Excursions in Electron Energy-Loss Spectroscopy

by

David Kordahl

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved May 2020 by the
Graduate Supervisory Committee:

Christian Dwyer, Chair
Peter Rez
John C.H. Spence
Maxim Sukharev

ARIZONA STATE UNIVERSITY

August 2020

ABSTRACT

Recent improvements in energy resolution for electron energy-loss spectroscopy in the scanning transmission electron microscope (STEM-EELS) allow novel effects in the low-loss region of the electron energy-loss spectrum ($\Delta E < 10$ eV) to be observed. This dissertation explores what new information can be obtained with the combination of meV EELS energy resolution and atomic spatial resolution in the STEM. To set up this up, I review nanoparticle shape effects in the electrostatic approximation and compare the "classical" and "quantum" approaches to EELS simulation. Past the electrostatic approximation, the imaging of waveguide-type modes is modeled in ribbons and cylinders (in "classical" and "quantum" approaches, respectively), showing how the spatial variations of such modes can now be imaged using EELS. Then, returning to the electrostatic approximation, I present microscopic applications of low-loss STEM-EELS. I develop a "classical" model coupling the surface plasmons of a sharp metallic nanoparticle to the dipolar vibrations of an adsorbate molecule, which allows expected molecular signal enhancements to be quantified and the resultant Fano-type asymmetric spectral line shapes to be explained, and I present "quantum" modelling for the charged nitrogen-vacancy (NV^-) and neutral silicon-vacancy (SiV^0) color centers in diamond, including cross-sections and spectral maps from density functional theory. These results are summarized before concluding.

Many of these results have been previously published in *Physical Review B*. The main results of Ch. 2 and Ch. 4 were packaged as "Enhanced vibrational electron energy-loss spectroscopy of adsorbate molecules" (99, 104110), and much of Ch. 5 appeared as "Prospects for detecting individual defect centers using spatially resolved electron energy loss spectroscopy" (100, 134103). The results from Ch. 3 are being prepared for a forthcoming article in the *Journal of Chemical Physics*.

ACKNOWLEDGEMENTS

Many people deserve to be thanked. Foremost among them is Christian Dwyer, who has supported me generously throughout my time at ASU. This dissertation would be much poorer without his extensive contributions and deep insights. I am also grateful for the help I received from Shery Chang, who taught me about microscopes, and for the varied contributions of Peter Rez, Maxim Sukharev, and John Spence. I also benefitted from the advice of my friends Lucas Madeira, Matt Voigts, and Jake Rapp, who became "doctors" before me, and from the commiserations of Ayush Saurabh and Jiawei Liu, who have trod the path beside me, as have my office mates Lance (Weiqing) Xu and Haotian Wen. My family also deserves a great deal of credit. My parents encouraged my return to graduate school and, year after year, gifted me a parking pass. My sister Angela has helped me to parse my complaints. Finally, my wife Holly and daughters Iris and Beatrix have given me a reason to keep at it, and I am thankful for the rich life that we share.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

FIGURE Page

FIGURE Page

Chapter 1

INTRODUCTION

1.1 The Story of the Elephant

Children around the globe are taught some version of the fable of the blind men and the elephant. This is the fable in which a few blind men (or, in another version, a few sighted men in a dark room) all encounter an elephant, and, as elephants are typically bigger than men, each man happens upon a different part of the animal. Depending on the part that each man encounters, each comes to a different conclusion about its nature, with the man who finds the elephant's leg concluding that the elephant is like a tree, the man who finds the elephant's trunk concluding it is like a snake, and so on. Usually, this fable ends with a call to epistemic humility. In versions striving toward profundity, the elephant is sometimes identified as God.

But in a scientific version of this fable, the elephant might stand in for whatever physical systems we might wish to study, and the blind men's differing grasps might stand in for all our different methods of studying such systems—from passive observations of the night sky, to the active bombardment of materials with electron beams. In the scientific version, we might further imagine that the blind men are able to collaborate. This modified version of the story has a different moral than the original one, since, by each sharing their findings, these local elephant investigators are collectively able to form a stable global picture of the underlying elephant.

Outside the academic setting, but during the time the research contained in this dissertation was conducted, I have argued, variously, that science alone cannot give evidence for whether or not God exists (Kordahl (2017)), that the "truth" of a scien-

tific theory may depend on its historical context (Kordahl (2018)), and that physics should not be seen as driving toward a single, fixed goal (Kordahl (2019)). Regardless, here I discuss systems whose underlying forms have been the subject of shared research interest by many investigators, preventing much of the radical doubt that some scientific questions invite. Since the basic forms of these systems are not in question, the relevant questions then become higher-order ones of what sorts of signals such systems might generate, and of whether such signals are detectable.

This dissertation is concerned with just one method of the proverbial elephant-bothering—namely, that of electron energy-loss spectroscopy in the scanning transmission electron microscope (STEM-EELS). STEM-EELS is a versatile technique for studying materials, and can probe both their collective and atomistic properties. Chapters 2 and 3 discuss effects that arise from the collective properties of nanostructures—in particular, effects that can be captured by a local dielectric description, in the electrostatic (Ch. 2) and electrodynamic (Ch. 3) versions of the theory. Chapters 4 and 5 discuss atomistic properties—in particular, the way that vibrational signals from an adsorbate can be enhanced by a nanoparticle substrate (Ch. 4), and the way that individual point defects might be located in bulk materials (Ch. 5).

To set this up, in this introduction I briefly review the motivation for this dissertation, summarizing a bit of history, a bit of theory, and a bit about our elephants.

1.2 A Brief History of STEM-EELS Microscopes

In modern tellings, the idea of electron microscopy is sometimes presented as an obvious extension, on the one hand, of Abbe's theory of how waves can combine to form images, and, on the other, of de Broglie's insight that electrons have a wavelike nature. In a paper on the modern theory of electron optics, Rose (2008) recounts that Abbe himself suspected there might be some way to overcome the wavelength-based

limits on resolution mandated by imaging with visible light, but as Mulvey (1962) reveals, history is often more tangled than any conceptual reconstruction. When Hans Busch discovered that a solenoid's magentic field could act on electrons in much the same way as a convex lens acts on light, he connected his theory to geometric optics, not Abbe theory, and Max Ruska (1986) recounted that when he and Max Knoll built their electron microscope prototype, they were not aware of resolution limits for electrons. "As engineers we did not know yet the thesis of the 'material wave' of the French physicist de Broglie that had been put forward several years earlier."

Nevertheless, by the end of the 1930s, the transmission electron microscope (TEM) had been developed to the point where the basic theory was understood, and where achievable magnifications on commercially available TEMs surpassed those of the best light microscopes. Not long thereafter, the first recognizable results in the style of STEM-EELS were presented by Hillier and Baker (1944). In this work, Hillier and Baker discussed their efforts to focus an electron beam on a selected sample region, and to separate outgoing electrons using a magnetic prism according to the how much energy each had lost in transit, which was identified as the energy each electron had given up to promote electronic transitions of elements in the sample.

While important work on developing the TEM continued in the years that followed, including the foundational work on resolution limits by Scherzer (1949), early histories (e.g., Big (1956)) mark the electron microscope mainly as a tool for biologists. The development of STEM-EELS as a technique for materials characterization is the often identified as having been spurred on by the microscope design improvements pioneered by Crewe *et al.* (1968). A first-person account of the changes in instrumentation and their consequences in the decades that followed is given by Egerton (2012), who briefly reviews how changes like the shift from photographic to digitized spectra opened up new opportunities for increased rigor in the interpretation of results.

Then, as now, STEM-EELS applications were often characterized as either "low loss," with energy losses in the tens of eVs, or "core loss," with energy losses in the hundreds of eVs. This dissertation is focused on low-loss applications, although interest in the STEM-EELS traditionally has been heavily weighted toward core-loss applications. At the end of the last century, a review article on core-loss EELS (Silcox (1998)) could claim an impressive range of accomplishments, in terms of instrumentation (with probe sizes down to a few ångstroms and energy resolution down to a few eV), theory (with atomic cross-sections under various approximations available for quantitative comparison), and applications (ranging from the characterization of high-T_c superconductors to detailed studies of grain boundaries in alloys).

But this was far from the end of the story. Real electron lenses, like real glass lenses, are plagued with various imperfections, from the failure to converge waves from single points in the object plane to single points in the image plane, as in spherical aberration; to the failure for waves of equal momenta in the object plane to reach single points in the diffraction plane, as in chromatic aberration; to the failure of a lens to maintain circular symmetry, as in coma or astigmatism. In recent decades, such issues in electron optics have been confronted directly. Rose (1971) established early on that the spatial resolution limits marked by Scherzer (1949) for circularly symmetric lenses would not apply to multipole lenses, but it took decades of development for this insight to become useful. Spherical aberrations were first corrected by Haider *et al.* (1998), and soon thereafter Batson *et al.* (2002) reported the first computer-aided aberration-corrected setup with sub-ångstrom resolution, leading to a period of intense interest in atomically-resolved STEM-EELS.

Likewise, energy resolution limits were pushed by the incorporation of monochromators, which spatially disperse the electron beam according to kinetic energy and select only those electrons with appropriate beam energies by passing them through a

narrow slit. Kimoto (2014) reviews the significant commercial designs. FEI designed a monochromator using a single Wein filter (a region of crossed magnetic and electric fields whose Lorentz force vanishes only for a specific energy of electron passing through it) which disperses the electron beam before it goes through the acceleration tube, and uses a selector slit after acceleration. JEOL improved on this with a model whose selector slit was before the acceleration tube, sandwiched between two Wein filters, correcting for spatial chromaticity (i.e., for energy dispersion in the final image). CEOS's "omega-type" monochromator, so called for the bent path electrons take through it, introduces spatial dispersion with electrostatic elements, and manages to correct not only for spatial but also for angular chromaticity resulting from electrons of different directions having different energies.

These designs all have their spatially dispersive elements before the acceleration tube, which allows any instabilities in the accelerating voltage to be reflected in the electron beam. To fix this, Krivanek *et al.* (2009) developed an "alpha-type" monochromator, again named for the beam electron's path, whose design differs markedly from the others in that its elements are all implemented after the beam has already been accelerated. This requires extra multipole lenses to make the slit selection feasible for the high-energy beam, and allows the monochromator to function at ground voltage. For beam energies of 100-200 keV, this monochromator was designed to attain $\sim$10 meV resolution for a beam focused to $\sim$1 Å^2, a goal that was achieved in microscopes built by the Nion company, allowing vibrational STEM-EELS signals to be observed for the first time (Krivanek *et al.* (2014)).

This dual achievement of high spatial and energy resolution in STEM-EELS forms the backdrop, and the motivation, for the work documented in this dissertation, and has led to many groundbreaking STEM-EELS studies on the vibrational properties of materials, including the work of Miyata *et al.* (2014), Nicholls *et al.* (2015), Dwyer

et al. (2016), Rez *et al.* (2016), Lagos *et al.* (2017), Hage *et al.* (2019), Venkatraman *et al.* (2019), Senga *et al.* (2019), and Hage *et al.* (2020). This renewed interest in low-loss phenomena has reopened once-settled questions to a new era of investigation.

1.3 A Brief History of Low-Loss Theory

Much of the interest in STEM-EELS as an experimental technique arises from its ability to aide in atomic-scale materials characterization. The quantitative theory of atomically-resolved imaging in STEM-EELS has been reviewed in some detail by Dwyer (2013) and Allen *et al.* (2015). This dissertation, however, with its focus on low-loss STEM-EELS, needs only to brush against that literature long enough to establish how the (mostly classical) low-loss models and the (mostly quantum) core-loss models can be seen as two complimentary ways of tickling the same elephant.

An early classical calculation in the modern style was carried out by Fermi (1940), who wished to calculate the energy loss per distance traveled for a charged particle passing through water or air. He treated the loss medium as a uniform material with a dielectric constant and treated the charged particle as having a fixed speed. After introducing an imaginary tube around the beam (effectively a spatial cutoff, keeping the potential energy of the point-like classical electron from diverging), he calculated the energy flowing outward by integrating over the Poynting vector.

Scattering of electrons by many-body quantum systems is, in principle, a difficult problem, and it is not immediately obvious how the quantum-mechanical description of an electron's interaction with various materials might relate to the classical description. The early work of Bohm and Pines (1953) on plasma oscillations in metals showed that normal modes of many-body systems could be related to the underlying quantum theory. But it was not until Hubbard (1955) showed how the long-ranged dynmaics of conduction electrons could be captured in the parameters of the Drude

model for metals that the well-developed dielectric models of classical electrodynamics could be interpreted in the light of quantum theory. This allowed the process of calculating spectra to be simplified, and led Ritchie (1957) to predict the existence of surface plasmons in metallic foils.

These early treatments laid the foundation for most subsequent work in the classical style, whose schematic outline is easy enough to present. The beam, in this treatment, is treated as a point charge running along z at a constant speed v:

$$\rho_e^c(\mathbf{x}, z, t) = -e\delta^2(\mathbf{x} - \mathbf{x}_0)\,\delta(z - vt) \tag{1.1}$$

The sample gives rise to an electric field $\mathbf{E}^{\text{ind}}$ (i.e., the total electric field minus the electric field associated with the beam electron). To solve for $\mathbf{E}^{\text{ind}}$, it is convenient to Fourier-transform the field of the beam electron into frequency-space before applying spatial boundary conditions. After the inverse transform, the work done on the beam electron by the z-component of $\mathbf{E}^{\text{ind}}$ is found using "force times distance":

$$-W = \frac{e}{2\pi}\int_{-\infty}^{\infty} dz \left[\int_{-\infty}^{\infty} d\omega \exp(-i\omega t) E_z^{\text{ind}}(\mathbf{x}, z, \omega)\right]_{\mathbf{x}=\mathbf{x}_0, z=vt}. \tag{1.2}$$

To obtain an expression for the EEL spectrum, we can switch the order of the ω and z integrals and rewrite the ω integral over a positive domain

$$-W = \int_0^{\infty} d\omega(\hbar\omega)\text{Re}\left[\frac{e}{\pi\hbar\omega}\int_{-\infty}^{\infty} dz \exp(-i\omega z/v) E_z^{\text{ind}}(\mathbf{x}_0, z, \omega)\right], \tag{1.3}$$

which allows the integrand to be interpreted as a STEM-EELS spectrum:

$$\frac{dP}{d\omega} = \frac{e}{\pi\hbar\omega}\text{Re}\left[\int_{-\infty}^{\infty} dz \exp(-i\omega z/v) E_z^{\text{ind}}(\mathbf{x}_0, z, \omega)\right]. \tag{1.4}$$

Decades of subsequent applications have proven the utility of this approach. A summary of this work was undertaken by Wang (1996), and more recently by García de Abajo (2010). Recent developments in this vein have been reviewed by

Talebi (2018). Calculations in this style will take up much of Ch. 2 and Ch. 4 below, so it is important here to linger on how this approach is consistent with the quantum-mechanical approach usually applied in modeling core-losses.

The "quantum" approach to low-loss STEM-EELS involves writing out the singly-occupied normal modes of the sample $|1_n\rangle$, scaled to have energy $\hbar\omega_n$, allowing for conceptually similar methods to be used in low-loss and core-loss STEM-EELS. The possibility of writing down such modes was pointed out in the classic work of Fuchs and Kliewer (1965), who worked out the modes of the dielectric slab in the electrostatic limit, then in a fully relativistic treatment (Kliewer and Fuchs (1966a,b)). Englman and Ruppin (1966) showed how these results could be generalized to any canonical geometry, and Lucas and Šunjić (1972) showed how the dielectric modes of the slab could be probed using STEM-EELS, clearing the way for investigations of many other geometries.

So how, roughly, does one obtain a result of the sort given by Eq. 1.4 in this context? Early on in the development of scattering theory, Frame (1931) established that an α-particle exciting an atomic transition could be treated either as a quantum object or as a classical charge center without altering the transition probability obtained, given the large mass of the α-particle relative to that of the atomic electrons. For typical beam energies considered with STEM-EELS (10-300 keV), a similar observation applies for inelastic electron scattering from low-loss normal modes (0-10 eV), since the beam energy is over a thousand times that of the considered losses. Treating the beam classically leads to the correct approximation for the first-order transitions, even as we conceede that this cannot be the entire story.

As before, it is simple enough to outline the quantum approach schematically. Suppose that a sample is initially in its ground state $|0\rangle$. To construct the STEM-EEL spectrum, we would like to calculate its probability of transition to the single-

occupancy state $|1_n\rangle$ in the presence of the beam, where n is a label for the state's quantum numbers. If each state has an associated classical potential $\Phi_n(\mathbf{r})\exp(i\omega_n t)$ whose scaling matches the energy normalization, we can write the beam-sample interaction potential in terms of undetermined coefficients c_n as

$$V_I(t) = \sum_n c_n \int d^3\mathbf{r}\, \rho_e^c(\mathbf{r},t)\Phi_n(\mathbf{r})\exp(i\omega_n t) \tag{1.5}$$

where $\rho_e^c(\mathbf{r},t)$ is just the expression given by Eq. 1.1. We can then use first-order perturbation theory to solve for $c_n^{(1)}$

$$c_n^{(1)} = \frac{1}{i\hbar}\int_{-\infty}^{\infty} dt \int d^3\mathbf{r}\, \rho_e^c(\mathbf{r},t)\Phi_n(\mathbf{r})\exp(i\omega_n t), \tag{1.6}$$

which, upon squaring, leads to a probability of transition to state $|1_n\rangle$:

$$P_n^c(\mathbf{x}_0) = \left(\frac{e}{\hbar v}\right)^2 \left|\int_{-\infty}^{\infty} dz \exp(i\omega_n z/v)\Phi_n(\mathbf{x}_0,z)\right|^2. \tag{1.7}$$

Since energy is conserved overall, the energy gained by the sample is the energy lost by the beam, and the STEM-EEL spectrum is written as

$$\frac{dP}{d\omega} = \sum_n P_n^c(\mathbf{x}_0)\delta(\omega-\omega_n), \tag{1.8}$$

where the delta functions can be replaced by distributions whose width is related to the imaginary part of ω_n when states have a finite lifetime.

Even if we recast the transition probabilities P_n in terms of the electric field

$$P_n = \left(\frac{e}{\hbar\omega_n}\right)^2 \left|\int_{-\infty}^{\infty} dz \exp(i\omega_n z/v)E_{z,n}^{\mathrm{ind}}(\mathbf{x}_0,z)\right|^2, \tag{1.9}$$

it is not obvious from the notation that the expression for the classical expression (Eq. 1.4) must match the quantum expression (Eq. 1.8) in the limit of a real $\epsilon(\omega)$. In Sec. 2.4, however, we show how this equivalence can be established for a specific case,

that of the sphere with a model dielectric function that respects causality, a result supported by more general arguments in Wang (1996).

In core-loss simulations, the initial and final states of a system both must be explicitly considered, a point that is revisited in Sec. 5.2 below. This issue is often ignored in low-loss simulations, as the ground state $|0\rangle$ contributes no macroscopic electric field. For much of the work that follows, the level of approximation presented above is already sufficient. Lucas and Šunjić (1972) considered the possiblity of multiple occupancy states in the dielectric slab (i.e., $|2_n\rangle$, $|3_n\rangle$,...), but concluded that probabilities were very small for electrons at normal incidence. Corrections accounting for the probe size have been found to have only modest effects if one supposes that all inelastically scattered electrons are collected. Using the first Born approximation, Ritchie (1981) established that broad-beam illumination would merely reproduce the results for a classical beam, only averaged over the area of illumination.

Later, Ritchie and Howie (1988) extended this analysis to investigate a beam whose wavefunction profile in the sample plane is described by $\phi(\mathbf{x})$. They found that the corrected quantum-beam transition probabilities could be expressed in terms of the classical-beam transition probabilities as

$$P_n^q(\mathbf{x}_0) = \int d^2\mathbf{x}\, |\phi(\mathbf{x} - \mathbf{x}_0)|^2\, P_n^c(\mathbf{x}_0), \tag{1.10}$$

that is, in terms of a 2D convolution with the probability density of the electron in the sample plane. Hence, if the features we wish to probe are larger than the beam size (as now is often the case, given the impressive strides in instrumentation outlined above), it is reasonable to use the delta-beam approximation.

1.4 A Preview of Our Elephants

This introduction has barely scratched the surface of the possible subjects, even within the constrained scope of STEM-EELS. There exist many fine overviews of high-resolution electron microscopy from which one may contextualize the current state of research, including the books by Kohl and Reimer (2008) and Spence (2009), and and a full-volume treatement of EELS technique by Egerton (2011). Introductory treatments of STEM-EELS by Egerton (2008) and Eljarrat *et al.* (2019) also provide helpful overviews, and the first volume of *Science of Microscopy* (Hawkes and Spence (2007)) gives a long view on the history of microscopy in its introduction, and specific STEM-EELS discussions in Ch. 2. But what we have is enough already to proceed.

Here is a brief overview of the topics discussed in the chapters that follow.

In the second chapter, the simple physics of frequency-dependent dielectrics is discussed in terms of the Born-Huang model, and the surface modes of canonical dielectric geometries are treated in the electrostatic approximation. The electrostatic surface modes are solutions to Laplace's equation, and occur at frequencies for which the real part of the dielectric function is negative, providing a restoring force. The "classical" and "quantum" approaches are carried through for the slab, and are shown to be equivalent for the case of the sphere. This analysis allows the important surface modes to be identified for various canonical geometries, including those of oblate and prolate spheroids, the latter of which is reused extensively in the fourth chapter.

In the third chapter, I discuss guided light modes of semi-infinite nanostructures, modes that are not captured by the electrostatic approximation. Here I consider models with the beam running parallel to the material interface, leading to expressions of $d^2P/dz\,d\omega$. As before, one can study these modes either through "classical" methods (i.e., using boundary conditions) or through "quantum" methods (i.e., using normal

modes). For the dielectric ribbon, these modes split into the familiar transverse electric and transverse magnetic (TE/TM) designations, which are conveniently studied through "classical" methods. For the dielectric cylinder, the TE/TM modes become become hybridized, which makes them more convenient to study through "quantum" methods. The modal analysis for cylinders finds promising agreement with the exprimental results of Flauraud and Alexander (2019), whose data I explore.

In the fourth chapter, I return to the electrostatic approach, but use it to develop a model for the coupling of adsorbates to metallic nanoparticles. This model captures the electromagnetic signal effects described by surface-enhanced Raman scattering (SERS), except in the context of STEM-EELS. In this model, the vibrational signal from the surface molecule appears as a Fano-type resonance atop the nanoparticle's background signal. To achieve experimentally relevant results for signal and signal-to-noise enhancements, the model molecule is positioned on the tip of a metallic prolate spheroid, which enhances the electric field of the passing beam electron, not unlike the effect of a lightning rod in a external field. This leads to signal enhancements $O(10^4)$ and signal-to-noise enhancements $O(10^2)$, with even greater effects possible when the nanoparticle is "tuned" to the frequency of the molecular vibration.

Finally, in the fifth chapter, I present results for the possibilities of imaging point defects in diamond, based on the outcome of density functional theory calculations. The point defects under consideration, the charged nitrogen-vacancy and neutral silicon-vacancy defects, have spatial maps that can be intuitively interpreted using the tight-binding approximation. These calculations give reason to believe that STEM-EELS measurements should be able to localize such defects to about a nanometer, and point the way forward toward single-defect detection for defects buried in bulk samples, acting as a case study in first-principles quantum modeling for the simulation of STEM-EELS signals. These results are briefly summarized before concluding.

Chapter 2

STEM-EELS OF ELECTROSTATIC MODES

2.1 Introduction

In *Science Since Babylon*, the historian of science de Solla Price (1975) distinguished between the prescientific attitudes of the Babylonians and of the Greeks. The Babylonians, he argued, were adept at numerical calculations, accumulating detailed schemes for precisely predicting the apparent motion of the planets, while never proposing any compelling models to give their calculations meaning. The Greeks, on the other hand, were advanced in conceptual argumentation, proposing models for the planetary motions, even as their calculations were comparatively clumsy. Modern science, argued de Solla Price, is the peculiar synthesis of these two styles of activity, where precise numerical calculations meet conceptually satisfying arguments.

The first purpose of this chapter is to present an overview of two different ways of calculating low-loss STEM-EELS spectra in the dielectric approach, under the electrostatic approximation. The first of these approaches, the "classical" approach, is more Babylonian in style, and the spectrum arises from the boundary conditions as a more or less automatic result, without any necessary insight into the material causes. The second of these approaches, the "quantum" approach, is more Greek in style, requiring increased upfront effort, with interpretive insight as the dividend. As we shall see, both approaches give equivalent results in the undamped limit, despite their seeming conceptual differences, and I will use each where convenience serves.

The second purpose of this chapter is to present a unified approach to shape effects in nanoparticles. The slab and the sphere act as two limiting cases, with

the slab having the widest possible range of frequencies for its surface states, and with the sphere having the narrowest possible range of frequencies for such surface states. Using spheroids as intermediary geometries, it is possible to understand how the surface states of the sphere evolve to those of semi-infinite geomtries, with the sphere evolving toward the slab via the surface states of the oblate spheroid, and the sphere evolving toward the cylinder via the surface states of the prolate spheroid.

A brief summary of the work in this chapter has been presented in Kordahl and Dwyer (2019), which also contains much of the material developed in Ch. 4. In that later chapter, it will become apparent how an understanding of these surface states can lead to specific applications for nanoparticles.

This chapter is organized as follows. First, I review the simple model of dielectric response proposed by Born and Huang (1954), a model that covers both the classical Drude and Lorentz dielectric models. In contrast to the classic treatment of Born and Huang, who limited their treatment to bulk samples, here the model is considered in finite particles, for which suface effects are important. I discuss STEM-EELS scattering from a dielectric slab, both in the "classical" and "quantum" calculations. After showing how these methods can be generalized to other geometries, I show how the two approaches are equivalent for aloof scattering from the sphere (that is, for scattering where the beam does not penetrate the sphere itself), and argue that this will be true generally where the imaginary part of the dielectric function is small. Finally, I summarize how the important modes develop from one geometry to another, emphasizing how shape effects arise from interactions between the beam and oscillating surface charge.

2.2 Born-Huang Model for Dielectrics

In modern texts on solid-state physics (cf. Cohen and Louie (2016)), the dielectric function is typically presented as an instance of *linear response* theory. For a weak enough perturbation, the electric potential Φ^{in} inside a material is screened linearly with respect to the "external" potential Φ^{ext} that has been introduced into the system while not being contained within the dielectric response itself. The inverse dielectric function $\epsilon^{-1}(\mathbf{r}, \mathbf{r}', t - t')$ is introduced via a general two-point function as

$$\delta\Phi^{\text{in}}(\mathbf{r}, t) = \int d\mathbf{r}' dt' \epsilon^{-1}(\mathbf{r}, \mathbf{r}', t - t')\, \delta\Phi^{\text{ext}}, \tag{2.1}$$

which allows $\epsilon^{-1}(\mathbf{r}, \mathbf{r}', t - t')$—or, equivalently, for $\epsilon(\mathbf{r}, \mathbf{r}', t - t')$ with the roles of Φ^{in} and Φ^{ext} reversed—to capture a wide range of physical effects, including spatial inhomogeneities and local anisotropies, as well as non-local quantum correlations, as in the famous Lindhard dielectric function for the free electron gas.

My discussion of dielectrics will be restricted to the *local* dielectric description, where the displacement field $\mathbf{D}(\mathbf{r}, t)$ is related to the electric field $\mathbf{E}$ in a material as

$$\mathbf{D}(\mathbf{r}, t) = \int \epsilon(t - t')\mathbf{E}(\mathbf{r}, t')dt', \tag{2.2}$$

which, having the form of a convolution in time, makes it convenient to solve many dielectric problems in Fourier space:

$$\mathbf{D}(\mathbf{k}, \omega) = \int d^3\mathbf{r}\, dt\, e^{-i(\mathbf{k}\cdot\mathbf{r} - \omega t)} \epsilon(t - t')\mathbf{E}(\mathbf{r}, t') = \epsilon(\omega)\mathbf{E}(\mathbf{k}, \omega). \tag{2.3}$$

Implicitly, Eq. 2.3 defines the form of the form of Fourier transforms used throughout this document.

Given the macroscopic Maxwell equations (Jackson (1999))

$$\begin{aligned} \nabla \cdot \mathbf{D} &= 4\pi\rho_{\mathrm{f}} \\ \nabla \cdot \mathbf{B} &= 0 \\ \nabla \times \mathbf{E} &= -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} \\ \nabla \times \mathbf{H} &= \frac{4\pi}{c}\mathbf{J}_{\mathrm{f}} + \frac{1}{c}\frac{\partial \mathbf{D}}{\partial t}, \end{aligned} \tag{2.4}$$

and the constitutive relations for isotropic linear media

$$\begin{aligned} \mathbf{D} &= \mathbf{E} + 4\pi\mathbf{P} = \epsilon\mathbf{E} \\ \mathbf{B} &= \mathbf{H} + 4\pi\mathbf{M} = \mu\mathbf{H} \end{aligned} \tag{2.5}$$

one can roll the effects of "bound" charge and current into ϵ and μ, respectively. The conventional electric boundary conditions at dielectric interfaces

$$\begin{aligned} \left(\mathbf{D}^{\mathrm{out}} - \mathbf{D}^{\mathrm{in}}\right) \cdot \hat{\mathbf{n}} &= 4\pi\sigma_{\mathrm{f}} \\ \left(\mathbf{E}^{\mathrm{out}} - \mathbf{E}^{\mathrm{in}}\right) \times \hat{\mathbf{n}} &= 0 \end{aligned} \tag{2.6}$$

become especially simple in the absence of free charge, with $\sigma_{\mathrm{f}} = 0$, and similarly for the magnetic boundary conditions and $\mathbf{J}_{\mathrm{f}}$.

I will further restrict our attention to non-magnetic materials, where $\mu = 1$, and will work in the electrostatic approximation (that is, the approximation that $c \to \infty$), which simplifies Maxwell's equations brutally, setting the last three to zero on the right-hand side. For uncharged particles characterized entirely by the dielectric description, our problem reduces to solving Laplace's equation:

$$\nabla^2\Phi(\mathbf{r},\omega) = 4\pi\rho_{\mathrm{f}}(\mathbf{r},\omega)/\epsilon(\omega). \tag{2.7}$$

In STEM-EELS simulations, the "free charge" is typically just the beam electron.

Here I would like to develop a model dielectric in the framework of Born and Huang (1954), whose theory comprises a pair of equations relating the ionic displacement $\mathbf{u}$,

the polarization density $\mathbf{P}$, and the macroscopic electric field $\mathbf{E}^{\text{in}}$ inside a particle. The equations include four physical parameters—the transverse resonant frequency ω_0; the effective ionic charge Z_0^*; the high-frequency dielectric constant ϵ_∞; and a damping parameter η—and are most legible in their explicitly time-dependent form:

$$\begin{aligned} Z_0^* \mathbf{E}^{\text{in}} &= \ddot{\mathbf{u}} + 2\eta\dot{\mathbf{u}} + \omega_0^2 \mathbf{u}, \\ \mathbf{P} &= Z_0^* \mathbf{u} + \frac{\epsilon_\infty - 1}{4\pi} \mathbf{E}^{\text{in}}. \end{aligned} \tag{2.8}$$

This model is manifestly *local*, and each equation has a clear physical meaning. The first treats the ionic displacements as a field of damped harmonic oscillators driven by the macroscopic electric field. The second stipulates that these displacements cause the polarization density, although the effects of fast-moving charges are rolled into ϵ_∞ and provide electronic screening.

More precisely, $\mathbf{u}$ is the displacement of ions from their equilibrium positions multiplied by $(\mu/\Omega)^{1/2}$, where μ is the reduced mass and Ω is the volume of the of the ionic pair, and the effective charge Z_0^* is a measure of the split between the low- and high-frequency limits of the dielectric function (ϵ_0 and ϵ_∞, respectively), or equivalently of the the LO-TO split, where ω_ℓ is the bulk longitudinal frequency:

$$4\pi (Z_0^*)^2 = (\epsilon_0 - \epsilon_\infty)\omega_0^2 = \epsilon_\infty(\omega_\ell^2 - \omega_0^2). \tag{2.9}$$

If we suppose that $\mathbf{u}$, $\mathbf{P}$, and $\mathbf{E}$ are continuous and smooth functions inside the particle with a time dependence of $\exp(-i\omega t)$, the two Born-Huang equations and the electrostatic equations lead us to find an expression for $\epsilon(\omega)$:

$$\epsilon(\omega) = \epsilon_\infty \frac{\omega(\omega + 2i\eta) - \omega_\ell^2}{\omega(\omega + 2i\eta) - \omega_0^2}. \tag{2.10}$$

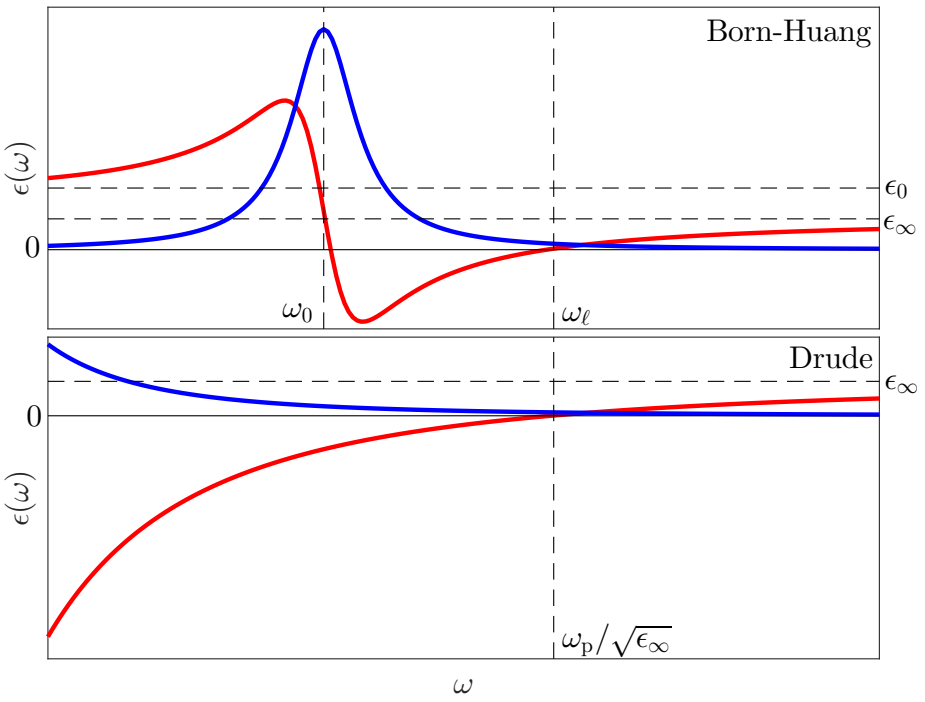

Figure 2.1: The dielectric function $\epsilon(\omega)$ for the Born-Huang and Drude models, with real (red) and imaginary (blue) parts. In the Born-Huang model a vanishing real part specifies the longitudinal frequency ω_ℓ, while in the Drude model it relates to the model plasma frequency ω_p.

