delta}{2} \pm \sqrt{\left(\frac{\lambda E_p - \epsilon_0 - i\delta}{2}\right)^2 + 4|t|^2}\],
\[Z^\pm(\mathbf{p}) = \frac{1}{2} \pm \frac{E_p - \epsilon_0 - i\delta}{2\sqrt{(\lambda E_p - \epsilon_0 + i\delta)^2 + 4|t|^2}}\],

where the renormalization factors \(Z^\pm(\mathbf{p})\) are the residues of the poles, i.e. \(Z^\pm(\mathbf{p}) = \text{Res}[G^R_{\mathbf{p}}(\mathbf{p}, \epsilon)]\). The energies \(\epsilon^\pm(\mathbf{p})\) are shown in the center panel of figure 2. Due to the hybridization between the graphene bands and adatoms there is a level repulsion around \(\epsilon_0\), close to which the hybridization is strong and gives rise to a finite lifetime of the order \(\delta^{-1}\) to the energy bands. Far from the level splitting, the energy states approach their uncoupled behaviors. The right panel of figure 2 shows the total density of states of the coupled system which exhibits a significant deviation close to the level splitting compared to the pristine graphene case. Specifically, there is a significant increase in the density of states close to \(\epsilon_0\) due to the coupling to the adatoms. The change of the bands compared to pristine graphene opens up new possible electronic transitions that alter the conductive and plasmonic properties of graphene, as discussed in the following.

The conductivity of the system can be computed using the Green’s function in equation (7a). This is achieved by calculating the current response to an electric field \(E(x, t) = E_\mathbf{q} e^{i(q.x - \omega t)}\) (\(t\) is time in this paragraph, not to be confused with the coupling above). We restrict our analysis to the response to a longitudinal electric field which in the temporal gauge, \(\phi(x, t) = 0\), is given by \(E(x, t) = -i\omega A(x, t)\). We set the electric field, \(E\), and the momentum, \(\mathbf{q}\), to be parallel to the \(x\)-axis. According to minimal substitution \cite{30}, the perturbation given by the field \(E\) is \(\delta H_x = e\gamma_F A_x \hat{j}_x(\mathbf{q}) e^{i(q.x - \omega t)}\) where \(\hat{j}_x(\mathbf{q}) = e\gamma_F \sum_{\mathbf{p}} \hat{\psi}_p \hat{\psi}_{\mathbf{p} - \mathbf{q}} \hat{\psi}_{\mathbf{p}}\) is the longitudinal current operator \cite{45,46,48}. The diamagnetic current is zero in the Dirac approximation \cite{48}. The conductivity of the system can be obtained from the current–current response function (longitudinal susceptibility) which relates the average value of the current to the vector potential to linear order \(\langle \hat{j}_x(\mathbf{q}, \omega) \rangle = \gamma_{j_x(\mathbf{q}, \omega)} A_x\) \cite{30,49}. From this expression and the relationship between the vector potential and the electric field, the conductivity can be seen to be \(\sigma(\mathbf{q}, \omega) = \frac{1}{\omega} \chi_{j_x(\mathbf{q}, \omega)}\). The current–current response can be expressed in terms of the Green’s function as

![Figure 2](attachment:image.png)

**Figure 2.** Left: linear energy bands of pristine graphene (solid magenta) and the momentum independent adatom energy level (dashed magenta). Center: the real part of the Green’s function poles \(\epsilon^p(\mathbf{p})\) as function of momentum, inset the imaginary part. The adatoms introduce a splitting of the positive energy graphene band. The splitting introduces a hybridization of the bare graphene and bare adatom energy levels, resulting in the level repulsion between the states. Right: density of states for clean graphene in dashed magenta and black solid line in the presence of the adatoms with \(\rho_0 = \rho(E_F)\). The peak in the density of states reflects the presence of the flat adatom energy level. The parameters in all panels are \(\epsilon_0 = 0.8E_F\), \(\delta = 0.001E_F\) and \(t = 0.1E_F\).
The new features that are present in the conductivity, as shown in Figures 3 and 4, arise from the modification of the graphene bands caused by the presence of adatoms which is visible in Figure 2. This modification of the conduction band around $\epsilon_0$ changes the possible electronic transitions and in particular the allowed intraband transitions within the conduction band. These new transitions start playing a role around energies

$$X_{j, ij}(q, \omega) = \int \frac{dp}{2\pi} \text{Tr}[\sigma_j G_0^R(p, \epsilon) \sigma_i G_0^A(p - q, \epsilon - \omega)]$$

where $g g_j$ is the total spin-valley degeneracy factors, see appendix C for details. In pristine graphene, $\Sigma^R_{\text{imp}}(\epsilon) = -i\delta$, the energy integral in equation (8) can be performed analytically to arrive at the expression considered in [50–53]. In the temporal gauge the zero frequency component of the response is unphysical and has to be removed to avoid having a response to a static vector potential [53]. In the following we present results for electron doped graphene and when adatoms are present we take $\epsilon_0$ to be positive (with respect to the Dirac point). However, our model can also be applied for hole doped graphene and/or negative $\epsilon_0$.

Before analyzing the conductivity in the presence of the adatoms it is useful to examine the conductivity for pristine graphene. In the pristine case there exists a Pauli blocked triangle, inside which plasmon losses vanish at zero temperature due to the real part of the conductivity being identically zero [3, 6, 50, 51]. This triangle in $(q, \omega)$-space is shown in the inset of Figures 3 and 4. For non-zero temperatures or when momentum relaxation is included, e.g. through a finite relaxation time, the triangle is no longer completely lossless but for moderate temperatures and relaxation times it is still the region where plasmons with low losses are expected to exist [5]. In this article we take the momentum relaxation time ($1/\tau$) to be 1 ps, as reported in [54]. This relaxation time accounts for all intrinsic relaxation channels of the graphene and we include this in a number conserving way following the Mermin prescription [5, 29]. As already introduced, we refer to graphene with the finite relaxation time as clean graphene to distinguish from graphene together with the adatoms (and also a relaxation time).

The calculated conductivity in the presence of adatoms is shown in Figure 3 for various values of the hopping parameter $t$ and an impurity energy fixed at $\epsilon_0 = 0.8E_F$, i.e., the energy level of the adatoms is close to the Fermi energy. The presence of adatoms with energies close to the Fermi energy has an effect on the conductivity for frequencies close to the transition frequency $|E_F - \epsilon_0|$ between the adatom energy level and the Fermi energy. In particular there are peaks that appear inside the originally lossless triangle, which will give rise to larger plasmons losses, see equation (4). The imaginary part of the conductivity is also changed which will lead to changes in the plasmon dispersion as can be seen from equation (3).

Figure 4 shows the conductivity for an increasing adatom energy level detuning from the Fermi energy and a fixed density of adatoms. As the energy level moves further from the Fermi energy the conductivity becomes more and more like the conductivity of clean graphene and the effect of the adatoms becomes negligible. From this we conclude that for the adatoms to have a large effect, the energy level of the adatoms needs to be close to the Fermi energy.

The new features that are present in the conductivity, as shown in Figures 3 and 4, arise from the modification of the graphene bands caused by the presence of adatoms which is visible in Figure 2. This modification of the conduction band around $\epsilon_0$ changes the possible electronic transitions and in particular the allowed intraband transitions within the conduction band. These new transitions start playing a role around energies.

**Figure 3.** The nonlocal conductivity $\sigma(q, \omega)$ as a function of frequency $\omega$ with $q = |E_F - \epsilon_0|/\omega$ for different values of the hopping parameter $t$. The plots are for a constant wave vector $q = 0.2k_F$, shown by the dashed line in the inset, where the white triangle shows the Pauli blocked region for zero temperature. Real part of the conductivity is shown in the left panel and the imaginary part is shown in the right panel. Note that the scale on the x axes is not linear. The deviation of the conductivity from the clean case is hardly visible for $t < k_F \bar{T}$, but for larger values of $t$ there are structures appearing that significantly modify the conductivity. The conductivity is shown in units of $\sigma_0 = e^2/(4\hbar)$.

\[ \sigma_0 = 0.8E_F \]
\[ \delta = 10^{-3}E_F \]
\[ \Gamma_T = 10^{-2}E_F \]
\[ T = 0.01E_F \]
4. Graphene plasmons in the presence of a dilute density of adatoms

For concreteness, in this section we set the adatom parameters to be \( \bar{t} = 7.5 \text{ eV}, \epsilon_0 = 0.16 \text{ eV}, \) and \( \delta = 10^{-3}E_F \), see appendix B for more details. These parameters are extracted from [40] for hydrogen adsorbed on graphene. The individual hopping parameter, \( \bar{t} \), is large compared to the other energies but in the dilute adatom case the relevant parameter \( \bar{t}^2 = t_{\text{imp}}^2 \) is small for the densities we consider. Hydrogen is chosen since it is a simple atom and can serve as a typical atom adsorbed on the graphene. We emphasize that our model can be applied for other types of atoms and even simple molecules as well.

As stated previously, the loss function exhibits peaks where the dispersion equation has solutions. Figure 5 shows the loss function for an adatom density on the graphene of \( m \) approximately 300 adatoms per \( \mu \text{m}^2 \). The magenta line shown in figure 5 is the plasmon dispersion obtained by solving equation (3) for clean graphene. The peak of the loss function and the solution to the dispersion equation are in good agreement. The exception is close to the gray dashed line where the deviation is significant and a level splitting occurs. The level splitting between the bare plasmon and the bare adatom is caused by the coupling of adatoms to the graphene surface. The energy around which the level splitting occurs is represented in the figure as the horizontal dashed gray line. This energy represents the energy needed to excite an electron from the adatom energy to above the Fermi energy.

