Spectroscopic Near-Field Imaging in the Mid-Infrared

William S. Hart

September 2018

Experimental Solid State Group
Department of Physics
Imperial College London
United Kingdom

Thesis submitted for the degree of Doctor of Philosophy of Imperial College London and for the Diploma of the Imperial College
Declarations

Declaration of Originality

The work contained within this thesis is my own, and all else has been appropriately referenced.

Copyright Declaration

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.
Abstract

This thesis focuses on two different approaches of beating the diffraction limit in the mid-infrared. One approach uses scattering-type scanning near-field optical microscopy (s-SNOM) to map the optical response of a sample surface with < 10 nm resolution. By utilising the recent development of widely tuneable quantum cascade lasers (QCL), this work experimentally investigates two distinct applications. The first maps out the dispersion of plasmons in graphene to determine the local Fermi energy with high accuracy. In addition, novel plasmon focusing phenomena are presented at interfaces between single and bilayer graphene.

The second experimental application uses infrared s-SNOM to image, for the first time, the ultrastructure inside individual human cells with ~ 8 nm resolution. Furthermore, nanoscale infrared spectroscopy is used to map out the location of a clinically relevant anti-cancer drug within a single cell without chemical labelling, for the first time.

The final part of this thesis comprises a theoretical study of metamaterial superlenses, which allow for deeply sub-wavelength spatial resolution to be preserved. The insights gained here led to the development of a novel infrared superlens that is shown to improve the distance at which sub-diffraction limited (< λ/15) features are preserved by a factor of 25 over previous designs.
Acknowledgements

There are many people without whom this thesis would not have possible, not least my supervisor Chris Phillips. His “hands off” approach has undoubtedly pushed me to be a better researcher, yet he always managed to have an open door in case I needed advice.

When it comes to experimental knowledge, mechanical skills, variegated conversation, and the ability to get things done, none have been more helpful than Hemmel Amrania. Similarly, were it not for Megumi Ito’s organisational prowess and optics expertise, many a setback would have left me out in the cold.

Much of the work in this thesis would not have been possible without the technical wizardry of Vishal Panchal and Christos Melios, or the illuminating insight of Alex Bak. I would also have had a much rockier start without Ned Yoxall’s practical knowledge.

Last, but certainly not least, I would like to thank my parents, my brother David, and Rob, whose unwavering support helped me to stay on the path that I have long dreamed of.
# Contents

1 Introduction and Motivation ................................................. 1
   1.1 Near-field imaging .................................................. 2
   1.2 Scattering-type SNOM ................................................. 3
   1.3 Metamaterial super-resolution ...................................... 4
   1.4 Thesis outline ....................................................... 4

2 Background and Theory ..................................................... 6
   2.1 Infrared spectroscopy ................................................ 6
      2.1.1 Infrared spectroscopy for chemical analysis .............. 6
      2.1.2 Limitations of far-field infrared spectroscopy ........... 8
      2.1.3 Raman spectroscopy ........................................... 9
   2.2 Beating the diffraction limit ...................................... 11
      2.2.1 Fluorescence techniques and their limitations ............ 11
      2.2.2 Electron microscopy .......................................... 11
      2.2.3 Atomic force microscopy ..................................... 12
      2.2.4 Photothermal infrared microscopy ............................ 13
      2.2.5 The optical near-field ....................................... 14
   2.3 Near-field microscopy .............................................. 16
      2.3.1 Scattering-type scanning near-field optical microscopy (s-SNOM) . 16
      2.3.2 Chemical contrast and spectroscopic imaging .............. 19
      2.3.3 Contrast from intrinsic optical fields ..................... 19
   2.4 Metamaterial superlenses ......................................... 20
      2.4.1 A brief introduction to metamaterials ....................... 20
      2.4.2 Anisotropic materials ....................................... 21
      2.4.3 Super-resolution from strong optical anisotropy ........... 22
# Mid-Infrared Plasmon Spectroscopy

## 3.1 Graphene plasmons
- An introduction to graphene
- A primer on surface plasmon-polaritons

## 3.2 Correlative nano-imaging of graphene
- Fermi energy measurement of graphene
- Graphene sample preparation
- Measuring $E_F$ with Kelvin probe force microscopy
- Measuring $E_F$ via the Hall effect
- Plasmon nano-imaging
- Dispersion relation of plasmons in homogeneous graphene
- High accuracy measurement of $E_F$ with widely tuneable real-space imaging of plasmon dispersion
- Plasmon reflection from 1LG/2LG interfaces

## 3.3 Conclusions

# Spectroscopic Infrared Nano-Imaging in the Life Sciences

## 4.1 Existing super-resolution techniques
- Limitations of electron microscopy
- Limitations of atomic force microscopy
- Limitations of fluorescence microscopy
- Limitations of photothermal microscopy

## 4.2 The key advantages of infrared nano-imaging
- Label-free chemical information
- Imaging in atmospheric conditions

## 4.3 Optical nano-imaging of human cells
- Testing biological s-SNOM with human red blood cells

## 4.4 Mid-infrared chemical nano-imaging for intra-cellular drug localisation
- Bortezomib as a test drug for intra-cellular chemical analysis
- Cell preparation
- Imaging sub-cellular ultrastructure with s-SNOM
- Chemical-specific mapping of Bortezomib in a single cell

## 4.5 Conclusions
5 Super-Resolution Imaging with Low-Loss Metamaterial Superlenses 57
  5.1 Creating anisotropy with periodic layers .................................. 57
  5.2 Tuning anisotropy with doped semiconductors ................................. 58
    5.2.1 The loss figure of merit .................................................. 60
  5.3 Angular dependence of losses ................................................... 62
  5.4 Requirements for a “good” superlens .......................................... 63
    5.4.1 How to achieve large anisotropy ......................................... 64
    5.4.2 How to achieve low absorption .......................................... 65
    5.4.3 Optimisation constraints .................................................. 65
  5.5 Novel design for a low-loss superlens ....................................... 66
    5.5.1 Tuneability of the superlens .............................................. 68
  5.6 Full-wave simulations ............................................................. 69
    5.6.1 The finite-difference time domain method ................................ 69
  5.7 Confirming low-loss super-resolution ....................................... 71
    5.7.1 Validity of the effective medium approximation .......................... 73
  5.8 Conclusions ........................................................................... 74

6 Conclusion and Further Work 75
  6.1 Conclusion ............................................................................. 75
  6.2 Further Work .......................................................................... 76
## Nomenclature

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1LG</td>
<td>Single-layer graphene</td>
</tr>
<tr>
<td>2LG</td>
<td>Bilayer graphene</td>
</tr>
<tr>
<td>BTZ</td>
<td>Bortezomib</td>
</tr>
<tr>
<td>EM</td>
<td>Electron microscopy</td>
</tr>
<tr>
<td>EMA</td>
<td>Effective medium approximation</td>
</tr>
<tr>
<td>EMM</td>
<td>Elliptical metamaterial</td>
</tr>
<tr>
<td>ENP</td>
<td>Epsilon-near-pole</td>
</tr>
<tr>
<td>ENZ</td>
<td>Epsilon-near-zero</td>
</tr>
<tr>
<td>FDTD</td>
<td>Finite-difference time domain</td>
</tr>
<tr>
<td>FFPE</td>
<td>Formalin-fixed, paraffin-embedded</td>
</tr>
<tr>
<td>FTIR</td>
<td>Fourier transform infrared (spectroscopy)</td>
</tr>
<tr>
<td>HMM</td>
<td>Hyperbolic metamaterial</td>
</tr>
<tr>
<td>IDC</td>
<td>Iso-frequency dispersion contour</td>
</tr>
<tr>
<td>IFL</td>
<td>Interfacial layer</td>
</tr>
<tr>
<td>KPFM</td>
<td>Kelvin probe force microscopy</td>
</tr>
<tr>
<td>LSPR</td>
<td>Localised surface plasmon resonance</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate-buffered saline</td>
</tr>
<tr>
<td>PML</td>
<td>Perfectly matched layer</td>
</tr>
<tr>
<td>QCL</td>
<td>Quantum cascade laser</td>
</tr>
<tr>
<td>QFS</td>
<td>Quasi-free-standing</td>
</tr>
<tr>
<td>RBC</td>
<td>Red blood cell</td>
</tr>
<tr>
<td>s-SNOM</td>
<td>Scattering-type scanning near-field optical microscopy</td>
</tr>
<tr>
<td>SNOM</td>
<td>(Aperture) scanning near-field optical microscopy</td>
</tr>
<tr>
<td>SPM</td>
<td>Scanning probe microscopy</td>
</tr>
<tr>
<td>STED</td>
<td>Stimulated emission depletion (microscopy)</td>
</tr>
<tr>
<td>TE/M</td>
<td>Transverse electric/magnetic</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction and Motivation

Light has always been the primary way of viewing the world. To get a closer look, the first optical magnifying elements were invented in the second century BC [1]. Ever since, there has been a continual effort to image smaller and smaller objects. The boundaries of microscopy were pushed ever downwards, from small insects to individual cells, until a seemingly hard limit was hit upon. This limit is caused by diffraction, whereby waves, such as classical light, spread out upon passing through some aperture (see Figure 1.1), and scales in proportion to the wavelength, $\lambda$, of the light [2]. This led to the development of a variety of new techniques that either reduce or break this limit. One approach is to reduce the wavelength of the imaging wave, by using higher frequency electromagnetic radiation. The ultimate limit of this approach is to use electrons, which can achieve sub-Ångstrom wavelengths with sufficient energy [3], in electron microscopy (EM). However, EM samples suffer from the harsh requirement of an ultra-high vacuum [4]. Furthermore, EM loses out on the deep information contained in the light spectrum, such as the rich chemical information of the mid-infrared [5].

Other approaches seek to break the diffraction limit altogether by artificially attaching fluorophores to the sample. One such method, stimulated emission depletion microscopy (STED) uses a second light source to selectively deplete emission in all but a small, sub-diffraction limited region [6], while others rely on the temporal and/or spatial separation between fluorescence events to digitally confine them to the centre of their imaged spot, such as in photo-activated localisation microscopy (PALM) [7].

A third class of approach is ascribed the name of scanning probe microscopy (SPM). Here, a sharp probing tip is brought into contact with the sample and raster scanned
Figure 1.1: A wave passing through an aperture diffracts, creating a smeared-out diffraction pattern in the image plane.

across its surface, in a manner analogous to the needle of a vinyl record player, to form an image with resolution defined by the size of the tip apex (typically $< 10$ nm) [8]. These techniques offer the possibility to measure many properties, from topography in an atomic force microscope (AFM) to surface potential in a Kelvin probe force microscope, and many more [9].

### 1.1 Near-field imaging

A particularly interesting application of SPM is to place a probe in the optical “near-field” of a sample. The near-field is the region very close to the sample relative to the wavelength of the illuminating light, i.e. $d \ll \lambda$. In this region, the fine spatial information that is lost in the far-field ($d \gg \lambda$), survives [10]. In order to collect this image, and beat the diffraction limit, light must typically be collected within 10’s of nanometres of the sample surface, as the near-field is populated by exponentially decaying, or “evanescent” waves [10]. In fact, this is not a new idea. In 1928, in a series of correspondence with Albert Einstein [11, 12], Edward Synge first proposed bringing a “very small hole in an opaque plate” into the near-field, which would allow only a sub-diffraction limited region of light to pass through. Lesser known is that Synge went on to suggest that:

A better way could be, if one could construct a little cone or pyramid of quartz glass having its point P brought to a sharpness of order $10^{-6}$ cm. One could then coat the sides and point with some suitable metal [...] and then remove the metal
1.2. Scattering-type SNOM

from the point, until P was just exposed.

This remarkably prescient suggestion predicted the development of aperture scanning near-field optical microscopy (SNOM) in 1986 by Betzig et al. [13], which uses a sharp, metal-coated optical fibre with a small aperture at its tip. SNOM became a commonly used technique for mapping the optical response of a sample surface, although its resolution is generally limited to \( \sim 100 \) nm by the decrease in the amount of light that can be collected as the aperture becomes smaller [13]. Indeed, this means that the resolution still depends inversely on the light wavelength, albeit to a lesser degree.

1.2 Scattering-type SNOM

Scattering-type SNOM (s-SNOM) removes this remaining barrier to achieving the sub-10 nm resolution of AFM by replacing the optical fibre with an entirely metal-coated probe without an aperture [14]. Rather than collecting the near-field light, s-SNOM scatters it into the far-field, which provides more than just an improvement in resolution, but also increases the strength of the interaction with the near-field because the metal tip acts as a “lightning rod” [15].

This strong interaction with the near-field, coupled with genuinely wavelength-independent resolution, allows s-SNOM to operate in the mid-infrared. This allowed Huth et al. [16] to demonstrate chemical specific mapping with \( \sim 10 \) nm spatial resolution in 2012, by using a thermal infrared source to perform nanoscale infrared spectroscopy. However, spectroscopic s-SNOM using thermal sources is limited by the poor signal-to-noise of incoherent light, in addition to the difficulties of focusing a broadband beam [17].

Since then, the development of s-SNOM has been closely tied to the development of new light sources. CO\(_2\) lasers provided a coherent source with high signal-to-noise ratio, but suffer from limited tuneability to discrete spectral lines [17]. Now, the maturing technology of the quantum cascade laser (QCL) is allowing, for the first time, broad tuneability of a coherent source in the mid-infrared [17]. These new lasers are just now ready to make spectroscopic infrared s-SNOM a powerful tool for chemical mapping of real life science systems.
1.3 Metamaterial super-resolution

The optical response of a material is determined by the average response of the arrangement of the atoms that comprise it. The limited range of naturally occurring materials ultimately limits the optical properties that can be found. This is why the vast majority of natural materials have an isotropic and positive refractive index [18].

Metamaterials create geometric structures out of bulk materials on length scales smaller than the light wavelength to create new and unusual optical responses. One highly unusual property that was first theorised by Veselago [19] in 1968, was that of a negative refractive index. One astounding consequence of negative refractive index is that the metamaterial would result in amplification the near-field, in contrast to the exponential decay in positive index materials. It was shown by Pendry [20] in 2000 that this would result in a “perfect” lens, which would allow imaging below the diffraction limit. Unfortunately, such perfect lenses have proven to be extremely challenging to fabricate, and subject to high absorptive losses [21].

An altogether different approach to super-resolution relies on creating metamaterials with highly anisotropic permittivity. This allows for the near-field to propagate through the material without diffraction, so preserving deeply sub-wavelength spatial detail [22]. A particularly simple arrangement to create such a “superlens” is that of a 1D-periodic multilayer comprising alternating dielectric and metal layers. However, multilayers involving metals are not tuneable and suffer from heavy losses [23]. An alternative method uses alternately doped and undoped semiconductors, which are tuneable through doping [24], but once again have so far been hampered by losses.

1.4 Thesis outline

This thesis begins with a review of the principles IR spectroscopy, which is widely used for chemical identification [25], before summarising the existing methods of beating the diffraction limit. Then, the theory and experimental details of s-SNOM, and theory of metamaterial superlenses, are laid out to set the stage for the rest of this work.

The first experimental work on graphene plasmonics is presented in chapter 3. Here, spectroscopic s-SNOM is shown for the first time to provide a more accurate platform for measuring the Fermi energy of graphene, than existing techniques, such as the Hall
effect. Furthermore, a new phenomenon is discovered, whereby plasmons are focused by single-layer/bilayer interfaces.

The following chapter (chapter 4) develops a new direction for spectroscopic s-SNOM by applying it to real-world life sciences problems. In particular, an answer is provided for the long unanswered question in medicine - where do drugs actually end up? For the first time, a clinically relevant anti-cancer drug is mapped within a single cell without the need for fluorescence labelling. This new technique offers a host of advantages over existing imaging techniques in the life sciences, such as the ability to image in atmospheric conditions, simple sample preparation protocols, and the ability to image the environment and the chemical of interest at the same time.

Chapter 5 considers, for the first time, the key parameters of anisotropic superlens performance. The resulting insights allow for a new superlens to be designed that maximises sub-diffraction imaging performance. Finally, using full-wave electromagnetic simulations, this new design is shown to give more than a factor of twenty-five improvement in deeply sub-wavelength image preservation.
Chapter 2

Background and Theory

In order to tackle the interesting challenges and opportunities raised in the previous chapter, this chapter must first provide an overview of the key background and theory of super-resolution imaging. In section 2.1, the infrared spectroscopy will be reviewed to show why it is such a powerful technique for studying chemical composition, and to understand its limitations in regards to spatial resolution due to the diffraction limit. Once the existing techniques for beating the diffraction limit have been considered in section 2.2, sections 2.3 and 2.4 and review two alternative methods based on imaging the optical near-field – scattering-type near-field optical microscopy (s-SNOM) and metamaterial superlenses.