As has been recognized by others (e.g., Lucas and Šunjić (1972); Lourenço-Martins and Kociak (2017)), this model is formally equivalent to the Drude model for metals in the case where the oscillators have no restoring force, and with limits

$$\omega_0 \to 0, \qquad 4\pi(Z_0^*)^2 \to \omega_\mathrm{p}. \tag{2.11}$$

Just as the Born-Huang model describes isotropic insulators (e.g., lithium flouride or cubic boron nitride), the Drude model describes noble metals (e.g., silver or gold). Hence plasmons function as longitudial optical phonons in the long-wavelength limit, and many results for plasmons translate to the vibrational context. The Born-Huang and Drude dielectric functions are compared in Fig. 2.1. In the next section, I apply Eq. 2.10 to the slab geometry in two different methods of solution.

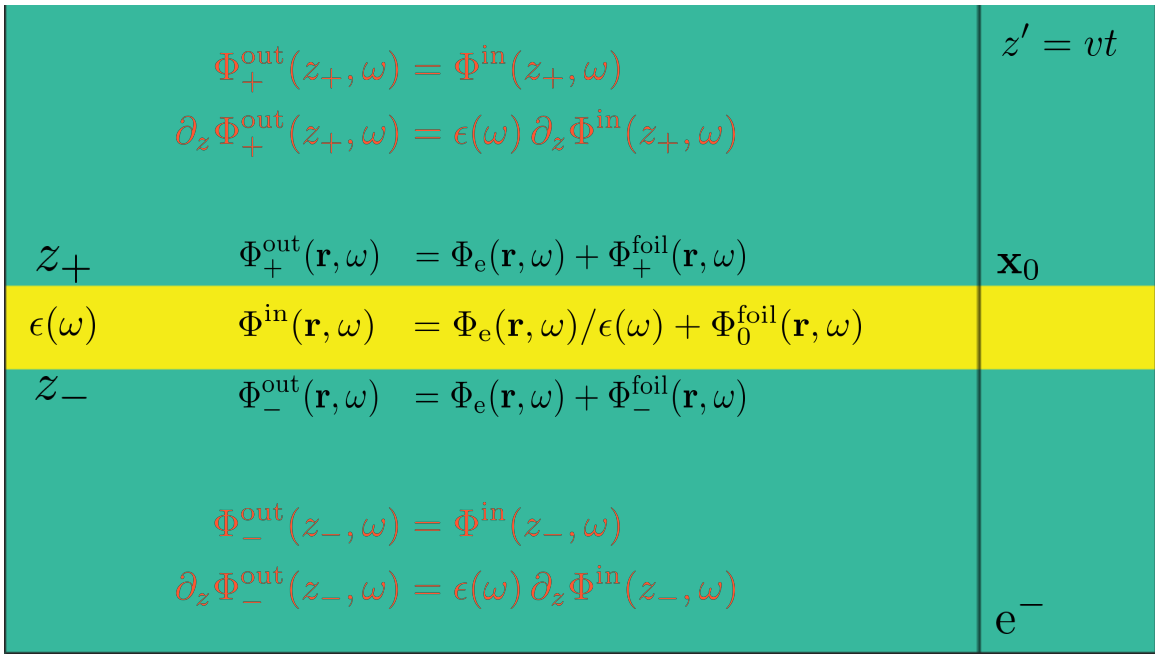

Figure 2.2: Slab geometry, with potential expansions labeled above and below the slab (green) and within the slab (yellow). Boundary conditions for the upper and lower surfaces are written above and below the slab, respectively.

2.3 Case Study: The Semi-Infinite Slab

2.3.1 "Classical" Approach

The electrostatic approximation is not a particularly good one for the semi-infinite slab (the foil's infinite transverse extent makes retardation effects important), but I have chosen this geometry for its pedagogical ease. In the electrostatic approximation, magnetic fields do not contribute, and we can describe all effects in terms of the electric potential. We impose periodic boundary conditions on the foil's induced potential in x and y and and assume it will approach zero as $z \to \pm\infty$. Since

this induced potential must follow Laplace's equation, we can expand it in harmonic functions above (Φ_+^{foil}), inside (Φ_0^{foil}), and below (Φ_-^{foil}) the foil:

$$\begin{aligned}
\Phi_+^{\text{foil}}(\mathbf{r},\omega) &= \sum_{\mathbf{k}_\perp} U_{\mathbf{k}}(\omega)\exp(-kz)\exp(i\mathbf{k}_\perp\cdot\mathbf{x}) \\
\Phi_0^{\text{foil}}(\mathbf{r},\omega) &= \sum_{\mathbf{k}_\perp} \left(A_{\mathbf{k}}(\omega)\exp(-kz) + B_{\mathbf{k}}(\omega)\exp(+kz)\right)\exp(i\mathbf{k}_\perp\cdot\mathbf{x}) \\
\Phi_-^{\text{foil}}(\mathbf{r},\omega) &= \sum_{\mathbf{k}_\perp} D_{\mathbf{k}}(\omega)\exp(+kz)\exp(i\mathbf{k}_\perp\cdot\mathbf{x}),
\end{aligned} \tag{2.12}$$

where now $\mathbf{x} = (x, y)$, $\mathbf{k}_\perp = (k_x, k_y)$, and $k = |\mathbf{k}_\perp|$. This particular expansion has no specific content, except to insist that the foil potential Φ^{foil} must obey Laplace's equation and must remain finite in each region. Fig. 2.2 illustrates this geometry alongside its associated expansions and boundary conditions.

For this expansion to help us to determine the unfixed coefficients, we need to expand our electron potential in terms of the surface functions—that is, in terms of $\exp(i\mathbf{k}_\perp\cdot\mathbf{x})$. For the beam electron traveling along $z = v_0 t$ and fixed at $\mathbf{x}_0 = (x_0, y_0)$, we can easily write down the unretarded potential as a function of time:

$$\Phi_{\text{e}}(\mathbf{r},t) = (-e)\int d^3\mathbf{r}'\frac{1}{|\mathbf{r}-\mathbf{r}'|}\delta^2(\mathbf{x}'-\mathbf{x}_0)\delta(z'-vt). \tag{2.13}$$

This peculiar form allows the spatial transform of the Coulomb part and the delta-function part of this expression to be taken separately, by the convolution theorem, such that

$$\Phi_{\text{e}}(\mathbf{k},t) = (-e)\frac{4\pi}{\mathbf{k}_\perp^2 + k_z^2}\exp(-i\mathbf{k}_\perp\cdot\mathbf{x}_0)\exp(-ik_z vt), \tag{2.14}$$

which in turn makes the time transform straightforward:

$$\Phi_{\text{e}}(\mathbf{k},\omega) = (-e)\frac{8\pi^2}{\mathbf{k}_\perp^2 + k_z^2}\exp(-i\mathbf{k}_\perp\cdot\mathbf{x}_0)\delta(\omega - k_z v) \tag{2.15}$$

If we transform back to position space, the unperformed integrations over $\mathbf{k}_\perp$ give an

expansion in the same terms as Eq. 2.12:

$$\begin{aligned}\Phi_{\mathrm{e}}(\mathbf{r},\omega) &= (-e)\int \frac{d^2\mathbf{k}_\perp}{(2\pi)^2}\frac{4\pi}{k^2+(\omega/v)^2}\exp(i\omega z/v)\exp(i\mathbf{k}_\perp\cdot(\mathbf{x}-\mathbf{x}_0)) \\ &= (-e)\sum_{\mathbf{k}_\perp}\frac{1}{L^2}\frac{4\pi}{k^2+(\omega/v)^2}\exp(i\omega z/v)\exp(i\mathbf{k}_\perp\cdot(\mathbf{x}-\mathbf{x}_0)).\end{aligned} \tag{2.16}$$

We suppose that this form will be adequate to describe the beam potential above and below the foil, but that it will need to be screened by $1/\epsilon(\omega)$ inside. Adding the response potential to the beam potential, we find forms for the *total* electric potential above (Φ_+^{out}), inside (Φ^{in}), and below (Φ_-^{out}) the foil:

$$\begin{aligned}\Phi_+^{\mathrm{out}}(\mathbf{r},\omega) &= \Phi_{\mathrm{e}}(\mathbf{r},\omega)+\Phi_+^{\mathrm{foil}}(\mathbf{r},\omega) \\ \Phi^{\mathrm{in}}(\mathbf{r},\omega) &= \Phi_{\mathrm{e}}(\mathbf{r},\omega)/\epsilon(\omega)+\Phi_0^{\mathrm{foil}}(\mathbf{r},\omega) \\ \Phi_-^{\mathrm{out}}(\mathbf{r},\omega) &= \Phi_{\mathrm{e}}(\mathbf{r},\omega)+\Phi_-^{\mathrm{foil}}(\mathbf{r},\omega).\end{aligned} \tag{2.17}$$

Applying the electric boundary conditions (Eqn. 2.6) to Φ^{in} and Φ^{out} leads us to four independent equations for each $\mathbf{k}_\perp$. We can solve these to give us the coefficients of our induced field. In principle, this gives us the electrical potential that acts back on the beam electron. If we want the electron energy-loss, we only need to find the integrated work performed on it by the induced potential, which we interpret as energy transferred to the dielectric, as in Eq. 1.4.

Notice that the induced potential excludes Φ_{e} outside the slab, but it includes $(\epsilon^{-1}-1)\Phi_{\mathrm{e}}$ inside the slab as the source of the bulk modes, which leads us to the same question that Fermi avoided by neglecting a small cylinder of radius b_0 surrounding the beam path—the question of how to avoid infinities. In the electrostatic context, we address this by imposing a cutoff of $K = 2\pi/b_0$ in Fourier-space. In general, we should expect such explicit cutoffs whenever the beam passes through the dielectric.

The discussion above reproduces the basic structure of many papers in dielectric EELS. In fact, just such an exposition is given by Lucas and Kartheuser (1970) for

the dielectric slab, where they find neat expressions for the bulk and the surface losses in the electrostatic limit. I will not reproduce their expressions for the surface losses, as they are equivalent to the expressions I will derive below using the "quantum" method. But finding the bulk losses using this method is very simple, compared to the quantum method, so I briefly derive it here.

If we find the induced bulk electric field associated with $(\epsilon^{-1} - 1)\Phi_{\mathrm{e}}$ using the integral expression in Eq. 2.16 and substitute it into the classical loss expression (Eq. 1.4), we quickly obtain, for a slab of thickness T, that

$$\frac{dP_B}{d\omega} = \int_0^K dk \frac{e^2 T}{\pi \hbar v^2} \frac{2k}{k^2 + (\omega/v)^2} \mathrm{Im}\left(-\frac{1}{\epsilon}\right). \tag{2.18}$$

Performing the integral and using the Born-Huang expression for the dielectric function, this leads, in the limit as the damping $\eta \to 0$, to the conclusion that

$$\frac{dP_B}{d\omega} = \frac{e^2 T}{2\pi \hbar v^2} \frac{\omega_\ell^2 - \omega_0^2}{\epsilon_\infty \omega_\ell} \log\left(K^2 \frac{v^2}{\omega_\ell^2} + 1\right) \delta(\omega - \omega_\ell), \tag{2.19}$$

i.e, that we obtain a spike in the spectrum at ω_ℓ, proportional to the sample thickness.

2.3.2 "Quantum" Approach

The quantum approach to such problems as the STEM-EELS scattering from a dielectric slab often requires more effort upfront, for greater dividends down the road. The primary roadblock for applications of this sort of model is the apparent need for an underlying physical model in the form of a Hamiltonian that keeps track of the various degrees of freedom that contribute to the dielectric response.

In unpublished work, Christian Dwyer has developed such a model for the Born-Huang model, starting with the ionic particle Hamiltonian

$$H = \int_V dV \left(\frac{1}{2}\dot{\mathbf{u}}^2 + \frac{1}{2}\omega_0^2 \mathbf{u}^2 + \frac{1}{2}(\rho - \rho_\infty)\Phi\right), \tag{2.20}$$

where ρ and ρ_∞ are the total and electronic polarization charge densities, respectively, and Φ is the electric field potential. Because surface charges may be present in the third term, the integral over the particle's volume must include the particle surface.

The Helmholtz-Hodge decomposition (Van Bladel (1958)) allows any $\mathbf{u}$ to be written as the sum of unique transverse, longitudinal, and harmonic components.

$$\begin{aligned} & & \nabla \cdot \mathbf{u}_t &= 0 & \nabla \times \mathbf{u}_t &\neq 0 \\ \mathbf{u} \propto \mathbf{u}_t + \mathbf{u}_\ell + \mathbf{u}_h & & \nabla \cdot \mathbf{u}_\ell &\neq 0 & \nabla \times \mathbf{u}_\ell &= 0 \\ & & \nabla \cdot \mathbf{u}_h &= 0 & \nabla \times \mathbf{u}_h &= 0 \end{aligned} \tag{2.21}$$

In the undamped, electrostatic Born-Huang model, it is straightforward to show that the transverse modes have frequency ω_0 but no associated electric fields. The longitudinal modes have frequency ω_ℓ, where $\omega_\ell^2 = \epsilon_0 \omega_0^2 / \epsilon_\infty$, and indeed have electric fields that contribute to scattering. However, in the following discussion, I will restrict the discussion to the harmonic modes, whose frequencies lie between ω_0 and ω_ℓ. These harmonic components only matter for finite particles, as their functional extrema occur at the boundary of the region being decomposed.

In particular, what form will said scalar potentials will take? Periodic boundary conditions will constrain the acceptable wavevectors in the xy-plane, where once again $\mathbf{x} = (x, y)$, $\mathbf{k}_\perp = (k_x, k_y)$, and $k = |\mathbf{k}_\perp|$. The form outside the slab will be restricted as we expect the potential to be finite as $z \to \pm\infty$. Above the slab, we expect

$$\phi_+ \propto \exp(i\mathbf{k}_\perp \cdot \mathbf{x}) \exp(-kz), \tag{2.22}$$

and below the slab, we expect

$$\phi_- \propto \exp(i\mathbf{k}_\perp \cdot \mathbf{x}) \exp(kz). \tag{2.23}$$

Within the slab, ϕ can take forms that are symmetric and antisymmetric in z:

$$\begin{aligned} \phi_c &\propto \exp(i\mathbf{k}_\perp \cdot \mathbf{x}) \cosh(kz), \\ \phi_s &\propto \exp(i\mathbf{k}_\perp \cdot \mathbf{x}) \sinh(kz). \end{aligned} \tag{2.24}$$

The (lower energy) cosh-type and (higher energy) sinh-type modes are sometimes referred to as the FK^- and FK^+ modes (for instance, by Lagos *et al.* (2018)), given that these modes were first identified by Fuchs and Kliewer (1965).

Merely insisting on this form is enough to solve for the harmonic frequencies, using the dielectric function Eq. 2.10 and the electrostatic boundary conditions. In the damped limit as $\eta \to 0$, we find that the cosh-type modes have lower frequencies

$$\omega_c = \omega_0 \left(\frac{\epsilon_0 + \coth(kT/2)}{\epsilon_\infty + \coth(kT/2)} \right)^{1/2}, \tag{2.25}$$

and the sinh-type modes have higher frequencies

$$\omega_s = \omega_0 \left(\frac{\epsilon_0 + \tanh(k_s T/2)}{\epsilon_\infty + \tanh(kT/2)} \right)^{1/2}. \tag{2.26}$$

In the limit of large k, each of these converges to the limiting frequency

$$\omega_{\mathrm{mid}} = \omega_0 \left(\frac{\epsilon_0 + 1}{\epsilon_\infty + 1} \right)^{1/2}, \tag{2.27}$$

such that $\omega_0 < \omega_c < \omega_{\mathrm{mid}}$, and $\omega_{\mathrm{mid}} < \omega_s < \omega_{\mathrm{mid}}$. A physical interpretation of ω_{mid} is that it occurs where $\epsilon(\omega_{\mathrm{mid}}) = -1$, where surface charge variations are quick enough to make the electric potential effectively local, uncoupling the top and bottom surfaces of the slab.

For modes written in terms of harmonic potentials ϕ_h, I adopt the notation that $\mathbf{u}_h = \nabla \phi_h$. This notation is somewhat confusing, since the displacements for the Born-Huang theory $\mathbf{u}$ have different units from this $\mathbf{u}_h$, which has units of inverse distance. Going forward, I will only use $\mathbf{u}_h$ as having units of inverse distance, and keep in mind that $\mathbf{u}_h$ is just a mathematical convenience.

To calculate our mode scattering, we need to normalize each mode such that it has energy $\hbar\omega_h$, where ω_h can stand for either ω_c or ω_s. For a dielectric particle whose total energy is given by Eq. 2.20, this can be achieved by using potentials that are

normalized such that

$$\int_V dV\,\bar{\mathbf{u}}_h \cdot \mathbf{u}_{h'} = \int_S d\mathbf{S} \cdot \bar{\phi}_h^{\text{in}} \mathbf{u}_{h'}^{\text{in}} = \delta_{h,h'}, \tag{2.28}$$

where the overbars denote complex conjugates. The scalar potentials are allowed to extend beyond the slab, with the overall form above and below following Eqs. 2.22 and 2.23, respectively, and the scale above and below the slab being set by continuity. The electric potential associated with this state can then be written both inside and outside the particle as

$$\Phi_h(\mathbf{r}) = \left(\frac{\hbar}{2\omega_h}\right)^{1/2} \frac{\omega_h^2 - \omega_0^2}{\omega_0} \phi_h(\mathbf{r}). \tag{2.29}$$

To make calculations easier, one can make the infinite slab periodic over length L, a length that is presumed "large," and which will not factor into the final expressions. This will allow us to deal with discrete k-states, with each discrete state labeled by $\mathbf{k}_\perp = (k_x, k_y)$ having a probability dependent on experimental parameters and $k = |\mathbf{k}_\perp|$. Once the potentials of Eq. 2.24 are normalized over one spatial period via Eq. 2.28 and scaled by continuity with the potenitals above and below the slab (Eqs. 2.22 and 2.23), the associated *electrostatic* potentials can be found from Eq. 2.29, and the transition probabilities can be calculated by Eq. 1.7, where the projection integral is split into parts for the ranges $(-\infty, -T/2)$, $(-T/2, T/2)$, and $(T/2, \infty)$, for the different forms for the potential below, within, and above the slab.

For a slab of thickness T, the transition probability for a single cosh-type k-state

$$P_c(k, L) = \frac{e^2}{v^2\hbar} \frac{8\pi k L^{-2}}{(k^2 + \omega_c^2/v^2)^2} \frac{(\omega_c^2 - \omega_0^2)^2}{\omega_c \omega_0^2 (\epsilon_0 - \epsilon_\infty)} \frac{\exp(kT)}{\sinh(kT)} \cos^2\left(\frac{\omega_c T}{2v}\right), \tag{2.30}$$

is found in terms of $\omega_c(k)$ as given by Eq. 2.25. To integrate over k-states, we use polar coordinates and multiply by

$$\text{“1”} = \frac{\mathrm{d}k_x \mathrm{d}k_y}{(2\pi/L)^2} = \frac{2\pi k \mathrm{d}k}{(2\pi/L)^2} = \frac{kL^2 \mathrm{d}k}{2\pi}, \tag{2.31}$$

which allows us to extract dP_c/dk as a quantity without any dependence on L:

$$\frac{dP_c}{dk} = \frac{e^2}{v^2\hbar}\frac{4k^2}{(k^2+\omega_c^2/v^2)^2}\frac{(\omega_c^2-\omega_0^2)^2}{\omega_c\omega_0^2(\epsilon_0-\epsilon_\infty)}\frac{\exp(kT)}{\sinh(kT)}\cos^2\left(\frac{\omega_c T}{2v}\right). \tag{2.32}$$

Inverting Eq. 2.25 to find the density of states, the EEL spectrum is derived as

$$\frac{dP_c}{d\omega} = \frac{dP_c}{dk}\left|\frac{dk}{d\omega_c}\right| \tag{2.33}$$

such that

$$\frac{dP_c}{d\omega} = \frac{e^2}{v^2\hbar}\frac{\exp(k_cT)}{\sinh(k_cT)}\frac{16T^{-1}k_c^2}{(k_c^2+\omega^2/v^2)^2}\frac{(\omega^2-\omega_0^2)^2}{(\omega^2\epsilon_\infty-\omega_0^2\epsilon_0)^2-(\omega^2-\omega_0^2)^2}\cos^2\left(\frac{\omega T}{2v}\right), \tag{2.34}$$

which can be plotted using $k_c(\omega)$ as solved by inverting Eq. 2.25

$$k_c(\omega) = \frac{1}{T}\log\left(\frac{\omega^2(\epsilon_\infty-1)-\omega_0^2(\epsilon_0-1)}{\omega^2(\epsilon_\infty+1)-\omega_0^2(\epsilon_0+1)}\right). \tag{2.35}$$

Of course, this expression only applies for the range of ω_c, from ω_0 to ω_{mid}.

The form for the transition probability of a single sinh k-state is very similar to the form for a single cosh k-state, using ω_s as defined by Eq. 2.26:

$$P_s(k,L) = \frac{e^2}{v^2\hbar}\frac{8\pi kL^{-2}}{(k^2+\omega_s^2/v^2)^2}\frac{(\omega_s^2-\omega_0^2)^2}{\omega_s\omega_0^2(\epsilon_0-\epsilon_\infty)}\frac{\exp(kT)}{\sinh(kT)}\sin^2\left(\frac{\omega_s T}{2v}\right). \tag{2.36}$$

Using the same steps as for the cosh-type states, we can find the spectrum contribution

$$\frac{dP_s}{d\omega} = \frac{e^2}{v^2\hbar}\frac{\exp(k_sT)}{\sinh(k_sT)}\frac{16T^{-1}k_s^2}{(k_s^2+\omega^2/v^2)^2}\frac{(\omega^2-\omega_0^2)^2}{(\omega^2-\omega_0^2)^2-(\omega^2\epsilon_\infty-\omega_0^2\epsilon_0)^2}\sin^2\left(\frac{\omega T}{2v}\right), \tag{2.37}$$

where, this time, the contributions are between ω_{mid} and the frequency of the longitudinal modes ω_ℓ, and can be plotted using $k_s(\omega)$ as found by inverting Eq. 2.26

$$k_s(\omega) = \frac{1}{T}\log\left(\frac{\omega^2(\epsilon_\infty - 1) - \omega_0^2(\epsilon_0 - 1)}{-\omega^2(\epsilon_\infty + 1) + \omega_0^2(\epsilon_0 + 1)}\right). \tag{2.38}$$

These contributions are plotted below in Fig. 2.4. This plot does not include the contribution of the longitudinal modes (Eq. 2.19), which would overwhelm the sinh-type mode contribution. The surface EEL spectrum spikes around ω_{mid} due to the large density of k-states surrounding that frequency, and would diverge were it not for the explicit cutoff K, which applies to both k_s and k_c, and exempts a small range of frequencies above and below ω_{mid} from contributing to the spectrum.

2.4 The "Classical" and "Quantum" Equivalence

For the slab, it is not transparently obvious that the "classical" and "quantum" methods of EELS calculation give the same results. In this section, I present a particular case for which the equivalence of the two methods can be seen explicitly.

For an electron beam at impact parameter b outside a dielectric sphere of radius a, the STEM-EEL spectrum has been calculated by by Ferrell *et al.* (1987) using the "classical" method. For a sphere surrounded by vacuum, their result becomes

$$\frac{dP}{d\omega} = \sum_{l=1}^{\infty}\sum_{m=-l}^{l}\frac{4e^2}{\pi\hbar}\frac{1}{(l+|m|)!(l-|m|)!}\frac{a^{2l+1}\omega^{2l}}{v^{2l+2}}K_{|m|}^2\Big(\frac{\omega b}{v}\Big)\text{Im}\left(\frac{l(\epsilon-1)}{\epsilon l + (l+1)}\right). \tag{2.39}$$

To see how this recovers the quantum result, we need to establish, for one thing, how this continuum expression recovers discrete states, and, for another, how the imaginary part of the dielectric leads to a real result. Here I will briefly derive the result using the "quantum" method, and proceed to show how the same result is the undamped limit of this "classical" expression.

Only harmonic states can be excited by an aloof beam, and such states are straightforward to write down using the spherical harmonics. If we require the same normal-

ization condition as in Eq. 2.28 and the same electrical potential scaling as in Eq. 2.29, we find the relevant transition potentials inside and outside the sphere as

$$\begin{aligned}\Phi^{\rm in}_{\ell m}(r,\theta,\phi) &= \left(\frac{2\pi\hbar}{\ell a^{2\ell+1}(\epsilon_0-\epsilon_\infty)\omega_{\ell m}}\right)^{1/2}\frac{\omega^2_{\ell m}-\omega_0^2}{\omega_0}r^\ell Y^m_\ell(\theta,\phi)\\ \Phi^{\rm out}_{\ell m}(r,\theta,\phi) &= \left(\frac{2\pi\hbar a^{2\ell+1}}{\ell(\epsilon_0-\epsilon_\infty)\omega_{\ell m}}\right)^{1/2}\frac{\omega^2_{\ell m}-\omega_0^2}{\omega_0}\frac{Y^m_\ell(\theta,\phi)}{r^{\ell+1}}\end{aligned} \tag{2.40}$$

with the associated harmonic frequencies

$$\omega^2_{\ell m} = \frac{\epsilon_0+(\ell+1)/\ell}{\epsilon_\infty+(\ell+1)/\ell}\omega_0^2. \tag{2.41}$$

Using the projection integral from Ferrell *et al.* (1987)

$$\int_{-\infty}^{\infty} r'^{-(l+1)}P_l^m(\cos(\theta'))e^{i\omega z'/v}dz' = \frac{2i^{l+m}K_m(\frac{b}{v}|\omega|)(\omega/|\omega|)^{l-m}|\omega|^l}{(l-m)!v^l}, \tag{2.42}$$

one can calculate the transition probabilities for $\Phi^{\rm out}_{\ell m}$ using Eq. 1.7 as

$$P_{\ell m} = \frac{e^2}{\hbar}\frac{2(2l+1)}{l(l+|m|)!(l-|m|)!}\frac{a^{2l+1}}{v^{2l+2}}\frac{(\omega^2_{\ell m}-\omega_0^2)^2\omega^{2l-1}_{\ell m}}{\omega_0^2(\epsilon_0-\epsilon_\infty)}K^2_{|m|}\Big(\frac{b}{v}\omega_{\ell m}\Big), \tag{2.43}$$

leading to the undamped spectrum in terms of Eqs. 2.41 and 2.43

$$\frac{dP}{d\omega} = \sum_{\ell=1}^{\infty}\sum_{m=-\ell}^{\ell}P_{\ell m}\delta(\omega-\omega_{\ell m}). \tag{2.44}$$

How might one obtain a result in this form from Eq. 2.39? The key is to begin by using the *damped* form of the dielectric function (Eq. 2.10), and to take the limit as $\eta\to 0$. Specifically, upon inserting Eq. 2.10 as the form for $\epsilon(\omega)$, one can, after a snarl of algebra, write the imaginary part of $\frac{\ell(\epsilon-1)}{\epsilon\ell+(\ell+1)}$ with a suggestive form:

$$\begin{aligned}\mathrm{Im}\Big(\frac{\ell(\epsilon-1)}{\epsilon\ell+(\ell+1)}\Big) &= \frac{\pi\ell(2\ell+1)(\epsilon_\infty\omega^2-\epsilon_0\omega_0^2)}{(l\epsilon_\infty+\ell+1)}\\ &\quad\times\lim_{\eta\to0}\frac{1}{\pi}\frac{\frac{\eta(\ell+1)\omega}{(\ell\epsilon_\infty+\ell+1)}}{\left(\omega^2-\frac{\epsilon_0+(\ell+1)/\ell}{\epsilon_\infty+(\ell+1)/\ell}\omega_0^2\right)^2+\left(\frac{\eta(\ell+1)\omega}{(\ell\epsilon_\infty+\ell+1)}\right)^2}\\ &= \frac{\pi\ell(2\ell+1)(\epsilon_\infty\omega^2-\epsilon_0\omega_0^2)}{(\ell\epsilon_\infty+\ell+1)}\delta(\omega^2-\omega^2_{\ell m})\\ &= \frac{\pi\ell(2\ell+1)(\epsilon_\infty\omega^2-\epsilon_0\omega_0^2)}{(\ell\epsilon_\infty+\ell+1)}\frac{\delta(\omega-\omega_{\ell m})}{|2\omega|}\end{aligned} \tag{2.45}$$

where the delta-function replacement has been made by identifying the limiting form of a Cauchy distribution, centered around $\omega = \omega_{\ell m}$

$$\delta(\omega - \omega_h) = \lim_{\Delta \to 0} \frac{1}{\pi} \frac{\Delta}{(\omega - \omega_h)^2 + \Delta^2}. \tag{2.46}$$

Putting the above form back into Eq. 2.39, one can recover the same form as the quantum version, showing that the "classical" and "quantum" versions are equivalent in the undamped limit. However, the classical version includes losses that are *not* included when when a complex dielectric function is merely approximated as real. This topic will be taken up again in Ch. 3, beyond the electrostatic approximation.

2.5 STEM-EELS of Canonical Dielectric Geometries

In this section, I discuss more generally how five canonical geometries—the five geometries corresponding to respective coordinate systems for which Laplace's equation can be solved most easily via separation of variables (Smythe (1950); Landau and Lifshitz (1960))—can be treated via the electrostatic approximation in a unified way. This development shows how both the "classical" and "quantum" versions of this problem are straightforward once the mathematical machinery is in place.

This machinery is summarized in Tables 2.1-4. In Table 2.1, I summarize the scale factors for each system as

$$h_i = \sqrt{\left(\frac{\partial x}{\partial \xi_i}\right)^2 + \left(\frac{\partial y}{\partial \xi_i}\right)^2 + \left(\frac{\partial z}{\partial \xi_i}\right)^2}, \tag{2.47}$$

which help us to find the gradient

$$\nabla = \sum_{j=1}^{3} \frac{\hat{\xi}_j}{h_j} \frac{\partial}{\partial \xi_j}, \tag{2.48}$$

differential volume element

$$dV = h_1 h_2 h_3 d\xi_1 d\xi_2 d\xi_3 \tag{2.49}$$

geometry	coordinates, ξ_i	coordinate ranges	scale factors, h_i
foil	$x = \xi_2$	$-\infty < \xi_1 < \infty$	$h_1 = 1$
	$y = \xi_3$	$-\infty < \xi_2 < \infty$	$h_2 = 1$
	$z = \xi_1$	$-\infty < \xi_3 < \infty$	$h_3 = 1$
circular cylinder	$x = \xi_3$	$0 \le \xi_1 < \infty$	$h_1 = 1$
	$y = \xi_1 \cos(\xi_2)$	$0 \le \xi_2 < 2\pi$	$h_2 = \xi_1$
	$z = \xi_1 \sin(\xi_2)$	$-\infty < \xi_3 < \infty$	$h_3 = 1$
sphere	$x = \xi_1 \sin(\xi_2)\cos(\xi_3)$	$0 \le \xi_1 < \infty$	$h_1 = 1$
	$y = \xi_1 \sin(\xi_2)\sin(\xi_3)$	$0 \le \xi_2 \le \pi$	$h_2 = \xi_1$
	$z = \xi_1 \cos(\xi_2)$	$0 \le \xi_3 < 2\pi$	$h_3 = \xi_1 \sin(\xi_2)$
oblate spheroid	$x = c((1+\xi_1^2)(1-\xi_2^2))^{1/2}\cos(\xi_3)$	$0 \le \xi_1 < \infty$	$h_1 = c((\xi_1^2+\xi_2^2)/(\xi_1^2+1))^{1/2}$
	$y = c((1+\xi_1^2)(1-\xi_2^2))^{1/2}\sin(\xi_3)$	$-1 \le \xi_2 \le 1$	$h_2 = c((\xi_1^2+\xi_2^2)/(1-\xi_2^2))^{1/2}$
	$z = c\,\xi_1\xi_2$	$0 \le \xi_3 < 2\pi$	$h_3 = c((\xi_1^2+1)(1-\xi_2^2))^{1/2}$
prolate spheroid	$x = c\,\xi_1\xi_2$	$1 \le \xi_1 < \infty$	$h_1 = c((\xi_1^2-\xi_2^2)/(\xi_1^2-1))^{1/2}$
	$y = c((\xi_1^2-1)(1-\xi_2^2))^{1/2}\cos(\xi_3)$	$-1 \le \xi_2 \le 1$	$h_2 = c((\xi_1^2-\xi_2^2)/(1-\xi_2^2))^{1/2}$
	$z = c((\xi_1^2-1)(1-\xi_2^2))^{1/2}\sin(\xi_3)$	$0 \le \xi_3 < 2\pi$	$h_3 = c((\xi_1^2-1)(1-\xi_2^2))^{1/2}$

Table 2.1: Convenient coordinates in which Laplace's Equation is separable. Contra tradition, here the cylinder and prolate spheroid lie along x rather than z, since our electron beam travels along z.

geometry	mode labels, h	inner potentials, $\phi_h^{\rm in}$	outer potentials, $\phi_h^{\rm out}$
foil	$k_2 = 2\pi m/L_2$	$\phi_c = \cosh\left(\lvert\mathbf{k}\rvert\xi_1\right)\exp\left(i\mathbf{k}\cdot\xi\right)$	$\phi_{+z} = \exp\left(-\lvert\mathbf{k}\rvert\xi_1\right)\exp\left(i\mathbf{k}\cdot\xi\right)$
	$k_3 = 2\pi n/L_3$	$\phi_s = \sinh\left(\lvert\mathbf{k}\rvert\xi_1\right)\exp\left(i\mathbf{k}\cdot\xi\right)$	$\phi_{-z} = \exp\left(+\lvert\mathbf{k}\rvert\xi_1\right)\exp\left(i\mathbf{k}\cdot\xi\right)$
	$m, n = \pm1, \pm2, \pm3\ldots$	$\mathbf{k} = (k_2, k_3), \quad \xi = (\xi_2, \xi_3)$	$\mathbf{k} = (k_2, k_3), \quad \xi = (\xi_2, \xi_3)$
circular cylinder	$m = 0, \pm1, \pm2\ldots$	$I_{\lvert m\rvert}\left(\lvert k\rvert\xi_1\right)\exp\left(i(m\xi_2 + k\xi_3)\right)$	$K_{\lvert m\rvert}\left(\lvert k\rvert\xi_1\right)\exp\left(i(m\xi_2 + k\xi_3)\right)$
	$k = 2\pi n/L$		
	$n = \pm1, \pm2, \pm3\ldots$		

Table 2.2: Modes and potentials for the semi-infinite nanoparticle geometries, with spatial periods in infinite directions fixed by lengths L. Potentials need to be scaled to ensure continuity at particle boundaries for each mode.

and differential surface element

$$d\mathbf{S} = \hat{\xi}_\mathbf{1} h_2 h_3 d\xi_2 d\xi_3 \tag{2.50}$$

for geometries defined by constants $\xi_1 = \xi_1^0$ (e.g., a sphere is defined by a constant radius in spherical coordinates).