To examine the effect of adatoms on the plasmons in more detail, the left panel of figure 6 shows a zoom of the loss function from figure 5 which is centered on the splitting. The red dashed line in the figure is the solution to the dispersion equation, equation (3), for the same density of adatoms as the loss function and the solid red line is the obtained loss, \( \sigma_i \), from equation (4). The level splitting is accompanied by large plasmons losses and an emergence of two separate plasmon branches. The low energy plasmon branch exhibits low losses and the high energy branch has a large amount of accompanying loss. The larger loss in the upper branch can be understood by considering the Fermi golden rule, a new loss channel is opened for the plasmons in this branch. The loss channel is the excitation of a single electron around the adatom energy (of which there are many, see the DoS in figure 2) to above the Fermi energy where there are unoccupied electron states. The plasmons in the low energy branch do not have enough energy to lose energy through this channel. On resonance with this transition there is a very pronounced plasmons loss which separates the plasmon branches.

The right panel of figure 6 shows the evolution of a loss function cut at \( q/k_F = 0.013 \) as the density of adatoms is varied. This particular cut is chosen to show the loss function evolution for this particular adatom species (hydrogen) as clearly as possible. For small values of adatom density, there is only one plasmon peak visible, but as the density increases, the splitting into two branches is visible in the two emergent peaks. The separation between the two peaks grows as the adatom density is increased even further.
To explore the adatom effect on plasmons in a large frequency range, figure 7 shows the plasmon propagation length along the dispersion until it crosses into the single-particle continuum. The plasmon propagation length (defined as the distance covered until the intensity of the plasmon falls by $e^{-1}$) in units of the plasmon wavelength, $L_p = q_p \lambda_p$, can be obtained from the ratio $q_1/q_2$ in equation (4) as $L_p/\lambda_p = q_1/(4\pi q_2)$. Note that the plasmon wavelength is $\lambda_p = 2\pi/q_p$.

Both panels of figure 7 show the plasmon propagation length as a function of frequency for different adatom densities on the graphene. The left panel is calculated for $E_F = 0.2$ eV, $T = 30$ K, adatom density $300 \text{ m}^{-2}$, and $\xi = \xi_2 = 1$. The solid gray line shows the edge of the Pauli blocked triangle and the dashed gray line shows the value of $\omega \approx |E_F - E_0| + \Gamma/2$ around which there is a level splitting.

Both panels of figure 7 show the plasmon propagation length as a function of frequency for different adatom densities on the graphene. The left panel is calculated for $E_F = 0.2$ eV, $T = 30$ K, adatom density $300 \text{ m}^{-2}$, and $\xi = \xi_2 = 1$. The solid gray line shows the edge of the Pauli blocked triangle and the dashed gray line shows the value of $\omega \approx |E_F - E_0| + \Gamma/2$ around which there is a level splitting.

To explore the adatom effect on plasmons in a large frequency range, figure 7 shows the plasmon propagation length along the dispersion until it crosses into the single-particle continuum. The plasmon propagation length (defined as the distance covered until the intensity of the plasmon falls by $e^{-1}$) in units of the plasmon wavelength, $L_p = q_p \lambda_p$, can be obtained from the ratio $q_1/q_2$ in equation (4) as $L_p/\lambda_p = q_1/(4\pi q_2)$. Note that the plasmon wavelength is $\lambda_p = 2\pi/q_p$.

Both panels of figure 7 show the plasmon propagation length as a function of frequency for different adatom densities on the graphene. The left panel is calculated for $E_F = 0.2$ eV, $T = 30$ K, adatom density $300 \text{ m}^{-2}$, and $\xi = \xi_2 = 1$. The solid gray line shows the edge of the Pauli blocked triangle and the dashed gray line shows the value of $\omega \approx |E_F - E_0| + \Gamma/2$ around which there is a level splitting.

Both panels of figure 7 show the plasmon propagation length as a function of frequency for different adatom densities on the graphene. The left panel is calculated for $E_F = 0.2$ eV, $T = 30$ K, adatom density $300 \text{ m}^{-2}$, and $\xi = \xi_2 = 1$. The solid gray line shows the edge of the Pauli blocked triangle and the dashed gray line shows the value of $\omega \approx |E_F - E_0| + \Gamma/2$ around which there is a level splitting.
and thus discriminate between different adatoms. We stress that different adatoms imp 5 t 0.4 e m 2 3 0.2 e 10imp 7 2 10imp
be probed by for example light scattering in a grating environment densities of adatoms on the graphene surface. The induced plasmon losses and the level splitting that we
In this article we have focused on the graphene plasmon properties and how they are in
5. Discussion
The left and right panel of figure 7 are obtained for identical parameters except for the Fermi energy. The left panel clearly shows a larger effect for the same density of adatoms compared to the right panel. The conclusion is that for sensing purposes the Fermi energy should be tuned close to the adatom energy level for the sensitivity to be large.

5. Discussion
In this article we have focused on the graphene plasmon properties and how they are influenced by small densities of adatoms on the graphene surface. The induced plasmon losses and the level splitting that we find can be probed by for example light scattering in a grating environment [11], or on patterned graphene microribbons [1]. For a properly dimensioned grating or microribbon array, it would be possible to perform a laser frequency sweep and measure optical signatures of the presence of the adatoms as seen in figure 6. For the doping levels considered in this article, the typical dimension of the grating periodicity or the microribbon arrays needed are on the order of a few hundred nanometers. An alternative route to investigate the plasmons is by nanotip experiments such as in [9, 10], where the number of plasmon oscillations are measured. Obtaining such data for different frequencies could reveal the presence of small amounts of adatoms and their energy levels.

Our analysis is restricted to the presence of a single kind of adatom, i.e., all the adatoms on the surface are characterized by the same $\epsilon_0$, $i$, and $\delta$. In our model, the values of $\epsilon_0$, $i$, and the adatom density determine the characteristic level splitting that separates the two plasmon branches, see figures 6 and 7. Also, the sensitivity of the plasmons to the adatoms is found to be large when the adatom energy is close to the Fermi energy. Hence, by measuring the plasmonic properties and taking advantage of the tunability of $E_F$ offered by graphene, it is possible to determine $\epsilon_0$ and thus discriminate between different adatoms. We stress that different adatoms exhibit different values of $\epsilon_0$ as indicated by DFT calculations, see [37–42, 55], and that our model is general enough to handle various adatoms and simple molecules. Our model thus enables selective sensing of various adatoms and molecules by probing the plasmonic properties of graphene. It should be noted that the adatom densities involved, 30–300 $\mu$m$^{-2}$, is enough to increase plasmon losses and create a level splitting, making sensing of minute amounts of substances possible using graphene plasmons coupled to electron energy levels of adatoms. This is an increase of 2–3 orders of magnitude in sensitivity compared to experimental results for biomolecules obtained in [36], where plasmon coupling to vibrational modes of molecules was utilized for sensing. In our model, the adatom densities needed to produce a measurable plasmon response depends on the adatom coupling strength, $i$, which is considered as an input in our model. Thus, the sensitivity of the proposed sensing scheme is different for different adatoms. Adatoms that couple strongly to graphene will give rise to larger plasmon response for a given density than weakly coupled adatoms.

In this article we have taken a view towards sensing of the adatoms on the graphene surface. However, the adatoms may also be considered as imperfections on the graphene that impedes electron propagation by allowing the electrons to tunnel onto the adatom. For the purpose of plasmonics, long propagation lengths are often sought after and such damping is unwanted. We find that even small amounts of adatoms may have a significant effect on the plasmon damping. This is potentially one of several mechanisms that induces the large plasmon damping found in experiments [9, 10].

![Figure 7. Plasmon propagation length, $L_p/\lambda_p$, computed using equations (7) and (8), versus $\hbar\omega$ for different adatom densities. The left panel is calculated for $E_F = 0.2$ eV and the right panel for $E_F = 0.4$ eV. The magenta lines represent clean graphene, green line $n_{imp} = 10^{-7}$, orange line $n_{imp} = 10^{-6}$, and blue line $n_{imp} = 10^{-5}$. This corresponds to approximately 3 $\mu$m$^{-2}$, 30 $\mu$m$^{-2}$, and 300 $\mu$m$^{-2}$ respectively. The sharp drop around $1.3E_F$ is due to the plasmon dispersion crossing into the particle-hole continuum where the plasmon becomes heavily damped.](image-url)
6. Conclusions

We have investigated how graphene plasmons are affected by adatoms by comparing plasmons in realistic quality graphene with and without adatoms. We found that adatoms with energy levels close to the Fermi energy induces a strong level splitting between the bare plasmon mode and the adatom energy level. This level splitting is accompanied by large plasmon losses and depending on the adatom density may separate the plasmon mode into two separate branches, one low energy branch and one high energy branch. The low energy branch is virtually unaffected by the presence of the adatoms, whereas the high energy branch experiences larger losses. This is due to a new plasmon decay channel opening up, namely the excitation of an adatom electron to an unoccupied state above the Fermi energy.

Furthermore, we studied the sensitivity of the plasmon losses to the presence of adatoms. As a typical atom, we considered hydrogen and we found that a density of 300 adatoms per $\mu m^2$ is enough to give rise to a significant level splitting, and already 30 adatoms per $\mu m^2$ is enough to damp the upper branch. These effects could be measured in various light scattering experiments using dielectric gratings as well as using nanotips, making it possible to envision ultra-sensitive devices that measure the plasmon dispersion and losses to infer the presence of adatoms and molecules on the graphene surface.

Our results highlight the sensitivity of graphene plasmons to microscopic degrees of freedom and the possibility to use this effect in applications. Microscopic models for coupling various degrees of freedom to the plasmons is a very rich field and has the potential to further increase the already large sensing potential of graphene plasmons.

Acknowledgments

Financial support from the Knut and Alice Wallenberg foundation (KAW), the Swedish Research Council (VR), and the Swedish Foundation for Strategic Support (SSF) is gratefully acknowledged. We wish to thank Philippe Tassin, Daniel J Persson and Samuel L Avila for interesting discussions. G V and T W have contributed equally to this project.