2.1 Infrared spectroscopy

2.1.1 Infrared spectroscopy for chemical analysis

Every different chemical compound by definition comprises a unique combination of chemical bonds. Each of these bonds have a number of different vibrational modes [25], as shown in Figure 2.1. Each mode will resonantly absorb light at a particular frequency in the mid-infrared.

In order to develop some intuition here, it is illuminating to consider a simple mass-on-a-spring model [26], as shown in Figure 2.2. For example, if two atoms have a (stronger) double bond between them, this is akin to having two springs between two masses. This system will have a higher resonance frequency than equivalent system with only one spring of the same stiffness. Similarly, systems with heavier masses (atoms) will resonate at a
2.1. Infrared spectroscopy

Figure 2.1: The bending and symmetric stretch vibrational modes of a water molecule.

Figure 2.2: Mass-on-a-spring model of chemical bonds and how the resonant frequency, $\nu$, depends on size (mass) and type of bond.

lower frequency.

This very simple model does not cover all of the complexities of molecular vibrations [25], but the general qualitative results do hold true and form the basis for why different bonds oscillate at different frequencies.

The combination of bonds, and hence of their resonances, is unique to each chemical. Therefore, recording the absorption spectrum of a sample yields a unique “chemical fingerprint” that can be used to identify which molecules are contained within it [27]. A typical spectrum is shown in Figure 2.3. Techniques that use this principle are grouped together under the term “infrared spectroscopy”.

By raster scanning the sample through the beam path and taking a spectrum at each point, it is possible to build up “hyperspectral” absorption maps. Such chemical maps are commonly used for microscale chemical imaging [29]. However, the long wavelengths involved bring about a major drawback – poor spatial resolution. This will be discussed in detail in subsection 2.1.2.
2.1. Infrared spectroscopy

Figure 2.3: Typical absorbance spectrum from oesophageal tissue. Modified from [28].

Other applications of infrared spectroscopy

Chapters 3 and 5 contain examples of the non-chemical information contained in this part of the spectrum in the realms of plasmonics and metamaterials, respectively.

2.1.2 Limitations of far-field infrared spectroscopy

The resolution of conventional far-field techniques, including infrared spectroscopy, is constrained by the wave properties of light. Waves passing through an aperture diffract, effectively blurring any sub-wavelength features. This means that as two points in an image move closer together, they will increasingly overlap and become more difficult to resolve (see Figure 2.4).

In 1873, Ernst Abbe determined this minimum distance $d$ such that two features are just barely distinguishable to be given by [2]:

$$d = \frac{\lambda}{2NA},$$

(2.1)

where $\lambda$ is the wavelength of the imaging light, and $NA$ is the numerical aperture of the imaging system. This is the famous Abbe limit, also known simply as the diffraction limit, for optical imaging resolution.
2.1. Infrared spectroscopy

Figure 2.4: Diffraction through a circular aperture created “Airy discs”, i.e. interference patterns resulting from diffraction. At the critical separation $d_{\text{min}}$, the Abbe limit, close objects cannot be resolved in the far-field.

The inverse proportionality of resolution with wavelength in Equation 2.1 demonstrates the key issue with infrared microscopy – the wavelengths are large, so the achievable resolution is poor. Traditional techniques for spectroscopic mapping, such as Fourier transform infrared spectroscopy (FTIR) [30], are diffraction limited to the order of $\sim 10 \mu m$. In practice, the poor signal-to-noise ratios of the broadband sources used, such as glowbars, mean that even diffraction-limited resolution is rarely achieved with FTIR systems.

If we wish to study the chemistry of individual animal cells ($\sim 10 \mu m$), bacteria ($\sim 1 \mu m$), or viruses ($\sim 100 \text{ nm}$), we must find some way of beating the diffraction limit in the mid-infrared.

2.1.3 Raman spectroscopy

As discussed in the previous section, conventional IR spectroscopy is limited to approximately ten micron resolution by the diffraction of the large wavelengths involved. Raman spectroscopy provides an alternative to IR spectroscopy that is less severely hampered by diffraction.

As illustrated in Figure 2.5, Raman scattering is closely related to IR absorption [31]. Consider a single vibrational mode with energy equal to the difference between states 0 and 1. This mode can be excited directly by the absorption of infrared light. Alternatively, this mode can be probed by Stokes Raman scattering, whereby a high energy photon (blue arrow) is absorbed causing a transition from state 0 to a virtual energy state, after which a lower energy photon (green arrow) is emitted, leaving the system in the excited energy
2.1. Infrared spectroscopy

Figure 2.5: Energy level transitions for a given vibrational mode. The transition from state 0 to state 1 can be probed directly by absorption (IR), or indirectly by Stokes and anti-Stokes Raman scattering. The arrow thicknesses give an indication of the relative interaction strengths.

state 1. The energy difference between the absorbed and emitted photons is then the same as the energy of a corresponding absorbed infrared photon. Therefore, these processes can probe the same vibrational modes. Similarly, anti-Stokes Raman scattering can occur, whereby a photon can be absorbed from an already excited state and a higher energy photon can be emitted, returning the system to state 0. Combined, these form the basis for Raman spectroscopy.

As the Raman effect relies on the difference between higher energy photons, Raman spectroscopes can use much shorter wavelength light sources. The shorter wavelengths significantly decrease the diffraction-limited resolution to the sub-micron level. This roughly ten-fold improvement in spatial resolution comes at a substantial cost – Raman scattering is a much weaker effect than IR absorption [31], as it involves an extra transition to a virtual energy state. This means that much higher powers must be used to get a useable signal in Raman spectroscopy, which can ultimately damage delicate samples [32]. This significantly limits its usability in studies of biological matter.

The final difference of note here is that the Raman effect relies on a change in the polarisability of the molecule, whereas infrared absorption requires a change in the molecule’s dipole moment [31]. This means that the IR spectroscopy and Raman spectroscopy are sensitive to different vibrational modes. See Sathyanarayana [31] for an in-depth discussion of their differences and similarities. Ultimately, the two techniques offer complemen-
2.2. Beating the diffraction limit

2.2.1 Fluorescence techniques and their limitations

Conventional light microscopes operate by illuminating a sample and measuring the reflected or transmitted light. Fluorescence microscopes take a more indirect approach. First, the sample is illuminated with light of one colour. This light then induces fluorescent emission in any fluorophores in the sample and it is this fluorescent light (a different colour) that is imaged.

In 1994, Hell and Wichmann [33] introduced the theory of Stimulated Emission Depletion (STED) microscopy, which was first experimentally demonstrated in 1999 [34]. STED is able to beat the diffraction limit by selectively depleting fluorophores around a constrained area of emission, which enhances the possible resolution of the image [35].

Since then, a number of other fluorescence-based super-resolution techniques have been developed, such as PALM [7] and STORM [36]. In theory, the ultimate resolution of these techniques could reach the single molecule level, but they are limited to tens of nanometres by signal to noise constraints, such as photon collection efficiency and fluorophore photo-bleaching [37]. Many issues also surround the fact that fluorescence techniques inherently require artificial labelling with fluorophores. For further discussion on the limitations of fluorescence-based super-resolution, see subsection 4.1.3.

2.2.2 Electron microscopy

Whilst electron microscopy (EM) is technically limited by diffraction, it is still capable of nanometric resolution. As shown by Equation 2.1, spatial resolution is limited by the wavelength of the imaging source. Wave-particle duality means that electrons also have a characteristic wavelength, known as the de Broglie wavelength $\lambda_e = \frac{h}{p_e}$, where $p$ is the momentum of the electron [38]. By accelerating the imaging electrons to sufficiently high energies (∼1 to 100+ keV), sub-Ångstrom wavelength can be achieved.

However, electron microscopy is typically limited to ∼1–10 nm resolution by the width of the electron beam and the electron interaction volume within the sample [39]. For
2.2. Beating the diffraction limit

Figure 2.6: Schematic diagram of an AFM operating in tapping mode.

A discussion of the limitations of EM with respect to recovering chemical information, especially in biological applications, see subsection 4.1.1.

2.2.3 Atomic force microscopy

An atomic force microscope (AFM) is comprised of a very sharp probing tip held very close to the surface of a sample that has been mounted onto a piezoelectric three-dimensional scanning stage [8]. The stage is capable of positioning the sample with nanometre precision in the horizontal directions and Angstrom precision in the vertical. The probe tip is attached to a cantilever, which is driven to vibrate at a known resonant frequency, \( f \), in the so-called “tapping mode” [40]. This causes the probe to oscillate vertically with a typical “tapping” amplitude between 10 and 100nm, such that the probe is only in contact with the sample at the lowest point of its oscillation, which lessens the damage to the surface and tip compared with the non-oscillating “contact mode” [41].

The experiments in this thesis measure the probe position using a laser that is reflected from the top of the cantilever before passing a knife-edge and finally hitting a detector. The role of the knife-edge is to make the measured detector signal vary with cantilever deflection by blocking out more or less of the light due to the changing angle of reflection. The detector signal then forms a feedback loop with the sample stage to maintain the probe-sample separation (see Figure 2.6) and the probe’s oscillation amplitude. The topography can be mapped through this procedure by recording the vertical position of the sample stage as it is raster scanned horizontally across the sample in discrete steps.
2.2. Beating the diffraction limit

This pixel-by-pixel approach means that the potential resolution of an AFM map is only limited be the size of the probe tip (typically 5–10 nm).

2.2.4 Photothermal infrared microscopy

There are a number of techniques that aim to probe infrared absorption thermally. The common principle of these techniques is that pulsed IR light is focused onto a sample, which absorbs light according to the chemical composition (see section 2.1), and subsequently causes a thermal expansion of the sample [42]. By tuning the wavelength of the IR source, a measure of the IR absorption spectrum can be recovered. In order to breach the diffraction limit, each technique attempts to measure this thermal expansion, rather than the outgoing IR light itself.

Optical photothermal infrared microscopy

One such method is called optical photothermal IR (O-PTIR) spectroscopy [43]. As shown in Figure 2.7, O-PTIR focuses a second, visible laser onto the sample co-aligned with the tuneable IR laser focus. Any thermal expansion of the sample, caused by absorption of the IR light, will cause the refractive index of the sample to change at visible wavelengths [44]. This refractive index change can be imaged by the visible laser. As the visible light has a comparatively short wavelength, its diffraction limit is smaller (∼0.5 – 1 µm) than that of the mid-IR laser (∼5 – 10 µm). Therefore, some information of the IR absorption, and hence the chemistry, of the sample can be recovered with spatial resolution below the IR diffraction limit. In practice, however, thermal diffusion limits the achievable resolution (see subsection 4.1.4).

AFM-IR

Atomic force microscope infrared (AFM-IR) spectroscopy instead uses an AFM to probe the thermal expansion mechanically (see Figure 2.8) [45]. Once again, pulsed IR light from a tuneable laser incident on a sample will cause a brief thermal expansion. This fast expansion causes the AFM probe to oscillate resonantly with an amplitude proportional to the sample’s IR absorption. As the measurement is made by an AFM, diffraction-beating images of the IR absorption can be achieved. However, as opposed the resolution that can be achieved with AFM, the actual resolution achieved by AFM-IR is limited typically to
2.2. Beating the diffraction limit

Figure 2.7: Schematic diagram of an O-PTIR. The incident IR pulse causes a thermal expansion in the sample. The concurrent visible pulse experiences a short-lived change in refraction that depends on the strength of the IR absorption.

Figure 2.8: Schematic diagram of an AFM-IR. The incident IR pulse causes a thermal expansion in the sample that oscillates the probe in proportion to the strength of the IR absorption.

$\sim 0.5 - 5 \, \mu m$, and at best $\sim 0.3 \, \mu m$ [46]. The causes of this limit on achievable resolution, such as thermal conduction, are discussed in subsection 4.1.4.

2.2.5 The optical near-field

The angular spectrum representation

To understand the manner in which near-field techniques achieve super-resolution, one must find a way of describing optical fields (in homogeneous media). The angular spectrum representation is one such method, which is understood as the series expansion of a complex optical field in terms of plane and evanescent waves with varying amplitudes and propagation directions [47]. In the paraxial limit, for example, this representation becomes particularly useful in describing the propagation of laser beams. In this form,
2.2. Beating the diffraction limit

Figure 2.9: The relationship between imaginary and real $k_z$ (left) and the behaviour of the wave’s electric field $E$ (right).

it is much easier to determine how the fields will be affected after passing some disturbance (the object to be imaged), including which components decay exponentially from the object, known as evanescent waves.

**Evanescent waves**

Consider a single plane wave with frequency $\omega$ and wavenumber $k$, i.e. one part of the total field in the angular spectrum representation:

$$E(r) = E_0 e^{i\mathbf{k}\cdot\mathbf{r} - \omega t}. \quad (2.2)$$

Assuming homogeneity, isotropy and linearity of the medium through which the light propagates, the time-harmonic wave must satisfy the Helmholtz equation:

$$(\nabla^2 + k^2)E(r) = 0, \quad (2.3)$$

where $k = \frac{\omega}{c} n$, $n$ is the refractive index of the medium and $c$ is the speed of light in free space. Taking the general beam propagation to be in the $z$-direction and defining $k_z = \sqrt{k^2 - k_x^2 - k_y^2}$ gives two possible solutions [47] (illustrated in Figure 2.9):

1. $e^{i(k_x x + k_y y)} e^{\pm i|k_z|z}$, i.e. $k_z \in \mathbb{R} \ \forall \ k_x^2 + k_y^2 < k^2$
2. $e^{i(k_x x + k_y y)} e^{-|k_z||z|}$, i.e. $k_z \in i\mathbb{R} \ \forall \ k_x^2 + k_y^2 > k^2$

In case (1), one recovers plane waves whose transverse wavenumbers must always be less than $k \sim 1/\lambda$. This wavelength dependence is the physical origin of the diffraction
limit in traditional light microscopy [2]. These plane wave solutions will propagate to the far-field, as usual.

In case (2), the magnitudes of the transverse components are greater than $k$, leaving an imaginary $k_z$ and a solution that decays exponentially in the $z$-direction – an evanescent wave. It is these waves (along with the corresponding magnetic fields) that form the optical near-field within a few nanometres of the object.

By virtue of comprising imaginary $k_z$, these evanescent waves contain arbitrarily large transverse wavevectors, $k_x$ and $k_y$, for any given $k$. Consequently, the near-field contains the high spatial frequency information that is otherwise lost in conventional far-field imaging. This reveals a method for imaging below the diffraction limit – the high spatial frequency, and hence high spatial resolution, information of the near-field must be recovered.

2.3 Near-field microscopy

By bringing a probe into the optical near-field, evanescent waves can be scattered into the far-field and the high spatial frequency information can be recovered. This is the general principle behind all forms of near-field microscopy, including scattering-type scanning near-field optical microscopy (s-SNOM), which forms the basis for much of this thesis.

2.3.1 Scattering-type scanning near-field optical microscopy (s-SNOM)

The fundamental difference between AFM and s-SNOM is the addition of a second laser to the system. This second beam is focused not onto the back of the cantilever, but onto the probe tip itself, from which the light is scattered. More concretely, there is a scattering coefficient, $\sigma$, which dictates how incident fields $E_{\text{in}}$ are transformed upon scattering into outgoing fields $E_{\text{out}}$, i.e.:

$$E_{\text{out}} = \sigma E_{\text{in}}. \quad (2.4)$$

The nature of this scattering is determined by the near-field interaction between the sample and the tip. Therefore, the magnitude and phase of the scattered light are sensitive to the dielectric constant of the sample surface in accordance with a simple mirror-dipole
2.3. Near-field microscopy

Figure 2.10: Schematic diagram of an s-SNOM operating in the pseudoheterodyne detection scheme. The AFM laser and detector have been omitted for clarity.

description of charges in the tip [15, 48, 49], as well as to any variation in the local optical fields, such as surface plasmon resonances. It is these variations that form the image contrast in s-SNOM.

A metal coating (platinum in this thesis) and the geometry of the tip produce a large field enhancement at its apex, in a nano-scale analogue of the “lightning rod” effect [50]. This field enhancement makes the near-field probe-sample interaction strong enough to be detectable in the far-field. When the illuminating light is polarised in a plane parallel to the probe’s axis (p-polarised), the free electrons within it oscillate with the electric field of the light. Electrons are forced together at the apex of the probe, creating a large, highly confined electric field. When the light is polarised in a plane perpendicular to the probe’s axis (s-polarised), this effect is greatly reduced and measurements lose a great deal of sensitivity [51]. This can actually be advantageous in some scenarios, where the sample itself is expected to play a role in the field enhancement and so the impact of the tip on these fields is to be minimised (see [52]). This thesis only considers cases in which strong interaction with the sample is an advantage, so all further discussion will assume p-polarised light.

The experiments presented here used a commercial s-SNOM (NeaSNOM, Neaspec GmbH, Germany) illuminated by a continuous-wave (CW) quantum cascade laser (QCL; MIRcat, Daylight Solutions, USA) that is tuneable between $\lambda \sim 5.5 - 11.4 \text{ \mu m}$).