In Tables 2.2 and 2.3, I give the harmonic functions for the five geometries. For the strictly finite geometries, I provide the Coulomb expansion coefficients C_h, as described below. For the spheroidal harmonics, Hobson (1931) is an authoritative

geometry	$R_h^{\rm in}(\xi_1)$	$R_h^{\rm out}(\xi_1)$	$S_h(\xi_2,\xi_3)$	C_h
sphere	$\left(\frac{\xi_1}{\xi_1^0}\right)^\ell$	$\left(\frac{\xi_1^0}{\xi_1}\right)^{\ell+1}$	$P_\ell^{\lvert m\rvert}(\cos(\xi_2))\exp(im\xi_3)$	$\frac{1}{\xi_1^0}\frac{(\ell-\lvert m\rvert)!}{(\ell+\lvert m\rvert)!}$
oblate spheroid	$P_\ell^{\lvert m\rvert}(i\xi_1)$	$\frac{Q_\ell^{\lvert m\rvert}(i\xi_1)}{P_\ell^{\lvert m\rvert}(i\xi_1^0)}$	$P_\ell^{\lvert m\rvert}(\xi_2)\exp(im\xi_3)$	$(2\ell+1)\frac{i(-1)^{\lvert m\rvert}}{c}\left(\frac{(\ell-\lvert m\rvert)!}{(\ell+\lvert m\rvert)!}\right)^2 P_\ell^{\lvert m\rvert}(i\xi_1^0)$
prolate spheroid	$P_\ell^{\lvert m\rvert}(\xi_1)$	$\frac{Q_\ell^{\lvert m\rvert}(\xi_1)}{P_\ell^{\lvert m\rvert}(\xi_1^0)}$	$P_\ell^{\lvert m\rvert}(\xi_2)\exp(im\xi_3)$	$(2\ell+1)\frac{(-1)^{\lvert m\rvert}}{c}\left(\frac{(\ell-\lvert m\rvert)!}{(\ell+\lvert m\rvert)!}\right)^2 P_\ell^{\lvert m\rvert}(\xi_1^0)$

Table 2.3: Potentials for finite nanoparticles with boundary ξ_1^0. For these potentials, $\ell = 0, 1, 2\ldots$ and $m = -\ell\ldots+\ell$.

geometry	**normalization,** N	$-\epsilon_h$
slab	$N = L^2 k\sinh(2ka)$ $k = (k_2^2+k_3^2)^{1/2}$	$-\epsilon_h^{cosh} = \coth(ka)$ $-\epsilon_h^{sinh} = \tanh(ka)$
circular cylinder	$2\pi L\Big(\lvert m\rvert I_{\lvert m\rvert}^2(\lvert k\rvert a) + \lvert k\rvert a I_{\lvert m\rvert}(\lvert k\rvert a)I_{\lvert m\rvert+1}(\lvert k\rvert a)\Big)$	$\frac{K_{\lvert m\rvert+1}(\lvert k\rvert a)/K_{\lvert m\rvert}(\lvert k\rvert a) - m/\lvert k\rvert a}{I_{\lvert m\rvert+1}(\lvert k\rvert a)/I_{\lvert m\rvert}(\lvert k\rvert a) + m/\lvert k\rvert a}$
sphere	$\frac{4\pi a\ell(\ell+\lvert m\rvert)!}{(2\ell+1)(\ell-\lvert m\rvert)!}$	$\frac{\ell+1}{\ell}$
oblate spheroid	$\frac{-4\pi c(\ell+\lvert m\rvert)!\bar{P}_\ell^{\lvert m\rvert}(ia/c)}{(2\ell+1)(\ell-\lvert m\rvert)!}\times$ $\Big(i(\ell-\lvert m\rvert+1)P_{l+1}^{\lvert m\rvert}(ia/c) + (\ell+1)aP_\ell^{\lvert m\rvert}(ia/c)/c\Big)$	$\frac{-Q_{\ell+1}^{\lvert m\rvert}(ia/c)/Q_\ell^{\lvert m\rvert}(ia/c) + ia(\ell+1)/c(\ell-\lvert m\rvert+1)}{P_{\ell+1}^{\lvert m\rvert}(ia/c)/P_\ell^{\lvert m\rvert}(ia/c) - ia(\ell+1)/c(\ell-\lvert m\rvert+1)}$
prolate spheroid	$\frac{4\pi c(\ell+\lvert m\rvert)!P_\ell^{\lvert m\rvert}(1/e)}{(2\ell+1)(\ell-\lvert m\rvert)!}\times$ $\Big((\ell-\lvert m\rvert+1)P_{l+1}^{\lvert m\rvert}(1/e) - (\ell+1)P_\ell^{\lvert m\rvert}(1/e)/e\Big)$	$\frac{-Q_{\ell+1}^{\lvert m\rvert}(1/e)/Q_\ell^{\lvert m\rvert}(1/e) + (\ell+1)/e(\ell-\lvert m\rvert+1)}{P_{\ell+1}^{\lvert m\rvert}(1/e)/P_\ell^{\lvert m\rvert}(1/e) - (\ell+1)/e(\ell-\lvert m\rvert+1)}$

Table 2.4: Factors that allow us to find the normalization and frequency for a given harmonic state. N factors have units of length, and ϵ_h, like the dielectric values in Gaussian units, are dimensionless. Although these factors involve spatial derivatives, I have put them in terms of the special functions themselves for ease of calculation.

reference, and associated Legendre functions of the second kind $Q_\ell^m(x)$ for $x > 1$ can be calculated using the method of Gil and Segura (1998).

The normalization factors N given in Table 2.4 are those needed to achieve Eq. 2.28. That is, to be used as the classical potentials for $\mathbf{u}_h$ and Φ_h, the potential ϕ_h must be normalized and continuous across the surface of the particle. If we calculate a normalization factor N for each of the unnormalized potentials

$$N = \int_V d^3x\,(\nabla\bar{\phi}_{\rm in})\cdot(\nabla\phi_{\rm in}) = \int_S d\mathbf{S}\cdot\bar{\phi}_{\rm in}\nabla\phi_{\rm in} \tag{2.51}$$

we can write normalized potentials in terms of it

$$\begin{aligned} \phi_h^{\text{in}} &= \frac{\phi_{\text{in}}}{\sqrt{N}} \\ \phi_h^{\text{out}} &= \frac{\phi_{\text{in}}\left(\xi_1^0\right)}{\phi_{\text{out}}\left(\xi_1^0\right)} \frac{\phi_{\text{out}}}{\sqrt{N}} \end{aligned} \tag{2.52}$$

where the expression for ϕ_h^{out} is scaled to ensure continuity of the potential, and, as defined above, $\mathbf{u}_h^{\text{in/out}} = \nabla \phi_h^{\text{in/out}}$, and overbars connote complex conjugates.

The harmonic frequencies can be understood easily in terms of these potentials. From the electrostatic boundary conditions we can obtain an expression for ϵ_h in terms of derivatives of the potentials at the surface

$$\epsilon_h = \left. \frac{\mathbf{u}_h^{\text{out}} \cdot \hat{\mathbf{n}}}{\mathbf{u}_h^{\text{in}} \cdot \hat{\mathbf{n}}} \right|_{\text{surface}} \tag{2.53}$$

where $\hat{\mathbf{n}}$ is the unit normal vector at the surface. This expression clarifies the geometric role of ϵ_h in the theory. For simple geometries, the coordinate system defines appropriate particle boundaries and separable harmonic functions of index h. The coordinate boundary establishes the particle shape and determines how ϕ_h^{in} and ϕ_h^{out} must be scaled to achieve continuity, and the function index h determines the potential derivative discontinuity at the particle surface, and the particular value it takes on sets the harmonic frequency via the dielectric function by $\epsilon(\omega_h) = \epsilon_h$.

Note that since $\epsilon(\omega)$ is negative for harmonic frequencies, the values for $-\epsilon_h$ are positive. In the undamped Born-Huang theory, the frequencies of the harmonic modes can be written as

$$\omega_h^2 = \frac{\epsilon_0 - \epsilon_h}{\epsilon_\infty - \epsilon_h} \omega_0^2. \tag{2.54}$$

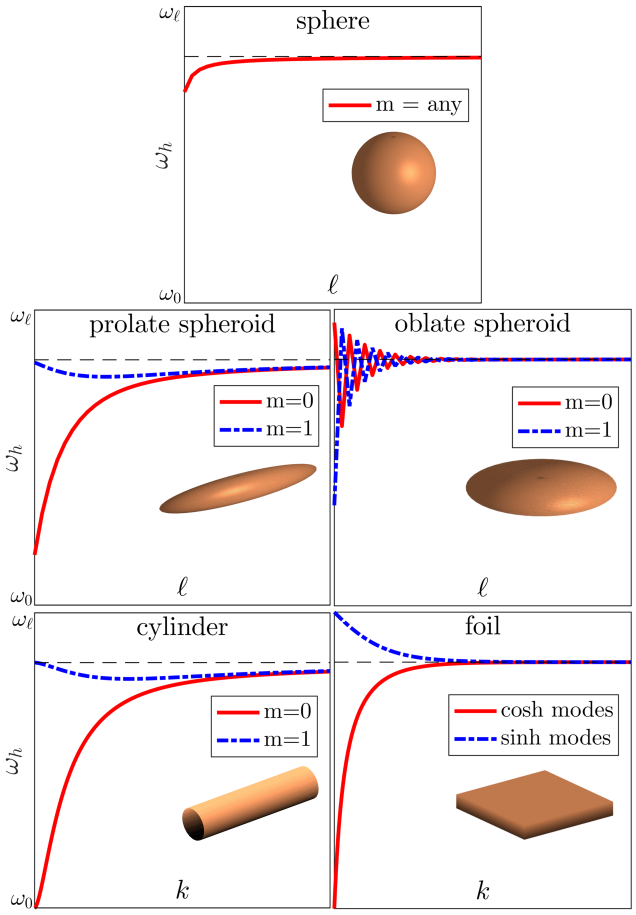

Figure 2.3: Comparison of harmonic frequencies across five canonical geometries. The dashed horizontal line is at the value of ω_h for which $\epsilon_h = -1$. Spheroid plots are for nanoparticles with cross-sectional eccentricites of 0.99.

This allows us to identify the transverse and longitudinal limits:

$$\begin{aligned} -\epsilon_h &\to \infty, \quad & \omega_h &\to \omega_0; \\ -\epsilon_h &\to 0, \quad & \omega_h &\to \omega_\ell. \end{aligned} \tag{2.55}$$

These frequencies are plotted for each of the geometries under consideration in Fig. 2.3, which shows how the frequencies spread out as the cross-sectional aspect ratio of the particle increases. But the way these frequencies manifest in spectra is quite distinct in the "classical" and "quantum" methods, as I will now review.

In the "classical" approach, the key realization is that the Coulomb potential of the beam potential can be expanded in terms of the harmonic potentials. (The

arguments given below are similar to those of Smythe (1950).) Suppose we have two points $\mathbf{r}_{\text{in}}$ and $\mathbf{r}_{\text{out}}$ whose radial ξ_1 components differ in that the former is smaller than the latter—i.e., $(\xi_1)_{\text{in}} < (\xi_1)_{\text{out}}$. We would like to write

$$\begin{aligned}\frac{-e}{|\mathbf{r}_{\text{out}} - \mathbf{r}_{\text{in}}|} &= \sum_h -eC_h \bar{\phi}_h^{\text{in}}(\mathbf{r}_{\text{in}})\phi_h^{\text{out}}(\mathbf{r}_{\text{out}}), \\ &= \sum_h -eC_h \phi_h^{\text{in}}(\mathbf{r}_{\text{in}})\bar{\phi}_h^{\text{out}}(\mathbf{r}_{\text{out}}).\end{aligned} \tag{2.56}$$

where the C_h coefficients are constants, the Coulomb coefficients given in Table 2.4, and the overbar notation denotes, as before, the complex conjugate.

The values C_h are easiest to find if we distinguish between the "radial" parts of the solutions to Laplace's equation $R_h^{\text{in/out}}(\xi_1)$, which are different inside and outside the particle, and the "surface" parts of the solutions $S_h(\xi_2, \xi_3)$, which are the same. We can write these functions symmetrically, where ξ_1^0 marks the particle boundary:

$$\begin{aligned}\phi_h^{\text{in}} &= R_h^{\text{in}}(\xi_1)R_h^{\text{out}}(\xi_1^0)S_h(\xi_2, \xi_3), \\ \phi_h^{\text{out}} &= R_h^{\text{in}}(\xi_1^0)R_h^{\text{out}}(\xi_1)S_h(\xi_2, \xi_3).\end{aligned} \tag{2.57}$$

The "radial" parts $R_h^{\text{in}}(\xi_1)$ and $R_h^{\text{out}}(\xi_1)$ are taken to be real-valued, while the oscillatory "surface" parts $S_h(\xi_2, \xi_3)$ are allowed to be complex. To expand a point charge in these functions, we first write it as a surface charge on the surface $\xi_1 = \xi_1^0$:

$$\sigma(\xi_2, \xi_3) = -e\frac{\delta(\xi_2 - \xi_2^0)}{h_2}\frac{\delta(\xi_3 - \xi_3^0)}{h_3}\bigg|_{\xi_1 = \xi_1^0}. \tag{2.58}$$

This allows us to find the potential due to the charge in terms of Gauss's Law, which allows the coefficients C_h to be calculated as

$$C_h = 4\pi\left(W(R_h^{\text{out}}, R_h^{\text{in}})\int_S \mathrm{d}S\,\frac{\bar{S}_h S_h}{h_1}\right)^{-1}, \tag{2.59}$$

where W is the Wronskian, the integral is over the nanoparticle's surface ξ_1^0, and h_1 is the scale factor of the radial coordinate.

Given this expansion, solving for the spectrum of an aloof beam becomes quite straightforward. The potential of the beam electron takes on the form

$$\Phi_{\mathrm{e}}(\mathbf{r},t) = \sum_h -eC_h\phi_h^{\mathrm{in}}(\mathbf{r})\bar{\phi}_h^{\mathrm{out}}\left(\mathbf{r}_{\mathrm{e}}(t)\right), \tag{2.60}$$

where $\mathbf{r}_{\mathrm{e}}(t)$ marks the straight-line path of the electron at speed v with along the path $z = vt$ with transverse position $\mathbf{x}_0$ outside the particle in the sample plane. If we transform the electron potential as a function of frequency, it becomes

$$\Phi_{\mathrm{e}}(\mathbf{r},\omega) = \sum_h -eC_hI_h(\mathbf{x}_0,\omega)\phi_h^{\mathrm{in}}(\mathbf{r}), \tag{2.61}$$

where $I_h(\mathbf{x},\omega)$ is the projection integral

$$I_h(\mathbf{x},\omega) = \int_{-\infty}^{\infty}\frac{dz}{v}\,\bar{\phi}_h(\mathbf{x},z)\exp(i\omega z/v). \tag{2.62}$$

Next we apply the electrostatic boundary conditions. The first requires the potential to be continuous across the particle surface, and the second requires $\mathbf{D}(\mathbf{r},\omega)\cdot\hat{\mathbf{n}} = \epsilon(\omega)\mathbf{E}(\mathbf{r},\omega)\cdot\hat{\mathbf{n}}$ to be continuous across the particle surface. These conditions give us the Fourier coefficients for the dielectric particle, and we can now find the spectrum *a la* Eq. 1.4, leading to a transparent expression

$$\frac{\mathrm{d}P}{\mathrm{d}\omega} = \frac{e^2}{\pi\hbar}\sum_h C_h\left|I_h(\mathbf{x}_0,\omega)\right|^2\mathrm{Im}\left(\alpha_h(\omega)\right) \tag{2.63}$$

where, following Ferrell *et al.* (1987), we introduce the "nanoparticle polarizability"

$$\alpha_h(\omega) = \frac{\epsilon(\omega)-1}{\epsilon(\omega)-\epsilon_h}, \tag{2.64}$$

in which ϵ_h is the same as it was in Eqn. 2.53. Notice that the harmonic frequencies appear in this expression via ϵ_h, and the height of each peak in the spectrum depends on the beam position via the projection integral $I_h(\mathbf{x}_0,\omega)$.

In the "quantum" version of this problem, we are interested in the transition $|0\rangle \rightarrow |1_h\rangle$, in which one quantum of mode h is created as a result of the electrostatic interaction with a beam electron in the STEM, the interaction energy has the form

$$-e\left(\frac{\hbar}{2\omega_h}\right)^{\frac{1}{2}}\frac{4\pi Z_0^*}{\epsilon_\infty - \epsilon_h}\bar{\phi}_h(\mathbf{r}). \tag{2.65}$$

In the projection approximation for inelastic scattering, the interaction energy is multiplied by the phase factor $\exp(i\omega_h z/v)$, integrated along the entire optic axis z, and multiplied by the interaction constant $1/\hbar v$. The squared modulus of the resulting expression then gives the scattering probability $P_h(\mathbf{x})$ for exciting the hth mode when the electron beam is positioned at the transverse coordinate $\mathbf{x}_0$:

$$P_h(\mathbf{x}_0,\omega_h) = \frac{e^2}{2\hbar\omega_h}\left(\frac{4\pi Z_0^*}{\epsilon_\infty - \epsilon_h}\right)^2 |I_h(\mathbf{x}_0,\omega)|^2 \tag{2.66}$$

where $I_h(\mathbf{x},\omega)$ is the projection integral of Eq. 2.62.

In the undamped limit, the spectrum for the harmonic states in this approach would appear as sum of delta functions, each with weight P_h and centered on ω_h. The simplest way to extend our formalism in the case of a nonzero damping parameter η is to replace ω_h in the projection integral by the damped value $\omega_h' = (\omega_h^2 - \eta^2)^{1/2}$. A spectrum is then built up by summing Lorentzians for each transition:

$$\frac{dP}{d\omega} = \sum_h \frac{\eta}{2\pi}\frac{P_h(\mathbf{x}_0,\omega_h)}{(\omega-\omega_h')^2 + (\eta/2)^2}. \tag{2.67}$$

This is the approach I have used to make plots in Section 2.6. If we restrict our attention to regions where the dielectric damping is negligible, this simply *reproduces* the results of the classical treatment, albeit with a different order of operations.

For constructing these plots, the projection integrals (Eqn. 2.62) present the greatest difficulty. Analytic solutions for these integrals are known for the sphere in Ferrell

et al. (1987) and for the spheroids in Illman *et al.* (1988) when the projection is along the nanoparticle axis of rotational symmetry, but for prolate spheroids oriented in the transverse plane, a different approach was needed. We found that the external prolate spheroidal harmonics can be expressed via a convolution with the spherical solid harmonics, and that by rotating these spherical solid harmonics (Gumerov and Duraiswami (2015)), one can reconstruct the projected prolate spheroidal potentials.

Although I have not treated them here, the classical and quantum treatments of bulk modes also compliment each other. For instance, many sources report on "Begrenzung" (boundary) effects for bulk modes in finite geometries. Roughly, this effect consists of the observation that if an electron travels through a finite dielectric, the energy loss probability at ω_ℓ is reduced relative to what one would expect for the same distance traveled in an infinite medium. In classical treatments, this manifests as a negative contribution of the harmonic modes to the loss spectrum at ω_ℓ. In quantum treatments (Lucas *et al.* (1970); Lucas and Šunjić (1972); Schmeits (1981)), this effect arises naturally as the orthogonality of modes shifts the density of states from bulk to surface modes.

2.6 Numerical Results for Canonical Geometries

Finally, I review here the EELS spectra obtained from bare nanoparticles with five canonical geometries (semi-infinite slab, semi-infinite circular cylinder, sphere, oblate spheroid, and prolate spheroid). This is useful for gaining an appreciation of how the frequencies plotted in Fig. 2.3 appear in the EEL spectra.

The nanoparticle spectra in this section were computed by applying Eqns. 2.66 and 2.67. We use experimentally measured parameters for cubic boron nitride (Eremets *et al.* (1995)), an isotropic material, but with an unrealistically small damping parameter ($\eta = 0.0005\omega_0$) to isolate the modes. For all plots going forward, we assume

a 60 keV electron beam and a 1 mrad collection semi-angle. Longitudinal mode contributions have not been included here, but for beam positions that pierce the nanoparticles, the longitudinal mode contribution would contribute a peak at ω_ℓ with an intensity proportional to the particle thickness.

Although the scattering probabilities here are small, and the predicted range of frequencies seems so narrow as to defy positive detection, the modes considered here also apply to the surface plasmon modes for metals, except with stretched frequency limits as given by Eq. 2.11. In the metallic context, the spectra arising from prolate spheroids, and the changes in these spectra as a function of cross-sectional eccentricity, will be especially relevant to the enhancement results presented in Ch. 4.

2.6.1 Foil, Cylinder, Sphere

Fig. 2.4 shows spectra for the three "limiting" geometries, namely, the slab, the cylinder, and the sphere (see Fig. 2.3). In each plot, three frequencies are marked: ω_0 as the lower bound for the harmonic frequencies, ω_ℓ as the upper bound, and a frequency between the two that picks out the intermediate value associated with $\epsilon_h = -1$ (which is the frequency associated with modes of high spatial frequency as $h \to \infty$; see Fig. 2.3).

As noted above, for the slab (Lucas and Kartheuser (1970)) the internal harmonic potentials take on symmetric and antisymmetric forms

$$\phi^{\text{in}}_{\mathbf{k}_\perp} \propto \exp(i\mathbf{k}_\perp \cdot \mathbf{x}) \begin{cases} \cosh(|\mathbf{k}_\perp|z) \\ \sinh(|\mathbf{k}_\perp|z) \end{cases} \tag{2.68}$$

where $\mathbf{k}_\perp = (k_x, k_y)$ are continuous quantum numbers; above and below the slab these potentials die off exponentially. As a general rule the cosh modes scatter strongly, whereas the sinh modes do not. This behavior is easily understood from the projection integral along the beam direction in Eqn 2.62: the antisymmetry of the sinh potentials

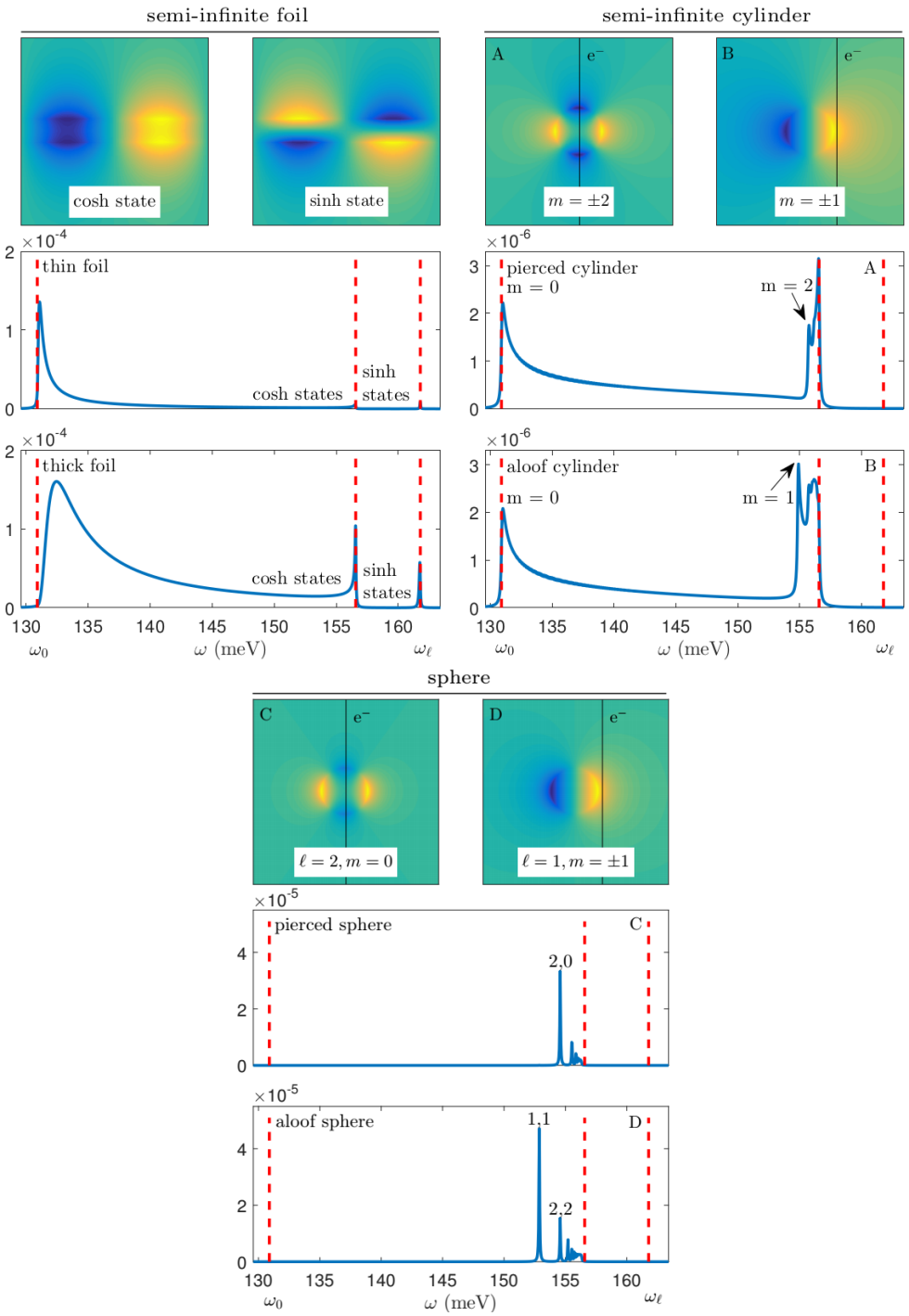

Figure 2.4: Surface spectra for the three "limiting" geometries. *Top Left*: thin (5 nm) and thick (50 nm) semi-infinite foil. *Top Right*: semi-infinite cylinder of diameter $2a = 50$ nm with (A) piercing beam at $b = 0$, and (B) aloof beam at $b = 1.01a$. *Bottom*: sphere of diameter $2a = 50$ nm with (C) piercing beam at $b = 0$, and (D) aloof beam at $b = 1.01a$. Spectra are plotted in units of meV^{-1}. Representative mode potentials are shown for each, with relevant electron beam trajectories indicated.

implies that their projection almost vanishes for slab thicknesses $\ll v/\omega$, which is the case here. For the thicker slab in Fig. 2.4, the peak around $\omega(\epsilon_h = -1)$ is due to a large density of states, although the states are individually very weakly scattering.

For the semi-infinite cylinder with its axis perpendicular to the electron beam (Zabala *et al.* (2001)), modes above $\omega(\epsilon_h = -1)$ are not present. The harmonic potentials inside and outside the cylinder

$$\begin{aligned} \phi_{km}^{\text{in}} &\propto \exp(ikx + im\varphi) I_{|m|}(|k|\rho) \\ \phi_{km}^{\text{out}} &\propto \exp(ikx + im\varphi) K_{|m|}(|k|\rho) \end{aligned} \tag{2.69}$$

are labeled by a discrete m and a continuous k, where I_m and K_m are modified Bessel functions of the first and second kind. The cylinder spectra depend on beam position. While the broad band extending down to ω_0 representing excitations of the $m = 0$ states can be seen in spectra both when the beam pierces the center of the cylinder and when the beam is aloof, the odd m bands are missing from the piercing beam spectrum, by symmetry.

For the sphere of radius a (Ferrell *et al.* (1987)), the potentials

$$\begin{aligned} \phi_{\ell m}^{\text{in}} &\propto (r/a)^{\ell} P_{\ell}^{m}(\cos\vartheta) \exp(im\varphi) \\ \phi_{\ell m}^{\text{out}} &\propto (a/r)^{\ell+1} P_{\ell}^{m}(\cos\vartheta) \exp(im\varphi) \end{aligned} \tag{2.70}$$

follow the familiar spherical harmonics. For the sphere, as for the cylinder, a beam through the center of the particle has difficulty exciting modes with an odd spatial parity, but as shown in Fig. 2.3 the range of harmonic frequencies available to the sphere is so reduced (and the degeneracy of states is so great) that the influence of beam position is less apparent for the sphere than for other geometries.

2.6.2 Oblate and Prolate Spheroids

To investigate beam-position effects in a slab-like geometry, we can use the oblate spheroid (Illman *et al.* (1988)). The oblate spheroid is flat and wide, an ellipse rotated

about its minor axis. In oblate spheroidal coordinates

$$\begin{aligned} x &= c\cosh(\varrho)\cos(\vartheta)\cos(\varphi) && 0 \leq \varrho < \infty \\ y &= c\cosh(\varrho)\cos(\vartheta)\sin(\varphi) && -\pi/2 \leq \vartheta \leq \pi/2 \\ z &= c\sinh(\varrho)\sin(\vartheta) && 0 \leq \varphi \leq 2\pi \end{aligned} \tag{2.71}$$

a surface of constant ϱ describes an oblate spheroid. Appropriate choices of c and ϱ describe a variety of particle surfaces—thick or thin, flat or spherical.

The oblate spheroidal harmonics

$$\begin{aligned} \phi_{\ell m}^{\text{in}} &\propto P_\ell^m(i\sinh\varrho)P_\ell^m(\sin\vartheta)\exp(im\varphi) \\ \phi_{\ell m}^{\text{out}} &\propto Q_\ell^m(i\sinh\varrho)P_\ell^m(\sin\vartheta)\exp(im\varphi) \end{aligned} \tag{2.72}$$

are labeled by two discrete quantum numbers, ℓ and m, and are written in terms of associated Legendre polynomials of the first and second kind. At large ϱ or with small c, these solutions approach the spherical harmonics. In Fig. 2.5, the maximal spheroid thickness remains constant at $2a = 50$ nm, but the eccentricity e changes, where $e \to 0$ for a sphere and $e \to 1$ for a slab. We can put the spheroid semi-major axis R, the projected disc radius, in terms of e and a:

$$R^2 = a^2/(1-e^2). \tag{2.73}$$

The symmetries of the oblate spheroid are such that only b, the beam's transverse distance from the origin, is relevant to spectral differences. Comparing the spectra for $b = 0$ and $b = 1.01R$ for in Fig. 2.5, we find that the aloof beam excites a wider range of energies than the piercing beam, and that there are roughly twice as many peaks in the aloof beam spectra as the piercing beam spectra. Specifically, the states with $\ell = 2, 4, 6, ...$ and $m = 0$ are most excited by the piercing beam, and the states with $\ell = 1, 2, 3, ...$ and $m = \pm\ell$ are most excited by the aloof beam.

Beam-position effects are even more pronounced with the prolate spheroid, a geometry that transitions smoothly between the sphere ($e \to 0$) and the cylinder

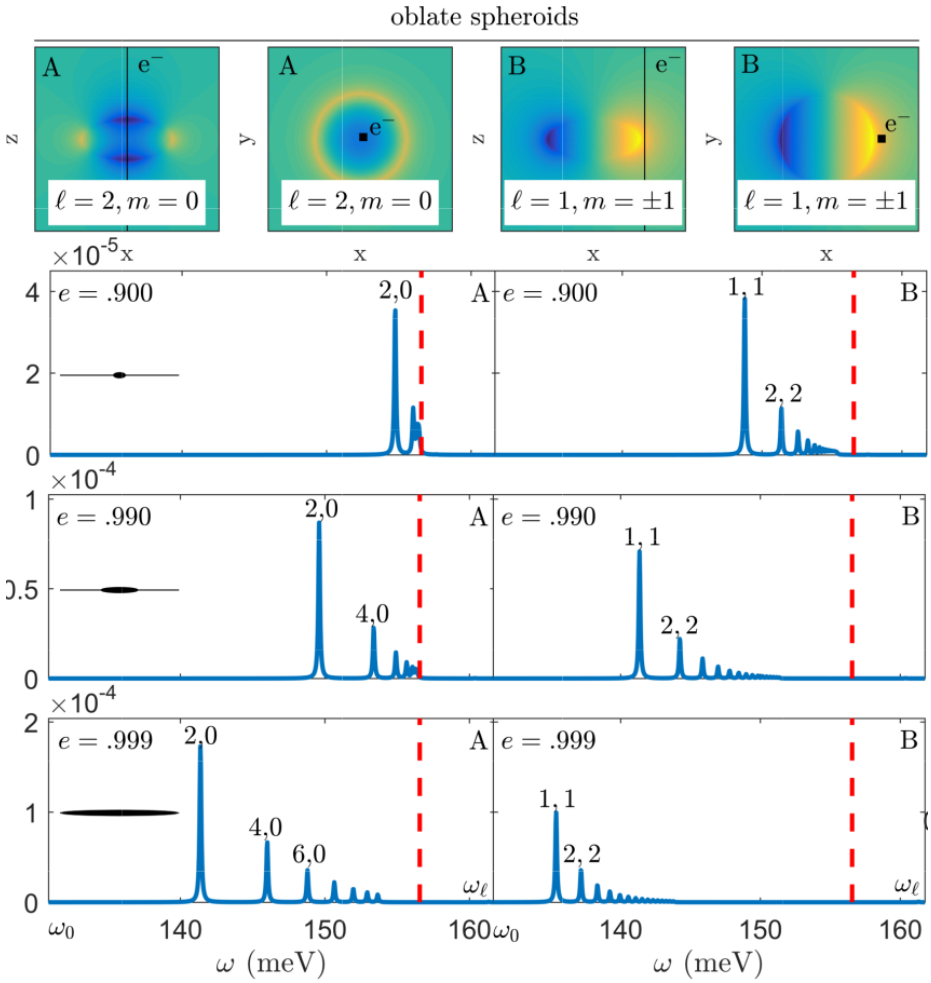

Figure 2.5: Oblate spheroids of maximal thickness $2a$ =50 nm over a wide range of eccentricity. Beam positions are (A) through the center $b = 0$, or (B) just past the edge R of the nanoparticle at $b = 1.01R$. Representative harmonic potentials in plan view and cross-section view are shown atop, with the electron beam trajectory indicated.