Appendix A. Derivation of equations (3) and (4) in the main article

This appendix details the derivation of equations (3) and (4) starting from equation (2). Equation (2) defines the plasmon dispersion relation in the non-retarded limit. In the presence of losses, the dispersion relation is obtained by allowing complex wave vectors, $q = q_1 + iq_2$, hence equation (2) becomes

$$ (\varepsilon_1 + \varepsilon_2) + \frac{i(q_1 + iq_2)\sigma(q_1 + iq_2, \omega)}{\varepsilon_0 \omega} = 0. \quad (A.1) $$

Under the assumption of low loss, $q_2/q_1$ is small, and equation (A.1) can be expanded to first order in $q_2/q_1$ and gives rise to:

$$ (\varepsilon_1 + \varepsilon_2) + \frac{i}{\varepsilon_0 \omega}(q_1 + iq_2) $$
$$ \times [\sigma(q_1, \omega) + i\sigma_2(q_1, \omega) + iq_2\partial_q\sigma_2(q_1, \omega) + i\sigma_2(q_1, \omega)] = 0. \quad (A.2) $$

The real part and zeroth order in $q_2/q_1$ of equation (A.2) is exactly equation (3), while the imaginary part and first order in $q_2/q_1$ of equation (A.2) gives equation (4).

Appendix B. Fano-Anderson model in graphene

This appendix gives details on the microscopic model used in the main text. In particular the full Green’s function in equation (7a) is derived.

We first consider pristine graphene, described in [45, 46, 48], coupled to adatoms by tunneling. The adatoms are modeled here as Fano-Anderson localized states [30] as described in the main text. To solve the system in the case of many adatoms, the following approximations are used: (i) all the impurities are identical, i.e., all of them are characterized by the same parameters $\epsilon_0$, $\delta$ and $\tilde{t}$, (ii) the adatoms are uncorrelated and far apart so an average on position can describe the system. These assumptions are also used in the $T$-matrix formalism for weakly interacting electron systems in the presence of low densities of impurities [30].

In the main text the Hamiltonian of graphene and a single adatom is presented in equation (6). The Hamiltonian for many adatoms is [45, 48]
\[ \hat{H} = \sum_l \epsilon_l \hat{d}_l \dagger \hat{d}_l + \sum_{p, \lambda} \left( \lambda E_p \hat{c}_p \dagger \hat{c}_p + \sum_l t_{l, p, \lambda} \hat{d}_l \dagger \hat{c}_{p, \lambda} + \text{h.c.} \right) , \]  

(B.1)

where \( \mathbf{p} \) is the momentum, \( t_{l, p, \lambda} = \lambda F e^{i \mathbf{p} \cdot \mathbf{x}_l} \), \( \lambda = \pm \) is the graphene band index and \( x_l \) is the position of the \( l \)th impurity. Hence, for many impurities with vanishing hopping between them, the total Hamiltonian is obtained by adding single impurity contributions which is performed by the sum on \( l \).

We introduce the notation \( P = (\mathbf{p}, \lambda) \) and \( E_p = \lambda E_p \). \( \hat{H} \) is quadratic so it can be diagonalized by a unitary transformation \( \hat{c}_p = \sum_{A} A_{p, i} \hat{\chi}_i \) and \( \hat{d}_l = \sum_{B} B_{l, j} \hat{\chi}_j \) with \( A \) and \( B \) matrices such that \( H = \sum_{l} \hat{E}_l \hat{\chi}_l \dagger \hat{\chi}_l \) and \( \{ \hat{\chi}_l, \hat{\chi}_j \} = \delta_{l, j} \). To find the matrices \( A \) and \( B \) one can compute commutators using the expression in equation (B.1) obtaining

\[ [\hat{d}_l, \hat{H}] = \epsilon_0 \hat{d}_l + \sum_{p} t_{l, p, \lambda} \hat{c}_p = \epsilon_0 \sum_{j} B_{l, j} \hat{\chi}_j + \sum_{p} t_{l, p, \lambda} \sum_{j} A_{p, j} \hat{\chi}_j , \]  

(B.2)

\[ [\hat{c}_p, \hat{H}] = E_p \hat{c}_p + \sum_{l} t_{l, p, \lambda} \hat{d}_l = E_p \sum_{j} A_{p, j} \hat{\chi}_j + \sum_{l} t_{l, p, \lambda} \sum_{j} B_{l, j} \hat{\chi}_j , \]  

(B.3)

where the second equality in both equations is obtained by substituting the expressions for \( \hat{c}_p \) and \( \hat{d}_l \) above. The same commutators computed using the diagonal expression of the Hamiltonian gives

\[ [\hat{d}_l, \hat{H}] = \sum_{j} \hat{E}_l \hat{\chi}_j B_{l, j} , \]

\[ [\hat{c}_p, \hat{H}] = \sum_{j} \hat{E}_l \hat{\chi}_j A_{p, j} , \]

and by matching with the previous expressions we obtain equations for \( A_{p, j} \) and \( B_{l, j} \). However, these equations still contain the operator \( \hat{\chi}_j \). This can be removed by performing additional commutations with \( \hat{\chi}_j \) giving

\[ [\hat{\chi}_j \dagger, [\hat{d}_l, \hat{H}]] \Rightarrow \epsilon_0 B_{l, j} + \sum_{p} t_{l, p, \lambda} A_{p, j} = \tilde{E}_l B_{l, j} , \]

\[ [\hat{\chi}_j \dagger, [\hat{c}_p, \hat{H}]] \Rightarrow E_p A_{p, j} + \sum_{l} t_{l, p, \lambda} B_{l, j} = \tilde{E}_j A_{p, j} , \]

from which

\[ (\tilde{E}_l - \epsilon_0) B_{l, j} = \sum_{p} t_{l, p, \lambda} A_{p, j} , \]  

(B.4)

\[ (\tilde{E}_j - E_p) A_{p, j} = \sum_{l} t_{l, p, \lambda} B_{l, j} , \]  

(B.5)

A formal solution to these equations is

\[ B_{l, j} = \frac{\sum_{p} \tilde{E}_j B_{l, j} \delta_{E_p, \epsilon_0} + \sum_{p} Z_{j, p}^{B} t_{l, p, \lambda} \delta_{E_p, \epsilon_0}}{\tilde{E}_j - E_p} , \]

\[ A_{p, j} = \frac{\sum_{l} t_{l, p, \lambda} B_{l, j} \delta_{E_p, \epsilon_0} + \sum_{l} Z_{l, j}^{A} t_{l, p, \lambda} \delta_{E_p, \epsilon_0}}{\tilde{E}_j - E_p} , \]

where \( Z_j^{A/B} \) are unknown coefficients and \( \delta_{a,b} = 1 - \delta_{a,b} \). Now, substitute these solutions into equations (B.4) and (B.5) obtain

\[ (\tilde{E}_l - \epsilon_0) B_{l, j} = \sum_{p} t_{l, p, \lambda} \left( \frac{t_{l, p, \lambda} A_{p, j} \delta_{E_p, \epsilon_0} + Z_{j, p}^{A} t_{l, p, \lambda} \delta_{E_p, \epsilon_0}}{\tilde{E}_j - E_p} \right) , \]  

(B.6)

\[ (\tilde{E}_j - E_p) A_{p, j} = \sum_{l} t_{l, p, \lambda} \left( \frac{t_{l, p, \lambda} B_{l, j} \delta_{E_p, \epsilon_0} + Z_{l, p}^{B} t_{l, p, \lambda} \delta_{E_p, \epsilon_0}}{\tilde{E}_j - \epsilon_0} \right) . \]  

(B.7)

To proceed further we take advantage of the assumption of identical impurities and we may thus perform the sums on \( l \). We assume that the positions of the impurities are independent and randomly distributed, so the average on the impurity position is \( \sum_l \frac{t_{l, p, \lambda} t_{l, p, \lambda}^*}{N_{\text{imp}}} = \frac{1}{N_{\text{imp}}} \sum_{p} |t_{l, p, \lambda}|^2 \delta_{p, p'} \) \( \delta_{l, l'} \), with \( N_{\text{imp}} \) the number of impurities per lattice site. This is standard procedure in \( T \)-matrix formalism for low density impurities [30]. Applying this average to perform the sums on \( l \) in equations (B.6) and (B.7) gives

\[ (\tilde{E}_l - \epsilon_0) B_{l, j} = \sum_{p} \frac{n_{\text{imp}} |t_{l, p, \lambda}|^2 B_{l, j} \delta_{E_p, \epsilon_0}}{\tilde{E}_j - E_p} + Z_{j, p}^{A} |t_{l, p, \lambda}|^2 \delta_{E_p, \epsilon_0} , \]  

(B.8)
From equation (B.9) we get \( \tilde{E}_j = E_p + \frac{|t|^2 \delta_{E_k \epsilon_0}}{\tilde{E}_j - \epsilon_0} + Z_0^R |t|^2 \delta_{E_k \epsilon_0} \), and hence the self energy for the graphene states is obtained: \( \Sigma_{\text{imp}}^{R}(\tilde{E}_j) = \frac{|t|^2 \delta_{E_k \epsilon_0}}{\tilde{E}_j - \epsilon_0} + Z_0^R |t|^2 \delta_{E_k \epsilon_0} \). Due to level repulsion the new eigenenergies are different from the bare adatom energy so \( \delta_{E_k \epsilon_0} = 0 \) and \( Z_0^R \) drops out of the equation.

The matrices A and B can be found by imposing the conditions \( \{ \tilde{A}_j, \tilde{d}_x^j \} = \{ \tilde{A}_j, \tilde{d}_y^j \} = \{ \tilde{A}_j, \tilde{d}_z^j \} = \sum_j B_{j} A_{j} \delta_{E_k \epsilon_0} \) and \( \{ \tilde{d}_x^j, \tilde{d}_y^j \} = \{ \tilde{d}_x^j, \tilde{d}_z^j \} = \{ \tilde{d}_y^j, \tilde{d}_z^j \} = 0 = \sum_j B_{j} A_{j} \delta_{E_k \epsilon_0} \). For our purposes, it is not necessary to find the values of A, B and Z. The existence of the unitary matrices A and B is enough to write down the Green’s function. Indeed, we know that the Green’s function in the frequency domain is \( G^{-}_{\epsilon_0}(\omega) = \delta_{E_k \epsilon_0} / (\omega + i\eta - \tilde{E}_j) \) since the Hamiltonian is diagonal in \( \alpha \).