The general bench-top optical setup is shown in Figure 2.10. The QCL light first passes
through a variable attenuator (not shown) to optimise the trade-off between maximising the light scattered from the tip and preventing laser intensity noise due to back-reflections into the laser cavity. Having vibrationally isolated the microscope with air-cushioning supports and an acoustic noise damping box, this laser noise is the dominant source of noise in this experimental setup, as discussed in detail in [53]. The beam is then passed through a non-polarising beamsplitter with one arm continuing on to the tip (the illumination beam) and the other arm directed to a reference mirror (the reference beam). The illumination beam is focused down onto the oscillating probe (frequency \( f \approx 280 \text{ kHz} \) in this thesis) by a piezoelectrically controlled parabolic mirror. The light from this path is scattered in all directions, but only the light that is scattered back along the path of the incident beam is collected – this light forms the signal beam. The mirror in the path of the reference beam is vibrated at a much lower frequency (\( M \approx 300 \text{ Hz} \)), introducing a phase modulation. This reference mirror modulation allows for a pure near-field signal to be reconstructed and is called pseudoheterodyne detection [54]. Briefly, this detection scheme relies on the fact that the near-field decays exponentially from a sample, whereas the far-field varies on length-scales of the incident wavelength (see [53] for a detailed discussion). As such, higher harmonics (\( n \in \mathbb{Z} \)) of \( f \), the measured signal has a greater contribution from the near-field. The ultimate choice of harmonic to consider is dependent on the tip oscillation amplitude and the sample composition, and is determined experimentally by ensuring that the signals at that harmonic disappear as the tip is taken out of contact with the sample. Finally, the reference and signal beams are recombined at the beamsplitter before detection in a HgCdTe detector.

The interferometric detection scheme bestows a further advantage, in that it allows for the near-field phase, \( \phi = \text{arg}(\sigma) \), to be recovered in addition to the near-field amplitude, \( s = |\sigma| \). The phase in particular is a key component that allows for the recovery of the local absorption, \( \alpha \), of a sample (see subsection 2.3.2).

The majority of the optical components are reflective to ensure that optical alignment does not change with the wavelength of the illuminating light. As a result, only the beamsplitter and the detector must be changed when working in different spectral regions. The ease or difficulty of focusing the illuminating beam onto the probe tip is determined by the diffraction-limited spot size of the light – the larger the wavelength (and hence spot), the easier to focus. This is a substantial advantage of mid-infrared sources.
2.3. Near-field microscopy

Ultimately, the radius of curvature of the tip apex determines the resolution of s-SNOM, as the enhancement is only strong enough directly beneath the apex of the probe to affect far-field scattering. As the typical radius of curvature of an s-SNOM probe is 5–20 nm, this gives rise to smallest observable features far below the far-field diffraction limit ($\lambda/1000$) in the mid-IR and beyond.

2.3.2 Chemical contrast and spectroscopic imaging

In a typical s-SNOM image, there are two possible sources of contrast. The first, chemical contrast, is formed by the dependence of the scattering coefficient, $\sigma$, on the sample’s local dielectric constant. The exact relationship between the two is too complex for an analytical solution. Therefore, approximating models must be used to build a sense of intuition. For details of commonly used models, refer to [53]. Beyond simply differentiating between materials in the same scan area with different optical properties, the wavelength dependence of the dielectric constant (and so by extension the scattering coefficient, $\sigma$) offers the opportunity to spectroscopically identify materials at nanometre scale.

As discussed in section 2.1, infrared spectroscopy provides a diffraction-limited method for the identification of arbitrary chemical compounds. S-SNOM offers a way to bring IR spectroscopy to the nanoscale. It has been shown [16] that, for a thin film of weak molecular oscillators, the imaginary component of the near-field scattering coefficient, $\sigma$, is directly proportional to the local absorption in a sample, $\alpha_{loc}$. Therefore, mapping $\text{Im}(\sigma) = s \sin(\phi)$ yields a nanoscale absorption map of a sample. By using a broadly tuneable mid-infrared source to recover this local absorption at different wavelengths, s-SNOM can identify chemicals with resolution better than $\lambda/1000$ [16]. This ability to chemically map samples at resolutions far below the diffraction limit, opens the door to ground-breaking applications in the chemical and biological sciences. An in depth discussion of a new method to achieve this and one of its exciting new applications is given in chapter 4.

2.3.3 Contrast from intrinsic optical fields

The second type of contrast in an s-SNOM image is formed by variations in the electromagnetic field across a surface that are set up via manipulation of the illuminating light by the sample itself. Typical causes of these modified field distributions are the
2.4 Metamaterial superlenses

2.4.1 A brief introduction to metamaterials

The way in which light interacts with a bulk material is determined by its constituent chemical elements and their structure on an atomic scale. This means that in order to create bulk materials with desirable optical properties their fundamental structure must be altered. Due to the difficulty of creating atomically arranged structures (e.g. positioning individual xenon atoms with a scanning tunnelling microscope [73]) on a large length scale, bulk optical elements can only make use of any materials properties that are found in nature, such as a positive index of refraction.

Metamaterials provide a way of customising optical properties, without a reliance on atomic-scale control. The key assumption is that a metamaterial will be constructed from components of length scale \( a \ll \lambda \), where \( \lambda \) is the wavelength for which the metamaterial should operate [74]. This means that light will see an effective medium that derives its properties from these sub-wavelength structures. Effectively, this adds an extra degree of freedom (i.e. geometry) with which novel and exotic optical properties can be achieved, such as optical cloaking [75], negative index of refraction [76], and anisotropy [77].

The following sections will introduce the theory of how metamaterials can be used to create a “superlens” that beats the diffraction limit that was introduced in subsection 2.1.2. A new low-loss superlens design, invented by the author, is discussed in chapter 5 along with validating simulations in subsection 5.6.1.
2.4. Metamaterial superlenses

2.4.2 Anisotropic materials

The way in which light interacts with any material is determined by two parameters in Maxwell’s equations:

- The electric permittivity \( \varepsilon = \varepsilon_r \varepsilon_0 \); and
- The magnetic permeability \( \mu = \mu_r \mu_0 \),

where \( \varepsilon_0 \) is the permittivity of free space and \( \varepsilon_r \) is the relative permittivity of that material (similarly for \( \mu \)). For the rest of this work, the permeability will always be assumed to be equal to the free space permeability \( \mu_0 \). This is a valid assumption here because most materials do not have a magnetic response, and in any case all of the arguments in the rest of this chapter could just as readily be applied to magnetic systems. To further simplify discussions, the relative permittivity \( \varepsilon_r \) will be used instead of \( \varepsilon \). Real-world values can then simply be recovered by multiplying \( \varepsilon_r \) by \( \varepsilon_0 \).

In general, permittivity is a rank-2 tensor of the form:

\[
\varepsilon = \begin{pmatrix}
\varepsilon_{xx} & \varepsilon_{xy} & \varepsilon_{xz} \\
\varepsilon_{yx} & \varepsilon_{yy} & \varepsilon_{yz} \\
\varepsilon_{zx} & \varepsilon_{zy} & \varepsilon_{zz}
\end{pmatrix}.
\] (2.5)

The vast majority of natural materials are isotropic (i.e. \( \varepsilon_{ij} = 0 \forall i \neq j \) and \( \varepsilon_{ij} = \varepsilon \forall i = j \)), which means that light travelling in different directions interacts with the material in the same way.

If this is not the case, the material is anisotropic, hence the optical properties experienced by the light depend on the direction of propagation and the orientation of polarisation. This is phenomenon can occur in the natural world [78], but metamaterials allow for this anisotropy to be custom designed to work in intriguing ways. Uniaxial anisotropy is the case in which one direction has permittivity \( \varepsilon_{zz} = \varepsilon_\perp \) and the other two orthogonal directions have \( \varepsilon_{xx} = \varepsilon_{yy} = \varepsilon_\parallel \). This gives a diagonal permittivity:

\[
\varepsilon = \begin{pmatrix}
\varepsilon_\parallel & 0 & 0 \\
0 & \varepsilon_\parallel & 0 \\
0 & 0 & \varepsilon_\perp
\end{pmatrix}.
\] (2.6)
2.4. Metamaterial superlenses

This is the case that will be examined in the following sections due to its interesting dispersion properties.

**Iso-frequency dispersion contours**

Before discussing the different regimes of uniaxial metamaterials, it is helpful to review the dispersion relation in general and to set out a method for gaining some predictive intuition of how a given metamaterial will behave.

The dispersion relation for a uniaxial metamaterial is given by [79]:

\[
\frac{k_\perp^2}{k_0^2} \frac{1}{\epsilon_\parallel} + \frac{k_\parallel^2}{k_0^2} \frac{1}{\epsilon_\perp} = 1,
\]

(2.7)

where \(k_\perp = k_z\), \(k_\parallel = \sqrt{k_y^2 + k_z^2}\) and \(k_0\) are the perpendicular, parallel and free-space wavevectors respectively. As before, the relative permittivity, \(\epsilon_r\), is used for simplicity.

For constant \(k_0 = 2\pi/\lambda_0\), and hence constant frequency, Equation 2.7 gives either the equation of an ellipse or a hyperbola. These are examples of *iso-frequency dispersion contours* (IDC) [80], which can be used as an aid to intuition.

The Poynting vectors in the metamaterial, \(S = E \times H\), are the normals to the IDC for a given input wave [81]. As can be seen in the IDC for air (\(\epsilon_\perp = \epsilon_\parallel = 1\)) in Figure 2.11, this gives a quick and easy way to visualise how light should behave in a given medium.

2.4.3 Super-resolution from strong optical anisotropy

**Hyperbolic metamaterials**

A *hyperbolic* metamaterial (HMM) is defined by having \(\epsilon_\perp \epsilon_\parallel < 0\). As per Equation 2.7, this gives two possible types of hyperbola [82]:

- Type I HMM: \(\epsilon_\perp < 0\), \(\epsilon_\parallel > 0\); and
- Type II HMM: \(\epsilon_\perp > 0\), \(\epsilon_\parallel < 0\).

As shown in Figure 2.12, hyperbolic dispersion curves have unbounded solutions for large lateral wavevectors, \(k_\parallel\). As such, hyperbolic metamaterials allow the plane wave propagation of optical fields with infinitely large \(k_\parallel\), thereby preserving the high spatial resolution of the optical near-field (discussed in subsection 2.2.5).
2.4. Metamaterial superlenses

Figure 2.11: Iso-frequency dispersion contour of an isotropic medium (air; black line) and corresponding Poynting vectors, \( \mathbf{S} \), (blue arrows).

(a) Type II HMM: \( \varepsilon_\perp = 1, \varepsilon_\parallel = -1 \)

(b) Type I HMM: \( \varepsilon_\perp = -1, \varepsilon_\parallel = 1 \)

Figure 2.12: Iso-frequency dispersion contours of the two different types of hyperbolic metamaterial (black lines) with their asymptotes (red dashed lines). The angle of incidence with such a material is defined by \( \theta_i \).
2.4. Metamaterial superlenses

Consider light propagating in the $\perp$-direction. In a Type II HMM (Figure 2.12a), all light with an angle of incidence, $\theta_i$, less than the angle between the hyperbola’s asymptote and the $k_\perp$-axis has no real solution to the dispersion curve, and therefore becomes evanescent (see section 2.2.5). Any fields with greater angle of incidence will propagate at fixed angles defined by the normal of the asymptotes. Unfortunately, this makes it unusable as a superlens, as most light is discarded at non-grazing incidence, i.e. light with angle of incidence, $\theta_i$, less than the angle of the asymptote, $\theta_A = \tan(\sqrt{\varepsilon_\parallel/\varepsilon_\perp})$.

On the other hand, a Type I HMM (Figure 2.12b) will allow all light to propagate. The light will again tend to propagate at fixed angles defined by the normal of the asymptotes. If $\varepsilon_\perp/\varepsilon_\parallel < 1$, the HMM can be used for sub-diffraction focusing of spatially separated light sources [82–85].

If $|\varepsilon_\perp/\varepsilon_\parallel| \gg 1$, the hyperbola flattens out and the metamaterial becomes effectively diffraction-free [82]. As shown in Figure 2.13, all of the light, including that with high lateral spatial frequency (i.e. large $k_\parallel$), propagates solely in the $\perp$-direction. This results in a hyperbolic superlens (also known as a hyperlens) that perfectly preserves the fine spatial information of the near-field.

One way of achieving this is to make a metamaterial with $\varepsilon_\parallel \approx 0$. There has been a substantial effort to create such epsilon-near-zero (ENZ) metamaterials in recent years [83, 84, 86–89].

![Figure 2.13: Type I HMM with large anisotropy ($\varepsilon_\perp = -100$ and $\varepsilon_\parallel = 1$) has a flat dispersion that makes it a hyperlens.](image)
2.4. Metamaterial superlenses

(a) Type I EMM: $\varepsilon_\perp = 3, \varepsilon_\parallel = 1$

(b) Type II EMM: $\varepsilon_\perp = 1, \varepsilon_\parallel = 3$

Figure 2.14: Iso-frequency dispersion contours of the two different types of elliptical metamaterial and corresponding Poynting vectors, $S$, (blue arrows).

The alternative method, is to make $\varepsilon_\perp$ very large in what are called \textit{epsilon-near-pole} (ENP) hyperbolic metamaterials. Comparatively little research has been undertaken to make $\varepsilon_\perp$ very large [90].

Both of these methods are largely limited by intrinsic losses, which are discussed in section 5.1.

Elliptical metamaterials

In contrast to the previous section, an \textit{elliptical} metamaterial (EMM) has both $\varepsilon_\perp > 0$ and $\varepsilon_\parallel > 0$ [91]. As with the hyperbolic case, EMMs can be split into different regimes:

- Type I EMM: $\varepsilon_\perp/\varepsilon_\parallel > 1$; and
- Type II EMM: $\varepsilon_\perp/\varepsilon_\parallel < 1$; and
- Isotropic Metamaterial: $\varepsilon_\perp/\varepsilon_\parallel \approx 1$, where we simply recover the isotropic case of Figure 2.11.

As shown in Figure 2.14, a type I EMM will tend to direct incident light in the $\perp$-direction, while a type II EMM will tend to direct light in the $\parallel$-parallel directions. This makes a type II EMM unsuitable as a superlens, as it suffers from enhance diffraction, which only acts to decrease the achievable resolution.

As with a hyperbolic metamaterial, an elliptical metamaterial can only preserve the high spatial frequency information of the near-field if $\varepsilon_\perp/\varepsilon_\parallel \gg 1$ [91]. However, there is
2.4. Metamaterial superlenses

Figure 2.15: A type I EMM with large anisotropy ($\varepsilon_\perp = 100$ and $/\varepsilon_\parallel = 1$) has a wide and flat dispersion allows near diffraction-less propagation.

an additional constraint, in that the real solutions for $k_\parallel$ are bounded by finite extent of the elliptical dispersion curve. This means that a truly effective EMM superlens must also have large $\varepsilon_\perp$, so that the maximum real solution $k_\parallel$ is also large. In other words, the more eccentric (or “flatter”) and “wider” the ellipse, the better the superlens will be. An ideal elliptical metamaterial superlens is illustrated in Figure 2.15. Until now, only one such elliptical superlens has been created, which utilised a multiple quantum well structure to create a large $\varepsilon_\perp$ resonance [92], achieving an eccentricity $\varepsilon_\perp/\varepsilon_\parallel \approx 3.5$ and $\varepsilon_\perp \approx 24$.

In chapter 5, a novel design for a low-loss epsilon-near-pole superlens, which exhibits both hyperbolic and elliptical dispersion, is introduced.
Chapter 3

Mid-Infrared Plasmon Spectroscopy

In this chapter, infrared s-SNOM will be used to study plasmons in graphene. Therefore, subsection 3.1.1 will give a brief overview of why graphene is interesting and what plasmons are. The remainder of the chapter is largely concerned with how s-SNOM can be used to launch and detect graphene plasmons. This will allow for a highly precise and accurate measurement of the graphene Fermi energy in subsection 3.2.7, as well revealing intriguing new plasmon reflection “hotspots” in section 3.2.8.