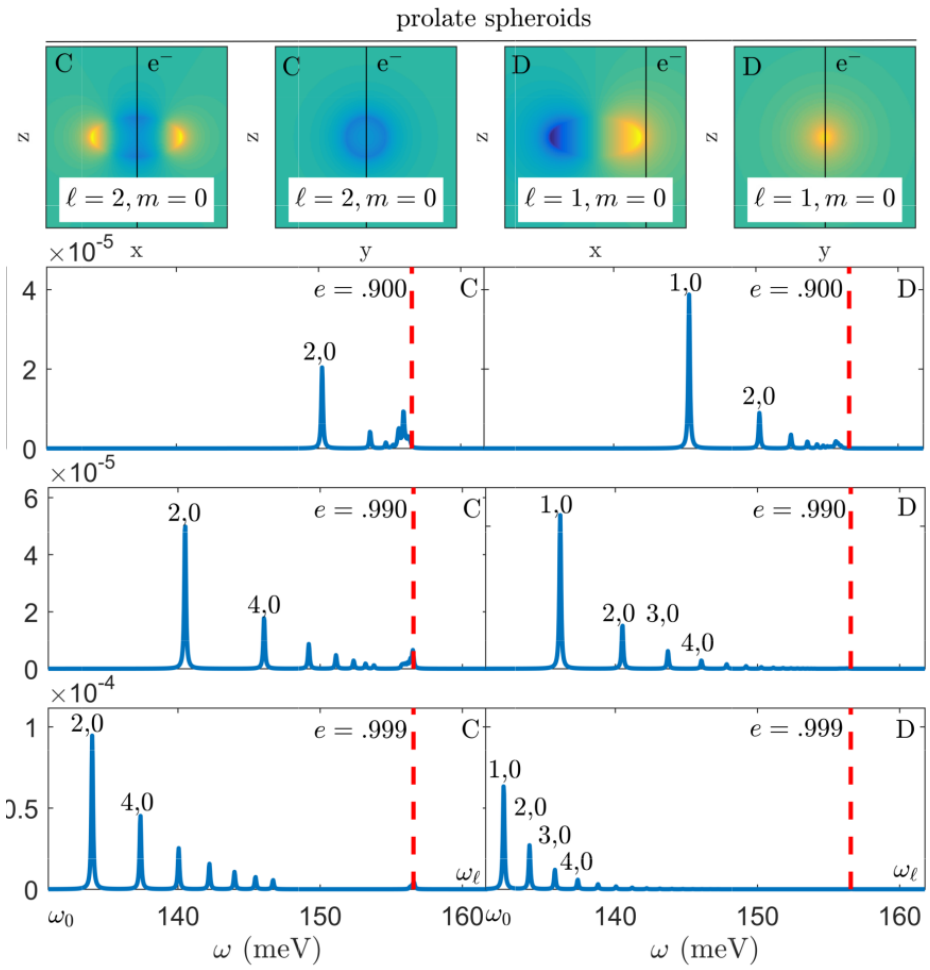

Figure 2.6: Prolate spheroids of maximal thickness $2a$ =50 nm over a wide range of eccentricity. For prolate spheroids of length $2R$, beam positions are either (C) piercing their centers, both in length and width, or (D) aloof just past the lengthwise tip, at $(x_0, y_0) = (1.01R, 0)$. Spectra are plotted in units of meV^{-1}. As in Fig. 2.5, representative harmonic potentials and electron beam trajectories are indicated.

$(e \to 1)$. As with oblate spheroidal coordinates, in prolate spheroidal coordinates

$$\begin{aligned} x &= c\cosh(\varrho)\cos(\vartheta) & 0 \leq &\varrho < \infty \\ y &= c\sinh(\varrho)\sin(\vartheta)\cos(\varphi) & 0 \leq &\vartheta \leq \pi \\ z &= c\sinh(\varrho)\sin(\vartheta)\sin(\varphi) & 0 \leq &\varphi \leq 2\pi \end{aligned} \tag{2.74}$$

a surface of constant ϱ defines a spheroid. As with the cylinder, we have shifted the long axis from z to x to make the long axis normal to the beam path.

Like the oblate spheroidal harmonics, prolate spheroidal harmonics

$$\begin{aligned}\phi^{\text{in}}_{\ell m} &\propto P^m_\ell(\cosh\varrho)P^m_\ell(\cos\vartheta)\exp(im\varphi)\\ \phi^{\text{out}}_{\ell m} &\propto Q^m_\ell(\cosh\varrho)P^m_\ell(\cos\vartheta)\exp(im\varphi)\end{aligned} \tag{2.75}$$

are labeled by ℓ and m, are written via associated Legendre polynomials, and reduce to the spherical harmonics. Again, we consider particles of maximal of maximal thickness $2a = 50$ nm. For this case, R in Eqn. 2.73 gives half the length of the (long) major axis.

As with the oblate case, for Fig. 2.6 we position one beam at the center of the spheroid, and one just outside the particle, past the tip at $(x, y) = (1.01R, 0)$. Again, we see that the aloof beam excites roughly twice as many frequencies as the piercing beam, but for the prolate spheroid we observe the *same* excited frequencies for our two chosen beam positions, only with half of the peaks missing from the spectra of the central piercing beam. Specifically, while the aloof beam strongly excites the states with $\ell = 1, 2, 3, ...$ and $m = 0$, the piercing beam only excites the states with $\ell = 2, 4, 6, ...$ and $m = 0$. This aligns to the general trend that those states whose projected charges have odd spatial parity are suppressed in the spectra of a centrally located beam.

2.7 Summary

In this chapter, I reviewed the theory of low-loss STEM-EELS in the electrostatic dielectric approximation (i.e., where $c \to \infty$). I presented the theory of STEM-EELS for the dielectric slab in both the "classical" approach, in which the beam potential and its consequences are foregrounded, and in the "quantum" approach, for which the normal modes of the dielectric itself are foregrounded. In the case of an undamped dielectric, the two approaches are equivalent, as I showed explicitly for an aloof beam passing a dielectric sphere. This equivalence of the "classical"

and "quantum" approaches, while expected, enables an expansion of interpretation, linking EELS with nanophotonics, plasmonics, quantum wells in semiconductors, etc.

A unified approach for canonical geometries was presented and applied for the surface states of the "limiting geometries" of the semi-infinite cylinder, the semi-infinite circular, the sphere, and of the "transitional geometries" of the oblate and prolate spheroids. This development leads to further applications of plasmonics in Ch. 4. In the next chapter, I will discuss how the theory of low-loss STEM-EELS changes when we press beyond the electrostatic approximation.

Chapter 3

STEM-EELS OF ELECTRODYNAMIC MODES

3.1 Introduction

In the last chapter, I showed how the STEM-EELS of several canonical dielectric geometries can be dealt with in the electrostatic limit of the Born-Huang theory. In this chapter, I incorporate effects that are not included in the simplified model. One complication is that the dielectric response of real materials may not be captured by the simple Born-Huang model, as I will discuss in the next section. But the more obvious complications arise from the breakdown of the electrostatic approach.

The importance of relativistic corrections to the core-loss theory have been demonstrated, for instance, by Schattschneider *et al.* (2005) and Dwyer and Barnard (2006); further discussions have been given by Sorini (2008) in his doctoral thesis. These results tend to introduce significant quantitative corrections to effects that could be found less precisely in the electrostaic approach. On the other hand, in the low-loss theory, Cherenkov-type effects can also change the qualitative nature of the outcomes of calculations, whether one adopts a "classical" or "quantum" style of calculation.

For the "classical" style of calculation, the reason can be seen in the electric field of the beam electron. For a classical electron traveling along $z = vt$ and centered at the origin of $\mathbf{R} = (x, y)$ in a material described by the dielectric function $\epsilon(\omega)$, the electric field at $b = |\mathbf{R}|$ can be written, following García de Abajo (2010), as

$$\mathbf{E}_e(\mathbf{r},\omega) = \frac{2e\omega \exp(i\omega z/v)}{v^2\epsilon(\omega)\gamma_\epsilon(\omega)} \left[\frac{i}{\gamma_\epsilon(\omega)} K_0\left(\frac{\omega b}{v\gamma_\epsilon(\omega)}\right) \hat{\mathbf{z}} - K_1\left(\frac{\omega b}{v\gamma_\epsilon(\omega)}\right) \hat{\mathbf{R}} \right] \tag{3.1}$$

where

$$\gamma_\epsilon(\omega) = \frac{1}{\sqrt{1 - \epsilon(\omega)v^2/c^2}}. \tag{3.2}$$

If we approximate $\epsilon(\omega)$ as real, this expression can tell us a great deal about the qualitative effects we should expect at different frequencies. For positive arguments x, the modified Bessel functions $K_\nu(x)$ decay for as $x \to \infty$, with high orders ν decaying more quickly than low orders. When $\epsilon(\omega) \leq 0$, it is obvious on inspection that $\gamma_\epsilon(\omega)$ will be positive, and that we should expect the electric field to decay as $\mathbf{R}$ becomes large. But for imaginary arguments, the functions $K_\nu(x)$ become proportional to the oscilliatory Hankel functions, leading to radiation. Simply from the form of $\gamma_\epsilon(\omega)$, one can see that this will happen when $v^2 > c^2/\epsilon(\omega)$, or, colloquially, when the speed of the electron is greater than the speed of light in the material.

The fact that charged particles lose energy to radiation when they move with $v > c/\sqrt{\epsilon}$ is well-known to physicists as Cherenkov radiation, famous as the source of the bluish light surrounding radioactive substances that undergo β-decay in a transparent medium. A standard physical interpretation of this effect (Jelley (1955)) is that when $v^2 < c^2/\epsilon(\omega)$, polarization dipoles in front of the electron and behind it can stretch one way and another along with the electron's Coulomb field as it passes, enforcing an approximate mirror symmetry. But when $v^2 > c^2/\epsilon(\omega)$, the material cannot respond quickly enough to maintain this symmetry, leading to an electric shock wave, a process analogous to the "sonic boom" that results when an object moves through air faster than the speed of sound.

This effect can also be described in terms of a "quantum" normal modal analysis. In the electrostatic analysis of the undamped dielectric slab, following Fuchs and Kliewer (1965), there exist just three types of modes: transverse modes at the lower limiting frequency ω_0, with $\epsilon(\omega_0) = -\infty$; longitudinal modes at the upper limiting frequency ω_ℓ, with $\epsilon(\omega_\ell) = 0$; and harmonic surface modes at frequencies ω_h lying

between the two, with $\omega_0 < \omega_h < \omega_\ell$, such that $-\infty < \epsilon(\omega_h) < 0$. Upon introducing retardation into the calculation of the normal modes, Kliewer and Fuchs (1966a) found that longitudinal modes maintain their same frequency ω_ℓ, and surface modes remain with in the range bounded by ω_0 and ω_ℓ, albiet at shifted frequencies. But the transverse modes require some conceptual refinement, shifting from modes for which $\mathbf{E}_t = 0$, to ones for which $\mathbf{E}$ is oscillatory and nonzero inside the slab and evanescent outside. (Longitudinal modes are also oscilliatory, but $\mathbf{E}_\ell = 0$ beyond the slab.)

In simple physics terms, these waveguide modes result from light bouncing back and forth between the surfaces of the slab via total internal reflection. As in the "classical" treatment, such modes can only be excited by an electron beam of sufficient speed to match the mode momentum. Such arguments from momentum conservation explain both why such modes can be excited when $v^2 > c^2/\epsilon(\omega)$, and why radiation cannot happen spontaneously in vacuum (once again, see García de Abajo (2010)).

This chapter is dedicated to the treatment of electrodynamic effects in silicon. After a discussion of local dielectric models, I present results for an electron running parallel to the interface of a silicon ribbon (i.e., a slab turned on its side) using the "classical" beam-first approach. Then I present results an electron running alongside the interface of a silicon cylinder using the "quantum" modes-first approach. Finally, I compare the cylinder model to experimental data on silicon nanodiscs taken by Flauraud and Alexander (2019). Combined results from the unbounded cylinder and a "thin disc" approximation yield lower and upper bounds on energies for the observed modes in real discs, and qualitatively describe the behavior of the low-order modes.

3.2 Dielectric Models of Silicon

The Born-Huang model for dielectric response, as discussed above, gives us a simple, local description of ionic or metallic materials. However, for a semiconductor

like silicon, this type of model is not the most efficient or accurate way to capture the important behaviors. In this section, I outline a few of the formal requirements that a local, causal dielectric model must obey, and discuss why these formal constraints are often ignored in practice for STEM-EELS simulations.

Some formal aspects of $\epsilon(\omega)$ have been tacitly used above without explicit acknowledgement. For instance, Eq. 2.2 writes the macroscopic displacement field $\mathbf{D}(\mathbf{r}, t)$ as convolution of the the local electric field $\mathbf{E}(\mathbf{r}, t)$ with the dielectric response $\epsilon(t)$. A real electric field should lead to a real displacement field, which implies that $\epsilon(t)$ is also real. But this means that

$$\epsilon(\omega) = \epsilon^*(-\omega) \tag{3.3}$$

by the Friedel symmetry of Fourier transforms of real-valued functions, meaning that $\mathrm{Re}[\epsilon(\omega)]$ is symmetric in ω and that $\mathrm{Im}[\epsilon(\omega)]$ is antisymmetric in ω.

The other widely cited formal requirement for $\epsilon(t)$ involves its *causal* structure, and leads to the Kramers-Kronig (KK) relations. (See Bohren (2010) for a brief history of the KK-relations, and Dethe *et al.* (2019) for a brief overview of their applications.) Since the past can influence the future but—presumably—the future cannot influence the past, the form of the dielectric function in terms of the electric susceptibility χ_e

$$\epsilon(t) = 1 + 4\pi\chi_e(t) \tag{3.4}$$

should not change if we slip in a Heaviside function:

$$\epsilon(t) = 1 + 4\pi\chi_e(t)\Theta(t). \tag{3.5}$$

One way to obtain the KK relations, championed as a "two-line proof" by Hu (1989) and as a derviation from systems theory by Schönleber *et al.* (2014), is simply to take the Fourier transforms of Eq. 3.4 and Eq. 3.5 and to set them equal. Splitting the dielectric response into its real and imaginary parts as $\epsilon(\omega) = \epsilon_1(\omega) + i\epsilon_2(\omega)$ and

using Eq. 3.3 leads to the KK relations:

$$\begin{aligned}\epsilon_1(\omega) &= 1 + \frac{2}{\pi}\int_0^\infty \frac{\omega'\epsilon_2(\omega')}{\omega'^2 - \omega^2}d\omega', \\ \epsilon_2(\omega) &= -\frac{2\omega}{\pi}\int_0^\infty \frac{\epsilon_1(\omega')}{\omega'^2 - \omega^2}d\omega'.\end{aligned} \tag{3.6}$$

These relations famously tell us that if we know $\epsilon_1(\omega)$, we can find $\epsilon_2(\omega)$, and vice versa. They also lead to one fruitful route toward modelling dielectrics. Adachi (1988), for instance, constructs an ingenious analytic model for crystalline silicon that takes all its parameters from the band structure. After some judicious approximations (parabolic bands, etc.), the model explicitly constructs $\epsilon_2(\omega)$ from the interband dipole matrix elements and the joint density of states, incorporating both direct and indirect transitions. The results of this model are shown in the upper left panel of Fig. 3.1. Adachi compared his results to the ellipsometry data of Aspnes and Studna (1983), whose results are similar to those of Jellison (1992), shown in the bottom left panel, and supplemented by the data from Aspenes and Studna at higher energies. Given the "sharp" features of the analytic approximation to $\epsilon_2(\omega)$, it is not surprising that the KK-constructed $\epsilon_1(\omega)$ in Adachi's model has unphysical singular points, labeled with dashed vertical lines at 3.38 eV and 4.27 eV.

A similar approach, where $\epsilon_2(\omega)$ is constructed and $\epsilon_1(\omega)$ is recovered by a KK analysis, is sometimes also undertaken in the context of density functional theory, at various levels of sophistication, as discussed by Zheng *et al.* (2017). A less physically grounded approach, which is nevertheless empirically adequate by construction, is to introduce various KK-consistent models and to see which one best fits the data. Leng *et al.* (1998) uses this method to fit dielectric response data for both crystalline silicon (c-Si) and amorphous silicon (a-Si), finding that different models were best suited to each material, and conceeding that any physical interpretations one might attach to such fits should be viewed with some suspicion.

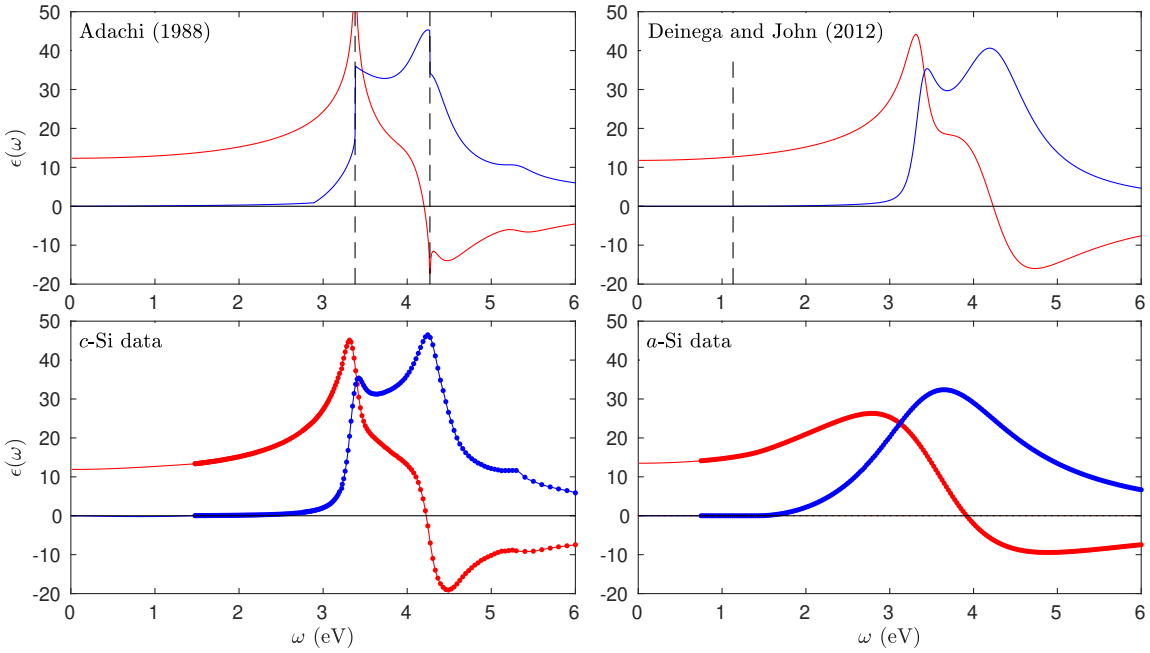

Figure 3.1: Models for $\epsilon(\omega)$ of silicon, with Re$[\epsilon(\omega)]$ plotted in red, and Im$[\epsilon(\omega)]$ plotted in blue. Counter-clockwise from bottom right: amorphous silicon (*a*-Si) from ellipsometery data; crystalline silicon (*c*-Si) data from Jellison (1992) for $\hbar\omega < 5.2$ eV and Aspnes and Studna (1983) for $\hbar\omega > 5.2$ eV; analytic model for *c*-Si from Adachi (1988); *ad hoc* model for *c*-Si from Deinega and John (2012)

.

Such difficulties in interpretation makes abandoning internal consistency an attractive possibility, as in Deinega and John (2012), who use an a use an *ad hoc* six-parameter fit to *c*-Si data, shown in the upper right of Fig. 3.1. This approach has its own limitations (their $\epsilon_2(\omega)$ becomes negative below 1.13 eV, below the range of their fitting data), but such issues are present even for a very direct model, as in the spline fit to the *c*-Si data shown as solid lines through the *c*-Si data points in Fig. 3.1.[1] And while KK-consistent models for amorphous semiconductors exist (e.g., the five-parameter model of Forouhi and Bloomer (1986)), when experimental data

[1] A spline should not extend beyond the range of avaliable data, but when low-frequency data is not available, one can interpolate over the range $(-\omega_{\max}, \omega_{\max})$ using $\epsilon(-\omega) = \epsilon^*(\omega)$ and the measured real-valued dielectric constant for $\epsilon(0)$.

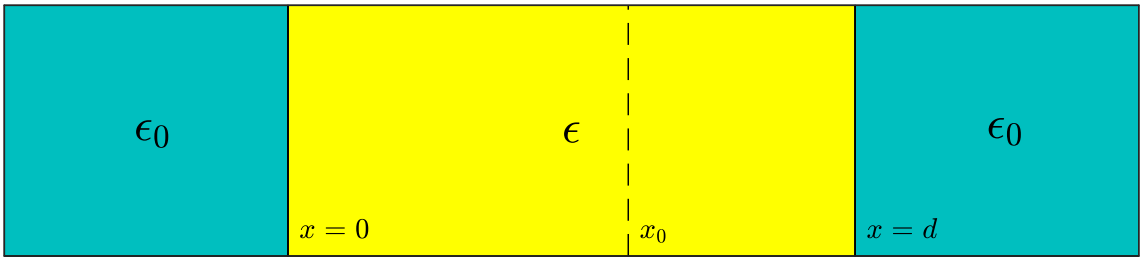

Figure 3.2: The geometry of interest: a material with a local frequency-dependent dielectric function $\epsilon(\omega)$ is flanked on both sides by vacuum ($\epsilon_0(\omega) = 1$). Meanwhile, the electron beam is positioned at $x = x_0$, where $0 < x_0 < d$.

is available, the temptation of simply using a smoothed spline is too great to resist.

This is the route I have taken in the following sections, of using interpolated results for $\epsilon(\omega)$ from measured $n(\omega)$ and $k(\omega)$ values from optical ellipsometry data ($\epsilon = (n + ik)^2$). It should be noted that such values could, in principle, be obtained directly from EELS data, using methods outlined in Potapov *et al.* (2009). But the relative ease of interpreting optical measurements may make such EELS measurements unattractive unless the spatial variation in $\epsilon(\omega)$ is itself being explored.

3.3 Modeling Silicon Ribbons

In this section, I present theoretical treatments of STEM-EELS losses for a classical electron beam penetrating a dielectric material with dielectric response $\epsilon(\omega)$, surrounded on either side by vacuum ($\epsilon_0 = 1$), as pictured in Fig. 3.2. In Sec. 3.3.1 I present the theory of such structures in the electrostatic approximation, and in Sec. 3.3.2 I review how this calculation is extended with an electrodynamic approach. In Sec. 3.3.3, I compare the results of the electrostatic and electrodynamic approaches, and show how the models evolve as ribbon width and beam energy increase. The contributions of the transverse electric (TE) and transverse magnetic (TM) modes are also isolated, along with the specific contributions of various TE/TM modes.

3.3.1 *Electrostatic Approach*

Concise expressions for the STEM-EELS spectra of a beam running parallel to a dielectric surface were first given by Echenique and Pendry (1975). Here I extend that calcuation for a beam running through a dielectric ribbon, with a derivation that runs parallel to theirs, but which adds an additional interface.

As is typical for electrostatic calculations, we begin with Laplace's equation for the beam electron, a classical point charge $-e$ in vacuum, traveling along $z = vt$ for which x_0 sets the beam position relative to the foil's surfaces:

$$\nabla^2 \Phi_e(\mathbf{r}, t) = -4\pi e \delta(z - vt)\delta(y)\delta(x - x_0). \tag{3.7}$$

To match boundary conditions at the surfaces of the foil, it will be useful to write this as an expression that is partly in real and partly in reciprocal space—viz., $\Phi_e(\mathbf{k}_\parallel, x, \omega)$, where $\mathbf{k}_\parallel = (k_z, k_y)$. Performing the spatial Fourier transforms yields

$$\Phi_e(\mathbf{k}_\parallel, x, t) = \frac{2\pi e}{|\mathbf{k}_\parallel|} \exp(-|\mathbf{k}_\parallel||x - x_0|) \exp(-ik_z vt), \tag{3.8}$$

and the t to ω transformation constrains the k_z momentum transfer

$$\Phi_e(\mathbf{k}_\parallel, x, \omega) = \frac{4\pi^2 e}{|\mathbf{k}_\parallel|} \delta(\omega - k_z v) \exp(-|\mathbf{k}_\parallel||x - x_0|), \tag{3.9}$$

which captures the x-dependence we will need for applying boundary conditions.

Now we consider how to apply this analysis to the three relevant areas of our foil: (1) $x \leq 0$, (2) $0 < x < d$, and (3) $x \geq d$. The total potential in each region $\Phi_{1,2,3}$, will include a part $\phi_{1,2,3}$ that is due to the background dielectric response:

$$\Phi_1(\mathbf{k}_\parallel, z, \omega) = \phi_1(\mathbf{k}_\parallel, z, \omega), \tag{3.10a}$$

$$\Phi_2(\mathbf{k}_\parallel, z, \omega) = \phi_2(\mathbf{k}_\parallel, z, \omega) + \Phi_e(\mathbf{k}_\parallel, z, \omega)/\epsilon(\omega), \tag{3.10b}$$

$$\phi_3(\mathbf{k}_\parallel, z, \omega) = \phi_3(\mathbf{k}_\parallel, z, \omega). \tag{3.10c}$$

These response potentials $\phi_{1,2,3}$ will not incorporate any free charge within their respective regions, so we can expect them to follow Laplace's equation

$$(i\mathbf{k}_\parallel)^2\phi + \frac{\partial^2\phi}{\partial x^2} = 0, \tag{3.11}$$

which, given that $\phi(x) \to 0$ for $x \to \pm\infty$, allows us to write

$$\phi_1(\mathbf{k}_\parallel, x, \omega) = C_1\phi(\mathbf{k}_\parallel, \omega)\exp(|\mathbf{k}_\parallel|x), \tag{3.12a}$$

$$\phi_2(\mathbf{k}_\parallel, x, \omega) = C_{2a}\phi(\mathbf{k}_\parallel, \omega)\exp(|\mathbf{k}_\parallel|x) + C_{2b}\phi(\mathbf{k}_\parallel, \omega)\exp(-|\mathbf{k}_\parallel|x), \tag{3.12b}$$

$$\phi_3(\mathbf{k}_\parallel, x, \omega) = C_3\phi(\mathbf{k}_\parallel, \omega)\exp(-|\mathbf{k}_\parallel|x), \tag{3.12c}$$

where $\phi(\mathbf{k}_\parallel, \omega) = -4\pi^2 e\delta(\omega - k_z v)/|\mathbf{k}_\parallel|$ for each term.

From these expressions, we can immediately find the relevant boundary conditions, first for the continuity of the potential at $x = 0$ and $x = d$ and next for the continuity of the x-component of the displacement field. From these, it is straightforward to solve for the coefficients C_{2a} and C_{2b} that predict the surface contribution to the potential induced by the beam when it is propagating inside a foil, and, summing the expressions, we find a neat form for the surface contributions:

$$C_{2a}e^{|\mathbf{k}_\parallel|x_0} + C_{2b}e^{-|\mathbf{k}_\parallel|x_0} = \frac{(\epsilon^2 - \epsilon_0^2)\left(\exp(2kx_0) + \exp(2k(d - x_0))\right) + 2\left(\epsilon - \epsilon_0\right)^2}{\epsilon\left[(\epsilon + \epsilon_0)^2 e^{2|\mathbf{k}_\parallel|d} - (\epsilon - \epsilon_0)^2\right]}. \tag{3.13}$$

Putting these together to find the response potential Φ_0, we can solve for the work W done on the particle, and extract our desired quantity from the integrand as

$$-\frac{dW}{dz} = q_e\frac{\partial\Phi_0}{\partial z}\bigg|_{beam} = \int_0^\infty d\omega\,(\hbar\omega)\frac{d^2P}{dz\,d\omega}, \tag{3.14}$$

leading us to an expression that can be calculated, for a cutoff k_y^{cut} to match the aperture size and/or to reflect the atomic scale of the material:

$$
\begin{aligned}
\frac{d^2P}{dz\,d\omega} &= \frac{e^2}{\pi\hbar v^2}\mathrm{Im}\left[\int_0^{k_y^{cut}} dk_y \left(\chi_{\mathrm{bulk}} + \chi_{\mathrm{surface}}\right)\Big|_{k_z=\omega/v}\right]; \\
\chi_{\mathrm{bulk}} &= -\frac{1}{|\mathbf{k}_\parallel|}\left(\frac{\epsilon-1}{\epsilon}\right), \\
\chi_{\mathrm{surf}} &= -\frac{\left(\epsilon^2-\epsilon_0^2\right)\left(\exp(2kx_0)+\exp(2k(d-x_0))\right)+2\left(\epsilon-\epsilon_0\right)^2}{|\mathbf{k}_\parallel|\epsilon\left[(\epsilon+\epsilon_0)^2e^{2|\mathbf{k}_\parallel|d}-(\epsilon-\epsilon_0)^2\right]}.
\end{aligned}
\tag{3.15}
$$

A few relevant formal features of this expression may be noted. The bulk loss term vanishes when $\epsilon = 1$, as there is no bulk loss in vacuum, and the surface loss term vanishes when $\epsilon = \epsilon_0$, as this reduces to the bulk loss case. Also, the surface losses depend on x_0 and $(d - x_0)$, the distances from the left and right surfaces of the ribbon to the beam. And these surface losses will be largest where the denominator of χ_{surf} is small, as revealed by factoring:

$$
\begin{aligned}
0 &\approx e^{|\mathbf{k}_\parallel|d}(\epsilon+\epsilon_0) - (\epsilon-\epsilon_0), \\
0 &\approx e^{|\mathbf{k}_\parallel|d}(\epsilon+\epsilon_0) + (\epsilon-\epsilon_0)
\end{aligned}
\tag{3.16}
$$

The other case with a small denominator, for $\epsilon \approx 0$, leads to a negative surface contribution, the so-called "Begrengzung" effect, from a shifting of the density of states away from the longitudinal bulk modes.

3.3.2 Electrodynamic Approach

The transition from the electrostatic to the electrodynamic (yet still classical) treatment of STEM-EELS is a fairly seamless one, for the very limited cases where analytical solutions are known. Garcia-Molina *et al.* (1985) presented a relativistic version for a beam running parallel to a dielectric surface. Later, Bolton and Chen (1995) extended the relativistic theory to the case of an arbitrary number of parallel layers. Wang (1996) provided a well-written and detailed review of these methods.

The innovation of Garcia-Molina *et al.*—which was carried forward by Bolton and Chen—was in solving for the dielectric response fields in terms of the Hertz electric vector potential. I will not present this solution in detail, but will describe its outline and quote the relevant result. The Hertz electric vector potential $\mathbf{\Pi}$ (see Stratton (1941); Essex (1977)) is related to the conventional vector and scalar potentials as

$$\begin{aligned} \phi &= -\nabla \cdot \mathbf{\Pi} \\ \mathbf{A} &= \mu\epsilon \frac{\partial \mathbf{\Pi}}{\partial t}, \end{aligned} \tag{3.17}$$

which fixes its relationship to the electric and magnetic fields

$$\begin{aligned} \mathbf{E} &= -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t} = \nabla\left(\nabla \cdot \mathbf{\Pi}\right) - \frac{\partial^2 \mathbf{\Pi}}{\partial t^2} \\ \mathbf{B} &= \nabla \times \mathbf{A} = \mu\epsilon \frac{\partial}{\partial t}\left(\nabla \times \mathbf{\Pi}\right). \end{aligned} \tag{3.18}$$

If we substitute these into Maxwell's equations and Fourier transform from t to ω, we find a neat relationship for $\mathbf{\Pi}$ that shows its utility in simplifying problems where there is a known current density $\mathbf{J}(\mathbf{r}, \omega)$:

$$\left(\nabla^2 + \frac{\epsilon\omega^2}{c^2}\right) \mathbf{\Pi}(\mathbf{r}, \omega) = \frac{4\pi}{i\omega\epsilon} \mathbf{J}(\mathbf{r}, \omega). \tag{3.19}$$

This allows solutions to the Helmholtz equation to be used in expressing the unknown electromagnetic response to the known beam current, in a way that is exactly analogous to the way that solutions to the Laplace equation were used in response to the beam potential in the electrostatic solution presented above. Just as above, the spatial Fourier transform is taken in the directions parallel to the interface—i.e., from (y, z) to $\mathbf{k}_\parallel = (k_y, k_z)$—and one then sets boundary conditions to find $\mathbf{\Pi} = (\Pi_x, 0, \Pi_z)$, where Π_y is taken to vanish by symmetry. In the end, the electric field is recovered, and the work done on the beam by the response field leads to the EEL spectrum as the integrand of the work in the usual way.

Of course, the mathematical solution to this is somewhat more involved than the electrostatic version, even if we maintain $\mathbf{H} = \mathbf{B}$ for non-magnetic materials. At each

boundary interface x_b, the continuity of H_x and H_z leads to the following boundary conditions for Π_z and Π_x

$$\begin{aligned} \epsilon_{\text{left}}\Pi_z^{\text{left}}(x_b, \mathbf{k}_\parallel) &= \epsilon_{\text{right}}\Pi_z^{\text{right}}(x_b, \mathbf{k}_\parallel) \\ \epsilon_{\text{left}}\Pi_x^{\text{left}}(x_b, \mathbf{k}_\parallel) &= \epsilon_{\text{right}}\Pi_x^{\text{right}}(x_b, \mathbf{k}_\parallel), \end{aligned} \tag{3.20}$$

where the cross product leads to a component flip (that is, the continuity of H_x leads to the Π_z condition, and the continuity of H_z leads to the Π_x condition). Likewise, the continuity of the perpendicular component of the displacement field $D_x = \epsilon E_x$ and of the parallel component E_z lead (when simplified by Eqs. 3.19 and Eq. 3.20 and using $\mathbf{J}(\mathbf{r}, \omega) = 0$) to boundary conditions at x_b for $\mathbf{k}_\parallel$

$$\begin{aligned} \epsilon_{\text{left}}\frac{\partial \Pi_z^{\text{left}}}{\partial x} &= \epsilon_{\text{right}}\frac{\partial \Pi_z^{\text{right}}}{\partial x} \\ ik_z\Pi_z^{\text{left}} + \frac{\partial \Pi_x^{\text{left}}}{\partial x} &= ik_z\Pi_z^{\text{right}} + \frac{\partial \Pi_x^{\text{right}}}{\partial x}. \end{aligned} \tag{3.21}$$

For a system with two boundaries, one needs to introduce expansion coefficients in the materials at the right and left, and two expansion coefficients in the middle layer (since none can be eliminated by the boundary condition at infinity), for both Π_x and Π_z. This leads to four unknowns for Π_x and four unknowns for Π_z, which can be solved with the four boundary conditions at the two interfaces ($4 + 4 = 8$ equations, to match the $4 + 4 = 8$ unknowns from Π_x and Π_z). The theory of Bolton and Chen casts these systems of equations in terms of invertible matrices, which generalizes this method to arbitrarily many parallel layers.

For the simple geometry considered here (Fig. 3.1), a concise expression for the loss per unit length is given by Couillard *et al.* (2008), as simplified from the theory of multilayered slabs developed by Bolton and Chen (1995). In our notation, it becomes

$$\frac{d^2P}{dz\,d\omega} = \frac{e^2}{\pi\hbar v^2}\text{Im}\left[\int_0^{k_y^{cut}} dk_y \chi_{\text{ribbon}}\Big|_{k_z=\omega/v}\right] \tag{3.22}$$

where

$$\chi_{\text{ribbon}} = \frac{1}{q\epsilon\mathbf{k}_\parallel^2}\left\{\epsilon k_y^2\left(\frac{v}{c}\right)^2\frac{\widetilde{\gamma^-}\widetilde{\zeta^+}}{\widetilde{L^+}\widetilde{L^-}} - q^2\frac{\gamma^+\zeta^-}{L^+L^-}\right\} \tag{3.23}$$

in which

$$\begin{aligned} q = \sqrt{\mathbf{k}_\parallel^2 - \epsilon(\omega/c)^2},\ q_0 = \sqrt{\mathbf{k}_\parallel^2 - \epsilon_0(\omega/c)^2}, \\ h^\sigma = q\epsilon_0 + \sigma q_0\epsilon,\ \widetilde{h^\sigma} = q + \sigma q_0, \end{aligned} \tag{3.24}$$

where σ sets the sign in these expressions (i.e., σ can be "+" or "-"), and

$$\begin{aligned} \gamma^\sigma &= h^+\exp(q(2d - x_0)) - \sigma h^-\exp(qx_0), \\ \zeta^\sigma &= h^+\exp(qx_0) + \sigma h^-\exp(-qx_0), \\ L^\sigma &= h^+\exp(qd) + \sigma h^-, \end{aligned} \tag{3.25}$$

and where the terms are in Eqn. 3.23 with tildes on top have the same form as their the counterpart expressions in Eqn. 3.25, except with h^σ replaced by $\widetilde{h^\sigma}$ as defined in Eq. 3.24. In the limit where $c \to \infty$, this expression reduces to the electrostatic expression of Eq. 3.15. Bolton and Chen remark that the term in χ_{ribbon} with $\widetilde{L^+}\widetilde{L^-}$ in the denominator is associated with the transverse electric (TE) waveguide modes, which travel in some direction $\hat{\mathbf{k}}_\parallel$ with $\mathbf{E}\cdot\hat{\mathbf{k}}_\parallel = 0$ and $E_x = 0$, and the term with L^+L^- in the denominator is associated with transverse magentic (TM) modes, which travel in some direction $\hat{\mathbf{k}}_\parallel$ with $E_z/E_y = k_z/k_y$ and $H_x = 0$.