Hence

\[
G^{-}_{\epsilon_0}(\omega) = \sum_{l,j} A_{l,j} A_{l,j}^{\dagger} \left( T \delta_j(t) \delta_{E_k} (0) \right) = \sum_{l,j} A_{l,j} A_{l,j}^{\dagger} \left( \delta_{E_k \epsilon_0} / (\omega + i\eta - \tilde{E}_j) \right)
\]

and from here equations (7a) follow.

Below we give further details on our model. The density of states \( \rho(\epsilon) \equiv -\frac{1}{\pi} \text{Tr} \text{ Im} (G^{-}_{\epsilon_0}(\epsilon) \rho) \), Tr is the trace on the sublattice indexes and \( \langle \gamma \rangle_{\rho} \) is the momentum averaged Green’s function defined as

\[
\langle G^{-}_{\epsilon_0}(\rho, \epsilon) \rangle_{\rho} = A_c \int d\rho G^{-}_{\epsilon_0}(\rho, \epsilon) = \frac{1}{2E_C} \int \left[ \frac{(\gamma)^2}{\epsilon} - E_C \right],
\]

where \( l \) is the unit matrix, \( \epsilon = -\Sigma_{\text{imp}}^{R}(\epsilon) \), \( A_c \) is the area of the unit cell in the lattice, \( d\rho = \frac{d\rho_\alpha}{2\pi} d\epsilon \) and \( E_C \) is the cut-off energy corresponding to the graphene bandwidth.

Even though the present model treats the electron hopping non-perturbatively, the Coulomb interaction between the adatom and graphene charge is up to now omitted. A full treatment of this interaction is beyond the aim of this work. Nevertheless, we consider the Coulomb interaction in the presence of charge fluctuations induced by the tunneling. This effect introduces a further relaxation channel, i.e., a finite lifetime to both the adatom states, \( \tau^{-1} \), and the electron states in the graphene \( \Gamma^{-}_{\epsilon_0} \). Following a method similar to [56], we evaluate them to be \( \Gamma^{-}_r \simeq 10^{-4}E_F, \Gamma^{-}_e \simeq 10^{-3}E_F \) at \( T = 0 \). Furthermore, we include a finite relaxation time for the graphene electrons which is caused by imperfections and phonons in the graphene lattice. From good quality graphene this number can be inferred to be approximately 1 ps, see [54]. For realistic doping levels of graphene, this relaxation time completely dominates \( \Gamma^{-}_{\epsilon_0} \) which is neglected. These values are used throughout the paper and the total relaxation time is included following the Mermin prescription [5, 29].

Appendix C. Derivation of equation (8) and simplified expression for the conductivity

In this appendix we give some details on the derivation of equation (8) and we report a simplified expression for the one-spin, one-valley susceptibility. Before proceeding, an observation needs to be pointed out. In the context of plasmons, it is common to describe the purely longitudinal electric field, along the x axes, with a potential \( \phi(x, t) = \phi_0 e^{i(qx - \omega t)} \) and \( A = 0 \) with \( E_0 = i\omega \phi_0 \) [30, 49–51]. The linear response of the system is then encoded in the density–density response function (polarisability \( \chi_{pp} \)) that expresses the density fluctuation induced by the potential as \( \langle \hat{\rho}(q, \omega) \rangle = \chi_{pp}(q, \omega) \). In the standard situation, the continuity equation relates \( \chi_{pp} \) to the conductivity. However, in the presence of the adatoms, the total charge density includes both the charge density in the graphene and on the impurity states \( \hat{\rho}_{\text{tot}}(q) = \hat{\rho}(q) + \hat{\rho}_{\text{imp}} \), with \( \hat{\rho}_{\text{imp}} = \epsilon \frac{d^2}{d^2} \). Also the last term, \( \rho_{\text{imp}} \), needs to be included to fulfill the continuity equation \( i\omega \hat{\rho}_{\text{tot}}(q) = -i\omega \hat{j}(q) \) which makes this approach more involved. Therefore, it is more convenient to evaluate the longitudinal current–current response function, \( \chi_{ji,ii}(q, \omega) \), in the temporal gauge since the adatoms carry no in-plane current and thus it is possible to avoid including terms related to the adatoms.

From the above we compute the average value of the longitudinal current \( \langle \hat{j}_j(q, \omega) \rangle = e\sigma_v \sum_{\alpha} \hat{\psi}^\dagger \psi \gamma_{s_{\alpha}} \phi_{s_{\alpha}} \), where \( s_{\alpha} \) is the first Pauli matrix, in presence of the perturbation \( \delta H_x = e\sigma_v A_{ij} \hat{j}_{j}(q)e^{-i\omega x} \). For \( E_0 = 0 \) and in equilibrium, the values of currents and density deviation are vanishing. In Keldysh formalism, the current to linear order in the perturbation is written \( \langle \hat{j}_j(q, \omega) \rangle = -\frac{1}{2} \text{Tr} \left[ \sigma_v \delta G^{R}(q, \omega) \right] \). As shown in [57]
\[ \delta G(q, \omega) = (G_0 \delta H, G_0)(q, \omega) = \int \frac{d\epsilon}{2\pi} G_0(p, \epsilon) \cdot \tau_{\epsilon} \delta H_{\epsilon} \cdot \delta G_0(p - q, \epsilon - \omega) \]  

\( \tau_{\epsilon} = \sigma_{\epsilon} \) in Keldysh space. \( \delta G_0 \) is the unperturbed Green’s in Keldysh space and \( \sigma \) stands for the sum on all internal degrees of freedom. The substitution of \( \delta G(q, \omega) \) in \( \chi_{k,j}(q, \omega) \) gives

\[ \chi_{k,j}(q, \omega) = \frac{i e^{2} \mu^{2}}{2} \int \frac{d\epsilon}{(2\pi)} \text{Tr}[\sigma_{\epsilon} G_{R}^{0}(p, \epsilon) \sigma_{\epsilon} G_{A}^{0}(p - q, \epsilon - \omega) \nonumber + \sigma_{\epsilon} G_{R}^{0}(p, \epsilon) \sigma_{\epsilon} G_{A}^{0}(p - q, \epsilon - \omega)] \]

which is identical to equation (8). The Keldysh component of the Green’s function at equilibrium is \( G_{R}^{0}(p, \epsilon) = f(\epsilon)(G_{R}^{0}(p, \epsilon) - G_{A}^{0}(p, \epsilon)) \) with \( f(\epsilon) = 1/(1 + e^{\beta(\epsilon - \mu)}) \), \( g_{R}^{0}(p, \epsilon) \) is \( g_{A}^{0}(p, \epsilon) \), and \( (G^{2}(q, \epsilon))^{R} = G_{R}^{0}(q, \epsilon) \). Substituting these relations and performing the trace in equation (C.2), the integrand becomes

\[ \text{Tr}[f(\epsilon - \omega)\sigma_{\epsilon} G_{R}^{0}(p, \epsilon)\sigma_{\epsilon} G_{A}^{0}(p - q, \epsilon - \omega) - G_{A}^{0}(p - q, \epsilon - \omega)] \]

\[ + f(\epsilon)\sigma_{\epsilon}(G_{R}^{0}(p, \epsilon) - G_{A}^{0}(p, \epsilon))\sigma_{\epsilon} G_{A}^{0}(p - q, \epsilon - \omega) \]

\[ = [f(\epsilon - \omega)g_{R}^{0}(p, \epsilon)(g_{R}^{0}(p - q, \epsilon - \omega) - g_{A}^{0}(p - q, \epsilon - \omega)) \]

\[ + f(\epsilon)(g_{R}^{0}(p, \epsilon) - g_{A}^{0}(p, \epsilon))g_{A}^{0}(p - q, \epsilon - \omega)]F_{xx}^{\lambda X}(q, p), \]

where

\[ F_{xx}^{\lambda X}(q, p) = 1 + \frac{\lambda(\omega \cos \theta - q \cos \theta)}{|p - q|} \]

is the square of the matrix element of the longitudinal current operator.

To simplify \( \chi_{k,j}(q, \omega) \) one may proceed to first perform the integral on \( \epsilon \). Note that in the clean case this integral can be performed following [30] to arrive at the corresponding expression used in [50, 51, 53]. The present case is more involved due to the finite self-energy but the integrand in equation (C.2) has known poles. The poles come from the Green’s function, i.e., from \( g_{R}^{0}(p, \epsilon) \), or from the Fermi distribution \( f(\epsilon) \), they are denoted \( \epsilon_{n}^{R,A}(p) \) and \( \epsilon_{n} = ik_{B}T(2m + 1) + \mu \) (integer) respectively. The \( \epsilon \) integral on the real axes in (C.2) can be closed in the complex plane in such a way that as few Green’s function poles as possible are included inside the path. Splitting the integral in two parts, the path of the term proportional to \( f(\epsilon - \omega) \) (first term in equation (C.4)) is closed in the upper half plane, to avoid the poles of \( g_{R}^{0}(p) \). Vice versa the integral of the term proportional to \( f(\epsilon) \) in equation (C.4), is closed in the lower half plane, to avoid the poles of \( g_{A}^{0}(p) \). Of course, all the poles of the Fermi distributions can not be avoided since they lie on both sides of the real axis. The energy integration therefore has two contributions \( \chi_{k,j}(q, \omega) = \chi_{k,j}^{P}(q, \omega) + \chi_{k,j}^{T}(q, \omega) \) with

\[ \chi_{k,j}^{P}(q, \omega) = \sum_{\lambda, \Lambda} e^{2} \mu^{2} \int dp[f(\epsilon_{n}^{R}(p - q))(\epsilon_{n}^{A}(p - q) + \omega, p) \nonumber + f(\epsilon_{n}^{R}(p))(\epsilon_{n}^{A}(p - q) + \omega, p - q)]F_{xx}^{\lambda X}(q, p), \]

\[ \chi_{k,j}^{T}(q, \omega) = \int_{0}^{\infty} \sum_{\lambda, \Lambda} e^{2} \mu^{2} \int dp[\epsilon_{n}^{R}(\epsilon_{n}^{R}(\epsilon_{n}^{A} - \omega, p - q) - \epsilon_{n}^{A}(\epsilon_{n}^{A} - \omega, p - q))]F_{xx}^{\lambda X}(q, p). \]

Here \( \chi^{P} \) and \( \chi^{T} \) are given by the poles of the Green’s function and Fermi function respectively. We underline that to separate \( \chi_{k,j}^{P}(q, \omega) \) and \( \chi_{k,j}^{T}(q, \omega) \) in these two terms, the only assumption is that the self-energy is a smooth function of \( \epsilon \). Hence, it can be extended to include also other contributions to the electronic self-energy such as the phonon and electron–electron interaction. We underline that the susceptibility contributions in equations (C.6) and (C.7) are for one valley and one spin and thus needs to be multiplied by \( g_{R}, g_{A} \), as is done in the main text, to obtain the total result for graphene.