3.1 Graphene plasmons

3.1.1 An introduction to graphene

Graphene is an allotrope of carbon arranged in a two-dimensional hexagonal lattice. When considered as an unbounded, free-standing sheet, graphene has some remarkable properties. For example, graphene is a zero-gap semiconductor with a linear, 2D electron dispersion relation in the region of the so-called “Dirac point”, K, [93] as illustrated in Figure 3.1. Note that, in contrast to the previous chapter, here the dispersion relation refers to the energy-momentum \((E-k)\) dispersion of electrons in graphene. The linear dispersion \((E \propto k)\) near a Dirac point yields an effective mass [94]:

\[
m^* = \hbar^2 k \left( \frac{\partial E}{\partial k} \right)^{-1} \propto k,
\]

(3.1)

for electrons in graphene. At \(k = 0\), the electrons can be thought of as effectively “massless”. As a result, the electron mobility, \(\mu_e \propto 1/m^*\), should be infinitely large.
3.1. Graphene plasmons

Figure 3.1: (a) The full dispersion relation for electrons in graphene has “Dirac points” (indicated by K) with a (b) linear ($E \propto k$) dispersion relation in the surrounding regions and with a band-gap of exactly zero. In undoped graphene ($n_e=0$), the Fermi energy $E_F$ is exactly at the point where the 2D isotropic momentum $k = k_x = k_y$, and hence effective mass $m^*$, is exactly zero (see Equation 3.1).

However, these properties are limited by the inability to produce pristine sheets of perfectly free-standing graphene. To be used in a device, the graphene must be attached to a substrate, such as silicon carbide (SiC), which breaks the assumption that it is free-standing. This introduces a doping to the graphene that shifts the Fermi level away from the $k = 0$ point. However, this doping can actually prove advantageous for certain applications, such as plasmonics (see subsection 3.2.5).

In particular, the Fermi energy, $E_F$, of single-layer graphene (1LG) and bilayer graphene (2LG) depend on the charge carrier concentration, $n_e$ [95]:

$$E_{F1LG}^e = v_F \hbar \sqrt{\pi n_e}, \quad (3.2)$$

and

$$E_{F2LG}^e = \frac{\pi \hbar^2 n_e}{2m^*}, \quad (3.3)$$

respectively, where $v_F$ is the Fermi velocity and $\hbar$ is the reduced Planck constant.

A further limitation is that current techniques, such as chemical vapour deposition (CVD), produce samples that comprise patches with different lattice orientations, resulting in grain boundaries that limit the mobility of electrons in graphene [96].
3.1. Graphene plasmons

Figure 3.2: Light of wavelength $\lambda_0$ propagating in a dielectric, with permittivity $\varepsilon_d$, incident on a conductor, with permittivity $\varepsilon_c$, can form a surface plasmon with a much smaller wavelength, $\lambda_{sp}$.

3.1.2 A primer on surface plasmon-polaritons

A surface plasmon-polariton, hereon referred to as a plasmon, is essentially an electromagnetic wave that has been confined to the interface between a conductor, usually a metal, and a dielectric, as illustrated in Figure 3.2. A plasmon is formed by the interaction of the electromagnetic fields of the incident light with the oscillation of free electrons in the conductor [97].

By solving Maxwell’s equations in the simple geometry of Figure 3.2, the plasmon wavelength, $\lambda_{sp}$, can be recovered [98]:

$$\lambda_{sp} = \frac{\lambda_0}{\sqrt{\varepsilon_d \varepsilon_c / \varepsilon_d + \varepsilon_c}}$$

(3.4)

where $\varepsilon_d$ and $\varepsilon_c$ are the dielectric constants of the isotropic dielectric and conductor respectively, and $\lambda_0$ is the free-space wavelength of the incident light. As the denominator on the right-hand side is typically considerably larger than one, this means that a plasmon confines the light down to length scales much smaller than the initial free-space wavelength, i.e. $\lambda_{sp} \ll \lambda_0$.

There are two distinct types of plasmon:

1. Propagating plasmons that travel along the conducting surface until they are entirely attenuated by absorptive losses in the conductor;

2. Localised surface plasmon resonances (LSPRs) that are confined by the spatial extent and geometry of their supporting structures.
3.2 Correlative nano-imaging of graphene

3.2.1 Fermi energy measurement of graphene

One of the key parameters in the properties of graphene is its Fermi energy, $E_F$. In graphene plasmonics, for example, $E_F$ strongly influences the dispersion of plasmons (see subsection 3.2.6). Common methods for measuring $E_F$ are the Hall effect and Kelvin probe force microscopy (KPFM). Each of these techniques brings advantages and disadvantages.

The Hall effect provides simple and direct measurement of the charge carrier density, which is then used to calculate $E_F$ for 1LG and 2LG using equations 3.2 and 3.3, respectively. However, the Hall effect method requires fabrication of Hall bar devices (see Figure 3.3(a)) with the typical sizes ranging from hundreds of nanometres to several millimetres [99, 100]. Alternatively, Hall effect measurements can be performed in the van der Pauw geometry (see Figure 3.3(b)) on continuous films. Both methods lead to averaging of the charge carrier measurement across a large sample area, which can misrepresent $E_F$ for inhomogeneous graphene samples. See references [95, 101] for further details on such Hall effect measurements of the Fermi energy.

KPFM is based on AFM (see subsection 2.2.3) and uses a conducting tip to probe...
the difference between the sample and tip work functions, $V_{CPD}$. A voltage, $V$, is applied between the tip and sample (see Figure 3.4) comprising a DC offset, $V_{DC}$, and an time-varying voltage, $V_{AC}\sin(f_{AC}t)$. During measurement, $V_{DC}$ is tuned to give zero overall DC component to the voltage, which thereby gives a measurement of $V_{CPD} = -V_{DC}$. By calibrating the probe’s work function to a known reference sample, the work function of the sample can be recovered. See reference [102] for further details. In this manner, KPFM can map the local work function of graphene, which can then be used to estimate $E_F$. However, KPFM is extremely sensitive to surface contamination [95, 103], and estimation of the sample’s work function demands precise calibration of the KPFM probe [104].

### 3.2.2 Graphene sample preparation

The graphene in the following work was grown on semi-insulating SiC substrates in a hot-wall reactor (VP508, Aixtron, Germany) by Włodek Strupiński at the Institute of Electronic Materials Technology in Warsaw, Poland. First, a complete interfacial layer (IFL) was grown under an argon laminar flow at 1600 °C. Following IFL growth, the sample was annealed in a hydrogen-rich environment. The hydrogen intercalates between the IFL and substrate, forming Si-H bonds, which transforms the IFL into quasi-free-standing (QFS) 1LG. Please see reference Ohta et al. [105] for further details. It is not possible to entirely avoid nucleation growth in certain areas, e.g. near the edges of SiC terraces, so these areas are often covered with 2LG. QFS graphene on SiC produced with this growth method typically has p-type doping [106, 107].

#### AFM measurements

Figure 3.5 shows the surface topography of the sample obtained with an AFM. The most prominent features are a gold contact in the upper right of the sample, and the small SiC terraces that are a few hundred nanometres wide and run along the larger SiC steps. There are also steps of 2–6 nm from the 1LG to the SiC substrate, as well as particulate remnants from fabrication (small white spots).

#### Raman spectroscopy measurements

Raman spectroscopy (see subsection 2.1.3) is particularly useful in the characterisation of graphene. See Malard et al. [108] for a comprehensive discussion of Raman spectroscopy.
applied to graphene. In particular, the “2D” Raman band at \( \sim 2700 \text{ cm}^{-1} \) contains information about the number of layers of graphene in a sample \([108]\). In 1LG, the 2D band typically has a full width at half maximum (FWHM) of \( \sim 24 \text{ cm}^{-1} \). Furthermore, the 2D peak can be fitted by a single Lorentzian for 1LG, while 2LG has a broader 2D peak with shoulders that requires four Lorentzians to fit \([108]\).

Figure 3.6 shows the mapping of the Raman 2D-peak width of the sample, as performed by Vishal Panchal and Christos Melios at the National Physical Laboratory in Teddington, UK. This map shows that there are strips of 2LG in the middle of the graphene sheet along the SiC terraces. The 2D-peak widths were determined to be \( \sim 60 \text{ cm}^{-1} \) for 2LG and \( \sim 20 \text{ cm}^{-1} \) for 1LG, which are consistent with the literature \([108]\). In addition, the measured 2D peak can be fitted by a single Lorentzian for 1LG (Figure 3.7), and four Lorentzians for 2LG, as expected.

### 3.2.3 Measuring \( E_F \) with Kelvin probe force microscopy

Using KPFM (see subsection 3.2.1), the sample’s work function, \( \Phi \), was mapped by Vishal Panchal and Christos Melios at the National Physical Laboratory in Teddington, UK. To do this, the KPFM probe’s work function, \( \Phi_{probe} \), was first calibrated against gold (\( \Phi_{Au} = 4.82 \text{ eV} \)) using \( \Phi_{probe} = \Phi_{Au} + eV_{SP}(Au) \), where \( V_{SP} \) is the surface potential measured on gold. Applying \( \Phi_{Sample} = \Phi_{probe} - eV_{SP}(Sample) \) to the surface potential image then
3.2. Correlative nano-imaging of graphene

Figure 3.6: 2D-peak width (see Figure 3.7) Raman map of QFS graphene on SiC.

Figure 3.7: Raman spectra for 1LG and 2LG showing typical features observed in this type of graphene. The inset shows the 2D-peak with its characteristic single-Lorentzian fit for 1LG and four-Lorentzian fit for 2LG.
3.2. Correlative nano-imaging of graphene

Figure 3.8: Work function of QFS graphene on SiC with blue arrows indicating islands of 2LG.

...yields the work function [104]. Figure 3.8 shows the work function of the sample. KPFM is only sensitive to changes in the surface potential, so it is able to clearly map out the structure of 1LG, 2LG, SiC and Au from their inherent differences in work function. For 1LG and 2LG, the work functions are $\Phi_{1LG} = 4.81 \text{ eV}$ and $\Phi_{2LG} = 4.56 \text{ eV}$, respectively. Assuming the intrinsic work function of graphene is $\Phi_0 = 4.47 \pm 0.05 \text{ eV}$ from previous work [103], the Fermi energy for 1LG can be estimated to be $E_F^{KPFM} = 340 \pm 10 \text{ meV}$.

3.2.4 Measuring $E_F$ via the Hall effect

Measurement of the Fermi energy of the sample via the Hall effect in the van der Pauw geometry (see Figure 3.3(b)) was made by Christos Melios at the National Physical Laboratory in Teddington, UK, giving an estimated, $E_F^{Hall} = 333 \pm 1 \text{ meV}$. While this measurement is precise, it suffers from the inaccuracy of measuring the patches of 2LG present in the sample, as well as the 1LG. Details of the protocol for such measurements are given in reference [101].

3.2.5 Plasmon nano-imaging

In 2012, s-SNOM was used by Chen et al. [109] in a collaboration between the groups of Javier García de Abajo, Rainer Hillenbrand and Frank Koppens in Spain to measure the wavelength of graphene plasmons by analysing the plasmon interference fringes that
Correlative nano-imaging of graphene

Figure 3.9: Incident light with free-space wavelength $\lambda_0$ is scattered by the s-SNOM probe tip into plasmon modes with wavelength $\lambda_{sp}$.

appear in the images. As illustrated in Figure 3.9, when the s-SNOM probe is brought within a few nanometres of the graphene, the Pt-coated tip radially scatters the incident laser light into the sample’s plasmon modes, which then propagate in the plane of the sample. Upon encountering a defect, such as the edge of the graphene sheet, plasmons are partially reflected and produce an interference pattern. The fringes produced have a spatial period of half the surface plasmon wavelength, $\lambda_{sp}$ [109]. The amount of light subsequently scattered back into the s-SNOM detection system is modulated by the field intensity at the probe, which allows the fringes to be mapped out in the s-SNOM image.

Using a first-order approximation of the dispersion relation of plasmons, Chen et al. [109] obtained an estimate of the Fermi energy, $E_F \approx 400$ meV, on exfoliated graphene. In the years since, this technique has been used often to study plasmons in graphene and other layered 2D materials, in particular in the study of chemical doping [110], grain boundaries [63], nano-resonators [111, 112], as well as hybrid systems with hyperbolic [84] or phonon-polaritons [66]. Until now, the energy ranges were constrained to discrete lines in a small spectral range $\lambda_0 \approx 9.2 - 10.6$ µm provided by the carbon dioxide lasers commonly used for s-SNOM [109, 113]. This limited potential precision and accuracy of extraction of $E_F$, which was further limited by the first-order approximated dispersion relation typically used.
3.2.6 Dispersion relation of plasmons in homogeneous graphene

The full dispersion relation of plasmons in homogeneous 1LG as a function of optical frequency, $\omega$, is given by [114]:

$$\frac{\varepsilon_{\text{sub}}}{\sqrt{\varepsilon_{\text{sub}} k_0^2 - k_{sp}^2}} + \frac{1}{\sqrt{k_0^2 - k_{sp}^2}} = -\frac{4\pi \sigma_c(\omega)}{\omega}, \quad (3.5)$$

where $\varepsilon_{\text{sub}}$ is the substrate permittivity, $k_0 \equiv \frac{2\pi}{\lambda_0} = \frac{\omega}{c}$ is the free space wavevector, $k_{sp} \equiv \frac{2\pi}{\lambda_{sp}}$ is the in-plane plasmon wavevector, and $\sigma_c(\omega)$ is the graphene conductivity. Under the random phase approximation, and at energies low enough to avoid interband transitions ($\hbar \omega \ll 2E_F$), $\sigma_c(\omega)$ has a linear dependence on the Fermi energy, $E_F$, of the form [114]:

$$\sigma_c(\omega) = \frac{e^2 E_F}{\pi \hbar^2} \frac{i}{\omega + i\tau^{-1}}, \quad (3.6)$$

where $e$ is the fundamental electron charge, and $\tau$ is the graphene charge carrier relaxation time.

3.2.7 High accuracy measurement of $E_F$ with widely tuneable real-space imaging of plasmon dispersion

The mid-IR s-SNOM amplitude, $s_2$, image in Figure 3.10 shows strong scattering from the 1LG. On top of this, plasmon interference fringes at the 1LG/SiC (white arrows) and 1LG/2LG (blue arrows) interfaces can be seen. The strongest fringes emanate from the right and top of the image, as indicated by the white arrows, arise from the strong plasmon reflection at the straight, lithographically defined edges of the graphene device (i.e. at the 1LG/SiC interface).

Tuning the QCL (see subsection 2.3.1) through $\lambda_0 = 8.93$–$10.43 \mu$m, in steps of 100 nm, the spectral profile of the fringe separation was mapped. This enabled a fuller and broader graphene plasmon dispersion curve to be obtained than previously reported. Figures 3.11(a) and 3.11(b) show fringes at $\lambda_0 = 10.03 \mu$m and $9.83 \mu$m, respectively. By extracting line profiles over the length of the red dashed rectangles, the graphene plasmon dispersion was measured over more than a decade in the surface plasmon wavelength, $\lambda_{sp} = 140$–$1700$ nm (Figure 3.11(c)). This is the first time such broad and dense measurement of graphene plasmon dispersion has been reported.
3.2. Correlative nano-imaging of graphene

Figure 3.10: S-SNOM amplitude, $s_2$, of QFS graphene on SiC at $\lambda_0 = 9.78 \, \mu m$. White and blue arrows indicate plasmons reflecting from 1LG/SiC and 1LG/2LG interfaces, respectively.

Figure 3.11: Plasmon interference fringes from a 1LG/SiC interface at (a) $\lambda_0 = 10.03 \, \mu m$ and (b) $\lambda_0 = 9.83 \, \mu m$. (c) Experimentally measured $\lambda_{sp}$ (red points) with a numerical fit (black line) using the dispersion relation of Equations 3.1 and 3.6. The fit yields a graphene Fermi energy $E_F^{sp} = 298 \pm 4 \, meV$. Errors are from the standard error of the mean of two successive fringe separations. The inset line profile in (c) shows is extracted from the red dashed rectangle in (a).
3.2. Correlative nano-imaging of graphene

As expected, the fringe separation, and so $\lambda_{sp}$, increases with $\lambda_0$. The larger error bar for the final data point stems from the poorer contrast of the peaks and troughs of the plasmon standing wave. The precise value of $E_F$ can then be extracted by combining Equations 3.5 and 3.6, and fitting the measurements with an error-weighted numerical minimisation using $E_F$ as the free parameter. The fit gives $E_{sp}^F = 298 \pm 4$ meV (Figure 3.11(c)) for the plasmon dispersion method.

The disagreement of this value with the Hall effect and KPFM measurements outside of the random errors quoted are derived from additional sources of systematic error. The larger estimation of the Fermi energy from the Hall effect, $E_{Hall}^F$, can be attributed to spatial averaging, which includes small quantities of 2LG, and tend to have a higher carrier density than 1LG [108]. On the other hand, the work function (KPFM) estimate of the Fermi energy, $E_{KPFM}^F$, has higher uncertainty than both the Hall effect method and the plasmon dispersion method due to the low signal-to-noise ratio of frequency-modulated KPFM in measuring surface potential [104]. A further consideration is that each method was performed in a different laboratory with slightly different environmental conditions, which could affect the doping of the graphene.

3.2.8 Plasmon reflection from 1LG/2LG interfaces

In addition to the reflections at the edge of the graphene device that were studied in the previous section, additional fringes were also present from plasmon reflection at the interfaces between 1LG and 2LG.