It is also worth quoting the expression for the bulk losses, where, as before,

$$\frac{d^2P}{dz\,d\omega} = \frac{e^2}{\pi\hbar v^2}\text{Im}\left[\int_0^{k_y^{cut}} dk_y\chi_{\text{bulk}}\Big|_{k_z=\omega/v}\right], \tag{3.26}$$

but for this case

$$\chi_{\text{bulk}} = \frac{\epsilon(v/c)^2 - 1}{\epsilon\sqrt{\mathbf{k}_\parallel^2 - \epsilon(\omega/c)^2}} \tag{3.27}$$

takes the place of χ_{ribbon} in Eqn. 3.22. This expression includes Cherenkov losses, as can be confirmed directly in the case where ϵ is real. In this case, the integral can be performed, and the expression reduces to the classic Frank-Tamm formula

$$\frac{d^2P}{dz\,d\omega} = \frac{e^2}{\hbar c^2}\left(1 - \frac{1}{\epsilon(v/c)^2}\right)\Theta\left(\epsilon(v/c)^2 - 1\right), \tag{3.28}$$

where $\Theta(x)$ is the Heaviside function. However, Eqs. 3.22 and 3.26 are not *limited* to Cherenkov losses, as the imaginary part of the dielectric function can account for energy absorbed by the material that is not tranlated into far-field radiation.

3.3.3 Numerical Results

For these numerical results, I have (somewhat artificially) turned the loss-per-length results into standard spectra by multiplying by a sample thickness of 100 nm. In all these results, I have also assumed a 150 μrad aperture, which alters the k_y cutoff used in the integration for each beam energy we under inverstigation. I have used a dielectric function for the ribbon matching the c-Si data where possible, and the model of Adachi (1988) in the far infrared region, to avoid negative interpolated values of $\epsilon_2(\omega)$, and hence to avoid negative values for the simulated spectra.

In Fig. 3.3, the electrostatic loss (Eq. 3.15) and electrodynamic loss (Eq. 3.22) are compared for different beam energies (10 keV, 30 keV, 300 keV), at the center of ribbons of varying widths (10 nm, 100 nm, 1000 nm, and bulk). The 10 keV beam does not reach the Cherenkov loss condition $v^2 > c^2/\epsilon$ for any frequency, and the electrostatic approach is empirically adequate for all the investigated widths. For the 30 keV beam, one can observe the emergence of sharp spikes in the spectrum atop a broad

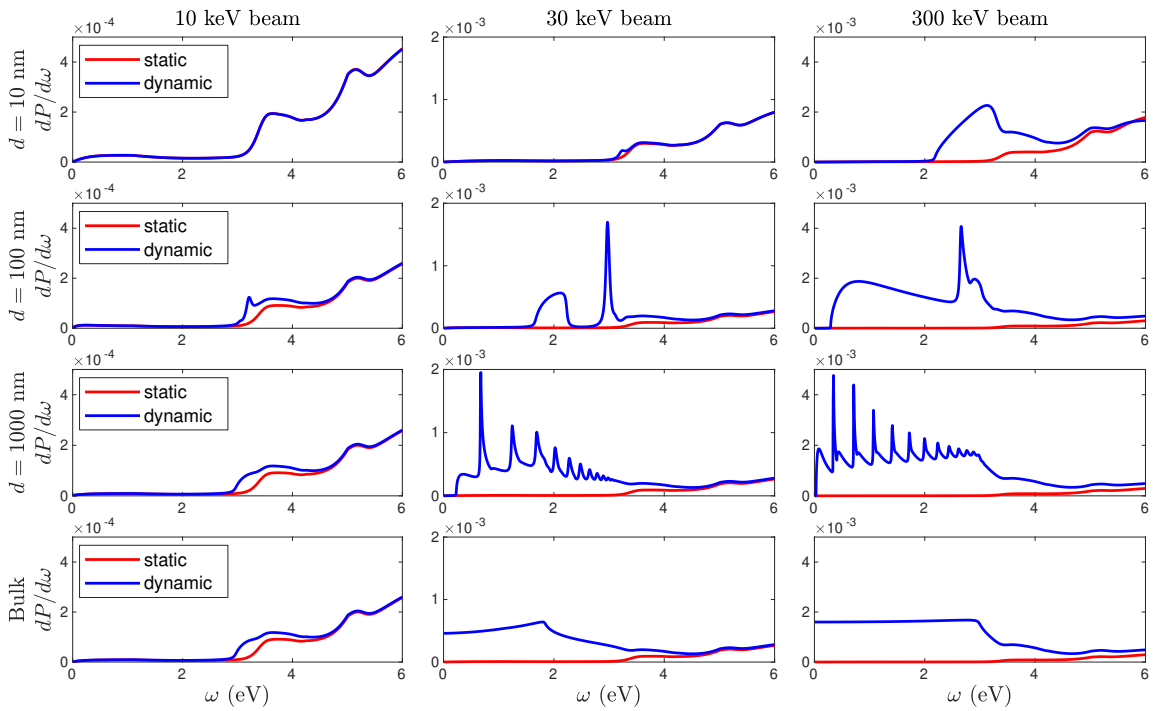

Figure 3.3: Electrostatic spectra (in red) and electrodynamic spectra (in blue) are shown for three beam energies (10 keV, 30 keV, 300 keV, for columns from left to right), and for four different sample widths (10 nm, 100 nm, 1000 nm, and bulk, from top to bottom), for a beam in the center of the ribbon. Where an electrostatic spectrum cannot be seen, it is indistinguishable from the electrodynamic spectrum.

background where $\mathrm{Re}[\epsilon(\omega)] > 0$ as the ribbon width increases. Even as this signal shifts to lower frequencies with more peaks for wider ribbons, however, the portion of the spectrum where $\mathrm{Re}[\epsilon(\omega)] < 0$ remains well-described by the electrostatic approach. The 300 keV beam makes the electrodynamic effects even more pronounced, shifting frequencies of the spectral spikes down further still. Where $\mathrm{Re}[\epsilon(\omega)] < 0$, the electrostatic approach now underestimates the intensity of the spectrum, but continues to qualitatively follow the electrostatic prediction. Given this (sometimes only qualitative) success of the electrostatic approach above ~ 3.4 eV, in the figures below I have restricted my inquiry to 0-3 eV.

To tease out the source of the spikes in the spectra of Fig. 3.3, I have separately plotted the TE and TM contributions to spectral maps across the ribbon in Fig. 3.4.

The sharp spikes in the spectra arise from the TE modes, and the broad background arises from the TM modes. This difference in character is explained by the fact that for TE modes, $E_z/E_y = -k_y/k_z$, and for TM modes, $E_z/E_y = k_z/k_y$, relationships whose consequences can be seen in Fig. 3.5 below. For each, the spectral maps across the ribbon clearly show standing wave-type patterns, with the onset of a modal type occuring at slightly lower energies for the TE modes, and at slightly higher energies for the TM modes. But at low energies, an extra band of TM contributions without any nodes presents itself, which means that for thin layers the EEL signal is dominated by TM mode contributions, as was noticed by Couillard *et al.* (2008).

Fig. 3.5 presents this TE/TM mode breakdown in more detail, showing the contrbutions of various modes as a function of k_y, where the value of k_z is set by ω/v as $\mathbf{k}_\parallel = (k_y, \omega/v)$. The main consideration of whether a mode contributes strongly or weakly to the spectrum is the projection of the mode's electric field along the z-direction. Since these various values for $\mathbf{k}_\parallel$ represent different directions that a mode might travel, the spectra reveal that TE modes are imaged most strongly at small values of k_y, when they are travelling at small angles relative to the optic axis, while TM modes are imaged fairly consistently over a larger range of k_y values, with less intense contributions over a wider angular range. This results from the fact that $E_z/E_y \approx -vk_y/\omega$ with $E_x = 0$ for imaged TE modes, and $E_z/E_y \approx v\omega/k_y$ with $E_x \neq 0$ for imaged TM modes, where v is the speed of the beam electron.

One intriguing feature of these plots is that the frequencies of observed modes decrease both when the ribbon's width increases, as might be predicted from the intuitive notion that longer-wavelength standing waves should have lower energy, and when the the beam energy increases, as might not be intuitively obvious. In the modes-first treatment of the silicon cylinder discussed below, it is much easier to see how this outcome results from physical constraints on momentum transfer.

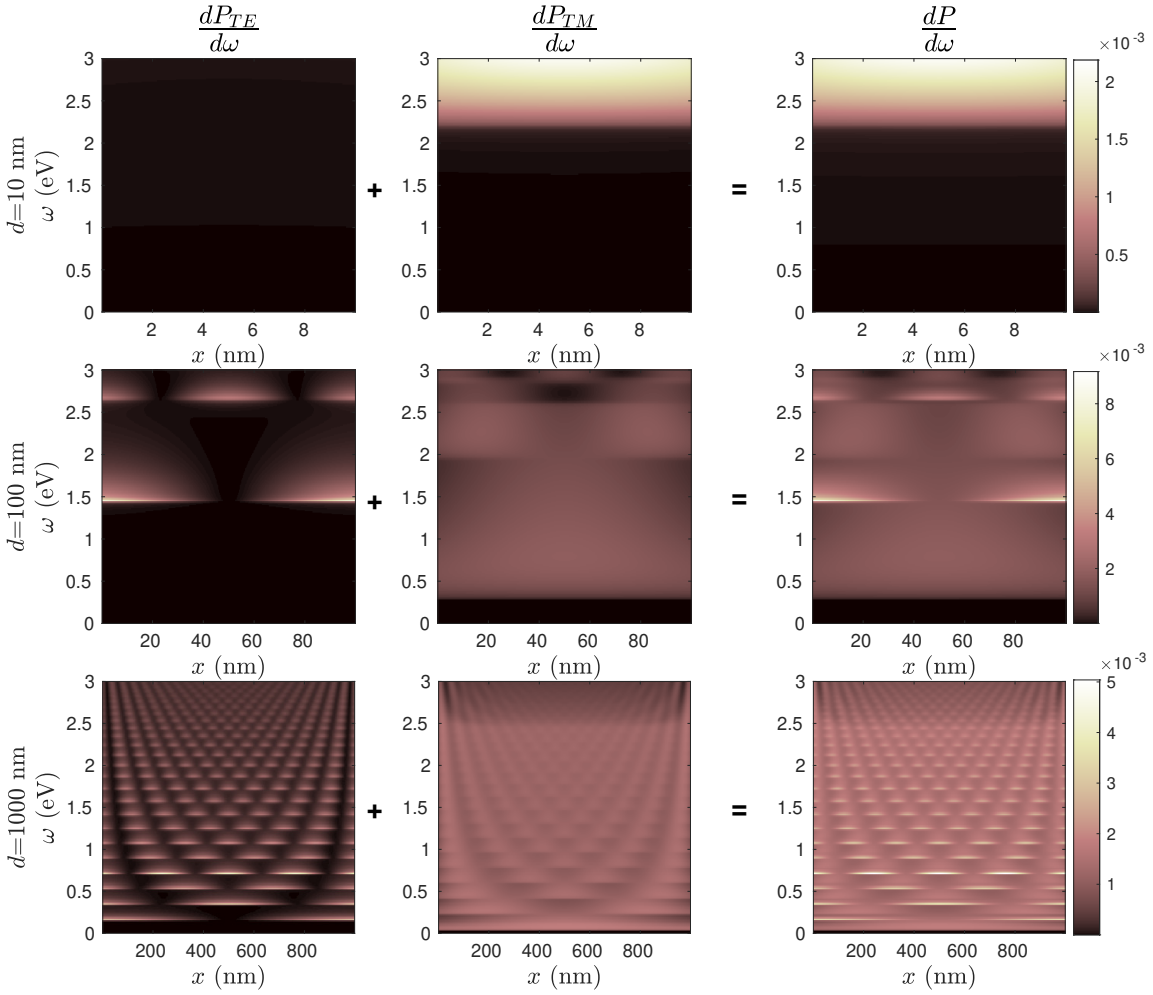

Figure 3.4: Breakdown of transverse electric (TE, leftmost column) and transverse magnetic (TM, center column) modes to the full STEM-EEL spectrum ($TE + TM$, rightmost column), for 300 keV electrons penetrating a 100 nm thick silicon ribbons with widths of 10 nm, 100 nm, and 1000 nm (top row, middle row, bottom row).

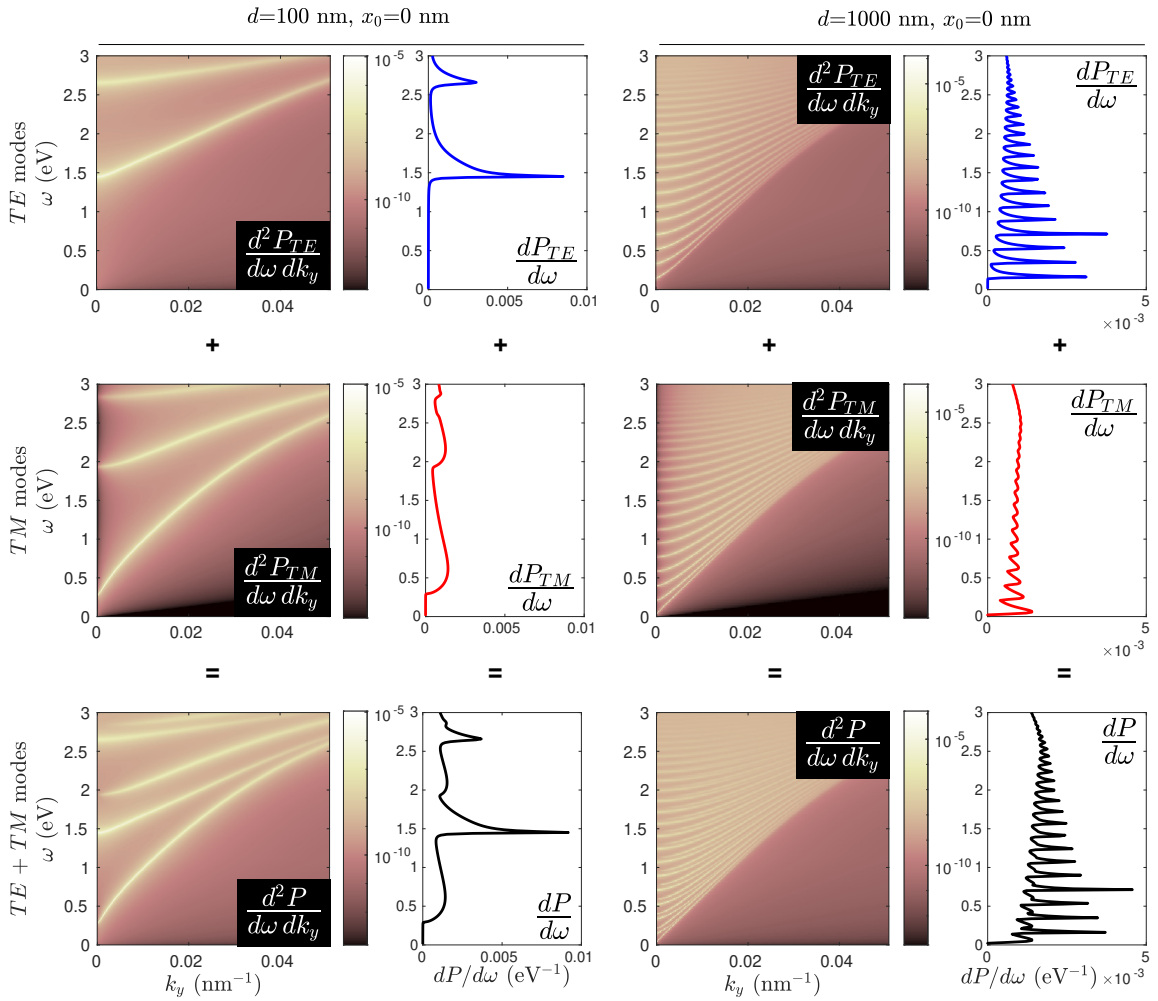

Figure 3.5: Mode contributions sorted by the wavenumber k_y for TE (top row) and TM (middle row) modes are shown alongside their resultant partial EEL spectra, for 300 keV electrons running along the edge of a silicon ribbon that is 100 nm-wide (left two columns) or 1000 nm-wide (right two columns). Each ribbon is assumed to be 100 nm thick. Summed TE and TM contributions are shown in the combined dispersion relations and spectra of the bottom row.

3.4 Modeling Silicon Cylinders

In the treatment of STEM-EELS of waveguide modes in dielectric ribbons, the derivation is consistent because the electric potential and the Hertz vector componets can be expanded in terms of terms of $\exp(\mathbf{k}_{\|} \cdot \mathbf{x}_{\|})$. In cylindrical coordinates, such an expansion does not help, so I take the route of solving for the waveguide modes in a dielectric cylinder and calculating their excitation probabilities per distance the electron travels within the cylinder. This omits the radiative mode contributions, which are expected to be much smaller than the contributions of waveguide modes.

In this treatment, the azimuthally symmetric waveguide modes split into TE and TM descriptions, where only TM modes contribute to the STEM-EELS spectrum for a beam running parallel to the cylindrical interface. Modes with an angular dependence do not cleave into the neat TE/TM categories, but they can be parametrized according to the relative proportion of H_z and E_z in each mode, allowing the hybrid modes to be categorized as EH (more like TE) or HE (more like TM).

In Sec. 3.4.1, this analysis is carried out using the simplifying assumption that $\epsilon(\omega)$ is real-valued. In Sec. 3.4.2, I discuss the difficulties in pushing this analysis forward for lossy dielectrics. And in Sec. 3.4.3, I apply the waveguide model to the experimental results of Flauraud and Alexander (2019) for amorphous silicon discs, showing that the modes developed in the waveguide theory can give us lower and upper energy limits for the modes observed in the disc data.

3.4.1 Undamped Guided Modes

One way to derive the waveguide modes of a cylindrical dielectric rod is to privilege the z-components of the electric and magentic fields, those components that point along the cylindrical axis. For quantized modes, one can introduce a fictitious length

L over which all components are periodic, such that $k_z = 2\pi\ell/L$, where ℓ is some integer. All modes are also described via an azimuthal quantum number m. Together, these lead to a spatial factor of $\exp(i(m\theta + k_z z))$ and a time dependence $\exp(-i\omega t)$, which are supressed for concision in writing out the field components below.

The Bessel $J_m(x)$ functions are used to describe the radial oscillations of modes within the cylinder, and the modified Bessel $K_m(x)$ functions are used to describe their radial decay outside the cylinder. To capture this concisely, I introduce notation to bundle these functions inside and outside the cylinder

$$B_m(k_r r) = \begin{cases} J_{|m|}(k_r r), & r \leq a \\ \frac{J_{|m|}(k_r^< a)}{K_{|m|}(k_r^> a)} K_{|m|}(k_r r), & r > a \end{cases} \tag{3.29}$$

where

$$k_r = \begin{cases} k_r^< = \sqrt{\epsilon(\omega)\omega^2/c^2 - k_z^2}, & r \leq a \\ k_r^> = \sqrt{k_z^2 - \omega^2/c^2}, & r > a \end{cases} \tag{3.30}$$

and where the frequency ω in question is assumed to be the eigenfrequency of the mode. The "$<$" and "$>$" superscripts are included on k_r for clarity only where ambiguity exists. (For instance, $k_r^<$ is always inside $J_m(k_r r)$, and $k_r^>$ is always inside $K_m(k_r r)$.) Factors in $B_m(k_r r)$ have been introduced to force its continuity at the cylindrical surface a. Below, I also introduce a factor to soak up any sign differences that occur between the inside and outside of the cylinder, a "sign function"

$$s_r = \begin{cases} 1, & r \leq a \\ -1, & r > a \end{cases} . \tag{3.31}$$

that is convenient as a result of my eccentric definition of k_r.

If we use the input of the E_z and H_z components, the other components can be found from the set form in t and z (i.e., $\exp(i(k_z z - \omega t))$). In a region with dielectric

function ϵ and a magnetic permeability $\mu = 1$, the $\hat{\theta}$ and $\hat{\mathbf{r}}$ components of the fields must follow

$$E_r = \frac{ik_z s_r}{k_r^2}\left(\frac{\partial E_z}{\partial r} + \frac{\omega}{k_z c}\frac{1}{r}\frac{\partial H_z}{\partial \theta}\right) \tag{3.32a}$$

$$H_r = \frac{ik_z s_r}{k_r^2}\left(-\frac{\omega\epsilon}{k_z c}\frac{1}{r}\frac{\partial E_z}{\partial \theta} + \frac{\partial H_z}{\partial r}\right) \tag{3.32b}$$

$$E_\theta = \frac{ik_z s_r}{k_r^2}\left(\frac{1}{r}\frac{\partial E_z}{\partial \theta} - \frac{\omega}{k_z c}\frac{\partial H_z}{\partial r}\right) \tag{3.32c}$$

$$H_\theta = \frac{ik_z s_r}{k_r^2}\left(\frac{\omega\epsilon}{k_z c}\frac{\partial E_z}{\partial r} + \frac{1}{r}\frac{\partial H_z}{\partial \theta}\right) \tag{3.32d}$$

to satisfy the Helmholtz equation at the mode eigenfrequency ω.

The transverse electric (TE) modes are defined by the fact that $E_z = 0$, and $H_z \propto B_0(k_r r)$, with no angular dependence (i.e., $m = 0$). Their frequencies are defined implicitly by the eigenvalue equation

$$\left(\frac{1}{k_r^<}\frac{J_0'(k_r^< a)}{J_0(k_r^< a)} + \frac{1}{k_r^>}\frac{K_0'(k_r^> a)}{K_0(k_r^> a)}\right) = 0 \tag{3.33}$$

where the derivatives of the special functions are with respect to their arguments

$$J_m'(x) = \frac{1}{2}\Big(J_{m-1}(x) - J_{m+1}(x)\Big), \tag{3.34a}$$

$$K_m'(x) = -\Big(K_{m-1}(x) + \frac{n}{x}K_m(x)\Big). \tag{3.34b}$$

The TE modes cannot be excited by an electron beam running parallel to the cylindrical axis, as $E_z^{TE} = 0$. Since they are not imaged I have not listed their field components here, although one can easily construct them using Eq. 3.32.

The transverse magnetic (TM) modes, on the other hand, have $E_z^{TM} \propto B_0(k_r r)$ and $H_z = 0$. Since $E_z^{TM} \neq 0$, TM modes can be imaged by an beam running parallel to the cylindrical axis. The TM eignevalue equation is

$$\left(\frac{\epsilon}{k_r^<}\frac{J_0'(k_r^< a)}{J_0(k_r^< a)} + \frac{1}{k_r^>}\frac{K_0'(k_r^> a)}{K_0(k_r^> a)}\right) = 0 \tag{3.35}$$

and allows mode frequencies to be found. Notice that the condition for these modes that k_r is real-valued both inside and outside the cylinder puts limits on the mode frequencies for a given k_z

$$\frac{k_z^2}{\epsilon} < \frac{\omega^2}{c^2} < k_z^2, \tag{3.36}$$

which also applies for the hybridized modes discussed below.

The electric field components for TM modes can be written as

$$E_z = B_0(k_r r), \qquad E_r = s_r \frac{ik_z}{k_r} B_0'(k_r r), \qquad E_\theta = 0 \tag{3.37}$$

and the magentic field components can be written as

$$H_z = H_r = 0, \qquad H_\theta = s_r \frac{i\omega\epsilon}{ck_r} B_0'(k_r r). \tag{3.38}$$

Other modes, modes without azimuthal symmetry, cannot be categorized according to the TE/TM distinction. The general eigenmode equation

$$\begin{aligned}&\left(\frac{1}{k_r^<}\frac{J_{|m|}'(k_r^< a)}{J_{|m|}(k_r^< a)} + \frac{1}{k_r^>}\frac{K_{|m|}'(k_r^> a)}{K_{|m|}(k_r^> a)}\right)\left(\frac{\epsilon}{k_r^<}\frac{J_{|m|}'(k_r^< a)}{J_{|m|}(k_r^< a)} + \frac{1}{k_r^>}\frac{K_{|m|}'(k_r^> a)}{K_{|m|}(k_r^> a)}\right) \\ &= \frac{m^2 k_z^2 c^2}{a^2\omega^2}\left(\frac{1}{(k_r^<)^2} + \frac{1}{(k_r^>)^2}\right)^2\end{aligned} \tag{3.39}$$

can be solved to find allowed mode frequencies. Following the notation of Snitzer (1961), the modes can be parameterized using a factor P_m that is related to the ratio $< H_z > / < E_z >$, a quantity that we should expect to diverge for the TE modes, and that we should expect to approach zero for the TM modes. For the hybridized modes EH and HE modes ($|m| \geq 1$), P_m can be calculated as

$$P_m = \frac{\frac{m}{a}\left(\frac{1}{(k_r^<)^2} + \frac{1}{(k_r^>)^2}\right)}{\left(\frac{1}{k_r^<}\frac{J_m'(k_r^< a)}{J_m(k_r^< a)} + \frac{1}{k_r^>}\frac{K_m'(k_r^> a)}{K_m(k_r^> a)}\right)}. \tag{3.40}$$

For the modes given below, $0 < |P_m^{HE}| < 1$, and $1 < |P_m^{EH}|$.

The electric field components for these modes can be written as

$$\begin{aligned} E_z &= B_m(k_r r) \\ E_r &= \frac{ik_z}{k_r} s_r \left(B'_m(k_r r) - P_m \frac{m B_m(k_r r)}{k_r r} \right) \\ E_\theta &= -\frac{k_z}{k_r} s_r \left(P_m B'_m(k_r r) - \frac{m B_m(k_r r)}{k_r r} \right) \end{aligned} \tag{3.41}$$

and the magnetic field components can be written as

$$\begin{aligned} H_z &= \frac{ik_z c}{\omega} P_m B_m(k_r r) \\ H_r &= -\frac{\omega\epsilon}{k_r c} s_r \left(P_m \frac{k_z^2 c^2}{\omega^2 \epsilon} B'_m(k_r r) - \frac{m B_m(k_r r)}{k_r r} \right) \\ H_\theta &= \frac{i\omega\epsilon}{k_r c} s_r \left(B'_m(k_r r) - P_m \frac{k_z^2 c^2}{\omega^2 \epsilon} \frac{m B_m(k_r r)}{k_r r} \right) \end{aligned} \tag{3.42}$$

keeping in mind, as always, implicit factors of $\exp(i(m\theta + k_z z))$ and $\exp(-i\omega t)$.

In modern texts on waveguides (for instance, in Yeh and Shimabukuro (2008)), modes are typically labelled with two quantum numbers—e.g., TM_{mn} or HE_{mn}—that refer resepctively to the azimuthal dependence (i.e., the factor of $\exp(im\theta)$) and the order of the solution (i.e., $n = 1$ labels the lowest-frequency solution for a for a given m and k_z, $n = 2$ labels the second-lowest-frequency solution, and so on). Solutions of different k_z but the same m and n often have the same qualitative features.

If we wish to simulate the STEM-EELS imaging of these modes, they must be correctly scaled so as to have physical units. The "field components" given in Eq. 3.37 and 3.38 or in Eq. 3.41 and 3.42 can be normalized using the methods outlined by Khrennikov *et al.* (2012), which gives them physical units such that scattering probabilities can be calculated. In the projection approximation for a point probe at $\mathbf{x} = (x_0, y_0)$, the EEL transition probability in terms of these normalized electric field modes for quantum numers ℓ (for k_z), m (for θ) and n (for solution number) is

$$P_{\ell mn}(\mathbf{x}) = \left(\frac{e}{\hbar\omega_{\ell mn}}\right)^2 \left| \int dz\, E^z_{\ell mn}(\mathbf{x}, z) \exp(i\omega_{\ell mn} z/v) \right|^2, \tag{3.43}$$

and the STEM-EEL spectrum becomes

$$\frac{dP}{d\omega} = \sum_{\ell mn} P_{\ell mn}(\mathbf{x}) \delta(\omega - \omega_{\ell mn}). \tag{3.44}$$

Since the normalization factor introduces a factor of L in $P_{\ell mn}$ for the box size, this leads to an expression of loss per unit length, $d^2P/dz\, d\omega$, where the integral over z picks out modes for which $k_z = \omega_{\ell mn}/v$.

The undamped approximation is a reasonable for $\hbar\omega < 3.1$ eV for c-Si and $\hbar\omega < 2.0$ eV for a-Si, the range for which $\epsilon_2(\omega) < 0.1\epsilon_1(\omega)$ for each material, though it is interesting to see what results we obtain if we push the model past its domain of apparent validity, as I do in Sec. 3.4.3. But first, to study the model without impertinence, in Fig. 3.6 the results of this model are shown for a cylinder with a 300 nm diameter and an unvarying dielectric dependence $\epsilon = 11.7$ (i.e., for crystalline silicon approximating $\epsilon(\omega) \approx \epsilon(0)$), mapped with a 60 keV beam and a 300 keV beam.

The leftmost column of Fig. 3.6 shows dispersion relations for the azimuthal quantum numbers $m = 0, \pm 1, \pm 2$. The dashed black lines indicate the frequency limits at each value of k_z, as given by Eq. 3.36. The dashed dispersion curves in these plots represent the TE modes (for $m = 0$) and the EH modes (for $|m| \geq 1$), which are either not imaged (the TE modes) or imaged with low intensity (the EH modes). The solid dispersion curves in these plots represent the TM modes (for $m = 0$) and the HE modes (for $|m| \geq 1$), which together dominate the spectra.

The rows of Fig. 3.6 show the contributions from the first three azimuthal quan-

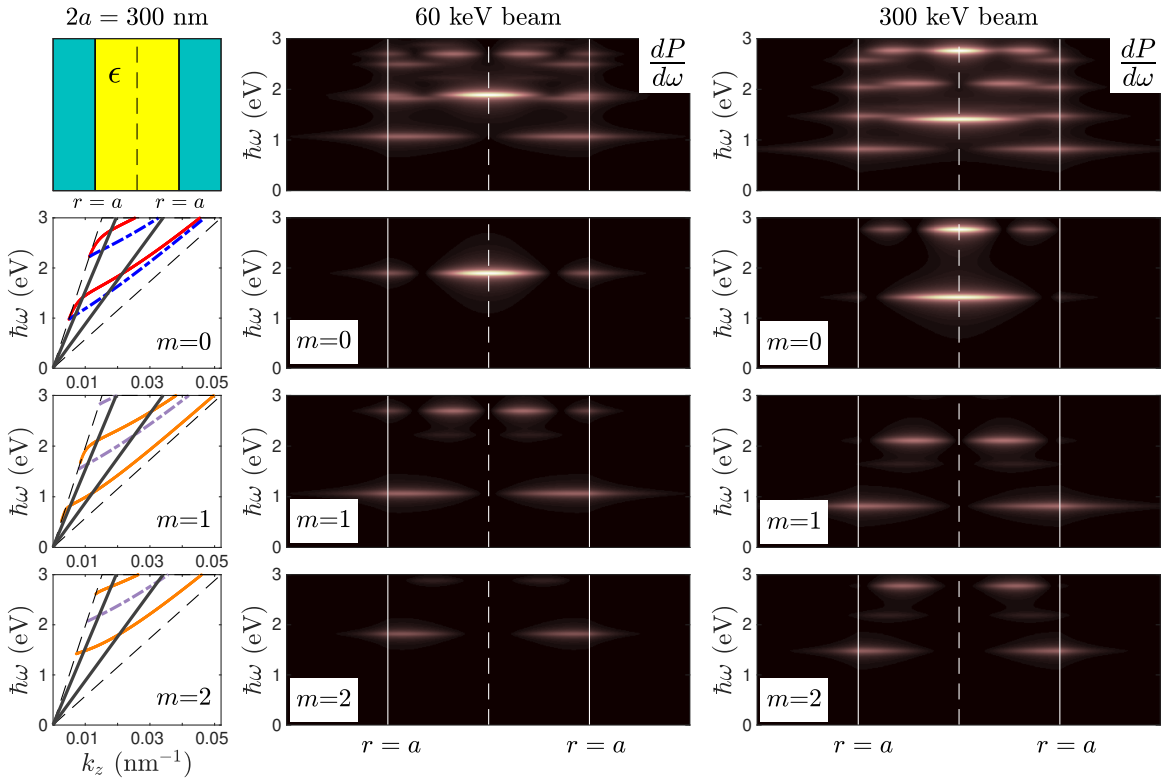

Figure 3.6: Dispersion relations and simulated spectra for an unbounded cylinder with diameter $2a$=300 nm for a cylinder with an unvarying dielectric constant $\epsilon = 11.7$. For the azimuthal quantum numbers $m = 0, \pm 1, \pm 2$, dispersion relations are displayed in the leftmost column, with the m=0 TE modes plotted as dashed blue lines and TM modes plotted as solid red lines, and with the EH and HE modes of the $m = \pm 1, \pm 2$ dispersion relations plotted, respectively, as purple dashed lines and solid orange lines. The grey lines intersect these curves to pick out the modes that are imaged for a 60 keV beam (the shallower slope) and a 300 keV beam (the steeper slope). The contributions to EEL spectra from modes with $m = 0, \pm 1, \pm 2$ are shown on the same rows as their dispersion relations for a 60 keV beam (center) and a 300 keV beam (right), and the summed contributions are shown on the in the top row, where it can be seen that the TM_{0n} and HE_{1n} contributions dominate.

tum numbers, imaged by beams of two different energies.[2] These beam energies are reflected in the dispersion relation plots as steep (300 keV) and shallow (60 keV) grey lines. Intersections of the grey lines with the dispersion curves indicate where the $k_z = \omega_{\ell mn}/v$ condition is fulfilled. These plots show why modes appear at lower energies for a higher energy beam. An intercomparison of modes between the 60 keV and 300 keV plots also shows how qualitative appearences of modes with the same m and n labels (e.g., the TM_{01} and HE_{11} modes for each case—the lowest-energy modes contributing to the spectra for $m = 0$ and $m = 1$, respectively) remain fairly similar, despite the slight wavelength broadening for patterns at lower energies. Notice that the summed contributions from $\pm m =$ lead to azimuthal symmetry in the measurements, even though the modes themselves vary azimuthally.