The momentum integrals in \( \chi_{k,j}^{P}(q, \omega) \) and \( \chi_{k,j}^{T}(q, \omega) \) are performed numerically using standard integration routines. It turns out that \( |\chi_{k,j}^{T}(q, \omega)| < 10^{-1} |\chi_{k,j}^{P}(q, \omega)| \) for the parameter range of interest. In particular, for \( \omega < \epsilon_{F} \),

\[ |\chi_{k,j}^{T}(q, \omega)| < 10^{-1} |\chi_{k,j}^{P}(q, \omega)| \]

and is therefore omitted in the analysis in the main text.

References

[34] Wu L, Chu H S, Koh W S and Li E P 2010 Highly sensitive graphene biosensors based on surface plasmon resonance Opt. Express 18 14395
Graphene plasmons: impurities and nonlocal effects

Giovanni Viola,1, *, Tobias Wenger,1 Jari Kinaret,2 and Mikael Fogelström1

1Department of Microtechnology and Nanoscience (MC2), Chalmers University of Technology, S-412 96 Göteborg, Sweden
2Department of Physics, Chalmers University of Technology, S-412 96 Göteborg, Sweden

(Dated: Draft of October 3, 2017)

This work analyses how impurities and vacancies on the surface of a graphene sample affect its optical conductivity and plasmon excitations. The disorder is analysed in the self-consistent Green’s function formulation and nonlocal effects are fully taken into account. It is shown that impurities modify the linear spectrum and give rise to an impurity band whose positions and widths depend on the two parameters of our model, the density and the strength of impurities. The presence of the impurity band strongly influences the electromagnetic response and the plasmon losses.

I. INTRODUCTION

Plasmons are electromagnetic fields resonantly enhanced by oscillations in the charge density. Due to the properties of graphene, in graphene plasmons show: low losses, tunable optical properties, strong optical field confinement, and environmental sensitivity. This makes graphene an attractive material for next generation technologies in sensing, photonics, and communication. To improve device design and performance, it is crucial to extend the microscopic theory of plasmons to include nonlocal effects together with the impact of defects and impurities in the sample as well as chemical compounds deposited on the surface. Defects and impurities may be due to the fabrication procedure, while chemical compounds can be deposited in a controlled fashion on the surface to functionalise the graphene substrate. Defects and impurities are inevitably sources of losses that must be understood in order to make high-performance samples and devices, mainly by circumventing their loss-producing effects.

The behaviour of plasmons in pristine graphene is by now rather well studied. The local transport properties are investigated in a series of articles, e.g., including effects due to the impurities, phonons, and localised charges. The nonlocal effects in the presence of impurities or adsatoms have been considered, among others, in Refs. Phonon- and electron-electron interaction has been studied in Refs.

First-principles studies have determined how crystal defects or atoms on the graphene surface influence the material properties. Defects are seen to give rise to new bands whose properties depend on the density and type of defects or adsorbates. This opens the possibility to engineer the band structure of the material.

While first-principle studies consider relatively small graphene supercells (on the order of 10^2 atoms), many-body techniques are more suitable to describe properties in μm size devices. In this work we include impurities in a simple self-consistent t-matrix treatment of elastically scattering impurities, and to explore how their presence modify the optical conductivity and plasmonic behaviour of graphene. In the microscopic model used here, as described in Sec. II, the nonmagnetic impurities are described as onsite, spin-preserving potentials and treated self-consistently. The nonlocal transport and optical properties are investigated in Sec. III. The optical response resembles the one obtained in the relaxation-time approximation in the case of dislocations in graphene, while novel features are observed if the impurity band is detuned from the Dirac point. In particular, it is observed that an impurity band far from the Dirac point enhances the plasmon losses. We note that the optical response and plasmonic behaviour can characterise the impurity itself and we also emphasize the potential of plasmon-based sensors and of contactless characterization of samples.

In the following the densities of electrons and impurities are given in units of 10^{12}/cm^2, the energies are measured in eV, 1eV=241.8THz=1239 nm and the conductance in units of $σ_0 = e^2/(4h) = 6.085 \times 10^{-5} \Omega^{-1}$.

II. MODEL

Longitudinal plasmons confined at a conducting interface between two dielectrics, with relative dielectric constants $\varepsilon_1$ and $\varepsilon_2$, satisfy the dispersion relation

$$\frac{(\varepsilon_1 + \varepsilon_2)}{q} + \frac{i\sigma(q,\omega)}{\varepsilon_0 \omega} = 0 \quad \text{(1)}$$

with the wave vector, $q$ ($q = |q|$), in the graphene plane and the angular frequency, $\omega$, of the electromagnetic field. Here we assume the non-retarded limit, $q \gg \sqrt{\varepsilon_1 \omega/c}$, as the light and the plasmons have a large momentum mismatch. To achieve an efficient coupling of light to plasmon modes is possible by introducing a dielectric grating or coupling via evanescent light modes to overcome the mismatch. The nonlocal longitudinal conductivity, $\sigma(q,\omega) = \sigma_1(q,\omega) + i\sigma_2(q,\omega)$, together with the dielectric environment, encodes the plasmon properties. As conductors in general are lossy, $\sigma_1(q,\omega) > 0$, we can read from Eq. (1) that either $q$ or $\omega$ is required to be complex to account for these losses. Connecting to scattering experiments e.g. in Refs. $\omega$ is the frequency...
We will use the Dirac approximation and in this scheme the conductivity is then given as

$$ \sigma(q, \omega) = \frac{i}{\omega} \chi_{j,j}(q, \omega). $$

The microscopic details of the material is encoded in the Green’s functions $G^R, G^A = (G^K)^\dagger$, and $G^K$ and the self-energies $\Sigma^{R,A,K}$. Here we consider a dilute ensemble of $s$-waves scatters, smooth on the atomic scale, included via a self-consistent $t$-matrix method. It is a two-parameter theory with $V_{\text{imp}}$, being the strength, and $n_{\text{imp}}$ the density of impurities respectively. The density $n_{\text{imp}} = N_{\text{imp}}/N$, is the number of impurities, $N_{\text{imp}}$, divided by the number of unit cells, $N$, in the crystal.

We will use the non-local conductivity $\sigma(q, \omega)$ is evaluated from the current-current response to an external electromagnetic field within RPA as

$$ \chi_{j,j}(q, \omega) = g_j g_0 \frac{i e^2 v_F^2}{2} \int \frac{dp \, d\epsilon}{(2\pi)} \text{Tr} \left[ \sigma_x G^R(p, \epsilon) \sigma_x G^K(p - q, \epsilon - \omega) + \sigma_x G^K(p, \epsilon) \sigma_x G^A(p - q, \epsilon - \omega) \right]. $$

The conductivity is then given as

$$ \sigma(q, \omega) = \frac{i}{\omega} \chi_{j,j}(q, \omega). $$

energy scale is set by a cut-off $\epsilon_c$ related to the band-width, we set $\epsilon_c = 8.2 eV$.

The band structure is obtained from the poles of the Green’s function

$$ G^R(p, \epsilon) = \sum_{\lambda = \pm} \frac{1/2}{\epsilon - E_{\lambda,p} - \Sigma_{\text{imp}}^R(\epsilon)} \left( \begin{array}{c} 1 \\ \lambda e^{i \phi_p} \end{array} \right) \right) $$

with the self-energy

$$ \Sigma_{\text{imp}}^R(\epsilon) = \frac{n_{\text{imp}} V_{\text{imp}}}{1 - V_{\text{imp}} \frac{\epsilon}{\nu_f} \Sigma \int d\epsilon G^R(p, \epsilon)}. $$

Here $E_{\lambda,p} = \lambda v_F p$ is the single particle energy for the pristine graphene, $\lambda = \pm$, the band index, $v_F$ the electron momenta, $|p| = p$, and $\phi_p = \arg(p_x + ip_y)$. $v_f$
evaluated analytically the self-energy $\Sigma$ of impurity scattering in Dirac materials. Is the Fermi velocity of graphene. The energies are measured from the Dirac point of pristine graphene. This single Dirac-cone approximation captures the physics in the regime of interest and the degeneracy number $g = 4$ will be included at the end to include spin and valley degeneracy (inter-valley and spin flipping process are omitted).

Equation (6) is derived as an average over a distribution of identical impurities as $\Sigma_{imp}^R(\epsilon) = n_{imp}T_{imp}(\epsilon)$. The scattering off an impurity is described by the single impurity T-matrix, $T_{imp}$. In the Dirac-cone approximation, where the momentum sum $\sum_{\mathbf{p}} G(p, \epsilon)$ can be evaluated analytically, we have

$$T_{imp}(\epsilon) = \frac{1}{V_{imp}} - \frac{1}{\epsilon_c^2} \ln \left( \frac{\epsilon_c}{\epsilon - z} \right).$$

with $z = \epsilon - \Sigma_{imp}^R(\epsilon)$ and $\epsilon_c$ is a cut-off that we set to the bandwidth of the graphene band structure. The poles of $T_{imp}$ describe the well-known impurity states which in the limit $V_{imp} \gg \epsilon_c$ and $n_{imp} \to 0$ (i.e. $\Sigma_{imp}^R \to 0$ so that $z \to \epsilon + i\gamma$) are localised states at energies

$$\epsilon_{imp} = \frac{\epsilon_c^2}{2V_{imp}} \ln \left| \frac{\epsilon_c}{2V_{imp}} \right|.$$ 

These low-energy impurity states are a generic feature of impurity scattering in Dirac materials. Turning to the self-energy $\Sigma_{imp}(\epsilon)$, the finite impurity density $n_{imp}$ leads to a narrow metallic band around the energy $\epsilon_{imp}$ with a width,

$$\gamma_{imp} \approx \sqrt{\frac{n_{imp}}{2V_{imp} \epsilon_c}}.$$ 

$\Sigma_{imp}^R(\epsilon)$ has a simple pole structure with a complex pole $\epsilon_{imp} = \epsilon_{imp} + i\gamma_{imp}$ indicating that a distribution of impurities induces scattering resonances rather than proper long-lived states.