Relative strength of reflections

Figures 3.12(a) and 3.12(c) show such 1LG/2LG interfacial fringes near the SiC terraces of Figure 3.5 at $\lambda_0 = 9.68$ µm and $\lambda_0 = 9.48$ µm, respectively. As can be seen in Figure 3.12(a), wherein representative linescans across fringes at different interfaces are shown, the reflections at the 1LG/2LG interfaces are generally weaker than those at 1LG/SiC interfaces.

The cause of the plasmon reflection at 1LG/2LG interfaces requires some discussion. It has been observed by other groups [113] that plasmon reflection can occur solely from terrace step edges in the SiC substrate. While there are additional complications of the 1LG/2LG also being present at SiC step edges in the present work, this is unlikely to
Figure 3.12: S-SNOM amplitude, $s_2$, images showing graphene plasmon reflection from various interfaces at (a) $\lambda_0 = 9.68 \mu m$ and (c) at $\lambda_0 = 9.48 \mu m$. (b, d) Magnified $s_2$ images showing “hotspots” of constructive plasmon interference near 2LG “strips”. (e) $s_2$ line profiles along the black and cyan rectangles shown in (a) indicating the relative strengths of plasmon reflections from SiC/1LG and 2LG/1LG interfaces, respectively. The two different interface types are illustrated schematically in (f). The blue dashed arrow in (d) indicates the line profiles (g) across “hotspot” site 2. For comparison, the line profiles are normalised to s-SNOM amplitude of the graphene background, $s_{2,\text{bg}}$. All scale bars are 1 $\mu m$ long.
be the primary cause of the 1LG/2LG reflections. This is because reflections are also observed around smaller 2LG islands (blue arrows in Figure 3.10), where the step height is just \(~0.35\) nm (Figure 3.5). This drop in surface height from 1LG to the 2LG islands is consistent with epitaxial growth of 2LG (whereby graphene layers are “grown” by removal of Si atoms in the SiC substrate), as opposed to holes in the 1LG, which would appear as a height gain. Further still, the plasmon interference fringes surrounding these islands are much weaker than those at the 1LG/SiC interfaces, which is consistent with the strength of the fringes at the Raman-confirmed (Figure 3.6) 1LG/2LG interfaces. Therefore, the plasmon reflection likely results from the change in carrier density across the 1LG/2LG interface (Figure 3.12(e)), rather than the morphology of the substrate.

See reference [115], for an investigation of the Fermi energy of 1LG, 2LG and tri-layer (3LG) graphene using plasmon nano-imaging.

**Plasmon reflection “hotspots”**

In addition to the simple linear fringes at straight 1LG/2LG interfaces, several plasmon focusing “hotspots” arise near the edges of 2LG “strips”, which have never before been reported. As shown in Figures 3.12(b) and 3.12(d), the relative strengths of these hotspots depend strongly on the \(\lambda_0\), and by extension on \(\lambda_{sp}\). These strips of 2LG commonly form, as in this case, on the terraces of the SiC substrate. This is illustrated in Figure 3.12(f).

Upon tuning the laser from \(\lambda_0 = 9.68\) µm (Figure 3.12(b)) to \(\lambda_0 = 9.48\) µm (Figure 3.12(d)), the “hotspot” at position 2 disappears. This is shown more clearly by the normalised \(s_2\) line profiles in Figure 3.12(g). To normalise the profiles, \(s_2\) was divided by the background s-SNOM amplitude of the graphene away from the fringes, \(s_{2,bg}\). Given the relatively small change (\(~2\)% to the illuminating wavelength, the drastic change in hotspot intensity shows that these hotspots are highly resonant in nature. Further examples of the wavelength tuneability of these resonant hotspots are the hotspots at positions 1 and 4, which significantly increase in strength. Interestingly, the hotspot at position 3 even splits up into three separate hotspots. This can be interpreted as a result of the increase plasmon wavelength of Figure 3.12(b) broadening the fringes, causing the three distinct hotspots to blur into one.
3.3 Conclusions

The principle aim of this chapter was to establish plasmon nano-imaging as a method of quantitatively determining the Fermi energy of graphene. To that end, experiments using s-SNOM and a widely tuneable QCL comprehensively mapped the dispersion of graphene plasmons, covering more than an order of magnitude range of the $\lambda_{sp}$. The uncommonly wide and dense spectral coverage allowed for the precise extraction of $E_F$ entirely optically. This new method is free from the error of including defects in the measurement, as is common with Hall effect measurements, as well as free from the meticulous calibration work of KPFM (see subsection 3.2.1). This work puts forward mid-IR nano-imaging as a technique for the quantitative measurement of local electronic properties of graphene in ambient conditions, and free from electrical contacts or other steps in preparation of the sample.

In taking these measurements, highly resonant, wavelength-tuneable hotspots were also observed. These hotspots are the result of the constructive interference of graphene plasmons reflections around geometrically non-trivial interfaces between 1LG and 2LG. The spectrally and spatially sensitive natures of these hotspots offer possibilities for bespoke design with appropriate manipulation of the 1LG/2LG geometry. The local patterning of electromagnetic fields in such custom-made hotspots could be used for enhanced chemical analysis, environmental monitoring, as well as plasmonic nanoantennas for boosting the sensitivity of fluorescence microscopy and vibrational spectroscopy.
Chapter 4

Spectroscopic Infrared
Nano-Imaging in the Life Sciences

This chapter is concerned with the challenge of mapping chemical composition at the nanoscale in the disciplines of the life sciences. To establish the case for a new technique to solve this problem, the limitations of existing techniques will be reviewed in section 4.1. Then, having established the advantages of chemical mapping using infrared s-SNOM in section 4.2, results will be presented in sections 4.3.1 and 4.4 that demonstrate, for the first time, that infrared nano-imaging can image nanoscale organelles inside single human cells and map the binding sites of a clinically relevant anti-cancer drug.

4.1 Existing super-resolution techniques

Before discussing the advantages of IR nano-imaging, this section will review the limitations of commonly used nanoscale imaging techniques in relation to the specific challenges of the life sciences.

4.1.1 Limitations of electron microscopy

Electron microscopy is typically considered the “gold standard” in nanoscale biological imaging for its high spatial resolution (see subsection 2.2.2). However, there are two major issues with EM in the life sciences. Firstly, biological material is typically composed of carbon, nitrogen and oxygen, which have similarly low atomic numbers, providing little electron scattering contrast [4]. Secondly, EM requires ultra-high vacuum conditions to
4.1. Existing super-resolution techniques

prevent electrons scattering from atmospheric gases [3]. These two issues place significant burdens on the preparation of biological samples for EM imaging. In order to provide sufficient contrast, samples must be stained with heavy metal compounds, such as the toxic osmium tetroxide [4]. Furthermore, to ensure samples are stable under vacuum, particularly harsh chemical fixation and dehydration protocols are necessary, which risk the integrity of the samples. The following is a typical transmission EM (TEM) preparation protocol for the type of system to be investigated later on in this chapter (cultured eukaryotic cell suspensions), as taken from reference [116]:

For the centrifugation procedure, prepare an initial fixative by mixing one part of either glutaraldehyde or glutaraldehyde/formaldehyde fixative (warmed to the same temperature as the culture) with an equal part of the cell culture. A recommended fixative is 4% formaldehyde/2.5% glutaraldehyde in PIPES buffer.

1. Centrifuge the culture/fixative mix at room temperature to generate a cell pellet. Excessive centrifugation must be avoided, but the pellet should be compact enough to remain a firm mass. Typically, mammalian cells are centrifuged at 1,000–2,000 rcf (relative centrifugal force) [...] for 5 to 10 min.
2. Carefully remove the supernatant and gently flow several milliliters of pure buffered fixative, at room temperature, on top of the pellet taking care not to disturb the pellet.
3. After 15 min, gently dislodge the pellet with either a pointed spatula or sharpened wooden applicator stick so that the opposing side of the pellet is exposed to the fixative for an additional 15 min at room temperature.
4. Remove the fixative with a pipet and replace with 0.1 M buffer.
5. Transfer the pellet into a Petri dish by flowing buffer into the tube and allow the intact pellet to move into a pool of buffer in the dish.
6. With a clean razor blade or scalpel, cut the pellet into small cubes no larger than 1 to 2 mm.
7. Transfer the cubes into a small vial containing several milliliters of buffer for 10 min. This is easily accomplished by tilting the Petri dish so that the cubes move into a common location. Lift and transfer the cubes with a sharpened applicator stick or piece of filter paper cut into a sharp point.
8. After two additional changes of buffer, each for 10 min, postfix the cubes in 1 to 2 mL of 1% buffered osmium tetroxide for 1 h at room temperature.
9. Rinse in distilled water and dehydrate the cubes in an ethanol series (25, 50, 75, 95% ethanol, each for 10 min).
10. As the specimens are undergoing dehydration, prepare two mixtures of propylene oxide:epoxy 812 embedding medium consisting of 1:1 and 1:3 parts, each. Use 10 mL of graduated, disposable polypropylene tubes with tight fitting lids. The propylene oxide:epoxy 812 embedding medium is best mixed by gentle aspiration and stirring with a pipet rather than inversion of the tubes because the caps may be dislodged as the result of pressures inside the sealed tube. Acetonitrile is a safer alternative than propylene oxide. After the two mixtures are prepared, store capped until needed in step 12.
4.1. Existing super-resolution techniques

11. After three changes of absolute ethanol, 10 min each, use three changes of propylene oxide (or acetonitrile), each for 15 min.

12. Remove most of the propylene oxide from the specimen cubes but leave a trace to prevent the specimens from drying out. Place 1 to 2 mL of 1:1 propylene oxide: epoxy 812 embedding medium over the specimen cubes and gently rotate the capped container five to six times during a period of 20 min.

13. Remove the 1:1 mixture and replace with 1:3 mixture of propylene oxide:epoxy medium and rotate the capped container five to six times during a period of 45 to 60 min.

14. Remove the 1:3 mixture and pour on several milliliters of pure epoxy 812 embedding medium and rotate the uncapped container five to six times during a period of 90 min. The specimens should gradually sink to the bottom of the tube. If not, leave them overnight, uncapped in a dry environment.

15. Fill several polypropylene embedding capsules half-way with epoxy 812 embedding medium. Hold the capsules upright in commercially supplied holders or produce a capsule holder by puncturing holes in a small cardboard box. Dislodge any air bubbles that form at the tip of the capsule using the sharp tip of a broken, wooden applicator stick or a dissecting needle. Use the wooden applicator stick to lift and transfer individual specimen cubes into the capsules. Allow the specimens to sink to the bottom tip of the capsules. This may take an hour or so. We often leave the capsules overnight at room temperature.

16. Place an identifying label in each capsule. Use pencil (ink will dissolve in resin) and small pieces of paper approximately 5 × 20 mm. Place the labels inside the capsules so that the numbers are legible. The easiest way is to roll the label and insert it so it is pressed against the inside wall of the capsule, the numbers facing out.

17. Transfer the capsules into a 60°C oven for 24 to 48 h until the epoxy resin is completely hardened and ready to section.

Particular issues with such protocols are the usage of toxic chemicals, such as osmium tetroxide (Step 8), and long incubation and baking times of up to two days. Even after this lengthy protocol has been completed, the resin must be sectioned into < 50 nm thick slices (this alone is a challenging procedure) and attached to a sample holder. Altogether, this means that EM analysis can often take weeks to perform. By contrast, the sample preparation used for s-SNOM in this thesis (see subsection 4.4.2) uses no toxic chemicals and takes less than one hour after the cells have been cultured and any dosing has been performed.

Another major drawback is that EM images do not have access to any chemical information contained within the sample, which is crucial for understanding the biochemistry of cells and mechanisms of medical drugs.
4.1. Existing super-resolution techniques

4.1.2 Limitations of atomic force microscopy

Atomic force microscopy (see subsection 2.2.3) has seen significant usage in the life sciences [117], as it solves many of the problems associated with EM by virtue of its ability to image in atmospheric conditions, while maintaining $< 10$ nm lateral resolution. However, the contrast mechanisms of AFM are the morphology and mechanical properties, such as stiffness and adhesion force [ref], of the sample to be imaged. This means that, as with EM, there is no chemical information to be extracted from the AFM signals. A further limitation is that the contrast of small height changes tends to be “washed out” by any large height changes in the sample. In biological samples, a single region often contains entities of very different size. This means that, for example, the relative contrast of mitochondria ($\sim 100$ nm) can be dwarfed by a cell nucleus ($\sim 5$ µm) [118].

4.1.3 Limitations of fluorescence microscopy

Fluorescence microscopy (see subsection 2.2.1) is able to both operate in atmospheric conditions and extract some degree of chemical information. The advantages of such techniques are that the signal recovered is highly specific, and imaging can be performed in three dimensions. These techniques have two major drawbacks, however. Firstly, the images produced only show the location of the predetermined fluorophores and not any unlabelled material. This means that fluorescence images must be correlated with other imaging techniques, so that the distribution of fluorophores can be matched with the incumbent biological structures. Figure 4.1 shows how the pure fluorescence information must be correlated to an image of the wider environment, in this case using TEM [119].

Secondly, all of these methods fundamentally rely on non-native fluorophores, such as dyes or proteins, being attached to the chemical of interest. This labelling means that the resulting images may not accurately reflect the behaviour of the original system. For example, fluorophores have been shown to affect the conformation and dynamics of certain proteins [120]. As such, extra care must be taken in the interpretation of any results from fluorescence microscopy. Furthermore, the very act of labelling requires prior knowledge of the biological problem, which limits the domain of study to fairly well understood systems. As such, it is used as a so-called “discovery technique”. To remove these problems, a label-free technique is required.
4.1. Existing super-resolution techniques

Figure 4.1: Typical correlative imaging used in fluorescence microscopy. (a) The STED fluorescence signal (pink) is overlaid upon the TEM image of human breast adenocarcinoma tissue. The yellow rectangle in (a) indicates the region of (b) the pure TEM image, (c) the pure fluorescence image, and (d) the superposition of the two. Note that the TEM correlative image is required to determine from which tissue structures the fluorescence signal is originating. This figure has been reproduced with permission from reference [119], as published by John Wiley & Sons.

4.1.4 Limitations of photothermal microscopy

In order for any photothermal effects (see subsection 2.2.4) to produce a signal significantly above the noise, the magnitude of the thermal expansion must be significant. This requirement sets two key limitations on this class of technique. The first, is that the material being imaged must have a sufficiently large thermal expansion coefficient. Therefore, photothermal techniques typically require soft samples, such as polymers and biological material [121]. The second limitation is that a sufficiently thick sample must be prepared, as the overall amount of thermal expansion is largely dependent on the amount of material to be expanded [122]. This presents a significant problem when attempting to image naturally very thin structures, such as mitochondria or deposits of medical drugs.

Another key limitation is that of thermal conduction. The principle on which photothermal techniques operate is that thermal expansion that is measured in the sample, either with an AFM or with a shorter wavelength laser, is directly proportional to the amount of absorption of the IR light pulse [45]. This is what allows a “chemical fingerprint” (see section 2.1) to be recovered. The crucial detail is that the spectrum is assumed to match the absorption spectrum of just the structure of interest. However, this is not the case. Consider a particular structure of interest that is heated by an IR pulse. As this structure is heated, according to its IR absorption spectrum, it will also cause its neighbouring structures to heat by thermal conduction [123], regardless of their own absorption spectrum. Explicitly, the spectrum recovered from any given object is not
4.2. The key advantages of infrared nano-imaging

4.2.1 Label-free chemical information

As discussed in subsection 4.1.3, super-resolution microscopy based on fluorescence labelling carries inherent disadvantages. By contrast, nanoscale imaging based on IR spectroscopy does not require chemical labelling, as it probes already present vibrational modes (see section 2.1). This greatly simplifies interpretation of results by ensuring the dynamic interactions of the chemical of interest with the host environment is relatively undisturbed (before chemical fixation and drying). An additional benefit of this label-free approach is that the resulting images can be used in isolation, as they provide some degree of contrast for all of the features in the sample.
4.2.2 Imaging in atmospheric conditions

The only requirements for a sample to be imaged with s-SNOM are that it is flat (\(<\sim 5 \mu m\) thick), dry, and dimensionally stable. No additional contrast agents are required, and samples do not need to withstand vacuum conditions, unlike for EM (see subsection 4.1.1). This means that sample preparation can be a relatively quick and cheap process that has less impact on the sample's integrity. Furthermore, the process of imaging itself is much less labour intensive, as there is no vacuum chamber within which experiments must be performed. This also makes it easy to make adjustments to the s-SNOM apparatus.

4.3 Optical nano-imaging of human cells

4.3.1 Testing biological s-SNOM with human red blood cells

Historically, s-SNOM has seen very little usage in biological systems. Any experiments have been limited to artificially produced test samples, such as isolated proteins on a Si substrate[124]. In order to establish IR nano-imaging as a scientifically and clinically useful tool, it was first necessary to attempt imaging of a “real-world” sample.