3.4.2 Damped Leaky Modes

As sets of solutions to the Helmholtz equation in cylindrical coordinates, the waveguide modes listed above are not, of course, unique. If one lifts the restriction that the arguments of the Bessel funtions are real-valued, for instance, one could, following Arnbak (1969), write the fields inside the cylinder as

$$E_z,\, H_z \propto J_{|m|}(q_{\text{in}} r) \tag{3.45}$$

and the fields outside the cylinder in terms of Hankel functions

$$E_z, H_z \propto H^{(1)}_{|m|}(q_{\text{out}} r), \tag{3.46}$$

where, as before, ω is determined by boundary conditions, and q_{in} and q_{out} are

$$q_{\text{in}} = \sqrt{\epsilon_{\text{in}}(\omega)\omega^2 - k_\ell^2}, \qquad q_{\text{out}} = \sqrt{\omega^2 - k_\ell^2}. \tag{3.47}$$

[2]The modal contributions have been broadened with $\eta = 0.1$ eV to build up spectra from summed Lorentzians, as in Eq. 2.67.

This form is consistent with the version presented above. Comparing Eq. 3.30 and Eq. 3.47, it is clear that $q_{\text{in}} = k_r^<$ and, when ϵ is real for $r < a$, and $q_{in} = ik_r^>$ for $r > a$. Outside the surface of the cylinder, $H_n^{(1)}(ix) \propto K_n(x)$, and the eigenvalue equation from this expansion

$$\left(\frac{1}{q_{\text{in}}}\frac{J'_{|m|}(q_{\text{in}}a)}{J_{|m|}(q_{\text{in}}a)} + \frac{1}{q_{\text{out}}}\frac{H^{(1)'}_{|m|}(q_{\text{out}}a)}{H^{(1)}_{|m|}(q_{\text{out}}a)}\right)\left(\frac{\epsilon}{q_{\text{in}}}\frac{J'_{|m|}(q_{\text{in}}a)}{J_{|m|}(q_{\text{in}}a)} + \frac{1}{q_{\text{out}}}\frac{H^{(1)'}_{|m|}(q_{\text{out}}a)}{H^{(1)}_{|m|}(q_{\text{out}}a)}\right) = \frac{m^2k_z^2c^2}{a^2\omega^2}\left(\frac{1}{q_{\text{in}}^2} + \frac{1}{q_{\text{out}}^2}\right)^2, \tag{3.48}$$

is equivalent to Eq. 3.39 above, again in the case that ϵ is real.

And what would such an alternative expansion give us? The primary reason in the literature for representing the external parts of the waveguide modes as Hankel functions is to makes clear that these can be "leaky" modes—that is, modes whose amplitudes die off exponentially in their direction of propagation, as the light leaks out. A compelling physical picture of such modes was developed in early papers for the case of a constant dielectric function (Snyder and Mitchell (1974); Sammut and Snyder (1976)). In this picture, the propagation constant k_z in $\exp(ik_z z)$ is allowed to be imaginary, implying a mode whose amplitude becomes arbitrarily large in the direction of its source, and arbitrarly small along the direction of its propagation. Radially, the amplitudes of such modes have a turning point outside the surface of the cylinder, reaching a minimum, then becoming arbitrarily large as r increases. This counterintuitive description has a straightforward physical interpretation. Within the cylinder, the EM wave propagates forward almost like a guided mode—except that part of it leaks out. For a wave traveling in the $+z$ direction inside the cylinder, the leaky part of the wave continues to propagate in the $+z$ direction, but also outward. Signal retardation leads to the area outside the cylinder where the amplitude decreases, and to the arbitrarily large growth as r increases from radiative contributions of the far-off $-z$-direction.

In the dispersion relations analogous to those shown in Fig. 3.6, such modes would be represented as passing through the $\omega = k_z c$ limiting line, with frequencies above those allowed by the particular k_z for guided modes. Such modes have been extensively treated for the slab waveguide, in Marcuse (1991) and Hu and Menyuk (2009). In the case of the dielectric cylinder, it would be helpful to be able to use a similar treatment to incorporate a complex dielectric function. This approach has been taken by Mansuripur *et al.* (2016), who calculate complex frequencies in place of the complex propagation constants for the case where the analytic form of $\epsilon(\omega)$ is known. In this context, Cruz y Cruz and Rosas-Ortiz (2015) have made an explicit connection between classical leaky modes and quantum mechanical resonances, but the meaning of such complex frequencies becomes unclear if the dielectric function is taken from data, for which frequencies can only be real.

In fact, many mathematical details for the cylindrical waveguide remain unclear. Poladian (2005) claims that leaky modes can be orthonormalized, by using a form of the radial integral that excludes the infinite contributions as $r \to \infty$. Yet beyond this, the questions of what particular branch cuts should be used in the square root function and the Hankel functions remains unclear. I have implemented Eq. 3.48 in various versions and have found the resultant mode frequencies to depend noticably on these choices. Yang and Song (2016) give the clearest existing analysis of these mathematical issues, but only deals with a the $m = 0$ case. Given these limitations, in the analysis of experiments given in the next section, I have limited my treatment to the approximation of a real dielectric function, despite its obvious shortcomings.

3.4.3 Numerical Results for Silicon Discs

In this section, I use the data on silicon cylinders collected and first reported by Flauraud and Alexander (2019). Using electron lithography, Flauraud and Alexander

fabricated amorphous silicon discs of varying diameters (100 nm, 200 nm, 300 nm, 400 nm, 500 nm) and collected STEM-EELS maps from them. A monochromated 300 keV beam was used (100-110 meV FWHM), and the discs were estimated to be 100 nm thick. The plotted data in this section is theirs, and has not been corrected except to normalize the spectrum at each spatial data pixel.

It is not obvious which dielectric function will be best to use for the simulations of the amorphous sillicon discs. In Fig. 3.7, two possible approximations are shown, alongside radially averaged experimental spectra, plotted as an idealized linescan across the disc for the disc whose diameter is 300 nm. The approximation of the middle row is that of a flat dielectric dependence $\epsilon = 13.5$, matching the low-frequency dielectric constant of amorphous silicon. The approximation of the bottom row is that of $\epsilon(\omega) = \mathrm{Re}[\epsilon(\omega)]$ for amorphous silicon, which will overestimate frequencies where the $\mathrm{Im}[\epsilon(\omega)]$ is large, though not as markedly as when the rise in $\mathrm{Re}[\epsilon(\omega)]$ is ignored. Though using only the real part of a complex dielectric function partially invalidates the normalization scheme, this seems preferable to ignoring the dispersion entirely. The approximation $\epsilon \approx \mathrm{Re}[\epsilon(\omega)]$ has been used in all remaining figures in this chapter.

The linescans of Fig. 3.7 for the disc with a 300 nm diameter reveal modes whose qualitative properties appear similar to those of the waveguide modes. However, plotting E_z for the two lowest-energy waveguide modes (HE_{11} and TM_{01}), reveals that the modes excited in the unbounded cylinder are not fully contained within the 100 nm thickness of the discs used experimentally. The finite thickness of the experimental discs also lifts the strict equality $k_z = \omega_{\ell mn}/v$ required for the unbounded cylinder. Indeed, looking at the plot of experimental data, it appears that the HE_{11}-type modes may be excited as a band, and not as a single discrete mode.

Still, plotting the four most prominent modes in the 300 nm disc as data maps, as has been done in Fig. 3.8 without radial averaging, shows that the qualitative patterns

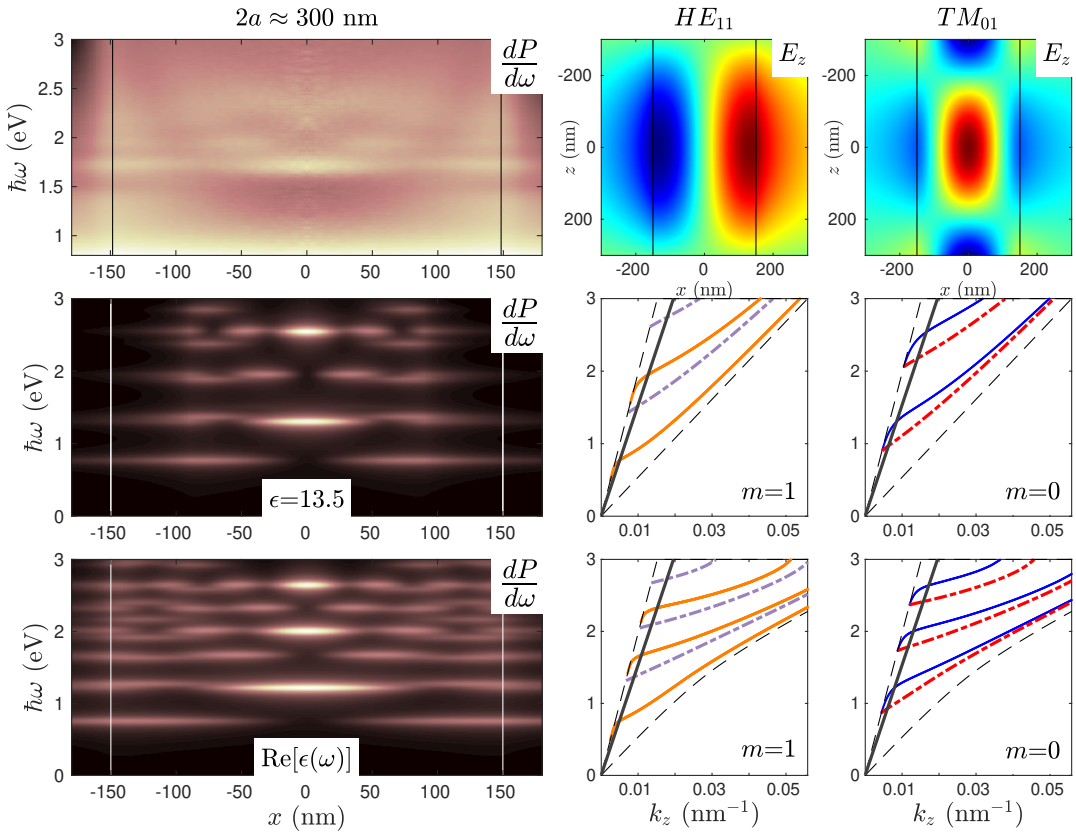

Figure 3.7: First comparison of theory and experiment. Top left shows the radially averaged (but otherwise uncorrected) spectral data from measurements on a silicon disc with $2a \approx 300$ nm, left/right reflected to represent an idealized linescan across the disc. Theoretical comparisons with the unbounded cylinder model are shown (middle row) for a flat dielectric function $\epsilon = 13.5$ matching $\epsilon(0)$ for amorphous silicon, and (bottom row) for a varying dielectric function whose real part matches the measured dielectric properties of amorphous silicon shown in Fig. 3.1, and whose imaginary part is artificially set to zero . Dispersion relations for $m = 0, \pm 1$ are shown for each, and follow the description of Fig. 3.6. On the top row, the z-component of the electric field is plotted in cross section for the two lowest-frequency modes.

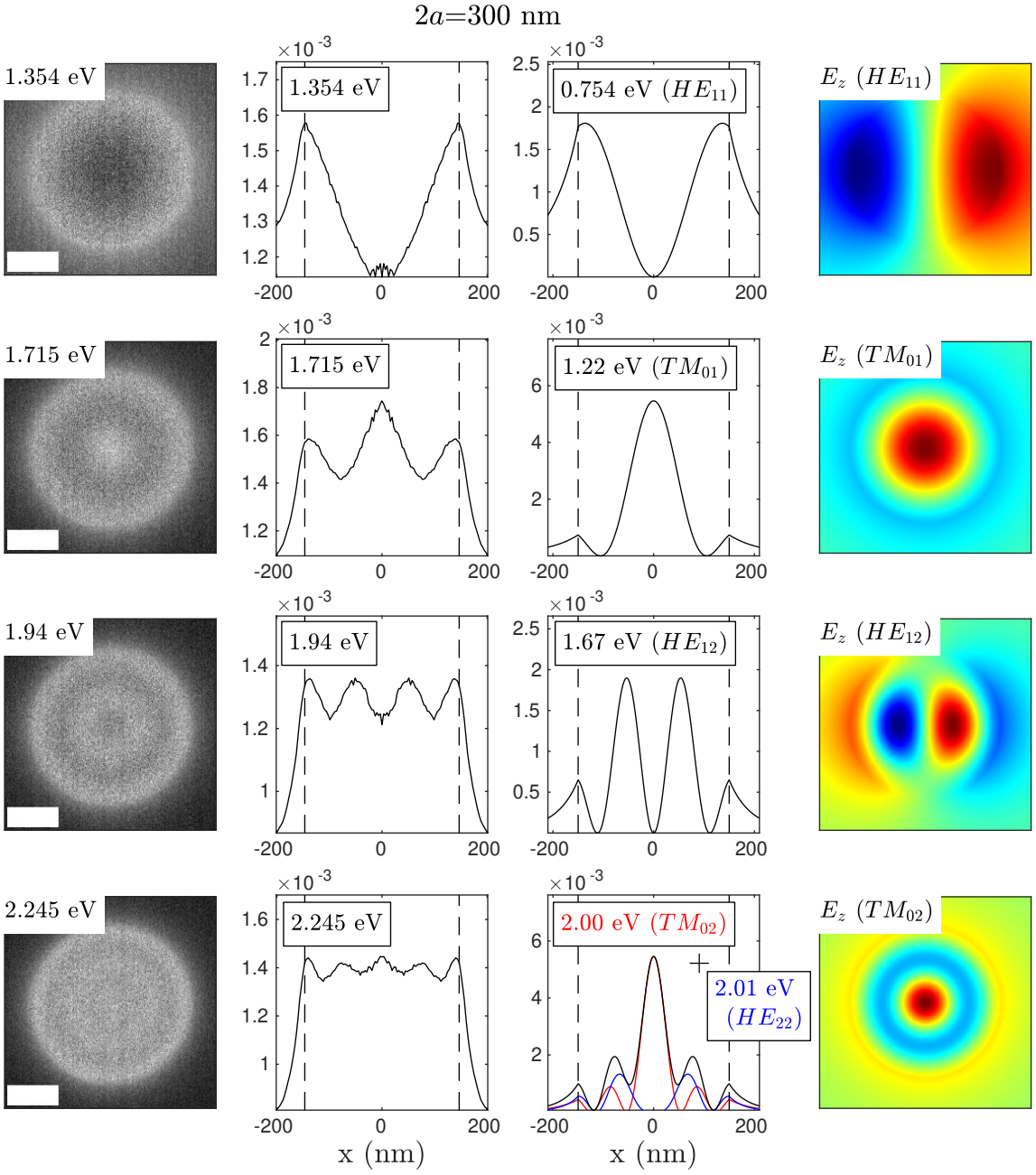

Figure 3.8: Low-frequency modes for the d=300 nm disc. Each row has a spectral map integrated over 0.1 eV from the unprocessed experimental data (far left), a reflected radial average of the experimental data (center left), a radial linescan of the predicted intensity from a theory mode (center right), and an in-plane map of E_z for the theory mode (far right). Dashed lines the center-right and center-left columns represent the cylinder edges, and scale bars on spectral maps represent 100 nm.

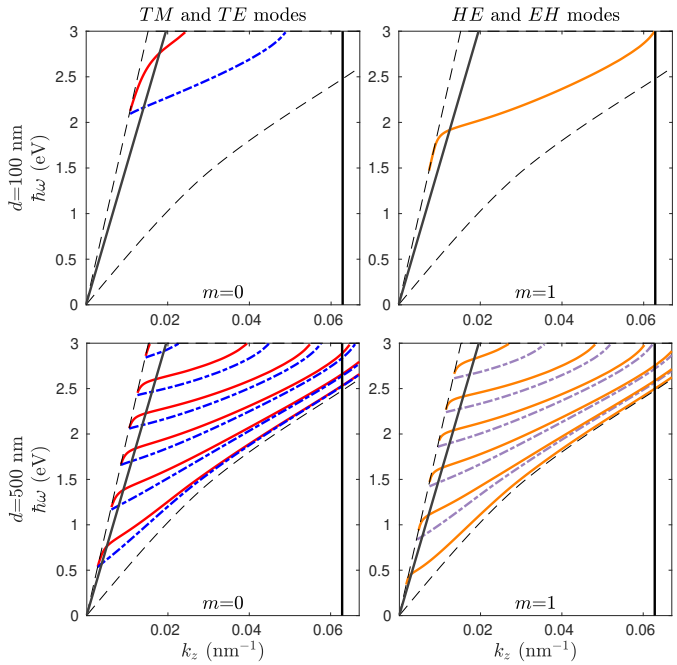

Figure 3.9: Dispersion relations for discs of diameter 100 nm (top) and 500 nm (bottom), for azimuthal quantum numbers $m = 0$ (right) and $m = \pm 1$ (left), with a color scheme matching that of Fig. 3.6. The sloped grey line on each subplot intersects modes that are excited by a 300 keV beam in the unbounded cylinder approximation, giving a low-energy limit for each associated mode, and the vertical black line intersects modes that are excited in the "thin disc" approximation for disc thickness T=100 nm, picking out the mode for which $k_z = 2\pi/(100\text{ nm})$.

observed in the data for low-order modes match the patterns expected from the unbounded cylinder quite well, although, as we might expect, mode frequenciies from the unbounded cylinder uniformly underestimate corresponding mode frequencies in the disc. But for both the experimental case and its theoretical counterpart, it may not be possible to separate contributions of the various modes in the spectrum at higher energies. The last row of the figure illustrates how contributions of nearly degenerate states can overlap, as the summed contributions of the TM_{02}- and HE_{22}-type modes in the waveguide model (bottom row, center left) compares more favorably with the experimentally observed map variations than either mode would on its own.

Energies of the lowest-order modes are sufficiently separated in the EEL spectra

that we can track them across discs of various sizes. Since experimental disc thicknesses must be finite, mode energies measured for fabricated discs should be expected to exceed those predicted by the waveguide model for corresponding modes. As an upper bound on mode energies, I have used the "thin disc" approximation of Zaretskaya *et al.* (2018), which looks for modes that have standing waves in z at $k_z = 2\pi/T$, where T is the disc thickness. These two limits are illustrated in Fig. 3.9. Notice that this upper bound is undefined for the $m = 0$ modes when $T = 100$ nm.

Fig. 3.10 shows that the same low-order modes appear in each of the discs, and that the waveguide and "thin disc" approximations correctly predict upper and lower bounds, respectively, for the HE_{11}-type and TM_{01}-type modes. In the bottom row, the "lower bound" is plotted in red, and the "upper bound" is plotted in blue. The experimental frequencies best match those predicted by the waveguide model for the discs whose diameter is smallest, which may be expected, as the $d = 100$ nm disc is the only one whose thickness is comparable to its diameter, and which may fairly be said to resemble the waveguide geometry.

3.5 Summary

In this chapter, I reviewed the theory of low-loss STEM-EELS beyond the electrostatic approximation. After a discussion of formal requirements of dielectric models, I derived the loss-per-distance spectrum in the electrostatic approximation. This led to a discussion of how this approach could be extended to include electrodynamic effects via the Hertz vector potential. Simulations for crystalline silicon nanoribbons in the "classical" approach indicate that it should be possible to observe spatial variation for TE and TM waveguide modes in crystalline silicon when the ribbon width is $d \sim 100$ nm, while ribbons with $d \ll 100$ nm have a sub-wavelength width, and ribbons with $d \gg 100$ nm approach bulk loss.

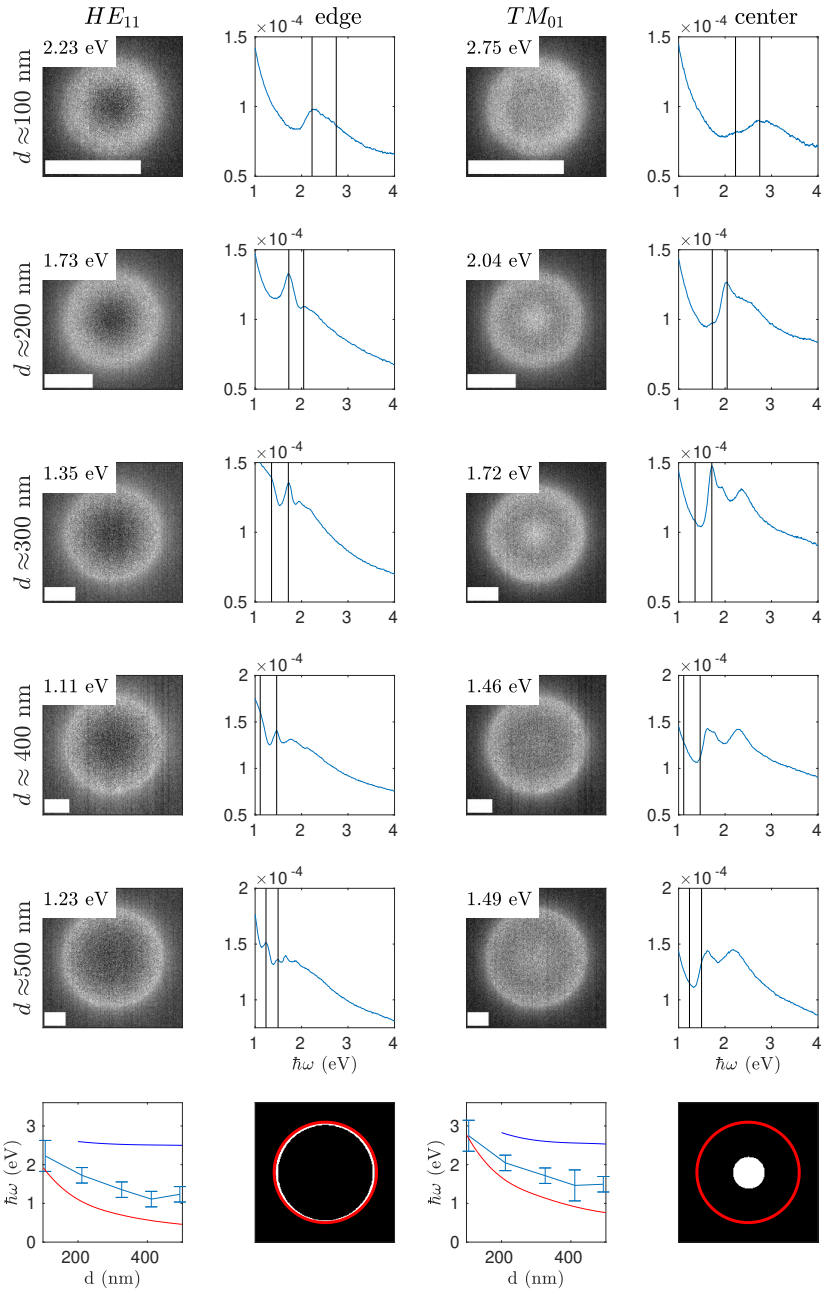

Figure 3.10: For five disc diameters, we show the HE_{11}-type mode map (far left); averaged spectra taken from the disc edge (center left); the TM_{01}-type mode map (center right); and averaged spectra taken from the disc center (far right). In the bottom row, the experimental frequencies are shown (with error bars to represent integration widths) that run between the theoretical predictions for the infinite rod and for the thin disc. Scale bars on mode maps represent 100 nm.

Simulations for amorphous silicon cylinders in the "quantum" approach capture key low-order modes appearing in STEM-EELS measurements on discs of amorphous silicon. The lowest-energy mode (resembling the HE_{11} waveguide mode, with the appearance of a propagating dipole mode) and the next-to-lowest-energy mode (resembling the TM_{01} waveguide mode, with azimuthal symmetry and an E_z maximum along the cylindrical axis) were tracked across discs of different diameters. For fabricated discs of thickness T, the unbounded cylinder model gives a lower bound on the experimental mode energy, and the "thin disc" approximation, which for which $k_z = 2\pi/T$, gives an upper bound on the experimental mode energy.

Chapter 4

PLASMON-ENHANCED SIGNALS FROM ADSORBATES

4.1 Introduction

This chapter returns to the electrostatic approach of Ch. 2, and presents a "classical" model for how nanoparticle shape effects might be harnessed to enhance the vibrational EELS signals from molecules adsorbed at the surface of a nanoparticle. The analogy here is with tip-enhanced Raman spectroscopy (TERS), where large signal enhancements can be achieved by bringing a sharp metallic tip near the sample (Verma (2017)). In that case, as in the EELS case, the molecular surroundings enhance the local electric field relative to the electric field from the probe, and the signal enhancement can be related to the electric field enhancement at the molecular adsorbtion site (Etchegoin and Le Ru (2010)). But in the EELS case, the intrinsically high spatial resolution may provide advantages over TERS, and it is this promise that ultimately motivates why such effects merit study.

In electromagnetic enhancement models, the reason for the shape-based field enhancement effect is essentially that of a lightning rod—or, for our modeling, of a prolate spheroid. Fig. 4.1 shows a simplified version of this mechanism, where the potential of a static electron in the presence of metallic particle is shown. A point charge in free space (top left) has a spherically symmetric potential; a few equipotential lines have been been plotted. When a metallic sphere is introduced (top right), the potential inside the sphere is constant (i.e., $\mathbf{E}_{\text{in}} \approx \mathbf{0}$), and the equipotential lines are deformed to produce an increased field at the surface of the sphere. When the metallic particle, keeping the volume constant, is increasingly deformed, the equipo-

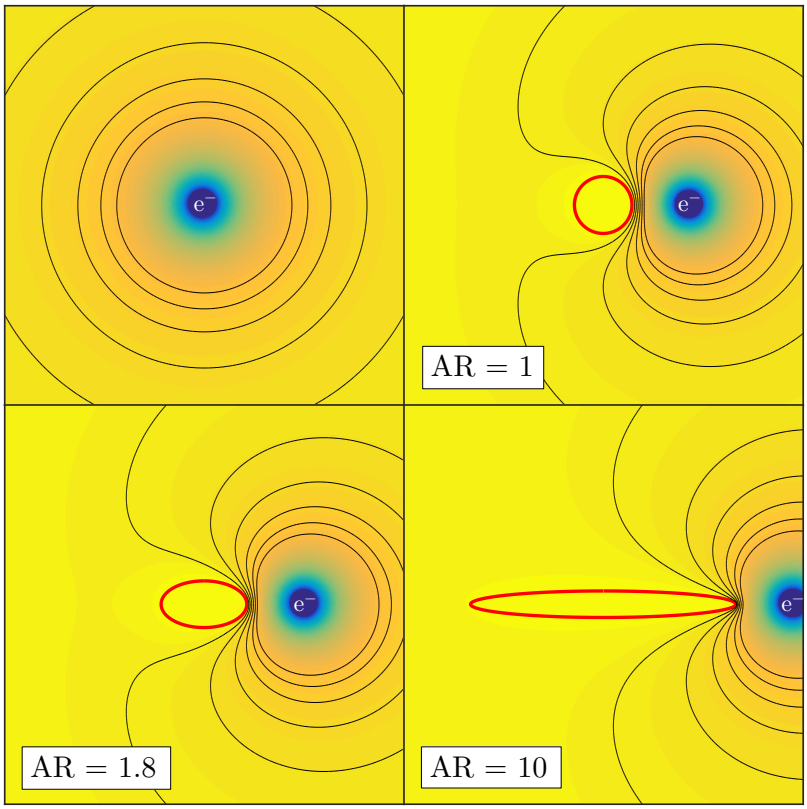

Figure 4.1: Electric potential and selected electric field lines for a stationary electron outside a metallic prolate spheroid. Clockwise from top left: bare electron; electron outside sphere; electron outside prolate spheroid with an aspect ratio=length/width of 1.8; and electron outside prolate spheroid with an aspect ratio of 10.

tential lines are increasingly bunched, leading to an increasingly strong electric field near the particle tip for particle aspect ratios of 1.8 (bottom left) and 10 (bottom right).

To translate this effect to the situation where a molecule is on the tip of a metallic particle and an electron is whizzing swiftly past, I introduce a classical model for EELS signals in the STEM from an adsorbed molecule (represented as a point dipole) coupled to the plasmonic excitations of a metallic nanoparticle (represented as a Drude-type dielectric). For a molecular frequency well separated from the particle's plasmonic frequencies the model predicts weaker enhancement with a symmetric spectral line shape, whereas for a resonance condition the model predicts stronger

enhancement which can exhibit Fano-type asymmetric line shapes due to multimodal contributions. The model predicts that the molecular signal enhancement is proportional to the square of the electric field at the adsorption site, and hence that the enhancement is greatest for a molecule adsorbed on a small nanoparticle with a sharp tip.

Such enhancement could directly increase the sensitivity of molecular vibrational STEM-EELS, and could provide a powerful tool for characterizing surface-functionalized nanoparticles and nanomaterials used for chemical sensing. The results presented here are in good agreement with the work of Konečná *et al.* (2018), who came to the same conclusion using the dielectric formalism to describe the STEM-EELS of a thin molecular layer on metallic nanorods. Our model predicts enhancement factors of several hundred times using silver nanoparticles with moderate aspect ratios, and of several thousand times by tuning the aspect ratio to a resonance condition.

In the time since the model in this chapter was first presented in Kordahl and Dwyer (2019), other investigators have experimentally realized systems that instantiate similar effects to those that our model describes. For instance, Smith *et al.* (2019) reports STEM-EELS observations of a Fano-type antiresonance in a gold disc-rod dimer, whose theoretical description is very similar to the one developed here. Of even greater relevance is the recent work of Tizei *et al.* (2020), which demonstrates a proof of principle for plamon enhancement of vibrational effects, not for adsorbates, but for thin h-BN flakes with silver nanowires resting on top.

This chapter is organized as follows. Sec. 4.2 provides a derivation which solves for the EEL spectrum for a molecule adsorbed on the surface of a dielectric nanoparticle. Sec. 4.3 presents detailed calculations of enhancement, with a discussion of the resulting Fano-type line shapes (Sec. 4.3.1). Adsorbed molecular signals for prolate spheroids are presented and discussed, including our method of quantifying signal

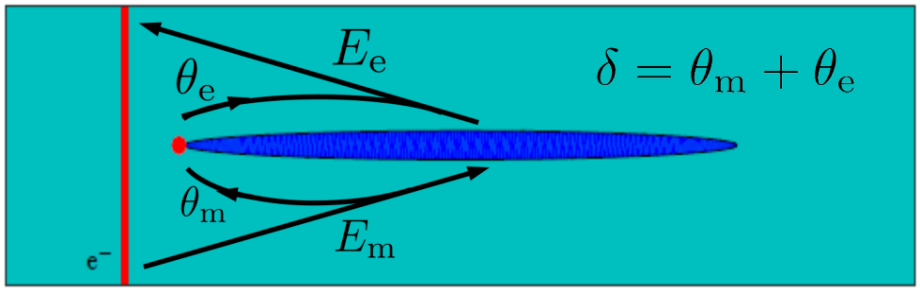

Figure 4.2: Route to molecular signal enhancement: the electron beam induces an electric field E_m from the particle (blue ellipse), which excites the molecule (red dot); then the molecule alters the particle's induced electric field E_e, which changes the electron beam's energy loss. The phase shift δ of the enhanced signal relative to that of the electron beam (a combination of the phase shifts θ_m from E_m and θ_e from E_e) strongly influences the molecule's observed contribution to the EEL spectrum.

enhancement (Sec. 4.3.2), and results for off-resonance (Sec. 4.3.3) and on-resonance (Sec. 4.3.4) enhancement.

4.2 Dielectric Nanoparticle with a Surface Molecule

In this section I obtain concise expressions for the expected STEM-EELS signal from an aloof electron passing a point dipole near the surface of a dielectric nanoparticle. The dipole models an adsorbed molecule, and the nanoparticle surface is defined by a single coordinate ξ_1 in an orthogonal separable coordinate system—e.g., in spherical, oblate spheroidal, or prolate spheroidal coordinates. The development here is redundant to that of Sec. 2.5, as our result includes that earlier result for bare nanoparticles as one part of the solution. I also reuse the notation for harmonic functions ϕ_h introduced above, assuming continuity between ϕ_h^in and ϕ_h^out at the nanoparticle surface, and I reuse the notation where $\mathbf{u}_h = \nabla\phi_h$.

As in the discussion in Sec. 2.5, this calculation begins with an expansion of the Coulomb potential in terms of the harmonic functions. For two points $\mathbf{r}_<$ and $\mathbf{r}_>$ with radial components $\xi_{1<} < \xi_{1>}$, the Coulomb potential can be written as

$$\frac{1}{|\mathbf{r}_> - \mathbf{r}_<|} = \sum_h C_h \bar{\phi}_h^{\text{in}}(\mathbf{r}_<)\phi_h^{\text{out}}(\mathbf{r}_>). \tag{4.1}$$

where the C_h are real coefficients given by Eq. 2.59. After a Fourier transform, this leads to an expression for the beam potential for a beam electron following a straight path along the optic axis with $z = vt$ and transverse coordinate $\mathbf{x_0}$, as

$$\Phi_{\text{e}}(\mathbf{r}, \omega) = \sum_h -eC_h I_h(\mathbf{x_0}, \omega)\phi_h^{\text{in}}(\mathbf{r}), \tag{4.2}$$

with $I_h(\mathbf{x}, \omega)$ being the projection integral of the harmonic potential defined by

$$I_h(\mathbf{x}, \omega) = \int_{-\infty}^{\infty} \frac{dz}{v}\, \bar{\phi}_h^{\text{out}}(\mathbf{r})e^{i\omega z/v}. \tag{4.3}$$

Eqs. 4.1, 4.2, and 4.3 are just reiterations of Eqs. 2.56, 2.61, and 2.62 given above.