Self-consistent solution of equations (5) and (6) is straightforward by simple iteration. The band structure, given by $E_{\lambda,p} = \epsilon \pm \text{Re}[\Sigma_{imp}^R(\epsilon)]$, is modified already for quite dilute impurity densities. The basic results are shown in figure 1. The impurity state, $\epsilon_{imp}$ is modified into a resonance that we define as $\text{Re}[\Sigma_{imp}^R(\epsilon_{res})] = 0$. The resonance $\epsilon_{res}$ is shifted towards smaller energies (absolute magnitude) compared to $\epsilon_{imp}$. There is also an impurity dependent shift of the Dirac point $E_{\lambda,p} = 0 = \epsilon_D - \text{Re}[\Sigma_{imp}^R(\epsilon_D)]$. In figure 2 we show how $\epsilon_{res}$ and $\epsilon_D$ depend on the inverse of the scattering strength.

In the Dirac point $\epsilon_D$ is a non-monotonic function of the inverse scattering strength $-1/V_{imp}$ and increases in magnitude with increasing impurity density. Finally, we plot $-\text{Im}[\Sigma_{imp}^R(\epsilon)]$ for different scattering strength in figure 1. For strong scatterers $-\text{Im}[\Sigma_{imp}^R(\epsilon)]$ has close to a Lorentzian lineshape around $\epsilon_{res}$ while for weaker scatterers $-\text{Im}[\Sigma_{imp}^R(\epsilon)]$ has a wider spread still with weak a maximum at $\epsilon_{res}$. As $-\text{Im}[\Sigma_{imp}^R(\epsilon)]$ gives an effective frequency-dependent single-particle scattering rate $\Gamma(\epsilon) = \frac{\hbar}{2\pi(\epsilon)}$, we predict that weak impurities will introduce plasmon losses at all frequencies along the plasmon dispersion while the losses incurred from strong impurities are pronounced when the plasmon mode interferes with $\epsilon_{res}$.

In this article, $V_{imp}$ is chosen to be negative, giving rise to an impurity state in the conduction band, which is a common scenario suggested by first-principle studies. All considerations can be repeated for $V_{imp} > 0$ for which the sign of $\epsilon_{res}$ and $\epsilon_D$ are reversed.

### III. Optical and Transport Properties

We now focus on the nonlocal graphene conductivity, $\sigma(q, \omega)$. With the knowledge of the band structure, and following the method presented in Refs. 32,49, a simplified expression for the conductivity is obtained. The results obtained below are found to be identical if the sign of $E_F$ and $V_{imp}$ are reversed, due to the particle-hole symmetry.

To start out we consider the impact of disorder on the local conductivity $\sigma(q = 0, \omega)$. This has been investigated in detail (see, e.g., Refs. 23,24,26,28). We revisit it nevertheless briefly with attention on how impurities modify the DC conductivity. The impurity contribution to the transport scattering time ($\tau_{DC}(E_F)$) is given by the relation

$$\frac{1}{\tau_{DC}(E_F)} = \frac{e^2 v_F^2}{2} \rho(E_F) \sigma(q = 0, \omega \to 0)$$

where $v_F$ is the Fermi velocity of graphene, $E_F$ is the Fermi energy, and $\rho(E_F)$ is the density of states at the Fermi energy.
in the limit $T \to 0$. The competition between the density of states and the DC conductivity gives rise to a non-trivial relation between Fermi energy and transport scattering time as shown in Fig. 3. The relaxation time is non-symmetric in $E_F$ for small $V_{\text{imp}}$ and a symmetric behaviour is recovered for large $V_{\text{imp}}$. Fig. 3 suggests that a chemical potential with the same sign as the more common impurity strength may increase $\tau_{\text{DC}}(E_F)$. For $|E_F| > 0.1 \text{eV}$ temperature effects on $\tau_{\text{DC}}(E_F)$ are on order of 1% or smaller up to 150K. The scattering time scales with the density of impurities as $\tau_{\text{DC}}(E_F) \propto n_{\text{imp}}$ as expected. In the following we focus mostly on the impurity density order of $n_{\text{imp}} \propto 10^{-3}$. This density of impurities gives a relaxation time that is suitable for plasmonic application in the THz regime.

Now we turn to the consequence of the impurity band on the nonlocal conductivity. According to the Fermi Golden rule, the lossy part of the conductivity ($\sigma_1$) gives the possibility of the electromagnetic field to release energy to the carriers in graphene by exciting electron-hole pairs. In pristine graphene, the $\sigma_1(q, \omega)$ is a triangle in the $(\omega, q)$-plane, given by $\hbar \omega_F q < \omega < 2\omega - \hbar \omega_F q$, where absorption is forbidden, i.e., $\sigma_1(q, \omega) = 0$ at $T = 0$. Absorption is allowed outside this Pauli-blocked triangle, i.e. for $\omega \leq \hbar \omega_F k$ and $2E_F - \hbar \omega_F k \leq \hbar \omega$. The absorption occurs in intraband and interband transitions respectively. It has also been shown that for pristine graphene the response depends only on one energy scale, the Fermi energy, which scales all energies $(\omega, E_{\lambda, p}$ and $k_B T)$.

In figure 4 we plot $\sigma_1(q, \omega)$ at $q/k_F = 0.15$ as function of frequency $\omega$. The impurity specific features we find are interband processes corresponding to transitions between the impurity band around $\epsilon_{\text{res}}$ and the states around Fermi energy. The transitions generate an extra peak in $\sigma_1(q, \omega)$ at frequencies $\hbar \omega \approx |\epsilon_{\text{res}} - E_F|$ added to the main peak at $\hbar \omega/E_F = 0.15$ as seen in figure 4. These transitions are single particle excitations. In the lower panel of Fig. 4 an extended range of density of impurities is considered. The extra peak in $\sigma_1(q, \omega)$ follows $\epsilon_{\text{res}}$ and is actually a band-like area in the $(\omega, q)$-plane around $\omega \simeq |E_F - \epsilon_{\text{res}}|$ where there are increased losses, independent of $q$. The width of this stripe is given mainly by $-\text{Im} \Sigma_R^{\text{imp}}(\epsilon)$. Temperature effects are also important but only at high temperatures. The lower panel of Fig. 4 shows how the density of impurities, and hence $-\text{Im} \Sigma_R^{\text{imp}}(\epsilon)$, affects transport properties. Since $\Sigma_R^{\text{imp}} \propto n_{\text{imp}}$, increasing the impurity density all features in the conductivity are broadened and $\sigma(q, \omega)$ tends to $\sigma_0 = e^2/4\hbar$. For a given $n_{\text{imp}}$, weaker impurities has a stronger effect on AC transport. For the range of parameters explored, the impurity induced peak emerges distinctly above the background at $T < 200K$. The new features in $\sigma_1$ reflect in a non-monotonic behaviour of $\sigma_2$, according to the Kramers-Kronig relations. Similar features are given by the impurity states in the presence of adatoms on the graphene surface.

![FIG. 3. (Color on line): The relaxation time induced by impurities obtained from equation (9) as function of the Fermi energy. In the plot, the density of impurities is $n_{\text{imp}} = 10^{-4}$, the relaxation time roughly scales as $1/n_{\text{imp}}$. In the inset a zoom for small value of $|E_F|$. Different colors and marks correspond to different value of $V_{\text{imp}}$ as indicated by the legend.](image)

![FIG. 4. (Color on line): Real part of the nonlocal conductivity $\sigma_1(q, \omega)$ as function of the frequency. In the top panel the impurity density is set to $n_{\text{imp}} = 10^{-4}$ and different lines correspond to different values of impurity strength $V_{\text{imp}}$. The fine magenta line is the pristine graphene case. Impurities induce peaks in the losses at frequencies $0.4 E_F \leq \hbar \omega \leq E_F$ for the values of $V_{\text{imp}}$ considered in the figure. The position of the peak depends on $V_{\text{imp}}$, and increasing the strength induces a blue shift of the feature. The lower plot shows the influence of the density of impurity for $V_{\text{imp}} = -100\text{eV}$. Here, $\sigma_1(q, \omega)$ is computed for $T = 30 \text{K}$, $E_F = 0.2 \text{eV}$ and $|q| = 2\pi/\lambda$ with $\lambda = 130 \text{nm} (q/k_F \simeq 0.15)^7$.](image)
In Fig. 5 the top row shows the plasmon dispersion relation \( \omega(q_1) \) and the propagation length \( L_P = 1/2q_2 \). In Fig. 5 the top row shows the plasmon dispersion relation \( \omega(q_1) \). The bottom row in Fig. 5 presents the propagation length \( L_P(\omega) \) in units of the plasmon wave length \( \lambda_P = 2\pi/q_1 \), for the same values of \( V_{\text{imp}} \) and \( n_{\text{imp}} \) as in the panel directly above. In each column two values of impurity density are shown, \( n_{\text{imp}} = 5 \times 10^{-5}, 10^{-4} \), and the impurity strength changes with the column. The dispersion relation obtained from the impurity-doped graphene is compared with results from a relaxation time approximation using the DC relaxation time value computed according to the finite temperature equivalent of Eq. (9). The analysis of the losses shows a disagreement between the two approaches as was observed in Refs. and here confirmed in a self-consistent \( \tau \)-matrix model. The effects of the impurities are fully considered also in the evaluation of the dispersion relation \( \omega(q_1) \). As seen in the figure there is quite a discrepancy between the relaxation time approximation and our impurity model. The impurity model show clear signature of the impurity resonance as the frequency is swept. This is particularly clear for the strong scattering case. For weaker scatterers we also see as structure in the propagation length at \( \epsilon_{\text{res}} \) as well as a shift in the plasmon dispersion at \( h\omega = \epsilon_{\text{res}} \). While the relaxation time approximation is able to show damping, it is only by using a microscopic model it is possible to obtain a quantitative description the damping processes of plasmons and of the optical response of graphene.