To that end, a sample comprising human red blood cells (RBCs) was prepared using the “formalin-fixed, paraffin-embedded” (FFPE) tissue preparation protocol. This protocol is the “gold-standard” in the field of histology of biological tissue and is described in detail in reference [125]. In short, this involves chemically fixing the tissue, in this case RBCs, with formalin, which is an aqueous solution of 37% formaldehyde, by mass. Next, the fixed tissue is embedded in paraffin wax, so that it can then be sectioned with a microtome into 1–4 \(\mu m\) thick slices, which are then deposited onto glass slides. To ensure adhesion of the tissue to the slides, the slides are heated to 60 °C for ten minutes and then allowed to cool at room temperature for five minutes. The final step is to remove the wax by immersion of the slides in a sequential series of two xylene, then three ethanol baths, each for five minutes. The end results are thin, dry sections of the tissue mounted on glass slides.

Figure 4.3(a) shows the AFM topography of a single RBC, as prepared with the protocol above. The shape clearly exhibits the familiar biconcave disc morphology characteristic of RBCs [118]. Figure 4.3(b) shows the first s-SNOM image of a single cell with spatial
resolution of 67 nm (∼\(\lambda_0/140\)). Given that the entire cell is smaller than the wavelength of the illuminating light, this is a strong demonstration of the capability of s-SNOM to image below the diffraction limit of Equation 2.1. This result laid the groundwork for attempts at nano-scale IR spectroscopy of cells based on mid-IR s-SNOM and was published as part of a paper entitled “New IR imaging modalities for cancer detection and for intra-cell chemical mapping with a sub-diffraction mid-IR s-SNOM” in Faraday Discussions in 2016 with Hemmel Amrania, Laurence Drummond, Charles Coombes, Sami Shousha, Laura Woodley-Barker, Kenny Weir, Ian Carter, and Chris Phillips [126].

4.4 Mid-infrared chemical nano-imaging for intra-cellular drug localisation

As discussed in section 2.1, IR absorption spectroscopy offers a robust method of chemical analysis, while s-SNOM provides a direct measure of the nanoscale absorption of a sample (see subsection 2.3.2). Furthermore, s-SNOM was shown by the author in subsection 4.3.1 to be capable of imaging cells, for the first time. In order to demonstrate the ultimate aim of chemical analysis within single cells, it is necessary to find a practical, yet medically relevant, test case.
4.4. Mid-infrared chemical nano-imaging for intra-cellular drug localisation

Figure 4.4: (a) The molecular structure of Bortezomib (BTZ) gives rise to (b) an IR absorption profile (blue) with a sharp, strong peak at 1020 cm\(^{-1}\) (black dashed line) in a region where typical human tissue gives an IR spectrum (red; reproduced with permission from [28]) with minimal features.

### 4.4.1 Bortezomib as a test drug for intra-cellular chemical analysis

Bortezomib (BTZ) is an anti-cancer drug that is clinically approved in the treatment of multiple myeloma [127]. Its accepted mechanism of action is to inhibit the function of proteasomes, which are sub-cellular structures that break down proteins [128]. Further to its medical relevance, BTZ is an especially promising candidate for IR analysis in cells. This is because its structure (see Figure 4.4(a)) includes a B-C moiety that gives rise to a particularly sharp and strong absorption peak (see Figure 4.4(b)) at wavenumber, \(\nu = 1020\ \text{cm}^{-1}\) (i.e. \(\lambda = 9.8\ \mu\text{m}\)), where human tissue absorption is typically featureless. This helps to ensure that the signal obtained is specific to BTZ and does not include any significant contribution from the intrinsic chemistry of the tissue.

**Details of Bortezomib FTIR measurements**

The following measurements were performed by Ainoa Rueda-Zubiaurre in the group of Matthew Fuchter at Imperial College London. A 99.94% pure sample of BTZ was acquired (Selleckchem, Catalogue No. S1013, Batch No. S101315) for use in all FTIR and nano-imaging experiments. For measurement IR absorption spectra, a commercial FTIR spectrometer (Frontier FT-IR/MIR, PerkinElmer), operating in an attenuated
4.4. Mid-infrared chemical nano-imaging for intra-cellular drug localisation

total reflectance (ATR) geometry, was used with a LiTaO$_3$ detector. Over a spectral range $4000 - 650 \, \text{cm}^{-1}$, an IR spectrum of BTZ was measured with $4 \, \text{cm}^{-1}$ spectral resolution and a scanning speed of $0.2 \, \text{cms}^{-1}$. Approximately $2 \, \text{mg}$ of solid BTZ was applied to the KBr ATR prism surface. A lever arm was used to apply pressure to the sample, in order to ensure there was no air gap between the ATR prism and the sample. Lastly, the spectrum of BTZ was recorded, and normalised against a background spectrum of the environmental air, to give the spectrum shown in Figure 4.4(b).

4.4.2 Cell preparation

Having identified a candidate drug (BTZ) in the previous section, it was necessary to test it in a real-world biological system. To that end, a sample of human cells dosed with BTZ was prepared by Alice Beckley in the group of Eric Aboagye at Hammersmith Hospital in London. RPMI-8226 myeloma cells were cultured in RPMI-1640 medium supplemented with $10\%$ FCS, penicillin (100 U/ml), $5\%$ glutamine, and streptomycin sulphate (100 µg/ml) at $37 \, \text{°C}$ in a humidified atmosphere containing $5\%$ CO$_2$. The cells were then harvested, washed two times with phosphate-buffered saline (PBS), and counted ($1 \times 10^7$ cells). The cells were incubated with $8 \, \mu\text{M}$ of BTZ in solution (0.1% dimethyl sulphoxide) for 1 hour at $37 \, \text{°C}$. Finally, the cells were washed two more times in PBS, and then pelleted for paraffin embedding and sectioning. See subsection 4.3.1 for details of the remainder of the sample preparation protocol. The one modification of that protocol was to slice thinner sections ($\sim 1 \, \mu\text{m}$). This better enabled the piezoelectric scanning stage of the s-SNOM to cover the entire height range of the cells. The end products are glass slides with sectioned cells that have been treated with the drug of interest (BTZ).

4.4.3 Imaging sub-cellular ultrastructure with s-SNOM

Figure 4.5 shows an overview of a single cell prepared in the manner described above. The topography (Figure 4.5(a)) shows that this cell has a relatively large ($\sim 7 \, \mu\text{m}$ diameter) nucleus. In contrast, the IR near-field amplitude, $s_2$, image (Figure 4.5(b)) reveals a plethora of additional sub-cellular features, collectively called “ultrastructure”. This is the first time such features have been imaged optically.

Figures 4.6(a-c) give a closer look at the ultrastructure within the black dashed square in Figure 4.5(b). The aspect to note is that the $s_2$ once again provides greater contrast of
4.4. Mid-infrared chemical nano-imaging for intra-cellular drug localisation

Figure 4.5: Overview of the (a) topography and (b) near-field scattered amplitude, $s_2$, at $\nu = 1020\text{ cm}^{-1}$ ($\lambda = 9.8\mu\text{m}$) of a single human myeloma cell dosed with Bortezomib (BTZ). The black dashed and pink squares are studied in detail in Figure 4.6, while the green square shows the region studied in Figure 4.7.

the nanoscale organelles. To show this more clearly, a linescan (Figure 4.6(e)) was taken across a sharp biological feature in Figure 4.6(d) (pink square in Figure 4.5(b)). Across just 8 nm, the near-field amplitude, $s_2$, drops by $\sim 70\%$, while the height only changes by $\sim 20\%$. This corresponds to an s-SNOM spatial resolution better than $\lambda/1,200$ at the 9.8 $\mu$m imaging wavelength. This is comparable with typical electron microscope performance (see subsection 2.2.2). Identification of the biological features in these images is complicated by the as yet undetermined degree of shrinkage that occurs during sample preparation (see subsection 4.4.2). Nevertheless, the white and blue arrows of Figure 4.6(b) indicate what are tentatively called mitochondria-like and vesicle-like organelles, respectively.

As will be discussed in the following sections, the phase information in Figure 4.6(c) offers the first glimpse of a chemical specific IR signal. In particular, the significant phase shift ($\sim 11^\circ$) of the vesicle-like structures relative to the glass substrate at the known 1020 cm$^{-1}$ IR absorption peak of BTZ gives a strong indication of the presence of the drug.

4.4.4 Chemical-specific mapping of Bortezomib in a single cell

To perform a quantitative spectroscopic analysis, in order to locate the BTZ, the area corresponding to the green square in Figure 4.5(b) is inspected more closely in Figure 4.7.
4.4. Mid-infrared chemical nano-imaging for intra-cellular drug localisation

Figure 4.6: The (a) topography, (b) near-field scattered amplitude, $s_2$, and (c) phase, $\phi_2$, of the sub-cellular ultrastructure indicated by the black dashed square in Figure 4.5(b) at 9.8 µm. The structures in (b) are provisionally designated mitochondria (white arrows) and vesicles (blue arrows). (d) shows a magnified view of the pink square region in Figure 4.5(b). (e) linescan along the pink arrow in (d) shows $s_2$ (red) and the height (black) across the edge of a biological feature. Errors are smaller than the plot markers.
4.4. Mid-infrared chemical nano-imaging for intra-cellular drug localisation

Figure 4.7: (a) The AFM topography of the magnified region indicated by the green square in Figure 4.5. The cyan and white dashed rectangles indicate the averaging regions for acquisition of the BTZ and glass near-field signals, respectively. The patch of BTZ measured is \( \sim 9 \) nm thick with radius \( \sim 60 \) nm. The near-field phase, \( \phi_2 \), images at 1020 cm\(^{-1}\) (b) and 1000 cm\(^{-1}\) (c) are normalised to give zero phase at the glass substrate. (d) shows the normalised imaginary component of the near-field scattering, \( \text{Im}(\sigma_2) \), (red circles) and the far-field absorption spectrum of bulk BTZ (black line). Error bars are calculated from the standard error of the mean values within each of the respective averaging areas.

The topography (see Figure 4.7(a)) a number of vesicle-like features, two of which (white discs at the top and bottom centre) are significantly larger than the rest. It is noteworthy to recall from subsection 4.4.1 that BTZ is thought to target proteasomes, which are a type of vesicle.

Figure 4.7(b) and Figure 4.7(c) show the near-field phase, \( \phi_2 \), on and off the known absorption resonance of BTZ shown in Figure 4.4(b), respectively. To account for variations in the path length of the s-SNOM interferometer (see subsection 2.3.1, these phase images have been shifted to give zero phase at the glass substrate, i.e. such that \( \phi_2 = \phi_2' - \phi_{2,\text{glass}} \), where \( \phi_2' \) is the original unshifted phase. Strikingly, the phase on the larger vesicle-like features drops significantly as the laser wavelength is tuned just 2% away from the BTZ absorption peak at 1020 cm\(^{-1}\). In accordance with subsection 2.3.2, this indicates a significant change in the absorption of these features, which suggests presence of BTZ.

Using the previously normalised near-field phase \( \phi_2 \) (described above), the imaginary component of the near-field, \( \text{Im}(\sigma_2) = s_2 \sin(\phi_2) \) can also be normalised relative to the glass substrate, such that \( \text{Im}(\sigma_2) = \text{Im}(\sigma_2') / \text{Im}(\sigma_{2,\text{glass}}) \), where \( \text{Im}(\sigma_2') \) is the original un-normalised signal. In this way, the imaginary component of the suspected BTZ-containing...
4.4. Mid-infrared chemical nano-imaging for intra-cellular drug localisation

Figure 4.8: (a) Near-field amplitude, $s_2$, and (b) phase, $\phi_2$, spectra of Bortezomib in a cell (as shown in Figure 4.7) show that the phase has dominates the local absorption spectrum of Figure 4.7(d).

vesicle-like feature (averaged over the cyan square in Figure 4.7(a)), normalised with respect to the glass substrate (averaged over the white dashed rectangle in Figure 4.7(a)), was measured (shown by the red circles in Figure 4.7(d)). The resulting $\text{Im}(\sigma)$ spectrum shows good qualitative agreement with the far-field IR spectrum of bulk BTZ (black line), as measured with a standard FTIR spectrometer (see subsection 2.1.1). In each spectrum there is a sharp and strong peak at 1020 cm$^{-1}$ with a FWHM $\approx 10 - 20$ cm$^{-1}$. Altogether this demonstrates, for the first time, the power of spectroscopic IR s-SNOM to localise chemicals in a sub-cellular biological feature with a volume of just $\sim 10^{-22}$ m$^3$, or $\sim 100$ zeptolitres (i.e. $10^{-19}$ litres).

**Relative contributions of amplitude and phase to local absorption**

Strictly speaking, it is the imaginary component of the near-field scattering, $\text{Im}(\sigma)$, that is proportional to local absorption [15]. In this case, however, the majority of the chemical contrast in the $\text{Im}(\sigma_2)$ spectrum of Figure 4.7(d) is determined by variations in the near-field phase, $\phi_2$. This is clearly demonstrated in Figure 4.8, which shows that the near-field amplitude, $s_2$, spectrum is relatively featureless in comparison with the phase spectrum. When these figures are compared with Figure 4.7(d), it is clear that the phase and imaginary spectra are in good agreement. Given that $s_2$ is approximately constant across the spectrum, and that $|\phi_2|$ is always below 10°, an approximation can be made that $\text{Im}(\sigma_2) \propto \sin(\phi_2) \approx \phi_2$, which explains the aforementioned agreement between imaginary component and phase of the near-field.
4.5 Conclusions

The purpose of this chapter was to introduce spectroscopic IR s-SNOM as a means for the mapping of chemicals inside single cells. First, the ability of s-SNOM to acquire cellular images with nanoscale resolution was established by the author of this thesis. This work (see subsection 4.3.1) was published in reference [126]. This foundational step allowed further work to demonstrate ability of spectroscopic IR s-SNOM to localise the anti-cancer drug Bortezomib to a nanoscale vesicle-like structure inside a single human myeloma cell (see section 4.4).

In summary, this chapter has introduced a new imaging technique that is capable of beating the diffraction limit (see subsection 2.1.2) by a factor of $\sim 1000$, which enables nanoscale chemical analysis of biological systems in ambient conditions. Furthermore, the techniques relies on widely used sample preparation methods (see subsection 4.4.2) that do not suffer from the limitations of chemical labelling (see subsection 4.1.3), and do not need to withstand vacuum conditions (see subsection 4.1.1).
Chapter 5

Super-Resolution Imaging with Low-Loss Metamaterial Superlenses

This chapter will begin with showing how a superlens based on optical anisotropy (see subsection 2.4.3) can be created with one-dimensionally periodic multilayer structures, as first introduced by Salandrino and Engheta [22]. In section 5.4, additional insight from the author will establish the requirements for a “good” superlens and show how these requirements can be fulfilled. Using these insights, a new a novel design, discovered by the author, to create such a superlens will then be introduced. Finally, the superlensing performance of the new design will be studied in section 5.7 using the finite-difference time-domain (FDTD) method (see section 5.6).

5.1 Creating anisotropy with periodic layers

Consider the 1D periodic structure shown in Figure 5.1 with period \( d \) and layer thicknesses \( d_1 \) and \( d_2 \). In this regime (\( \lambda \gg d \equiv d_1 + d_2 \)), it was shown by Ramakrishna et al. [129] that such a structure can be approximated by an effective medium with a uniaxially anisotropic permittivity tensor (see Equation 2.6), in which:

\[
\varepsilon_\parallel = \frac{d_1 \varepsilon_1 + d_2 \varepsilon_2}{d_1 + d_2}, \quad \text{and} \quad \varepsilon_\perp = \frac{(d_1 + d_2) \varepsilon_1 \varepsilon_2}{d_2 \varepsilon_1 + d_1 \varepsilon_2},
\]

(5.1) (5.2)
5.2 Tuning anisotropy with doped semiconductors

Consider a multilayer (Figure 5.1) composed of semiconducting materials, in which all of the layers are large enough to avoid quantum confinement effects. If layer 1 is doped with a charge carrier density, $n_e$, its permittivity, $\varepsilon_1$, will depend on the angular frequency, $\omega$. The permittivity of the doped semiconductor can be described by the effective medium approximation (EMA), which states that:

$$\varepsilon_{\text{eff}}(\omega) = \varepsilon_m(\omega) - \frac{n_e}{\varepsilon_0}$$

where $\varepsilon_m(\omega)$ is the permittivity of the undoped semiconductor and $\varepsilon_0$ is the permittivity of free space. The charge carrier density, $n_e$, can be expressed as:

$$n_e = \frac{e}{m^*} \langle \phi \rangle$$

where $e$ is the electron charge, $m^*$ is the effective mass, and $\langle \phi \rangle$ is the mean electron density.