If we suppose that the dipole $\mathbf{p}$ (the molecule) points along the nanoparticle surface normal such that

$$\mathbf{p}(\omega) = p(\omega)\hat{\mathbf{n}}, \tag{4.4}$$

then the dipole potential

$$\Phi_{\text{m}}(\mathbf{r}, \omega) = \mathbf{p}(\omega) \cdot \nabla_{\text{m}} \frac{1}{|\mathbf{r} - \mathbf{r}_{\text{m}}|} \tag{4.5}$$

can also be expanded using C_h:

$$\begin{aligned} \Phi_{\text{m}}^{\text{in}}(\mathbf{r}, \omega) &= \sum_h C_h p(\omega)\hat{\mathbf{n}} \cdot \bar{\mathbf{u}}_h^{\text{out}}(\mathbf{r}_{\text{m}})\phi_h^{\text{in}}(\mathbf{r}), \\ \Phi_{\text{m}}^{\text{out}}(\mathbf{r}, \omega) &= \sum_h C_h p(\omega)\hat{\mathbf{n}} \cdot \bar{\mathbf{u}}_h^{\text{in}}(\mathbf{r}_{\text{m}})\phi_h^{\text{out}}(\mathbf{r}). \end{aligned} \tag{4.6}$$

Using these expressions, the electric potential outside and inside the dielectric nanoparticle is expressed in terms of unknown coefficients $A_h(\omega)$ and $B_h(\omega)$:

$$\begin{aligned} \Phi_{\text{out}}(\mathbf{r}, \omega) &= \Phi_{\text{e}}(\mathbf{r}, \omega) + \Phi_{\text{m}}^{\text{in/out}}(\mathbf{r}, \omega) + \sum_h A_h(\omega)\phi_h^{\text{out}}(\mathbf{r}), \\ \Phi_{\text{in}}(\mathbf{r}, \omega) &= \sum_h B_h(\omega)\phi_h^{\text{in}}(\mathbf{r}). \end{aligned} \tag{4.7}$$

Notice that in applying the electrostatic boundary conditions $\Phi_{\mathrm{m}}^{\mathrm{in}}(\mathbf{r},\omega)$ is used for the molecular contribution to $\Phi_{\mathrm{out}}(\mathbf{r},\omega)$ since the dipole is positioned strictly on the outside of the nanoparticle's surface. Applying the boundary conditions leads to

$$A_h(\omega) = C_h \alpha_h(\omega) \left(e I_h(\mathbf{x_0},\omega) - p(\omega) \hat{\mathbf{n}} \cdot \bar{\mathbf{u}}_h^{\mathrm{out}}(\mathbf{r}_{\mathrm{m}}) \right), \tag{4.8}$$

reusing the nanoparticle polarizability $\alpha_h(\omega)$ (Eq. 2.64, Ferrell *et al.* (1987))

$$\alpha_h(\omega) = \frac{\epsilon(\omega) - 1}{\epsilon(\omega) - \epsilon_h}, \tag{4.9}$$

and the harmonic values of the dielectric function ϵ_h (Eq. 2.53)

$$\epsilon_h = \left. \frac{\mathbf{u}_h^{\mathrm{out}} \cdot \hat{\mathbf{n}}}{\mathbf{u}_h^{\mathrm{in}} \cdot \hat{\mathbf{n}}} \right|_{\mathrm{surface}}. \tag{4.10}$$

We suppose the molecular dipole $p(\omega)$ has a charge q, reduced mass μ, damping factor γ, and bare resonant frequency ω_{m}. If the dipole is driven by the electric field $\mathbf{E}(\omega)$ along $\hat{\mathbf{n}}$, its equation of motion

$$\frac{q^2}{\mu} \hat{\mathbf{n}} \cdot \mathbf{E}(\omega) = \left(\omega_{\mathrm{m}}^2 - \omega(\omega + 2i\gamma) \right) p(\omega) \tag{4.11}$$

allows us to find $p(\omega)$ in terms of two ancillary functions, $E_{\mathrm{m}}(\omega)$ and $\Delta\omega_{\mathrm{m}}^2(\omega)$:

$$p(\omega) = \frac{q^2}{\mu} \frac{E_{\mathrm{m}}(\omega)}{\omega_{\mathrm{m}}^2 - \Delta\omega_{\mathrm{m}}^2(\omega) - \omega(\omega + 2i\gamma)}. \tag{4.12}$$

The first of these ancillary functions

$$E_{\mathrm{m}}(\omega) = \sum_h e C_h I_h(\mathbf{x_0},\omega) \left(\mathbf{u}_h^{\mathrm{in}}(\mathbf{r}_{\mathrm{m}}) - \alpha_h(\omega) \mathbf{u}_h^{\mathrm{out}}(\mathbf{r}_{\mathrm{m}}) \right) \cdot \hat{\mathbf{n}}, \tag{4.13}$$

is the electric field along $\hat{\mathbf{n}}$ driving the molecular dipole (m). The $\mathbf{u}_h^{\mathrm{in}}$ terms encode the field contributions of the electron beam, and the $\mathbf{u}_h^{\mathrm{out}}$ terms encode the

contributions of the nanoparticle's induced field. For our model parameters in Section 4.3, at the adsorption site the induced field from the nanoparticle will be much larger than the direct field from the electron beam.

The function in the denominator of $p(\omega)$

$$\Delta\omega_{\mathrm{m}}^2(\omega) = \frac{q^2}{\mu}\sum_h C_h\alpha_h(\omega)\left|\mathbf{u}_h^{\mathrm{out}}(\mathbf{r}_{\mathrm{m}})\cdot\hat{\mathbf{n}}\right|^2 \tag{4.14}$$

arises from the electrostatic interaction of the dipole with its image charge, shifting the natural resonance ω_{m} to a lower frequency. Physically, given a molecular dipole $p(\omega)\hat{\mathbf{n}}$ outside the dielectric particle surface, we would expect an additional field $\mathbf{E}_{\mathrm{m}}^{\mathrm{m}}(\mathbf{r}_{\mathrm{m}})\cdot\hat{\mathbf{n}}$ induced by the molecule in the particle to act back on the molecule along its dipole axis. This field is

$$\mathbf{E}_{\mathrm{m}}^{\mathrm{m}}\cdot\hat{\mathbf{n}} = -p(\omega)\Delta\omega_{\mathrm{m}}^2(\omega). \tag{4.15}$$

Hence the frequency shift is proportional to the electric field per unit dipole at the molecular site induced by the molecule itself, a sort of self-interaction.

Now we can find the total potential $\Phi(\mathbf{r},\omega)$ that acts back on the beam electron. Given our solutions for $A_h(\omega)$ and $p(\omega)$, the only subtlety here is that we use the alternative dipole expansion, since the beam electron passes outside the molecule on the nanoparticle surface:

$$\Phi(\mathbf{r},\omega) = \Phi_{\mathrm{m}}^{\mathrm{out}}(\mathbf{r},\omega) + \sum_h A_h(\omega)\phi_h^{\mathrm{out}}(\mathbf{r}). \tag{4.16}$$

Next we find the work done on the beam electron by this potential, and by switching the integration order of ω and z, we extract the EEL spectrum:

$$\frac{\mathrm{d}P}{d\omega} = \frac{e}{\pi\hbar}\mathrm{Im}\left(\int_{-\infty}^{\infty}\frac{\mathrm{d}z}{v}\Phi(\mathbf{r},\omega)e^{-i\omega z/v}\right). \tag{4.17}$$

In the end, we find that the EEL spectrum splits neatly into two parts, one that matches that of the bare nanoparticle (without the adsorbed molecule), and one that captures how the molecule alters the spectrum:

$$\underbrace{\frac{\mathrm{d}P}{\mathrm{d}\omega}}_{\text{full}} = \underbrace{\frac{\mathrm{d}P_0}{\mathrm{d}\omega}}_{\text{bare nanoparticle}} + \underbrace{\frac{\mathrm{d}P_\mathrm{m}}{\mathrm{d}\omega}}_{\text{molecule}} . \tag{4.18}$$

The bare nanoparticle spectrum has already been quoted in Eq. 2.63

$$\frac{\mathrm{d}P_0}{\mathrm{d}\omega} = \frac{e^2}{\pi\hbar} \sum_h C_h \left| I_h(\mathbf{x_0}, \omega) \right|^2 \mathrm{Im}\left(\alpha_h(\omega)\right). \tag{4.19}$$

The molecular contribution to the spectrum depends on the dipole amplitude $p(\omega)$. It can be written as

$$\frac{\mathrm{d}P_\mathrm{m}}{\mathrm{d}\omega} = \frac{1}{\mu}\frac{q^2}{\pi\hbar}\mathrm{Im}\left(\frac{E_\mathrm{m}(\omega)E_\mathrm{e}(\omega)}{\omega_\mathrm{m}^2 - \Delta\omega_\mathrm{m}^2(\omega) - \omega(\omega + 2i\gamma)}\right), \tag{4.20}$$

where $E_\mathrm{m}(\omega)$ and $\Delta\omega_\mathrm{m}^2(\omega)$ are given by Eqns. (4.13) and (4.14), respectively, and $E_\mathrm{e}(\omega)$ is given by a function similar to but distinct from $E_\mathrm{m}(\omega)$:

$$E_\mathrm{e}(\omega) = \sum_h eC_h\bar{I}_h(\mathbf{x_0}, \omega)\left(\bar{\mathbf{u}}_h^\mathrm{in}(\mathbf{r}_\mathrm{m}) - \alpha_h(\omega)\bar{\mathbf{u}}_h^\mathrm{out}(\mathbf{r}_\mathrm{m})\right)\cdot\hat{\mathbf{n}}. \tag{4.21}$$

The function $E_\mathrm{e}(\omega)$ arises from the electric field that acts back on the passing electron (e), with the $\bar{\mathbf{u}}_h^\mathrm{in}$ terms encoding the direct field from the molecule, and the $\bar{\mathbf{u}}_h^\mathrm{out}$ terms encoding the field from the particle induced by the vibrating molecule. As with $E_\mathrm{m}(\omega)$, for our model parameters in Section 4.3 the induced field from the nanoparticle tends to dominate the direct field from the molecule.

Signal enhancement thus roughly takes the following route: the electron beam induces an electric field in the nanoparticle, which drives the molecular dipole, which

induces a response electric field in the nanoparticle, which then acts back on the electron beam. This process is illustrated schematically in Fig. 4.2 above, which indicates also that δ, the combined phase shift of the molecular signal relative to the beam potential, resulting from the summed phase shifts from $E_{\rm m}$ and $E_{\rm e}$, will have important effects in the observed line shapes for the adsorbate signal. These issues are further analyzed in Sec. 4.3.1.

4.3 Molecular Adsorbate Signals

This section presents detailed simulations of the enhancement of the EELS signal from an adsorbed molecule on the surface of a prolate-spheroidal Drude-metal nanoparticle. As we have seen in Ch. 2 (e.g., in Fig. 2.6) the mode frequencies of the prolate spheroids exhibit large redshifts with increasing eccentricity. For a Drude metal considered in this section the behavior will be similar, but with stretched high- and low-frequency limits $\omega_\ell \to \omega_{\rm p}/\sqrt{\epsilon_\infty}$ and $\omega_0 \to 0$. Thus the surface plasmonic excitations of prolate spheroidal metallic nanoparticles with large cross-sectional eccentricities are shifted into the infrared, and hence can be used to enhance molecular vibrations. In some cases, the nanoparticle excitations can be "tuned" to a resonance condition, giving a large degree of enhancement, as shown below. Generally, we expect that a molecule positioned at site of higher curvature on the nanoparticle surface to experience a larger electric field, and thus a larger degree of EELS enhancement.[1] Hence we consider molecules adsorbed on the tips of prolate spheroids.

[1]Some have been tempted to overinterpret the role of surface curvature in situations involving metallic surfaces. Price and Crowley (1985) show that while it is sometimes claimed that an isolated charged conductor will maximize the electric field strength wherever its surface curvature is maximized (an idea they call the "lightning-rod fallacy"), it is possible to construct geometries where the two do not coincide. On the other hand, Liu (1987) derives explicit expressions relating the surface charge to the surface curvature for charged metallic ellipsoids, so it is not unreasonable to expect

(Oblate spheroids yield similar but less dramatic trends, and are not presented here.)

We assume a 60 keV beam, but now we assume it has a finite energy width described by a Gaussian distribution with a full-width at half-maximum (FWHM) of 5 meV. The nanoparticle's Drude parameters are taken to be those of silver (Yang *et al.* (2015)): $\hbar\omega_{\rm p} = 8.9$ eV, $\hbar\eta = 0.10$ eV, $\epsilon_\infty = 5$. The molecule is defined via its charge q, reduced mass μ, resonant frequency $\omega_{\rm m}$, and damping parameter γ; we have used $q = e$, $\mu = 1$ amu, $\hbar\omega_{\rm m} = 530$ meV, and $\hbar\gamma = 12$ meV. This large natural damping γ compensates for another omission in our calculations.

For an adsorbed molecule realistically near the nanoparticle surface, the predicted frequency shift from $\Delta\omega_{\rm m}^2(\omega)$ (Eqn. (4.14)) is unrealistically large. One reason for this overestimation is that a point dipole (representing the molecule in our model) can interact with the arbitrarily fine spatial variations of the arbitrarily high-order harmonic modes, and if the molecule is allowed to approach the surface, this shift diverges. A model similar to ours (Ford and Weber (1983)) included finite-size effects but also overestimated the frequency shifts observed in IR studies.[2] To address this problem, we have used a "physicist's solution," and have set $\Delta\omega_{\rm m}^2(\omega)$ to *zero*. But this breaks self-consistency, so we need to introduce a large γ to maintain positivity for all spectra. The effect of setting $\Delta\omega_{\rm m}^2(\omega)$ to zero should be a systematic underestimate of the molecular signal (since signal is inversely related to frequency), and a system-

such relationships to hold approximately, for the physical reasons illustrated in Fig. 4.1.

[2]Shifts in the vibrational frequencies of adsorbate molecules, relative to corresponding frequencies of molecules in gas phase, tend to be less than 5%, and have been studied as a function of such variables as electric field orientation dependence (Cohen de Lara (1999)) and surface coverage (Kho *et al.* (2012)). But the questions involved are complex. For chemisorbed molecules, frequency shifts can be either positive or negative, depending on the specifics of the surface bonding. I have not found any studies of frequency shifts studied as a function of particle shape.

atic overestimate of the molecular signal-to-noise (since this neglects any additional particle-mediated peak broadening). The function $\Delta\omega_{\mathrm{m}}^2(\omega)$ captures the interaction of the molecular dipole with its own image dipole, so its neglect also eliminates the ability of more sharply curved absorption sites to encourage larger molecular resonance redshifts. I have included $\Delta\omega_{\mathrm{m}}^2(\omega)$ in the discussion of line shapes below, but it should be understood the such contributions are not included in the numerical results.

4.3.1 Off- vs. On-resonance – Fano Line Shapes

Since the molecular damping γ is small compared to the dielectric particle damping η, the phase of the molecular vibration varies quickly relative to phase of the nanoparticle's induced electric field. As shown below, when a resonance condition is approached—that is, when a nanoparticle mode frequency approaches the molecular frequency—this typically leads to a strong Fano-type asymmetric spectral signature in the molecular spectrum $\mathrm{d}P_{\mathrm{m}}/\mathrm{d}\omega$. (Limonov *et al.* (2017) points out that this is a general feature of coupled systems.) If the molecular frequency is far separated from any of the nanoparticle's harmonic frequencies, then $\mathrm{d}P_{\mathrm{m}}/\mathrm{d}\omega$ reverts to a weaker symmetric bump poking above the bare nanoparticle signal.

To see this, we might assume that $\Delta\omega_{\mathrm{m}}^2(\omega)$ is slowly varying, such that it can be absorbed into the other parameters of the model, writing

$$\begin{aligned} \omega_{\mathrm{m}}'^2 &= \omega_{\mathrm{m}}^2 - \mathrm{Re}[\Delta\omega_{\mathrm{m}}^2(\omega_{\mathrm{m}})], \\ \gamma' &= \gamma + \mathrm{Im}[\Delta\omega_{\mathrm{m}}^2(\omega_{\mathrm{m}})]/2\omega_{\mathrm{m}} \end{aligned} \tag{4.22}$$

such that the molecular contribution to the spectrum (Eq. 4.20) becomes

$$\frac{\mathrm{d}P_{\mathrm{m}}}{\mathrm{d}\omega} \approx \frac{1}{\mu}\frac{q^2}{\pi\hbar}\mathrm{Im}\left(\frac{E_{\mathrm{m}}^{\mathrm{e}}(\omega)E_{\mathrm{e}}^{\mathrm{m}}(\omega)}{(\omega_{\mathrm{m}}')^2 - \omega(\omega + 2i\gamma')}\right). \tag{4.23}$$

The terms $E_{\mathrm{m}}^{\mathrm{e}}(\omega)$ and $E_{\mathrm{e}}^{\mathrm{m}}(\omega)$ in Eq. 4.23 both contain information about the phase as well as the magnitude of the induced fields, but to understand their importance we

also need to appreciate the importance of the molecular resonance frequency in the denominator. This term is easier to think about in terms of the familiar Lorentzian distribution, so we can carry out the approximation that allows for such a replacement, throwing away small terms, and terms that are small near ω'_{m}:

$$\begin{aligned}\frac{1}{\omega'^2_{\mathrm{m}} - \omega^2 - 2i\gamma'\omega} &= \frac{-1}{(\omega + i\gamma' - \omega'_{\mathrm{m}})(\omega + i\gamma' + \omega'_{\mathrm{m}}) + \cancel{\gamma'^2}} \\ &\approx -\frac{1}{2\omega'_{\mathrm{m}}}\left(\frac{1}{\omega + i\gamma' - \omega'_{\mathrm{m}}} - \cancel{\frac{1}{\omega + i\gamma' + \omega'_{\mathrm{m}}}}\right) \\ &\approx -\frac{1}{2\omega'_{\mathrm{m}}}\left(\frac{(\omega - \omega'_{\mathrm{m}}) - i\gamma'}{(\omega - \omega'_{\mathrm{m}})^2 + \gamma'^2}\right) \\ &= \frac{1}{2\omega'_{\mathrm{m}}\gamma'}\left(\frac{i - \Omega}{1 + \Omega^2}\right).\end{aligned} \tag{4.24}$$

In the last line, a new dimensionless quantity $\Omega = (\omega - \omega'_{\mathrm{m}})/\gamma'$ has been introduced which gives us our position in the spectrum relative to the resonance, in terms of the resonance width. The magnitude of this term is small except around ω'_{m}, and the numerator forces a phase shift of π across the resonance.

Near the molecular resonance the fields $E_{\mathrm{m}}(\omega)$ and $E_{\mathrm{e}}(\omega)$ (Eqns. 4.13 and 4.21) can be estimated respectively by $|E_{\mathrm{m}}(\omega'_{\mathrm{m}})|\exp(i\theta_{\mathrm{m}})$ and $|E_{\mathrm{e}}(\omega'_{\mathrm{m}})|\exp(i\theta_{\mathrm{e}})$. Substituting these, and the expansion for the denominator, into Eq. 4.20 yields an approximate expression for the molecular contribution to the spectrum

$$\frac{\mathrm{d}P_{\mathrm{m}}}{\mathrm{d}\omega} \approx \frac{1}{\mu}\frac{q^2}{\pi\hbar}\frac{|E_{\mathrm{m}}(\omega'_{\mathrm{m}})E_{\mathrm{e}}(\omega'_{\mathrm{m}})|}{2\omega'_{\mathrm{m}}\gamma'}\left(\frac{\cos(\delta) - \Omega\sin(\delta)}{1 + \Omega^2}\right), \tag{4.25}$$

where $\delta = \theta_{\mathrm{m}} + \theta_{\mathrm{e}}$. In Eq. 4.25 the values $\delta = 0$ and π correspond to positive- and negative-symmetric line shapes, respectively, whereas $\delta = \pi/2$ and $3\pi/2$ correspond to asymmetric line shapes. For our model parameters, the fields $E_{\mathrm{m}}(\omega)$ and $E_{\mathrm{e}}(\omega)$, and hence the phase δ, are dominated by the nanoparticle response.

Well below the nanoparticle frequencies (off-resonance), the nanoparticle responds "in phase" with the Fourier-decomposed field of the electron beam. Hence $\delta = 0+0 =$

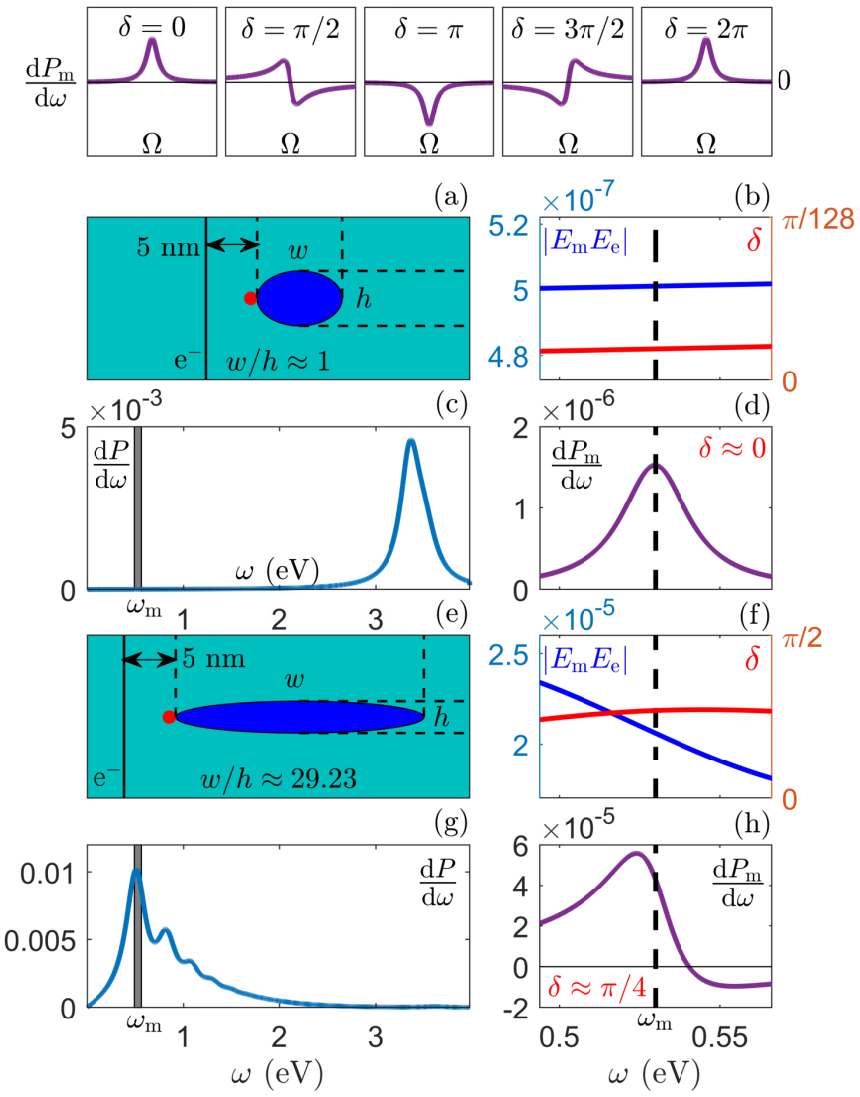

Figure 4.3: Relating enhanced molecular line shapes to δ, the cumulative phase shift from E_m and E_e (top row). In (a-d) the molecule adsorbed on a silver sphere gives an off-resonance condition (c) leading to $\delta \approx 0$ and hence a symmetric line shape (d). In (e-h) a long prolate spheroid gives an on-resonance condition (g) leading to $\delta \approx \pi/4$ and hence an asymmetric line shape (h). The grey strips in (c) and (g) indicate the plotting range used in (b), (d), (f), and (h). The red dots indicate the molecules in the cartoons (a) and (e), which are not to scale. The sphere and spheroid both have a central thickness $h = 50$ nm.

0, corresponding to a positive-symmetric line shape. At a nanoparticle frequency ω_h (on-resonance), the phase of the response due to mode h is $\pi/2$. Thus, in a simple picture consisting of just the single nanoparticle mode h, at resonance $\delta = \pi/2 + \pi/2 = \pi$, corresponding to a negative-symmetric line shape. The molecular spectrum $\mathrm{d}P_\mathrm{m}/\mathrm{d}\omega$ sits atop the bare nanoparticle spectrum $\mathrm{d}P_0/\mathrm{d}\omega$, so the condition that $\mathrm{d}P_\mathrm{m}/\mathrm{d}\omega < 0$ implies that the field causing the beam electron's energy loss is weaker in the presence of the molecule. But when the contributions of many nanoparticle modes are included, δ can deviate significantly from the "simple resonance value" of π. Hence at a resonance condition we can obtain Fano-type line shapes with significant asymmetry, which ultimately arises from multimodal contributions.

Fig. 4.3 compares line shapes for the molecule adsorbed on the surface of a sphere and a prolate spheroid. The sphere, with its low eccentricity, gives an off-resonance condition with $\delta \approx 0$, leading to a symmetric line shape. At the resonance condition $\omega_\mathrm{m} \approx \omega_{10}$ (i.e., the molecular frequency matches the harmonic frequency associated with $\ell = 1$, $m = 0$), the phase $\delta \approx \pi/4$ ($\neq \pi$ due to multimodal contributions), leading to a Fano-type asymmetric line shape. These spectral features are similar to those predicted in the recent work of Konečná *et al.* (2018), who modeled the spectral response of a thin dielectric layer, representing a molecular layer, on the tip of a metallic nanorod, using finite element analysis.

4.3.2 Quantifying Signal Enhancement

The intensity of the enhanced molecular signal is most easily characterized in the off-resonance limit that $\delta = 0$ and the line shape is symmetric. Analytically, if we assume that the electric field $\mathbf{E}^\mathrm{out}(\mathbf{r}_\mathrm{m}, \omega)$ varies slowly in ω near the bare molecular frequency ω'_m, and that the molecular damping factor $\gamma' \ll \omega'_\mathrm{m}$, the total enhanced signal obtained from integrating over the resonance is

$$P_{\mathrm{m}} \approx \frac{q^2}{2\hbar\omega'_{\mathrm{m}}\mu}|\mathbf{E}^{\mathrm{out}}(\mathbf{r}_{\mathrm{m}},\omega'_{\mathrm{m}})\cdot\hat{\mathbf{n}}|^2, \tag{4.26}$$

where, as before, γ' and ω'_{m} are the shifted parameters. This is the same as for a bare molecule, except the field $|\mathbf{E}^{\mathrm{out}}(\mathbf{r}_{\mathrm{m}},\omega'_{\mathrm{m}})\cdot\hat{\mathbf{n}}|^2$ now contains contributions from the dielectric particle, and not just from the beam electron.

Hence the enhanced signal is proportional to the square of the field at the molecular site. Intuitively, this is squared because one factor of the electric field arises from the molecular excitation, and one factor arises from the molecule acting back on the nanoparticle. This contrasts with the case of surface-enhanced Raman spectroscopy (SERS), in which the signal enhancement scales roughly as the electric field enhancement to the fourth power: two powers for photoabsorption, and two for photoemission, as discussed by Ru and Etchegoin (2006).

Fig. 4.4 illustrates how the signal has been quantified, using the same physical situations as in Fig. 4.3. The full EEL spectrum $dP/d\omega$ (plotted in blue) is dominated by the nanoparticle, but the presence of the molecule causes the spectrum to depart from the bare nanoparticle spectrum $dP_0/d\omega$. This departure constitutes the enhanced molecular spectrum $dP_{\mathrm{m}}/d\omega$ (plotted in purple). Compared to the bare molecular spectrum $dP^0_{\mathrm{m}}/d\omega$ (plotted in red), for which the nanoparticle is absent, the enhanced molecular spectrum exhibits greater intensity, since the adsorbed molecule is coupled to the nanoparticle, whose response to the adsorbed molecule ultimately causes the signal enhancement.

Fig. 4.4 defines three quantities: the background signal from the nanoparticle P_0, the enhanced molecular signal P_{m}, and the bare molecular signal P^0_{m}. We then define

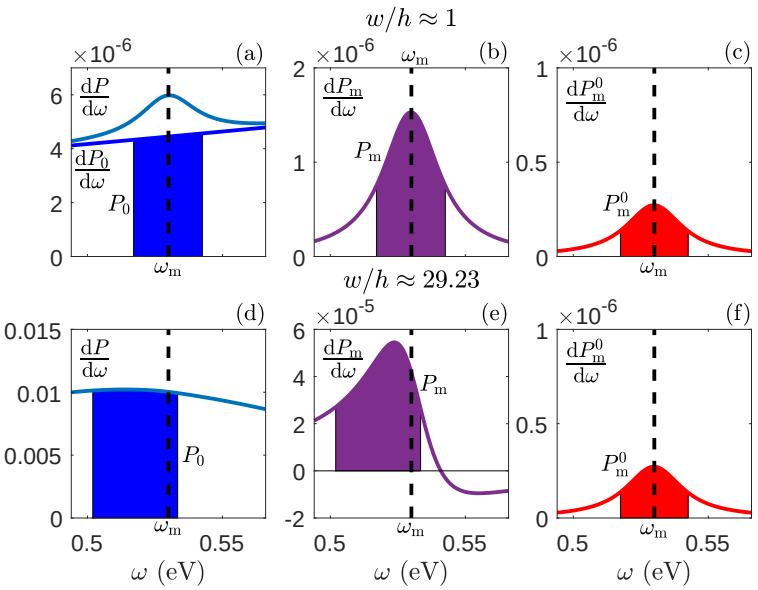

Figure 4.4: Quantifying signal enhancement from prolate spheroids with significantly different aspect ratios (w/h). Plots (a-c) show signals for the sphere of Fig. 4.3, and plots (d-f) show signals for the prolate spheroid of Fig. 4.3. The molecular signal P_{m} in (b) and (e) is estimated by finding the maximal feature in the enhanced spectrum and integrating the FWHM. The bare-nanoparticle signal P_0 in (a) and (d) is found by integrating the same spectral region for the background. The bare molecular signal P_{m}^0 in (c) and (f) is equal for both cases, as it only depends on the beam distance from the molecule (5 nm throughout).

$$\text{signal enhancement} = \frac{P_{\mathrm{m}}}{P_{\mathrm{m}}^0}. \tag{4.27}$$

The signal enhancement is defined relative to the bare molecular signal. The widths of the peaks in $dP_{\mathrm{m}}/d\omega$ and $dP_{\mathrm{m}}^0/d\omega$ are determined both by the particle damping γ and the beam's energy width. For each, we integrate over the FWHM so that the signal enhancement is insensitive to the beam's energy width.

The enhanced molecular spectrum sits atop the spectrum from the nanoparticle, so we also quantify the enhancement of the signal-to-noise ratio (SNR) of the molecular signal, assuming Poisson noise only:

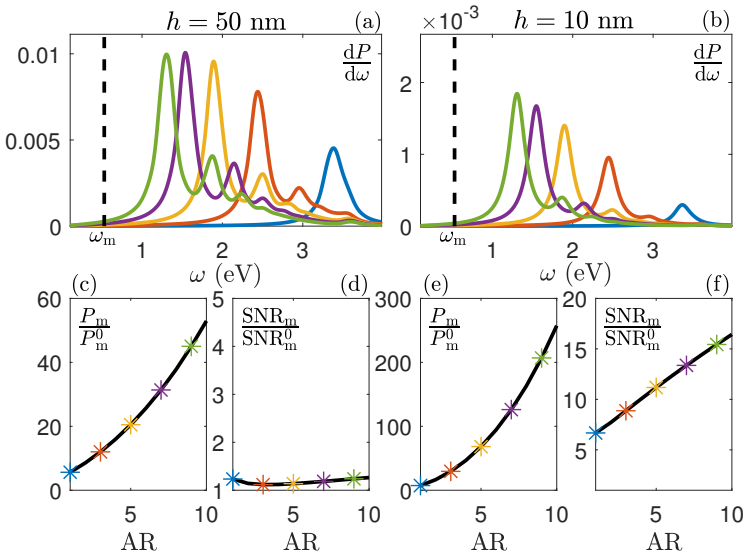

Figure 4.5: Full EEL spectra (a, b), signal enhancements (c, e), and SNR enhancements (d, f) for adsorbed moleules on the tip of prolate spheroids of aspect ratios (AR $= w/h$) between 1 and 10 (specific ARs are color coded). Cases of constant thickness are shown, both with $h = 50$ nm (left) $h = 10$ nm (right).

$$\begin{aligned} \text{SNR enhancement} &= \frac{\text{SNR}_\text{m}}{\text{SNR}_\text{m}^0} \\ &\approx \frac{P_\text{m}/\sqrt{P_0}}{P_\text{m}^0/\sqrt{P_\text{m}^0}} = \frac{P_\text{m}}{\sqrt{P_0 P_\text{m}^0}}. \end{aligned} \tag{4.28}$$

The SNR enhancement is defined relative to the SNR_m^0 of the bare molecular signal. The SNR enhancement depends on the background signal P_0 which is proportional to the integration interval, scaling inversely with the square root of the width of the molecular peak.

4.3.3 Off-resonance Enhancement

Fig. 4.5 shows how the signal enhancement and the SNR enhancement vary as the prolate spheroid's aspect ratio (AR = width/height) changes from 1 to 10. Both "thick" ($h = 50$ nm) and "thin" ($h = 10$ nm) prolate spheroids are considered. All of these are off-resonance cases where the nanoparticle mode frequencies lie well

above the molecule's resonance frequency ω_{m}. For AR = 10, we find the greatest signal enhancements of about 50 in the thick case, versus about 250 in the thin case. The greater enhancement for the thin nanoparticles comes from the sharper tip, and hence the stronger electric field at the adsorption site. The results here also give us our first confirmation that small, sharp nanoparticles will help the most with signal enhancement. Along with its sharper tip, a small nanoparticle contributes a lesser background signal, allowing the molecular signal to stand out. For AR = 10, the calculations in Fig. 4.5 predict a SNR enhancement barely above unity for the thick nanoparticle, versus about 16 for the thin nanoparticle. Still, without larger particle aspect ratios, adsorbate signals are unlikely to be detected.

Fig. 4.6 explores how the SNR enhancement varies more generally. First consider the red curves, which use the same nanoparticle/beam geometry as above. We investigate for a fixed thickness $h = 10$ nm how the molecular signal, the signal enhancement, and the SNR enhancement vary with nanoparticle length, specified by w, keeping the beam distance from the adsorbate constant at 5 nm. The signal enhancement increases with increasing nanoparticle length, which is readily understood in terms of the local electric field, as explained above. In the colormap figures, we notice that the SNR enhancement is more closely aligned with the thickness h of the nanoparticle than with the aspect ratio.