To analyse the dispersion relation in more detail we use a rather large value of the chemical potential \( E_F = 0.4 \text{ eV} \) and a temperature of 30K. The results are shown in Fig. 5. The purpose of this choice is to enhance the visibility of the effects of the impurity band. Thanks to the approximate scale invariance of the system, the main features remain valid also for smaller values of the Fermi energy. However, the position of the impurity resonance needs to be wisely rescaled and one must keep in mind that line shape of \( -\text{Im}\Sigma^R_{\text{imp}}(\epsilon) \) becomes broader the further away \( \epsilon_{\text{res}} \) is from the Dirac point. The left panel of Fig. 5 shows the case of strong impurities \( (V_{\text{imp}} = -10^4 \text{ eV}) \), this may represent a graphene lattice with dislocations or holes in it. The impurity resonance is expected around \( \epsilon_{\text{res}} \approx 0.02 \), 0.04eV for the densities used in the plot \( n_{\text{imp}} = 5 \times 10^{-5}, 10^{-4} \) (full and dashed lines) respectively. The signature of \( \epsilon_{\text{res}} \) is a marked drop in the propagation length, \( L_P/\lambda_P \) at \( h\omega/E_F \approx 0.96 \) for \( n_{\text{imp}} = 5 \times 10^{-5}(10^{-4}) \). This brings us to a first conclusion: holes and dislocations in the graphene crystal reduce the bandwidth of the plasmons to \( h\omega < E_F \) from the range \( h\omega < 1.3E_F \) that the relaxation time approximation approach suggests. We do not consider effects of phonons to underline pure impurity effects. According to the literature an optical phonon introduces an extra bound to the working frequency of graphene plasmons to \( h\omega < 0.2\text{eV} \). In Fig. 5, the second and third column, present the dispersion relation for impurities bands that lie away from the Dirac point. The column in the middle display the case when \( V_{\text{imp}} = -100 \text{eV} \) and corresponds to an impurity band around \( \epsilon_{\text{res}} \approx 0.082 \text{eV} \) and \( \epsilon_{\text{res}} \approx 0.093 \text{eV} \) for \( n_{\text{imp}} = 5 \times 10^{-5} \) and \( n_{\text{imp}} = 10^{-4} \) respectively. A clear increase in over-all losses appear and close to \( h\omega \approx E_F - \epsilon_{\text{res}} \) we see a signature of \( \epsilon_{\text{res}} \) as dip in \( L_P \). In the right column, the dispersion relation for impurities with strength \( V_{\text{imp}} = -60 \text{eV} \). Now the impurity band is even higher up in the conductance band compared to the case with \( V_{\text{imp}} = -100 \text{eV} \) and we find \( \epsilon_{\text{res}} \approx 0.15 \text{eV} \) and \( \epsilon_{\text{res}} \approx 0.16 \text{eV} \) for the two densities. This reflects the even more lossy conductance and \( L_P/\lambda_P \lesssim 1 \) for all frequencies. For large impurity densities, \( n_{\text{imp}} > 10^{-3} \), the longitudinal plasmons appear to be overdamped according to Eq. (2) and it may not be appropriate to speak about modes. This suggests one obvious reason why graphene of to low quality is not suitable for (longitudinal) plasmonic applications.

The comparison between temperature and finite momenta losses, considered in, and impurity losses reveals that the last are dominating up to room temperature for \( \omega/E_F < 1.2 \) and for \( n_{\text{imp}} > 5 \times 10^{-5} \). At lower density \( n_{\text{imp}} \approx 10^{-5} \) the two source of losses are comparable in the range of frequency \( 1 < \omega/E_F < 1.4 \) and impurity losses are dominant at lower frequency.

### IV. CONCLUSION

In this work, a microscopic model of plasmons that considers together impurities and nonlocal effects, has been developed and analysed. The present work also contributes to shaping a full microscopic picture of the plasmon in graphene with the long term aim to develop further the design of plasmon devices.

The first principle computations suggest that impurities on the graphene surface introduce an almost flat impurity band close to the Dirac point, whose width and position in energy depend on the type and density of impurities. This is the main feature that is newly included in theoretical description in this work. An impurity model based on the \( \tau \)-matrix formalism in the Green’s function framework is developed and analysed. The
The plasmon dispersion relation, $\hbar \omega_P(q)$, and propagation length, $L_P/\lambda_P = q_1/(4\pi q_2)$ are plotted for different values of impurity strength and density. The impurity strength from the left to the right takes the values $V_{\text{imp}} = -1000, -100, -60$ eV. The impurity density $n_{\text{imp}} = 10^{-4}$ and $n_{\text{imp}} = 5 \times 10^{-5}$ are the dashed and full black lines respectively. $q_{1,2}$ were computed according to Eq. (2). The plasmon dispersion is given in units of $E_F/\hbar$ (here $E_F = 0.4$ eV). The propagation length is scaled of plasmon wavelength. The results are compared with relaxation-time approximation results evaluated at the marked scattering times $\tau_{\text{eff}}$. $\tau_{\text{eff}}$ is chosen so that for each pair ($V_{\text{imp}}, n_{\text{imp}}$) the DC relaxation time is extracted from relation (9) and used to compute the dispersion relations according to Ref. 3. In the top row: the grey dashed line is the single particle continuum $\hbar \omega_P = h\nu_0 q$, and the vertical dash-dotted orange line correspond to the wavelength of $\lambda = 130$ nm for $E_F = 0.4$ eV. The plots show a strong non trivial behaviour as a function of the parameters $n_{\text{imp}}$ and $V_{\text{imp}}$. The dip in the ratio $L_P/\lambda_P$ is in correspondence with the difference $|\mu - V_{\text{imp}}|$.

The possibility to identify the type of surface impurities from their effect on the optical response is an avenue to explore. Indeed the behaviour of the optical conductivity indicates that it may be possible to extract the value of impurity strength and impurity density, $V_{\text{imp}}$ and $n_{\text{imp}}$, and so extract the impurity resonance $\epsilon_{\text{res}}$. Finally the energy of the impurity position can be compared with the result from DFT simulations to determine the chemical compound present on the sample. The proposal for a sensor can take in account realistic laboratory constraints, so that it works at fixed incident light frequency and grating periodicity.

**ACKNOWLEDGMENTS**

The authors wish to thank the Knut and Alice Wallenberg (KAW) foundation for financial support. We also thank Tomas Löfwander for the stimulating discussion.

---

1. *gioga.viola@gmail.com*
In this work the propagation length is defined according to Ref. 30 as the distance that the waves propagate until the intensity is reduced by a factor $1/e$. According to Ref. 45 as the distance that the waves propagate until the intensity is reduced by a factor $1/e$.

For $E > 0$, the relevant transitions are $\text{sign}(E) = \text{sign}(V_{\text{imp}})$ can improve DC transport properties. To finally confirm this statement the effect of other source of relaxation (e.g. phonon and electron electron interactions) should be included at a self-consistent level 31. For $\epsilon_{\text{res}} - E < 0$, the transitions between single particle states in the impurity band and states above the Fermi energy contribute to an enhancement of losses at this frequency (while for $\epsilon_{\text{res}} - E > 0$ the relevant transitions are from the Fermi energy to the impurity level).


In this work the propagation length is defined according to Ref. 45 as the distance that the waves propagate until the intensity is reduced by a factor $1/e \approx 0.37$.

Current-controlled light scattering and asymmetric plasmon propagation in graphene

Tobias Wenger, Giovanni Viola, Jari Kinaret, Mikael Fogelström, and Philippe Tassin

1Department of Microtechnology and Nanoscience (MC2), Chalmers University of Technology, SE-412 96 Goteborg, Sweden
2Department of Physics, Chalmers University of Technology, SE-412 96 Goteborg, Sweden

(Dated: October 3, 2017)

We demonstrate that plasmons in graphene can be manipulated using a DC current. A sourcedrain current lifts the forward/backward degeneracy of the plasmons, creating two modes with different propagation properties parallel and antiparallel to the current. We show that the propagation length of the plasmon propagating parallel to the drift current is enhanced, while the propagation length for the antiparallel plasmon is suppressed. We also investigate the scattering of light off graphene due to the plasmons in a periodic dielectric environment and we find that the plasmon resonance separates in two peaks corresponding to the forward and backward plasmon modes. The narrower linewidth of the forward propagating plasmon may be of interest for refractive index sensing and the DC current control could be used for the modulation of mid-infrared electromagnetic radiation.

I. INTRODUCTION

Graphene has recently emerged as an exciting tunable plasmonic platform, which exhibits large electromagnetic field enhancements at terahertz to mid-infrared frequencies. The tunability is achieved by applying a gate voltage that allows the charge carrier density in graphene to be controlled externally. This opens up many exciting possibilities for creating tunable photonic devices that can potentially help bridge the gap between photonics and electronics.

Graphene plasmons have also been considered for sensing purposes and the large field enhancement can be utilized to tune molecules close to graphene. Such graphene sensors have been shown to be very sensitive to small amounts of molecules and the tunable nature of graphene plasmons enables selective sensing of specific molecules.

Nonlocal effects of plasmons in graphene, due to the fact that the plasmon wavelength can become comparable with the Fermi wavelength, have previously been discussed in various contexts. The dispersion of long-wavelength plasmons can be treated in a local approximation. For short-wavelength plasmons, however, it may be important to also consider nonlocal effects to correctly describe the plasmon behaviour. Recently, experimental verification of the need of a nonlocal approach was achieved in Ref., where graphene plasmons were confined by a metallicized nanotip in the direction perpendicular to the graphene sheet. In this article, we use a nonlocal response function to capture these effects. In fact, for the asymmetric plasmon effects that we find, a nonlocal description is necessary. Without nonlocal effects there is nothing that distinguishes between parallel and anti-parallel to the current direction and because of symmetry, plasmons in the two directions behave identically.