The permittivity of the doped semiconductor can be written as:

$$\varepsilon_{\text{eff}}(\omega) = \varepsilon_m(\omega) - \frac{e}{m^*} \langle \phi \rangle \varepsilon_0$$

The effective permittivity of the multilayer structure can be calculated using the EMA and the principle of superposition. The effective permittivity of the multilayer structure can be expressed as:

$$\varepsilon_{\text{eff}}(\omega) = \sum_{i=1}^{N} \frac{n_i \varepsilon_i}{n_i + \varepsilon_i}$$

where $n_i$ is the characteristic length of the $i$th layer, $\varepsilon_i$ is the permittivity of the $i$th layer, and $N$ is the number of layers.

It is clear from Equations 5.3 and 5.4 that there are two possible regimes for anisotropic superlensing ($|\varepsilon_\perp/\varepsilon_\parallel| \gg 1$) in such a structure. Firstly, $\varepsilon_2 = -R\varepsilon_1$ will result in an epsilon-near-zero ($\varepsilon_\parallel = 0$) metamaterial superlens, as discussed in subsection 2.4.3. Such superlenses [22, 82, 130] have had limited success because they tend to occur in regions of high absorptive loss, as will be shown in subsection 5.4.2. The remaining option is to maximise $|\varepsilon_\perp|$ with the resonance condition $\varepsilon_1 = -R\varepsilon_2$. This will be investigated in the next section.

5.2 Tuning anisotropy with doped semiconductors

Consider a multilayer (Figure 5.1) composed of semiconducting materials, in which all of the layers are large enough to avoid quantum confinement effects. If layer 1 is doped with a charge carrier density, $n_e$, its permittivity, $\varepsilon_1$, will depend on the angular frequency,
5.2. Tuning anisotropy with doped semiconductors

Figure 5.2: (a) Drude response of doped In$_{0.53}$Ga$_{0.47}$As with doping $n_e = 1.07 \times 10^{18}$ cm$^2$/Vs and $\varepsilon_{01} = 12.15$. (b) Effective medium response the multilayer demonstrated by Hoffman et al. [24]. The doped layer is as per part (a) of this figure, and the undoped layer is Al$_{0.49}$In$_{0.52}$As ($\varepsilon_2 = 10.23$). The structure used equal layer thicknesses ($R = 1$), causing high losses.

$\omega = 2\pi c/\lambda$, of the illuminating light according to the Drude model [131]:

$$\varepsilon_1(\omega) = \varepsilon_{01} \left(1 - \frac{\omega_p^2}{\omega^2 + i\omega/\tau}\right),$$

(5.5)

where $\varepsilon_{01}$ is the background permittivity of layer 1, and $\tau$ is the charge carrier mean scattering time. The plasma frequency, $\omega_p$, of a doped semiconductor is given by:

$$\omega_p = \sqrt{\frac{e^2 n_e}{m^*\varepsilon_0}},$$

(5.6)

where $e$ is the fundamental electron charge, and $m^*$ is the effective mass of the free carriers. An example of the Drude response of a doped semiconductor is shown in Figure 5.2(a). Assuming layer 2 is an undoped semiconductor with positive and constant $\varepsilon_2$, the monotonically decreasing $\varepsilon_1$ will fulfil the resonance condition $\varepsilon_1 = -R\varepsilon_2$ at a resonance frequency:

$$\omega_{\text{res}} = \frac{\omega_p}{\sqrt{1 + R\varepsilon_2/\varepsilon_{01}}}.$$  

(5.7)

Such a Drude response can be inserted into Equations 5.3 and 5.4 to give a resonant effective medium response. In 2007, Hoffman et al. [24] demonstrated such a multilayer with $R = 1$ (see Figure 5.2(b)). The loss factor, $\alpha$, for a 60° angle of incidence shown in Figure 5.2(b) was calculated by the transfer matrix method [81, 132], and is defined by
5.2. Tuning anisotropy with doped semiconductors

Figure 5.3: The loss figure of merit (FoM) has a generally inverse relationship with the loss factor, $\alpha$.

the Beer-Lambert law [133]:

$$T = (1 - R)e^{-\alpha t}, \tag{5.8}$$

where $T$ and $R$ are the transmittance and reflectance of a sample of thickness $t$, respectively. A discussion of the polarisation dependence in subsection 5.2.1 will show that only the losses associated with the TM mode, $\alpha_{TM}$, are of interest here, as TE modes are not sensitive to $\varepsilon_{\perp}$, and so cannot propagate without diffraction.

In all possible regions of operation (large $|\varepsilon_{\perp}|$, and $\varepsilon_{\parallel} \approx 0$), this particular superlens suffers from large loss factors $\sim 600\text{mm}^{-1}$. These losses must be avoided, if a usable superlens is to be created.

5.2.1 The loss figure of merit

In the field of metamaterials, it has been useful to derive a standard figure of merit (FoM) to give an indication of the losses associated with a given structure. Originally, this was defined for the case with isotropic metamaterials with refractive index, $n$, as the relative size of the real part of the refractive index to the imaginary part, i.e. $\text{FoM} = \frac{\text{Re}(n)}{\text{Im}(n)}$ [24]. This definition has an inverse relationship with the loss factor, $\alpha$, so the higher the FoM, the better. See Figure 5.3 for a comparison of the two.

As anisotropic metamaterials were introduced, a more general form became necessary,
5.2. Tuning anisotropy with doped semiconductors

Figure 5.4: The geometry of an plane wave incident on a uniaxial material can be reduced to a 2D problem by setting the y-component, $k_y$, equal to zero.

as the refractive index is now a tensor. The figure of merit for the uniaxial case (see subsection 2.4.2) is then [24]:

$$\text{FoM} = \frac{\text{Re}(k_{\perp})}{\text{Im}(k_{\perp})},$$

(5.9)

where $k_{\perp} = k_z$ is the perpendicular wavevector component, as defined in section 2.4.2.

The (electric field) wave equation for light in a general anisotropic medium is derived from Maxwell’s equations as:

$$\nabla \times \nabla \times \mathbf{E} = \mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2},$$

(5.10)

where $\varepsilon$ is the permittivity tensor as defined in Equation 2.5. In momentum-space, a plane-wave solution ($\mathbf{E} \sim \exp[i(\mathbf{k} \cdot \mathbf{r} - \omega t)]$) gives:

$$\mathbf{k} \times \mathbf{k} \times \mathbf{E} + k_0 \varepsilon \mathbf{E} = 0,$$

(5.11)

with wavevector $\mathbf{k} = k_x \hat{x} + k_y \hat{y} + k_z \hat{z}$. Setting the y-axis as normal to the plane of incidence, $k_y$ can be set to zero without loss of generality (see Figure 5.4). Expanding this equation with $\mathbf{k} = k_{\parallel} \hat{x} + k_{\perp} \hat{z}$ and a uniaxial permittivity tensor (Equation 2.6) gives:

$$\begin{pmatrix}
  k_{\parallel}^2 \varepsilon_{\parallel} - k_{\perp}^2 & 0 & k_{\parallel} k_{\perp} \\
  0 & k_{\parallel}^2 \varepsilon_{\parallel} - k_{\perp}^2 & 0 \\
  k_{\parallel} k_{\perp} & 0 & k_{\perp}^2 \varepsilon_{\perp} - k_{\parallel}^2
\end{pmatrix}
\begin{pmatrix}
  E_x \\
  E_y \\
  E_z
\end{pmatrix} = 0.$$
There are two pairs of non-trivial (i.e. $E \neq 0$) solutions, found by setting the determinant of the above matrix to zero. These correspond to the forward/backward (+/−) propagating transverse electric (TE) and transverse magnetic (TM) polarised modes:

$$k_{\perp,\text{TE}} = \pm \sqrt{k_0^2 \varepsilon_\parallel - k_\parallel^2}$$  \hspace{1cm} (5.13)

$$k_{\perp,\text{TM}} = \pm \sqrt{k_0^2 \varepsilon_\parallel - k_\parallel^2 \left( \frac{\varepsilon_\parallel}{\varepsilon_\perp} \right)}.$$  \hspace{1cm} (5.14)

Anisotropic superlensing requires that either $\varepsilon_\perp \gg 1$ or $\varepsilon_\parallel \approx 0$ (i.e. $|\varepsilon_\perp / \varepsilon_\parallel| \gg 1$), as shown in subsection 2.4.3. As Equation 5.13 does not depend on $\varepsilon_\perp$, the only possibility for superlensing is to set $\varepsilon_\parallel \approx 0$. However, this results in a non-flat hyperbolic dispersion, $k_{\perp,\text{TE}} \approx -k_\parallel^2$, which does not allow for the diffractionless propagation required. Therefore, the incident light must have TM polarisation. As such, all discussions, simulations and measurements here assume TM polarised illumination.

Furthermore, as $|\varepsilon_\perp / \varepsilon_\parallel| \to \infty$, Equation 5.14 becomes $k_{\perp,\text{TM}} = \pm k_0 \varepsilon_\parallel$ for all but the largest $k_\parallel$ (care must be taken when multiplying zero by infinitely large terms). It is worth noting that, in this regime, the maximum possible value of the numerator in Equation 5.9 is determined by $\varepsilon_\parallel$. This more concretely justifies the assertion that ENZ ($\varepsilon_\parallel \approx 0$) superlenses have intrinsically higher losses (lower FoM) than ENP superlenses, which can have significant, non-zero values of $\varepsilon_\perp$.

In summary, the loss figure of merit for any uniaxial superlens can be found by assuming only TM polarisation:

$$\text{FoM} = \frac{\text{Re}(k_{\perp,\text{TM}})}{\text{Im}(k_{\perp,\text{TM}})}.$$  \hspace{1cm} (5.15)

5.3 Angular dependence of losses

This section will briefly discuss the effect of varying the angle of incidence, $\theta_i$, on the metamaterial response. To investigate this, the response of a typical semiconductor superlens with varying $\theta_i$ is shown in Figure 5.5.

The main difference is that the height of the absorption peak at the plasma frequency, $\omega_p$, ($\lambda \approx 8.8\mu$m; see Equation 5.5) increases with increasing $\theta_i$. Fortunately, as $\omega_p$ is
5.4 Requirements for a “good” superlens

Figure 5.5: Effective medium response of a doped/undoped semiconductor multilayer for varying angle of incidence, $\theta_i$.

spectrally separated from the $\varepsilon_\perp$ resonance frequency, $\omega_{\text{res}}$, this strong absorption does not significantly affect the losses in the spectral region that the superlens operates ($\omega \approx \omega_{\text{res}}$).

5.4 Requirements for a “good” superlens

Unfortunately, the complicated relationship between the components of permittivity, the losses associated with the structure as a whole, and all of the parameters involved (e.g. layer thicknesses) is too complicated to extract an analytical solution for an optimal structure. Even the concept of an “optimal” superlens is unclear – there are trade-offs between lowering absorptive losses, flattening the dispersion curve (see subsection 2.4.3), and the difficulty of sample growth.

Therefore, it is necessary to narrow the scope to only a few critical performance indicators and find the parameters that effect the biggest changes in them. For an ENP superlens (i.e. with large $|\varepsilon_\perp|$ – see subsection 2.4.3), the most important indicators of superlensing performance are:

1. Large $\varepsilon_\perp/\varepsilon_\parallel$ – The magnitude of $\varepsilon_\perp$ must be maximised, as $\varepsilon_\parallel$ only varies gently (see Figure 5.2(b)). This will allow for the high spatial frequency, and so high spatial resolution, light to propagate with minimal diffraction;

2. Low absorptive loss – The resonance of the perpendicular effective permittivity, $\varepsilon_\perp$,
5.4. Requirements for a “good” superlens

Figure 5.6: The resonance in perpendicular permittivity, $\varepsilon_\perp$, (solid black line) strengthens greatly upon increasing scattering time, $\tau$, from (a) $\tau = 75$ fs to (b) $\tau = 150$ fs.

must be spectrally separated from the regions of high loss factors, $\alpha(\omega)$, at the plasma frequency, $\omega_p$, and at ENZ point ($\varepsilon_\parallel \approx 0$ – see subsection 2.4.3). This will allow enough light to propagate through the superlens for detection.

The following subsections will discuss how to optimise each of these in isolation.

5.4.1 How to achieve large anisotropy

Consider the resonance condition for $\varepsilon_\perp(\omega)$ with the Drude model:

$$\varepsilon_1 = \varepsilon_{01} \left( 1 - \frac{\omega_p^2}{\omega_{res}^2 + i\omega_{res}/\tau} \right) = -R\varepsilon_2, \tag{5.16}$$

where $\omega_{res}$ is the resonance frequency defined by Equation 5.7. Given that $\varepsilon_2$ is chosen to be real, it is necessary to minimise the relative magnitude of the imaginary term in the denominator ($\omega_{res}^2 - \omega_{res}/\tau = \omega_{res}(\omega_{res} - 1/\tau)$) of Equation 5.16, as it dampens the desired resonance. Therefore, $\tau$ must be maximised regardless of resonance frequency. This is shown in Figure 5.6, which demonstrates the importance minimising free carrier scattering time, $\tau$.

Assuming the charge carriers in a doped semiconductor can be modelled as a non-degenerate electron gas in a parabolic band, the mean scattering time, $\tau$, itself is given by [134]:

$$\tau = \frac{m^*\mu_e}{e}, \tag{5.17}$$

where $\mu_e$ is the mobility of the carriers. Furthermore, $\mu_e$ decreases with the increasing
5.4. Requirements for a “good” superlens

Figure 5.7: The resonance in perpendicular permittivity, $\varepsilon_\perp$, (solid black line) shifts (a) to the right of $\varepsilon_\parallel = 0$ with larger layer thickness ratio, $R = 1.25$, and (b) to the left of $\varepsilon_\parallel = 0$ with a smaller ratio, $R = 0.8$.

carrier density, $n_e$, as the addition of dopant atoms causes defects in the crystal lattice, increasing scattering [135]. Conversely, experiments show that $m^*$ increases with $n_e$ [136].

5.4.2 How to achieve low absorption

The next challenge is to reduce absorptive losses in the superlensing ENP region. As the central frequency for the $\varepsilon_\perp$ resonance, $\omega_{\text{res}}$, is always at a lower frequency than the plasma frequency, $\omega_p$ (see Equation 5.7), the ENP region must be shifted to the low loss region between $\omega_p$ and the frequency at which $\varepsilon_\parallel = 0$ (see Figure 5.2(b)).

Equation 5.7 shows that the layer thickness ratio, $R$, offers an extra degree of freedom in the multilayer design. Figure 5.7 demonstrates that tuning $R$ to values larger than one shifts the ENZ region to the left of the $\varepsilon_\parallel = 0$ point, and vice versa. This suggests that minimal loss, $\alpha$, can be attained by designing a structure with $R < 1$.

It is important to note that the scale of the $\varepsilon_\perp$ resonance also decreases with $R$, which diminishes the flatness of the superlens IDC (see section 2.4.2). This means that there is a trade-off between loss minimisation and anisotropy maximisation, so an optimal $R$ for a given structure must be determined numerically.

5.4.3 Optimisation constraints

Intrinsic semiconductor properties and growth considerations constrain each of the parameters discussed in the previous sections. Semiconductors are chosen as they have
appropriate permittivities in the mid-infrared region. The main constraints are:

- Layers 1 and 2 cannot be any arbitrary combination of semiconductors. To prevent the formation of crystal defects, which would put into question the simple uniform multilayer model of Figure 5.1, the semiconductors must have matching lattice parameters;

- Real semiconducting materials typically have bulk relative permittivity between 7 and 16, and a limit to maximum possible doping of \( n_e \approx 10^{19} \text{ cm}^{-3} \) [137];

- The minimum layer thickness (to avoid quantum effects) is \( \sim 10 \) nm;

- The bilayer thickness, \( d \), must be sufficiently smaller than the incident wavelength, \( \lambda \), to ensure the effective medium approach is justified;

- The superlensing \( \varepsilon_\perp \) resonance must occur at the desired wavelength.

The following section will introduce a new superlens design that aims to find an optimal balance goals of loss minimisation and anisotropy maximisation, subject to the above constraints.

5.5 Novel design for a low-loss superlens

As discussed in subsection 5.4.1, a material with large scattering time, \( \tau \), must be used for the doped layers in the superlens. Table 5.1 shows the electron transport properties of commonly used semiconductors, as found in the literature.

Two strong candidates stand out from these data, InAs and InSb, with \( \sim 5 \times \) larger scattering times, \( \tau \), than GaAs and In\(_{0.53}\)Ga\(_{0.47}\)As used for superlenses in the literature [24, 91]. InAs has the highest overall scattering time, and so is a strong candidate for a superlens. To avoid lattice matching issues, a multilayer of alternating doped and undoped InAs is proposed.

From experimental measurements in the literature, the effective mass of electrons in InAs is given by [136]:

\[
\frac{m^*_{\text{InAs}}}{m_e} = 0.023 + 1.3 \times 10^{-14} n_e^{2/3},
\]  

(5.18)

66
5.5. Novel design for a low-loss superlens

<table>
<thead>
<tr>
<th>Material</th>
<th>(m^*/m_e)</th>
<th>(\mu_e) (cm(^2)/Vs)</th>
<th>(\tau) (fs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GaAs</td>
<td>0.067</td>
<td>2,592</td>
<td>98.7</td>
</tr>
<tr>
<td>In(<em>{0.53})Ga(</em>{0.47})As</td>
<td>0.046</td>
<td>3,807</td>
<td>99.6</td>
</tr>
<tr>
<td>AlAs</td>
<td>0.150</td>
<td>132</td>
<td>11.3</td>
</tr>
<tr>
<td>GaP</td>
<td>0.350</td>
<td>120</td>
<td>23.9</td>
</tr>
<tr>
<td>InAs</td>
<td>0.045</td>
<td>17,500</td>
<td>448</td>
</tr>
<tr>
<td>InP</td>
<td>0.080</td>
<td>2,139</td>
<td>97.3</td>
</tr>
<tr>
<td>InSb</td>
<td>0.033</td>
<td>23,000</td>
<td>432</td>
</tr>
</tbody>
</table>

Table 5.1: Effective mass, \(m^*\), mobility, \(\mu\), and scattering time, \(m^*\), for various semiconductors at room temperature with electron concentration \(n_e \approx 1.2 \times 10^{18}\) cm\(^{-3}\). Data compiled from references [135, 137].

where \(m_e\) is the electron rest mass. From a semi-empirical model, the mobility of InAs is given by [135]:

\[
\mu_{e,\text{InAs}} = 1000 + \frac{33000}{1 + (n_e/1.1 \times 10^{18})^{0.32}},
\]  

in units of cm\(^2\)/Vs, where \(n_e\) is in units of cm\(^{-3}\). The infrared background permittivity of bulk InAs is taken to be \(\varepsilon_{01} = \varepsilon_2 = \varepsilon_{\text{InAs}} \approx 12.3\) [138].

Choosing \(d_1 = 70\) nm and \(d_2 = 100\) nm (i.e. \(R = 0.7\)) to position the ENP resonance in the low loss region, the response of a doped \((n_e = 0.7 \times 10^{18}\) cm\(^{-3}\))/undoped InAs multilayer (60 repeated bilayers) is shown in Figure 5.8.

For comparison, the In\(_{0.53}\)Ga\(_{0.47}\)As/Al\(_{0.48}\)In\(_{0.52}\)As multilayer response of reference [24] (see Figure 5.2(b)) will be used as a control. Figure 5.8 shows that the \(\sim 4.5\times\) increase in scattering time \(\tau\) of InAs \((\tau = 345\) fs\) versus \(\text{In}_{0.53}\text{Ga}_{0.47}\text{As} (\tau = 99.6\) fs\) in the “control” system has had the desired effect of greatly increasing the maximum \(\varepsilon_\perp\) from \(\varepsilon_{\perp,\text{max}} \approx 135\) to \(\varepsilon_{\perp,\text{max}} \approx 250\). Furthermore, the increased \(\tau\) has sharpened the absorption peak at \(\lambda \approx 9\) \(\mu\)m.

By selecting \(R = 0.7\), the spectral window of low loss (between the sharp absorption peak and the high loss \(\varepsilon_\parallel < 0\) region at \(\lambda \approx 14\) \(\mu\)m) has broadened and attained further reduced absorption. Finally, the decreased \(R\) has centred the ENP superlensing region (see subsection 2.4.3) in the improved low loss region, resulting in a loss factor \(\alpha \approx 52\) mm\(^{-1}\) that is more than \(12\times\) smaller than in the In\(_{0.53}\)Ga\(_{0.47}\)As/Al\(_{0.48}\)In\(_{0.52}\)As control
Figure 5.8: Effective response of 60×(doped/undoped)-InAs multilayer structure. The doping concentration is \( n_e = 0.7 \times 10^{18} \text{ cm}^{-3} \) in the doped layer, and the layer thickness ratio is \( R = 0.7 \). Note that the large \( \varepsilon_{\perp} \) (solid black line) resonance is positioned in the centre of the region, where the loss, \( \alpha \), (red dot-dash line) is consistently low.

\( (\alpha \approx 650 \text{ mm}^{-1}) \). This corresponds to a 100× increase in the loss figure of merit (see subsection 5.2.1) from \( \text{FoM} \approx 0.4 \) to \( \text{FoM} \approx 40 \). This is the first superlens design to offer an excellent loss factor \( \alpha < 100 \text{ mm}^{-1} \) and \( \varepsilon_{\perp,\text{max}} > 50 \) at the same time.

### 5.5.1 Tuneability of the superlens

Since the plasma frequency, \( \omega_p \), depends on the charge carrier concentration, \( n_e \), (see Equation 5.6) \( n_e \) can be used to tune spectral region where superlensing occurs. By changing only the doping (within experimentally achievable levels), Figure 5.9 shows that the superlens retains low losses (\( \alpha \approx 10040 – 60 \text{ mm}^{-1} \)) and large permittivity resonance (\( \varepsilon_{\perp,\text{max}} > 100 \)) across a wide spectral range \( \lambda \approx 5 – 25 \text{ \mu m} \). An important observation is that regardless of the doping concentration (and hence operating wavelength) that is chosen, setting the layer thickness ratio \( R \) to 0.7 keeps the superlensing region in the low loss window. It is also worth noting that the strength of the \( \varepsilon_{\perp} \) resonance increases as the superlens is tuned to shorter wavelengths.
5.6. Full-wave simulations

Figure 5.9: By changing only the doping concentration, $n_e$, from (a) $43 \times 10^{17}$ cm$^{-3}$ to (b) $0.8 \times 10^{17}$ cm$^{-3}$, the doped/undoped InAs superlens can be tuned from $\lambda \approx 5$ to 25 µm.

5.6 Full-wave simulations

Once a superlens has been designed, it is crucial to simulate how it will perform before attempting fabrication. The main question to be answered, is how thick should the superlens be, such that two sub-diffraction limit features will remain resolvable on the far side. If it is made too thick, absorption will dissipate any possible image. If it is made too thin, the superlens will be fragile. In order to answer this question, the full set of Maxwell’s equations must be used to simulate light propagating through the structure. There are a number of techniques available to do this, but here the finite-difference time domain method will be used.

5.6.1 The finite-difference time domain method

As the name suggests, the finite-difference time domain (FDTD) method solves numerically Maxwell’s equations, in combination with the constitutive relations, to simulate the evolution of electromagnetic fields in time. The FDTD method has the the capability of calculating fields for a band of a frequencies simultaneously. A further advantage is that this time-evolution of fields can be monitored, allowing deeper insights and understanding of what is occurring in a given system to be developed. The information in this section is a brief summary of the most relevant parts of [139].
5.6. Full-wave simulations

The Yee algorithm

The main algorithm for FDTD simulations was first proposed by Kane Yee [140] in 1966 and comprises the following steps:

1. Approximate all derivatives in Maxwell’s equations with central finite differences, i.e. assume that:

\[
\frac{\partial f(x)}{\partial x} \approx \frac{f(x + \frac{\Delta x}{2}) - f(x - \frac{\Delta x}{2})}{\Delta x}; \tag{5.20}
\]

2. Stagger the electric and magnetic fields in both space and time on a discrete grid;

3. Solve the resulting equations to give “update” equations that relate future fields to past fields;

4. Calculate the magnetic fields in the next time step, followed by the electric fields;

5. Iterate step 4 for a desired duration, or until a certain condition is reached (e.g. until the total energy left in the system is less than a certain amount).

As mentioned above, the evolution of fields can be monitored by sampling the fields after every \( n \) time steps. To recover spectral information, the fields are Fourier transformed at each step, then added to the Fourier transforms of the previous steps.

Boundaries in FDTD

Special consideration must be given to the boundaries in FDTD simulations. For example, simply setting all fields outside of the simulation to zero (so-called Dirichlet boundary conditions) results in perfect reflection of all waves at the boundaries. For waves incident normal to the structure of interest, periodic boundary conditions are commonly used for the axes normal to the direction of propagation. This is useful for simulating structures that can be approximated as infinitely periodic, such as diffraction gratings.

The situation becomes more complicated when considering structures that are not periodic, such as a single nano-antenna, and when considering boundaries along the axis of incidence. This led to the development of the perfectly matched layer (PML) boundary [141]. Strictly speaking, the PML boundary is a region of perfect absorption just inside of a Dirichlet boundary. This combination means that any fields that are incident on a boundary will be perfectly absorbed, such that no fields remain to be reflected at
5.7 Confirming low-loss super-resolution

To investigate the capabilities of the new superlens design, the structure shown in Figure 5.10 has been simulated with finite-difference time domain method (see subsection 5.6.1). The patterned gold mask creates a sub-diffraction-limit image of two slits immediately after the mask. The preservation of that image depends on the performance of the superlens underneath (see section 5.4). The following simulations were performed using a commercial FDTD solver (FDTD Solutions, Lumerical) with PML boundaries on the z-axis boundaries, and periodic boundaries in the x- and y-directions. The periodic boundaries were placed sufficiently far from the slits to prevent any contribution from neighbouring cells. The incident light was modelled as a broadband Gaussian pulse with a central wavelength $\lambda = 11.8 \, \mu m$ to probe the system across a range of wavelengths. In order to simplify the interpretation of results, the source propagates normally to the sample surface.

Figure 5.11 shows the results of simulating both the $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}/\text{Al}_{0.48}\text{In}_{0.52}\text{As}$ con-
5.7. Confirming low-loss super-resolution

Figure 5.11: The imaging performance of (a) the new InAs/InAs superlens (see Figure 5.8) against (b) the In\textsubscript{0.53}Ga\textsubscript{0.47}As/Al\textsubscript{0.48}In\textsubscript{0.52}As control superlens (see Figure 5.2(b)). The operating wavelength is $\lambda = 11.8$ $\mu$m in both cases. (c) Demonstration that the 800 nm slit separation is still resolved by the InAs superlens after 84 $\mu$m, but not by the In\textsubscript{0.53}Ga\textsubscript{0.47}As/Al\textsubscript{0.48}In\textsubscript{0.52}As control. (d) The InAs superlens resolves the features above the Rayleigh contrast limit ($a/b > 0.26$) up to 85 $\mu$m from the gold mask.

trol (see Figure 5.2(b)) and the new doped/undoped InAs superlens (see Figure 5.8).

Figure 5.11(a) shows that the new InAs design increases the propagation distance of the image of the gold mask relative to the control (see Figure 5.11(b)). Figure 5.11(c) defines the contrast ratio, $a/b$, which is used to determine the maximum depth that the image is still resolvable according to the Rayleigh criterion, $a/b > 0.26$ [142, 143]. The InAs superlens still resolves the deeply sub-wavelength ($\lesssim \lambda/15$) slit separation, at distances of up to 85 $\mu$m. In comparison with previously proposed mid-infrared superlenses [24, 83, 84, 91, 92, 144, 145], this imaging depth represents more than a 25$\times$ improvement.
5.7. Confirming low-loss super-resolution

Figure 5.12: Comparison of the superlens simulated as (a) an effective medium, and as (b) an exact multilayer. The simulated wavelength is $\lambda = 11.8\ \mu m$ in both cases. (c) The field profiles of the effective medium approximation (EMA) and exact structure are in good agreement. (d) The EMA and exact multilayer simulations agree on the maximum resolvable distance to within $\sim 5\%$.

5.7.1 Validity of the effective medium approximation

Recently, there have been a number of studies that have cast doubt on the accuracy of the effective medium approximation (EMA) in multilayer metamaterials [146–148]. As such, it is important to confirm that the EMA used in this work provides an accurate representation of light propagation in the superlens. As a test of the validity of the EMA, for this structure, additional FDTD simulations were performed without the use of this model. Specifically, a multilayer structure comprising repeating layers of doped InAs ($d_1 = 70\ nm$) and undoped InAs ($d_2 = 100\ nm$) were simulated explicitly in the same geometry as Figure 5.10. All other parameters were kept the same as in section 5.5.

Figure 5.12 shows that the image is preserved at similar depths for both the EMA and the explicit multilayer. Figure 5.12(d) demonstrates that the EMA predicts maximum
5.8. Conclusions

This chapter began with a systematic study of what it means for a superlens based on repeating bilayer semiconductors (see section 5.2) to perform well, and how an optimal superlens can be designed (see section 5.4).

The first key insight discovered by the author, is that the mean charge carrier scattering time, $\tau$, must be maximised, as this maximises the magnitude of the perpendicular permittivity, $\varepsilon_\perp$, in the superlensing ENP region (see subsection 2.4.3). The second insight is that the ratio of doped layer thickness to undoped layer thickness, $R$, must be tuned to $R \approx 0.7$, as this shifts spectrally the superlensing region into a region of substantially lower absorptive loss.

Using these new insights and imposing some experimental constraints (see subsection 5.4.3), a novel InAs-based superlens design was introduced in section 5.5 with extremely large anisotropy ($\varepsilon_\perp \approx 250$, $\varepsilon_\parallel \approx 3.5$) and low absorptive losses ($\alpha \approx 52$).

The superlensing performance of this new design was then tested using FDTD simulations (see section 5.6) in section 5.7. Here, the new InAs superlens demonstrated resolution of an 800 nm sub-wavelength feature ($\sim \lambda/15$) at distances of up to 85 $\mu$m ($\sim 7\lambda$). This represents a large (25×) increase in the superlens’s imaging distance, when compared with a control from the literature [24]. The results of this new superlens design were published in a paper entitled “Ultra low-loss super-resolution with extremely anisotropic semiconductor metamaterials” published in AIP Advances in February 2018 with Alexey Bak and Chris Phillips [149]. Work is in progress to fabricate a device for experimental validation using s-SNOM, as used by Bak et al. [92].
Chapter 6

Conclusion and Further Work

6.1 Conclusion

The overarching aim of this work was to investigate the potential applications of spectroscopic near-field imaging in the mid-infrared. In particular, s-SNOM was used in conjunction with a widely tuneable quantum cascade laser to explore two distinct themes: surface plasmon-polaritons (a.k.a plasmons), and biological imaging. The first theme was studied in chapter 3, where it was shown that s-SNOM is capable of providing nanoscale measurement of the Fermi energy in graphene with higher accuracy than standard techniques, such as electron transport measurements and Kelvin probe force microscopy.

In order to thoroughly establish the case for spectroscopic s-SNOM in the life sciences, chapter 4 began by demonstrating, for the first time, that s-SNOM can take optical images of human red blood cells with nanoscale spatial resolution. With this foundation, the remainder of the chapter set about using spectroscopic s-SNOM to perform IR spectroscopy (see section 2.1) at the nanoscale. In the process, IR s-SNOM was demonstrated to have ability to image cellular ultrastructure as small as 8 nm across. Ultimately, this proved to be highly successful, as shown by the chemical-specific mapping of the anti-cancer drug Bortezomib within a single human myeloma cell. This is the first time any medical drug has been successfully mapped directly (i.e. without fluorophores) in any cell and introduces a highly advantageous new chemical imaging technique for the life sciences.

The remainder of this work considered an alternative method of beating the diffraction limit, namely metamaterial superlenses (see chapter 5). In order to create a superlens
that performs well, the author first had to introduce a framework of what would even constitute a “good” superlens. Using this framework, a novel InAs-based superlens was proposed, which promises to maximise the key ability for a sub-wavelength image to be preserved over large distances. Simulations demonstrated that the new design was able to image at distances more than 25 times greater than any previously proposed superlens designs.

6.2 Further Work

The most obvious next step would be to use spectroscopic mid-infrared s-SNOM to map different chemical compounds in different biological systems. There exists a broad range of possibilities here, running the spectrum from studying the mechanisms of antibiotics on individual bacteria to seeking out new endogenous biomarkers for diseases, such as cancer and diabetes, in human tissue. Indeed, early work is underway to explore novel treatments of tuberculosis.

A particularly important task for a new imaging paradigm such as this is to undertake thorough correlative studies using more well-established imaging techniques, such as electron and fluorescence microscopy. This will allow for more confident interpretation of the images produced by infrared s-SNOM, for which there is, as yet, only a very small body of work to rely on.

The way forward for the superlens research is also quite clear. Naturally, the new design must be fabricated for ultimate experimental validation using s-SNOM. This also hints at an exciting new application – to use the superlens as a dry buffer between s-SNOM and live cells, which would otherwise render s-SNOM imaging extremely difficult, if not impossible. Aside from keeping the s-SNOM dry, this would also mean that the scanning surface could be atomically flat, allowing for faster tip scanning. This would remove a significant barrier to high-speed s-SNOM and could usher in a new wave of possibilities.


Bibliography


Bibliography


Bibliography


Bibliography