In addition to tuning the carrier concentration in graphene, it is also a possible to induce a stationary current by a voltage bias between a source and a drain contact, see Fig. 1. This may enable even more control over graphene plasmons and by extension also over light at the nanoscale. Indeed, DC currents in graphene have previously been utilized to tune plasmon resonances in metal bow-tie nanoantennas.

In this article, we study graphene plasmons in the presence of a constant drift current in graphene. The drift current leads to a lifting of the forward-backward plasmon degeneracy along the current direction and creates asymmetric plasmon propagation. Plasmons propagating in the direction of the electrons are enhanced in terms of propagation length and plasmons propagating against the electrons experience a decreased propagation length. We show that already moderate drift currents allow for control over the graphene plasmon propagation and thus enables electrical control over mid-infrared electromagnetic waves in the subwavelength regime.

Non-equilibrium plasmons in graphene have recently been considered in the literature, such as in Ref., where...
the authors studied the plasmon dispersion in the presence of a drift current and found asymmetries in the plasmon dispersion. In Ref. 29, graphene plasmons were investigated for small drift currents and an increased lifetime was found for plasmons propagating parallel to the drift velocity. Furthermore, Ref. 29 found that in the regime of low doping, there can be plasmon gain. Plasmon gain was also reported in Ref. 30, where graphene was considered to be driven out of equilibrium with a strong laser pulse and the gain was achieved when the electrons relax towards equilibrium, releasing energy as they do so. Graphene out-of-equilibrium has also been considered for terahertz amplification 31, as well as terahertz emission and detection 32,33.

In this article, we investigate graphene plasmons for any drift current. In contrast to previous studies, we also calculate the plasmon propagation lengths, and compute the light scattering by the out-of-equilibrium plasmons. We find that the drift-current-induced plasmon control and propagation asymmetries produce clear light-scattering signatures which could be obtained in experiments on gratings 3,34, ribbons 7, or using nanotips 35,36.

II. NON-EQUILIBRIUM PLASMONS

Graphene transverse magnetic (TM) plasmons in the non-retarded limit satisfy the dispersion equation 23,24

\[ 1 + \frac{iq}{\varepsilon_0(\varepsilon_1 + \varepsilon_2)\omega} \sigma(q, \omega) = 0 \]  

where \( \sigma(q, \omega) \) is the nonlocal sheet conductivity of graphene, and \( q = q_1 + iq_2 \) is the complex wave vector. Expanding this equation for small \( q_2/q_1 \) i.e., small plasmon losses, yields

\[ \varepsilon_1 + \varepsilon_2 + \frac{q_1 \sigma_2(q_1, \omega)}{\omega \varepsilon_0} = 0 \]  

\[ \frac{q_2}{q_1} = \frac{\sigma_1(q_1, \omega)}{\sigma_0(q_1 \sigma_2(q_1, \omega))} \]

where \( \sigma(q_1, \omega) = \sigma_1(q_1, \omega) + i \sigma_2(q_1, \omega) \). The first of these equations can be solved for a given \( \omega \) to obtain \( q_1 \), which can be inserted in the second equation to give \( q_2 \), which is a measure of the plasmon damping. The conductivity of graphene can be obtained from a linear response calculation of the polarizability 20,21

\[ \Pi(q, \omega) = \lim_{\nu \to 0^+} \frac{g_0 g_e}{2} \int \frac{d^2 k}{(2\pi)^2} f_{k,k'} - f_{\nu k,k'} \times \left( 1 + \frac{\lambda' \cos(\phi_0,k,{k'})}{1 + i \lambda' \varphi_0} \right) \]

where \( k' = k + q \), and \( \phi_{0,k'} = \phi_{0,k} - \phi_0 \). The conductivity can then be treated using Mermin’s approach to include a phenomenological relaxation time \( \tau \) in a number-conserving way 24,37:

\[ \sigma(q, \omega) = \frac{i e^2 \omega (1 + i \Gamma)}{q^2} \frac{(1 + \frac{\Gamma}{\omega})\Pi(q, \omega + i \Gamma)}{1 + \frac{\Gamma}{\omega} \frac{\Pi(q, \omega + i \Gamma)}{\Pi(q, 0)}} \]

The phenomenological parameter \( \Gamma = \tau^{-1} \) accounts for intraband scattering that is known to exist from experiments and to cause significantly increased plasmon damping 24. We take \( \tau = 170 \) fs, which is achieved in experiments 38,39. The plasmon propagation distance in units of the plasmon wavelength can be obtained by \( \lambda_p/\gamma_0 = q_1/(4\pi q_2) \) 40,41, which gives the number of oscillations the plasmon makes before it decays.

The stationary electron distribution in the presence of relaxation and a drift velocity can be written as 42

\[ f_{\lambda,k} = \frac{1}{1 + e^{(\lambda E_k - \nu_F |\vec{\beta} - \mu|)/(k_B T)}} \]

where \( \vec{\beta} \) is the drift velocity vector scaled with \( \nu_F \) and \( E_k = \nu_F |k| \) is the linear dispersion for Dirac electrons. When \( |\vec{\beta}| = 0 \), Eq. (8) reduces to the ordinary Fermi distribution. Throughout the article we take \( \vec{\beta} \) to be either parallel or anti-parallel to the plasmon wave vector \( \vec{q} \). Eq. (8) describes an asymmetric electron distribution where the asymmetry is parametrized by the parameter \( |\vec{\beta}| \), which takes values between +1 and −1. The distribution given by Eq. (8) does not conserve the number of particles as \( \beta \) increases from zero. Since the system we
We next consider the effect of a drift current on the optical properties of graphene. As shown before\textsuperscript{7,8,34}, these optical properties are intimately linked with plasmons, and we evaluate them using COMSOL, a finite element method (FEM) solver. Graphene enters the solver as a perfectly matched layer (PML) region above $E_F$ — is the region with the strongest localization and is typically heavily damped\textsuperscript{40,43}. These highly localized plasmons have the potential to facilitate strong coupling to emitters in its vicinity\textsuperscript{1} and the drift current supplies a method by which their damping can be suppressed. However, care must be taken to not heat the graphene sample extensively as this leads to larger plasmon losses and, hence, smaller propagation lengths.

\section{Light Scattering from Non-equilibrium Plasmons}

We now turn our attention to the plasmon propagation and how it is affected by the drift current. Fig. 4 shows the plasmon propagation length as a function of the plasmon energy for different values of the drift velocity $\beta$. Plasmons propagating parallel to the drift velocity experience an enhanced propagation length, whereas plasmons propagating anti-parallel to the drift velocity have a suppressed propagation. This effect allows for external control and enhancement of the plasmon propagation without increasing the mobility of the graphene samples. In addition, the region where the effect of the drift current is most pronounced — the high-energy region above $E_F$ — is the region with the strongest localization and is typically heavily damped\textsuperscript{40,43}. These highly localized plasmons have the potential to facilitate strong coupling to emitters in its vicinity\textsuperscript{1} and the drift current supplies a method by which their damping can be suppressed. However, care must be taken to not heat the graphene sample extensively as this leads to larger plasmon losses and, hence, smaller propagation lengths.

\begin{align}
\sigma(x \rightarrow x', \omega) = \int \sigma(x \rightarrow x', \omega) E(x', \omega) \, dx'.
\end{align}

where the integration is over the graphene sheet and the electric field is the in-plane component along the graphene sheet. By considering the grating region to be infinite in extension, the system can be simulated using a unit cell containing one period of the grating and periodic boundary conditions. The space-dependent graphene conductivity is calculated from the nonlocal
Figure 5. Extinction curves as a function of incident wavelength for various values of $\beta$. As $\beta$ is increased the plasmon peak separates into two distinguishable modes, related with plasmons propagating parallel and anti-parallel to the drifting electrons. The separation of the modes grows as the magnitude of $\beta$ grows and for values around $\beta \geq 0.4$, the two peaks are visible.

The effects of the drift velocity on the plasmon propagation allows not only for increased propagation distances in experimental settings, but also creates new possibilities for creating plasmonic graphene devices with active control over the plasmons. The current creates a region of essentially unidirectional plasmon propagation that can be turned on and off by the current and the direction can also be reversed by reversing the current. This can potentially be used to create plasmonic modulators where the DC electric current modulates the plasmon propagation and thus the plasmon signal. This could be lead to compact photonic devices, with sizes comparable to the plasmon wavelength on the order of 100 nm, in the mid-infrared.

Direct measurements of the enhanced plasmon propagation distance should be possible using nanotip experiments similar to Refs. $^{35,36}$. In these experiments, only one nanotip was used and the plasmons were launched from the tip, reflected off a boundary, and returned to the tip for out-coupling to a photon detector. In such a setup, the plasmon necessarily travels both parallel and anti-parallel to the induced current making the results more difficult to interpret. However, since the effects of propagation enhancement and increased losses are not each other’s inverse, it should still be possible to observe the effects of the DC current in a one-tip experiment. We point out, however, that a two-tip experiment, where the plasmon is launched at one tip and outcoupled at the other, would provide a measurement signal that is more directly related to the results obtained in this article.

The mechanism discussed here could be important for the modulation of mid-infrared signals, since the drift control allows control over the extinction. It can also be used in experiments where control over the linewidth is desired.

V. CONCLUSIONS

We have shown that graphene plasmons can be controlled by a DC current in the graphene sheet. This leads to asymmetric plasmon propagation and, in particular, plasmons propagating in the same direction as the electrons in the current experience an enhanced propagation length. We find that optical signatures of this asymmetry can be detected by light scattering in a subwavelength grating environment, where the equilibrium plasmon peak separates into two separate peaks as the DC current increases. The two peaks correspond to plasmons propagating parallel and anti-parallel to the current.

Furthermore, we discussed the implications of asymmetric plasmon propagation. This leads to narrower linewidths of the plasmons, which could be useful for sensing applications utilizing graphene plasmons. The ability to control terahertz and mid-infrared wavelengths...
using DC current control over graphene plasmons could enable compact modulators at these wavelengths.

The effects we discussed in this article can be measured directly in optical scattering experiments, or, alternatively, the asymmetric plasmon propagation can be measured using nanotips.