The blue and grey curves in Fig. 4.6 investigate another possibility, as raised by Konečná *et al.* (2018). They reported the possibility of "ultraremote sensing," with the beam on the opposite side of the nanoparticle as the adsorbate, as shown in our cartoon by beam B. While the simulations of Konečná et al. differ from ours—they use a thin dielectric layer on the nanoparticle surface rather than a point dipole—our results are consistent. At far beam distances, the intrinsically small signal from the bare molecule allows for enormous signal and SNR enhancements. But when

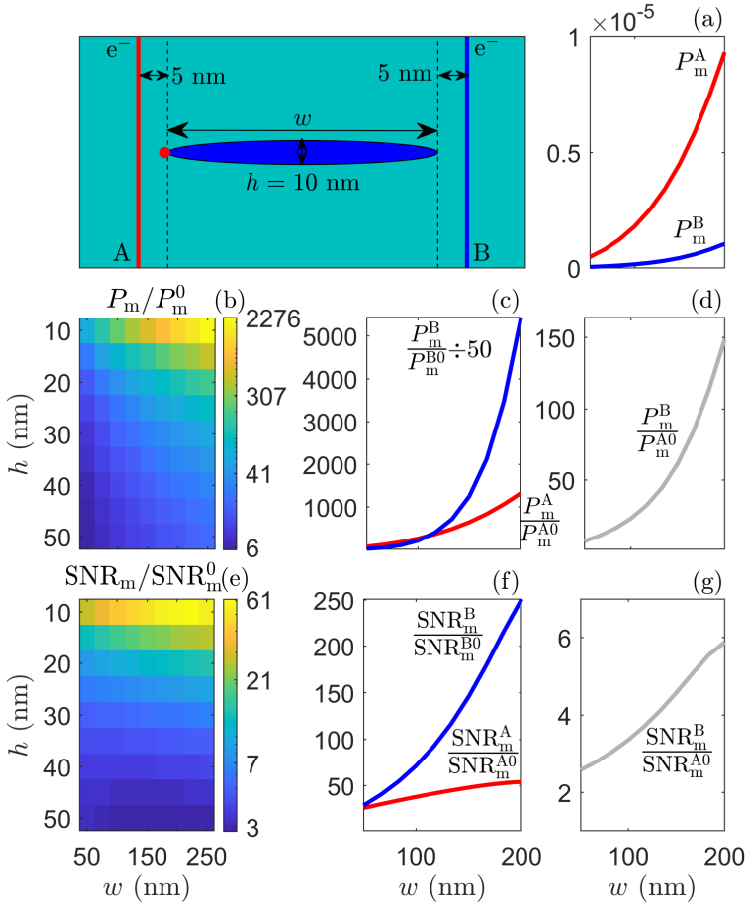

Figure 4.6: Results are given for beam positions adjacent to the molecule (A) and opposite to the molecule (B), with the beam 5 nm from the particle in each case. The colormap plots (b) and (e) summarize many different cases for the adjacent beam (A). Plots (a), (c), and (f) contrast outcomes from beams A and B in terms of the definitions given in the main text, while (d) and (g) compare the results of beam B to the benchmarks used for beam A.

the "ultraremote enhancement" is defined in terms of the expected signal for case A, as in the grey curves in Fig. 4.6, we find the small relative enhancements for the ultraremote beam B that one might expect.

4.3.4 On-resonance Enhancement

Pushing toward higher aspect ratios, Fig. 4.7 shows how particle resonances can be harnessed to improve enhancements further still. An obvious possibility is to

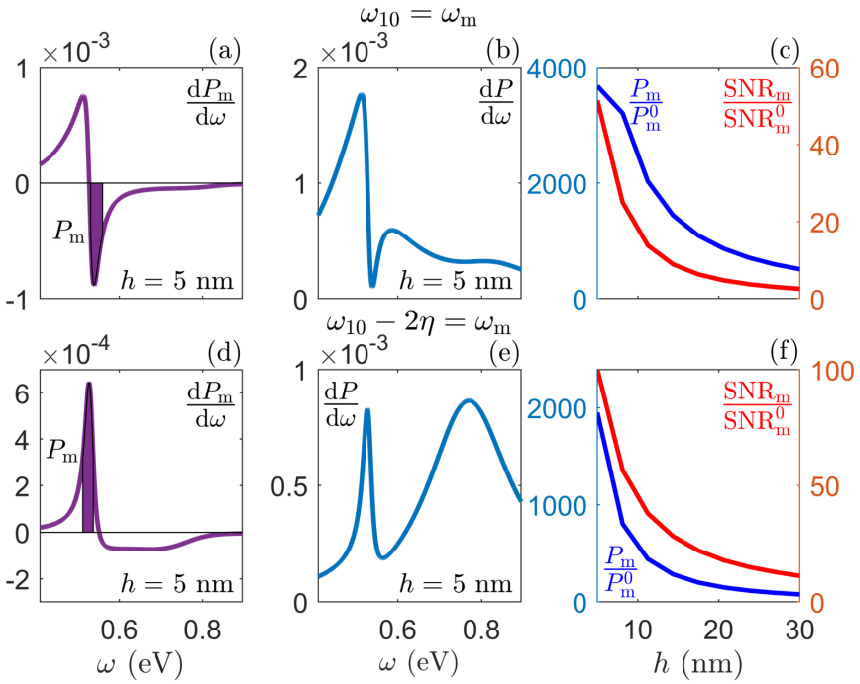

Figure 4.7: Calculations for adsorbates on prolate spheroids with fixed AR = 29.23 (a-c, with ω_{10} matching ω_m), and AR = 18.32 (d-f, with $\omega_{10} - 2\eta$ matching ω_m). For thickness $h = 5$ nm, (a) and (d) show the molecular spectra, while (b) and (e) show the total spectra. In (c) and (f), we quantify the signal enhancement (in blue) and the SNR enhancement (in red) for the two cases.

tune the aspect ratio of the particle such that the lowest particle resonance ω_{10} (that is, the resonance corresponding to mode $\ell = 1$, $m = 0$) aligns with the molecular resonance ω_m. This leads to enormous signal enhancements, up to several thousand times, with the signal manifesting as a Fano-type dip into rather than a bump atop the nanoparticle background. (Recall that the enhancement effect comes about as a change to the signal from the particle, so the fact that the "molecular signal" is negative here is not inconsistent, though it is perhaps semantically clumsy.) Such dips are sensitive to the beam energy resolution, and the position of the dip atop a particle resonance limits the achievable SNR enhancement.

One way both to lower the nanoparticle background signal and to regain an asymmetric bump in place of the resonant dip is to tune the particle resonance ω_{10} such

that it lies just above the molecular resonance ω_{m}—2η above, say, such that the molecular resonance lies on the bottom edge of the particle resonance. Fig. 4.7 shows that such particles will produce roughly half the signal enhancement of the particles tuned to overlap the molecular resonance directly, yet such off-resonant tuning is able to achieve somewhat better SNR enhancements, with the SNR enhancement approaching 100 for small particles with a $h = 5$ nm thickness.

4.4 Summary

In this chapter, I extended the Born-Huang theory of molecular vibrations to model the vibrational EEL signal from a molecule adsorbed on the surface of a metallic nanoparticle. The degree of signal enhancement in this model is proportional to the square of the local electric field at the adsorption site. Detailed expressions for the associated energy-loss spectrum split the spectrum into two parts, one from the bare nanoparticle, and the other from the molecule. The bare nanoparticle expression is the same as the expression derived in Ch. 2. Hence the molecular signal sits atop the bare nanoparticle signal, and can add to or subtract from the nanoparticle signal, depending on its phase relative to that of the surface plasmons.

Physically plausible parameters for a molecule adsorbed on the tip of a silver prolate spheroid yielded predicted signal enhancements of several hundred times for nanoparticles of suitably small thickness (e.g., 10 nm) and high aspect ratio (e.g., 10). Signal enhancements of several thousand times were predicted when the nanoparticle's lowest surface plasmon frequency was tuned to the molecular resonance frequency. Near resonance, the resultant spectral line shapes typically exhibit significant Fano-type asymmetry from the contributions of multiple surface plasmon modes.

Such enhancement effects could significantly increase the sensitivity of the STEM-EELS of molecular vibrations. Recent work by Tizei *et al.* (2020) using plasmonic

nanorods suggests that such enhancement effects may be useful for probing the local vibrational properties of dielectric materials. Combined with the simultaneous imaging capabilities of the STEM, such effects could also provide a powerful tool for characterizing surface-functionalized nanoparticles for chemical sensing applications.

Chapter 5

STEM-EELS MAPPING OF POINT DEFECTS

5.1 Introduction

Point defects in bulk materials define a "final frontier" for STEM-EELS imaging. Interfaces (2D defects) and dislocation lines (1D defects) have been well-studied at the atomic level, but the small volumes and weak signals associated with individual point defects (0D defects) makes them intrinsically hard to detect. The last decade has seen progress in the analysis of individual heavy dopant atoms within bulk materials using annular dark field imaging (ADF-STEM), as in the work of Gunawan *et al.* (2011), Mittal and Mkhoyan (2011), Phillips *et al.* (2012), and Johnson *et al.* (2017). Progress has also been reported in imaging lighter dopants or vacancies embedded in lower-dimensional materials by Liu *et al.* (2011) and Senga and Suenaga (2015). But locating lighter dopants or vacancies in 3D materials remains difficult.

Yet such point defects are crucial to the operation of many semiconductors, quantum computing materials, quantum sensing materials, etc. Locating and characterizing them typically relies on optical techniques such as photoluminescence, where the spatial resolution is diffraction-limited by the wavelength of the light, which typically requires their properties to be measured in ensembles. The hope of improving the performance of defect centers, both individually and as ensembles, now motivates the development of new techniques aiming toward higher spatial resolution, a capability that would be valuable to many physicists, materials scientists and chemists.

Of course, this chapter forwards STEM-EELS as a promising possiblity. EELS does not require the atomic number of a point defect to differ greatly from that of

the atoms surrounding it, making it applicable to important points defects such as nitrogen-vacancy (NV) centers in diamond. Using EELS, point defects have already been located at submicron levels, as with the characterization of NV concentrations in nanodiamonds (Chang *et al.* (2016)) and reports of poorly-understood defects in BAlN thin films (Wang *et al.* (2019)). On the other hand, we expect that the EELS signal (or any spectroscopic signal) associated with any *individual* point defect to be very weak. Moreover, apart from the weak signals, a grand challenge in detecting individual point defects via EELS is the degree of spatial delocalization given by v/ω: since it is inversely proportional to the energy loss $\hbar\omega$, the resolution of spatial mapping becomes poor (~ 10 nm) for the excitation energies of interest (~ 1 eV).

The style of calculation in this chapter is quite different from any of those that preceede it, in that it applies first-principles techniques to model physical systems from the ground up. Here density functional theory (DFT) is used to model two point defects in diamond, and transitions from the single-particle DFT wave functions are used to make predictions, from the corresponding STEM-EELS maps for different defect orientations, to the feasible experimental localization of such defects, and to the plane-wave cross-sections for the electronic transition under study. As will be shown, these defects can be intuitively understood using a tight-binding description, and the qualitative and quantitative aspects of this development support each other.

Given the enormous literature on DFT as a method for first-principles modeling, one could fill an entire thesis devloping of the underlying theory, including its strengths and shortcomings—which would inevitably lead through the familiar signposts of Hohenberg and Kohn (1964) and Kohn and Sham (1965), the replacement of a $3N$-dimensional many-body wavefunction with N 3-dimensional single-particle electron wavefunctions, and the development of various functional approximations. This material will not be reiterated here, as I am content to gesture toward the re-

view of Hasnip *et al.* (2014) demonstrating the variety of spectroscopic signals that have been successfully predicted using DFT, and to place this study in that lineage.

The prospects of using STEM-EELS to locate individual point defects in 3D materials are analyzed in the following pages. I consider the negatively-charged nitrogen-vacancy (NV^-) center in diamond, with a zero-phonon transition energy of $\sim$1.9 eV, and the charge-neutral silicon-vacancy (SiV^0) center in diamond, with a zero-phonon transition energy of $\sim$1.3 eV. The NV^- center has been studied for half a century (Doherty *et al.* (2013)), but continued interest is spurred by its relevance as a room-temperature qubit with well-established read-out procedures (Weber *et al.* (2010); Neumann *et al.* (2008)). When the robust spin coherence of SiV^0 is considered alongside advances in production and initialization (Rogers *et al.* (2014); Dhomkar *et al.* (2018)), it also emerges as a strong candidate for quantum information applications (Green *et al.* (2017)). Both defects show promise as single-photon emitters to be used as building blocks in quantum information science, quantum sensing, and metrology.

The results of this chapter were first presented in Kordahl *et al.* (2019). This work was completed in heavy collaboration with my co-authors. It takes the following order. First, I review how STEM-EELS maps can be constructed from DFT wavefunctions (Section 5.2) and describe our computational methods (Section 5.3). Then electronic structures are reviewed and spectral maps are presented for NV^- (Section 5.4) and SiV^0 (Section 5.5). These results are discussed (Section 5.6) before concluding.

5.2 STEM-EELS of Electronic Transistions

The theoretical approach presented in this chapter differs cosmetically from that of earlier chapters, insofar as the transitions from $|0\rangle \to |1_n\rangle$ elided the role of the "vacuum state" $|0\rangle$, while here the single-particle ground state has an obvious role. The Møller potential formalism developed by Dwyer (2005, 2015, 2017) covers both

cases, with the earlier version in some sense a special case of the general theory.

For a defect electronic transition $\varphi_0 \to \varphi_n$, the Møller potential is given by

$$\Phi_n(\mathbf{r}) = -e \int d^3\mathbf{r}' \, \frac{\varphi_n^*(\mathbf{r}')\varphi_0(\mathbf{x}')}{|\mathbf{r} - \mathbf{r}'|}. \tag{5.1}$$

This is similar to the electrostatic $\Phi_h(\mathbf{r})$ of Ch. 2 (Eq. 2.29), but it explicitly includes both the initial and final states. The quantity $-e\varphi_n^*\varphi_0$ is essentailly the charge density that gives rise to this transition potential, and we refer to it as the transition charge:

$$\rho_n(\mathbf{r}) = -e\varphi_n^*(\mathbf{r})\varphi_0(\mathbf{r}). \tag{5.2}$$

The Møller potential describes the instantaneous Coulomb interaction between the beam electron and the transition charge ρ_n. The only relativistic correction in this calculation is in the speed of the electron v, which is found using the relativistic version of the beam kinetic energy, $(1/\sqrt{1 - v^2/c^2} - 1)m_e c^2$.

The projected Møller potential is defined as

$$\Phi_n(\mathbf{x}) = \int_{-\infty}^{\infty} dz \, \Phi_n(\mathbf{r}) \exp(i\omega_n z/v), \tag{5.3}$$

where $\hbar\omega_n$ is the energy loss associated with the transition. The projected Møller potential is conveniently calculated via Fourier space as

$$\Phi_n(\mathbf{x}) = \int \frac{d^2\mathbf{q}}{(2\pi)^2} \tilde{\Phi}_n(\mathbf{q}) \exp(i\mathbf{q} \cdot \mathbf{x}), \tag{5.4}$$

where

$$\tilde{\Phi}_n(\mathbf{q}) = -\frac{4\pi e \tilde{\rho}_n(\mathbf{q})}{\mathbf{q}^2 + \omega_n^2/v^2}. \tag{5.5}$$

Here $\tilde{\rho}_n$ is the Fourier transform of the projected transition charge density

$$\tilde{\rho}_n(\mathbf{q}) = \int d^2\mathbf{x} \, \rho_n(\mathbf{x}) \exp(-i\mathbf{q} \cdot \mathbf{x}), \tag{5.6}$$

with the projected transition charge density itself being defined by

$$\rho_n(\mathbf{x}) = \int_{-\infty}^{\infty} dz\, \rho_n(\mathbf{r}) \exp(i\omega_n z/v). \tag{5.7}$$

As before, we adopt the projection approximation and assume that inelastic scattering is accurately described by a first-order interaction. Under these approximations, the inelastic scattering can be regarded as taking place within a plane coinciding with the point defect and perpendicular to the optic axis $\hat{\mathbf{z}}$. In this plane, the wave function describing beam electrons in the nth inelastic channel is given by

$$\psi_n(\mathbf{x}) = \frac{ie}{\hbar v}\Phi_n(\mathbf{x})\psi_0(\mathbf{x}), \tag{5.8}$$

where v is the speed of the beam electron, $\Phi_n(\mathbf{x})$ is the projected Møller potential, and ψ_0 is the wave function for the elastic channel.

For STEM-EELS, the elastic wave function ψ_0 is that of a focused beam. The strength of the EELS signal corresponds to the probability that beam electrons will be scattered inelastically and admitted by the EELS collection aperture in the diffraction plane. This probability is calculated by integrating the squared modulus of the Fourier transform of the inelastic wave function over those wave vectors admitted by the collection aperture, which reproduces the probability of Eq. 1.7 above. In general, the inelastic wave function will undergo elastic scattering from the atoms in the sample downstream from the defect plane before reaching the diffraction plane. However, for simplicity elastic scattering has been neglected, and all inelastically scattered electrons are presumed to be collected. The relatively large inelastic delocalization length v/ω and the consequent nanometer, rather than atomic-scale, spatial resolution mean that these simplifications do not significantly affect our conclusions in what follows.

5.3 Computational Methods

Our DFT calculations have been performed using the Vienna *Ab initio* Simulation Package (VASP), as documented in Kresse and Hafner (1993), Kresse and Hafner (1994), Kresse and Furthmüller (1996), and Kresse and Furthmüller (1996). We used a $4\times4\times4$ conventional supercell ($4^3\times8-1=511$ atoms) with Γ-point sampling and a 420 eV cutoff. Projector augmented-wave (PAW) pseudopotentials were used (Blöchl (1994); Kresse and Joubert (1999)) under the generalized-gradient approximation (Perdew *et al.* (1996)). Excited states were calculated using an occupation constraint. The pseudized wave functions in this approach match the all-electron wave functions outside the atomic cores where the pseudopotentials are deployed, so we should expect the wave functions for the defect states to accurately describe the defect states as independent electron states, localized as they are in the vacancy regions.

Real-space electron wave functions were constructed on a 3D grid using Fourier coefficients extracted via WaveTrans (Feenstra *et al.* (2013)). These symmetrized wave functions were used to compute the real-space transition charges (Eq. 5.2). Using the 3D transition charge densities, we calculated the projected transition charge densities on a 2D grid (Eq. 5.7), padded the grid as needed for convergence, and Fourier-transformed the results to construct the **q**-space projected Møller potentials (Eq. 5.5). Inverse Fourier transformation yielded the projected Møller potentials in real space (Eq. 5.4). Since the dielectric function for diamond is fairly flat for the frequencies of interest (Zhang *et al.* (2008)), screening was incorporated by dividing the projected Møller potentials by the static dielectric constant of diamond, $\epsilon_0 = 5.7$.

The wave functions output by VASP were initially complex-valued, but were made real-valued by similarity transformations among degenerate states. For non-degenerate wave functions, this only required multiplication by an overall phase. But

for degenerate wave functions, a unitary transformation was required, which corresponds to a similarity transformation between irreducible representations (irreps) of the relevant point group.

For instance, the $e_{x/y}$ wave functions (as in the NV^- case of Fig. 5.2 below) form a representation of the C_{3v} group. C_{3v} has the following as one possible irrep:

$$\begin{aligned} E &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, C_3 = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, C_3^2 = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \\ \sigma_1 &= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \sigma_2 = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \sigma_3 = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}. \end{aligned} \tag{5.9}$$

Physically, E represents no change to the wave function; C_3 and C_3^2 represent $2\pi/3$ and $4\pi/3$ rotations about the defect symmetry axis; and $\sigma_{1,2,3}$ represent reflections acoss the planes defined by the symmetry axis and the carbon nuclei of $C_{1,2,3}$ (see Fig. 5.1). To find the irrep M' implied by a particular pair of twofold degenerate DFT wave functions $|\psi_1'\rangle$ and $|\psi_2'\rangle$ with $e_{x/y}$ symmetry, the matrix elements are

$$M'_{ij} = \langle \psi_i' | \hat{O} | \psi_j' \rangle . \tag{5.10}$$

As a concrete example, if one wishes to calculate C_3', the operator $\hat{O}$ rotates $|\psi_j'\rangle$ by $2\pi/3$ about the symmetry axis before the inner product is made with $|\psi_i'\rangle$.

As an intermediary step to finding real-valued wave functions $|\psi_{1,2}\rangle$ from linear combinations of $|\psi_{1,2}'\rangle$, we find the transformation matrix U that makes a similarity transformation between the representation M of Eq. 5.9 and the representation M':

$$M = U^\dagger M' U. \tag{5.11}$$

General approaches exist for finding similarity transformations between irreps (see Mozrzymas *et al.* (2014)), but here we use a simplified approach that hinges on the

fact that all group elements in M' can be formed from products of σ_1' and C_3'. If these two follow Eq. 5.11, so will the rest.

Since a similarity transformation preserves eigenvalues, σ_1' will have eigenvalues of -1 and +1 with corresponding orthonormal (column) eigenvectors $\hat{\mathbf{e}}_-$ and $\hat{\mathbf{e}}_+$. The 2×2 matrix $U = (\hat{\mathbf{e}}_-,\, \delta\hat{\mathbf{e}}_+)$ will satisfy Eq. 5.11 for σ_1' for any $|\delta| = 1$ by construction. One can then use $C_3 = U^\dagger C_3' U$ to find an appropriate δ. This U will still contain an overall phase ambiguity, but the wave functions $|\psi_{1,2}''\rangle$

$$\begin{pmatrix} |\psi_1''\rangle \\ |\psi_2''\rangle \end{pmatrix} = U \begin{pmatrix} |\psi_1'\rangle \\ |\psi_2'\rangle \end{pmatrix} \tag{5.12}$$

will differ from our desired real-valued wave functions $|\psi_{1,2}\rangle$ at most by an overall phase, which can easily be corrected at the end of the calculation.

This approach is also works for wave functions with $e^g_{x/y}$ and $e^u_{x/y}$ symmetries under a D_{3d} point group, as pertains to the SiV^0 defect. The point group D_{3d} can be written as the direct product $D_{3d} = D_3 \times i$, where i is the inversion (parity) point group. Moreover D_3 is isomorphic to C_{3v} discussed above: the three two-fold rotations about the horizontal axes in D_3 replace the three vertical mirrors in C_{3v}. The product with the inversion point group means that there are twelve rather than six elements in each representation, and that representations differ for the gerade and ungerade cases. But the method itself is the same.

5.4 The NV^- Defect Center

This section contains our results for the NV^- defect center in diamond. This defect consists of a substitutional nitrogen atom immediately adjacent to a carbon vacancy, with the defect carrying an additional electron making it negatively-charged. For definiteness we choose the N-V axis to lie along the [111] crystallographic direction of the (nonprimitive) face-centered cubic diamond unit cell, as in Fig. 5.1.

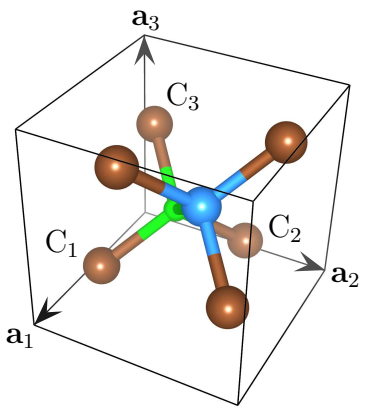

Figure 5.1: The atomic structure of the NV center in diamond, with brown carbon atoms, a blue nitrogen atom, and a green vacancy. Carbon atoms immediately adjacent to the vacancy are labeled $C_{1,2,3}$. This visualization, as with the isosurface plots, has been built using VESTA (Momma and Izumi (2011)).

5.4.1 Electronic Structure of NV$^-$

The electronic structure and excited states of the 1.945 eV transition of the NV$^-$ center have been extensively characterized, from both the directions of molecular-orbital and group theoretic analysis (Lenef and Rand (1996); Manson *et al.* (2006); Maze *et al.* (2011)), and of DFT (Gali *et al.* (2008); Larsson and Delaney (2008); Gali *et al.* (2009)), with recent DFT analyses going so far as to incorporate the vibrational effects (Deng *et al.* (2014); Alkauskas *et al.* (2014); Plakhotnik *et al.* (2015)) and surface effects (Hertkorn *et al.* (2019)). Later we will use DFT wave functions to calculate the EELS scattering probabilities based on the theory outlined in Section 5.2. But first we review the electronic structure of the NV$^-$ defect from a tight-binding perspective, which yields results that are less accurate than those of DFT, but whose results will guide our physical interpretation.

For the tight-binding description, let $c_{1,2,3}$ denote the dangling bond orbitals from the three carbon atoms $C_{1,2,3}$ immediately adjacent to the vacancy, as labeled in Fig. 5.1, and let n denote the dangling bond orbital from the N atom. Assuming no overlap between dangling bond orbitals, we can form the following mutually-

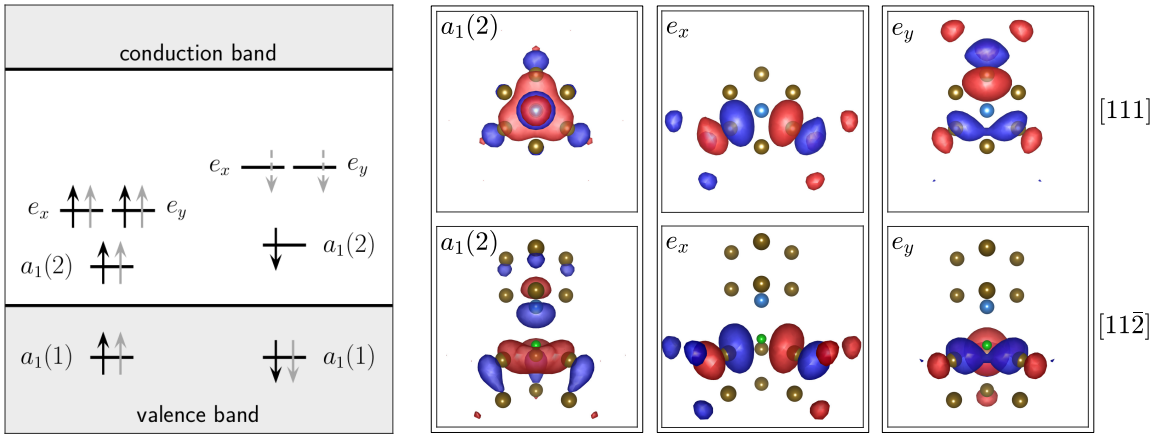

Figure 5.2: *Left:* The DFT (LSDA) predicted ground and excited states of the NV^- defect. Black and gray arrows represent the ground and excited states, respectively, while solid and dashed arrows denote full and half occupancies, respectively. The ground state has $^3A_2$ (triplet) electronic configuration. The optically-allowed excited state is the ^3E (triplet) configuration resulting from single-electron transitions from $a_1^{\downarrow}(2)$ to $e_x^{\downarrow}$ or $e_y^{\downarrow}$, as is discussed by Gali *et al.* (2009). *Right:* Isosurface plots of the wave functions involved in the excitation (red is negative, blue is positive), viewed both along and perpendicular to the [111] N-V axis (carbon atoms are brown, nitrogen atom is blue, vacancy is green).

orthogonal symmetry-adapted orbitals for the C_{3v} point group of the NV^- defect:

$$\begin{pmatrix} a_1(1) \\ a_1(2) \\ e_x \\ e_y \end{pmatrix} = \begin{pmatrix} \alpha/\sqrt{3} & \alpha/\sqrt{3} & \alpha/\sqrt{3} & \beta \\ -\beta/\sqrt{3} & -\beta/\sqrt{3} & -\beta/\sqrt{3} & \alpha \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 & 0 \\ 1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6} & 0 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ n \end{pmatrix}, \tag{5.13}$$

where α denotes the degree of mixing of carbon and nitrogen dangling bonds and $\beta = \sqrt{1-\alpha^2}$ preserves normalization. Each symmetry-adapted orbital can have spin up or spin down. For the NV^- defect, these orbitals are populated by six electrons: three from $C_{1,2,3}$, two from the nitrogen atom, and one excess dopant electron that makes the defect charged.

Fig. 5.2 shows the electronic configurations predicted by DFT (LSDA) for the ground state and for the optically-allowed excited state of interest here. Spin-orbit interactions lift the degeneracy of spin up and spin down orbitals, such that the

ground state is populated by an excess of spin up electrons (similar to Hund's rule from atomic theory). The excited state corresponds to the promotion of an electron from $a_1^{\downarrow}(2)$ to $e_x^{\downarrow}$ or $e_y^{\downarrow}$. The excited state is regarded as containing half-occupied $e_y^{\downarrow}$ and $e_y^{\downarrow}$ orbitals, which preserves the C_{3v} symmetry. The excitation preserves electron spin (i.e., there are no spin flips). The LSDA difference in system energy between the constrained excited state and ground state is 1.708 eV, which is to be compared with the accepted experimental value of 1.945 eV. The discrepancy of 0.23 eV is typical of such DFT predictions, and apart from slightly affecting the predicted energy position of the NV^- spectral signature, it does not cause any substantial inaccuracy in the present work. The relevant DFT wave functions are also shown in Fig. 5.2; they qualitatively resemble the symmetry-adapted orbitals given above.

It is also instructive to use the symmetry-adapted orbitals to construct the transition charges $\bar{e}_x a_1(2)$ and $\bar{e}_y a_1(2)$ and their associated dipole moments. Assuming for simplicity that the bond orbitals are completely separated in space, the transition charges have the following simple approximate forms, obtained from the tight-binding orbitals of Eq. 5.13:

$$\begin{aligned} \bar{e}_x a_1(2) &\approx \frac{\beta}{\sqrt{6}}(-|c_1|^2 + |c_2|^2), \\ \bar{e}_y a_1(2) &\approx \frac{\beta}{\sqrt{18}}(-|c_1|^2 - |c_2|^2 + 2|c_3|^2). \end{aligned} \tag{5.14}$$

For convenience, we introduce $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ perpendicular to the [111] N-V axis and $\hat{\mathbf{z}}$ parallel to it: $\hat{\mathbf{x}} = \frac{1}{\sqrt{2}}[\bar{1}, 1, 0]$, $\hat{\mathbf{y}} = \frac{1}{\sqrt{6}}[\bar{1}, \bar{1}, 2]$, $\hat{\mathbf{z}} = \frac{1}{\sqrt{3}}[111]$. We then use an ansatz for the dipole moments $\mathbf{p}_{1,2,3}$ of the dangling bonds $|c_{1,2,3}|^2$, consisting of parallel component $\mathbf{p}_{\parallel}$ and a perpendicular component $\mathbf{p}_{\perp}$:

$$\begin{aligned} \mathbf{p}_{1,2} &= p_{\perp}\left(\mp\frac{\sqrt{3}}{2}\hat{\mathbf{x}} - \frac{1}{2}\hat{\mathbf{y}}\right) - p_{\parallel}\hat{\mathbf{z}}, \\ \mathbf{p}_3 &= p_{\perp}\hat{\mathbf{y}} - p_{\parallel}\hat{\mathbf{z}}. \end{aligned} \tag{5.15}$$

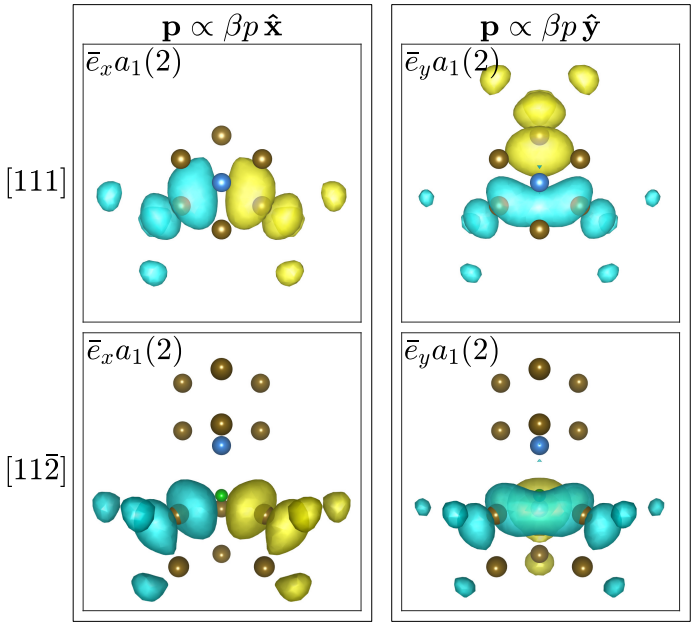

Figure 5.3: The two transition charges for the NV^- excitation shown in Fig. 5.2. Each transition charge is viewed both along and perpendicular to the N-V axis [111], and each has an obvious dipole character indicative of an optically-allowed excitation (cyan is negative, yellow is positive).

Applying this ansatz to Eq. 5.14, the transition dipole moments adopt especially simple forms given by

$$\begin{aligned} \mathbf{p}(\bar{e}_x a_1(2)) &\approx \frac{\beta}{\sqrt{6}}(-\mathbf{p}_1 + \mathbf{p}_2) = \frac{\beta p_\perp}{\sqrt{2}}\hat{\mathbf{x}}, \\ \mathbf{p}(\bar{e}_y a_1(2)) &\approx \frac{\beta}{\sqrt{18}}(-\mathbf{p}_1 - \mathbf{p}_2 + 2\mathbf{p}_3) = \frac{\beta p_\perp}{\sqrt{2}}\hat{\mathbf{y}}. \end{aligned} \tag{5.16}$$

Neither of these dipole moments has a parallel component. Their forms imply that, in the decay of the excited state to the ground state, the NV^- center emits unpolarized photons traveling along the N-V axis.

Fig. 5.3 shows the transition charges formed from the DFT wave functions in Fig. 5.2. The DFT transition charges show good correspondence with the tight-binding expressions given above. Viewed along the [111] N-V axis, the DFT transition charges posses an obvious dipole character indicative of an optically-allowed transi-

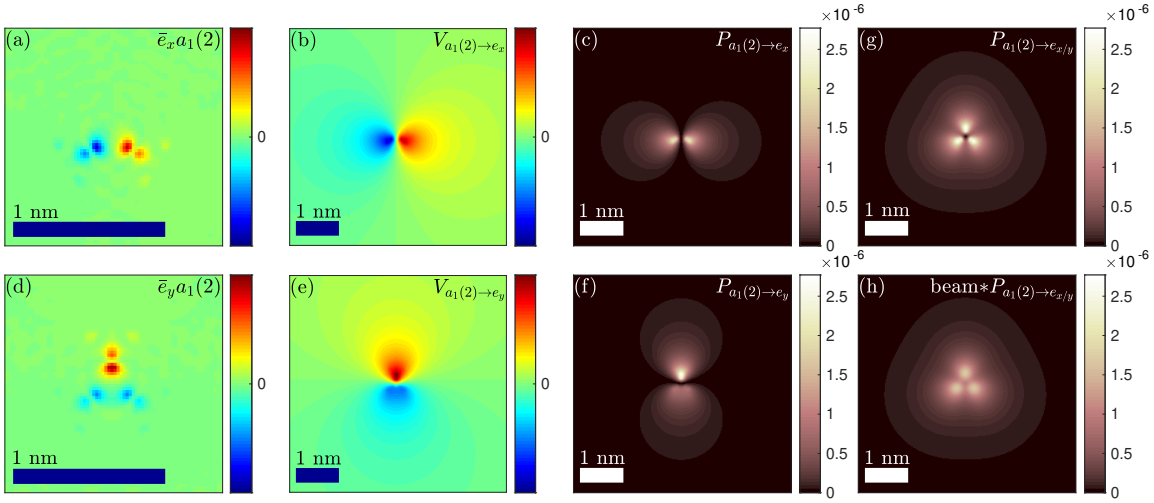

Figure 5.4: Calculation of STEM-EELS maps for NV^- viewed along the [111] N-V axis. (a–c) Projected transition charge, projected Møller potential and point-probe STEM-EELS map for transition $a_1^{\downarrow}(2) \to e_x^{\downarrow}$. (d–f) Analogous data for transition $a_1^{\downarrow}(2) \to e_y^{\downarrow}$. (g) Total point-probe STEM-EELS map (sum of (c) and (f)) showing three-fold rotational symmetry consistent with the C_{3v} point group. (h) Total STEM-EELS map for Gaussian probe with 1 Å FWHM.

tion. From the DFT transition charges we find $\beta p_\perp \approx 1.38$ Å, which, for comparison, has the same order of magnitude as that for a hydrogen $1s \leftrightarrow 2p$ transition.

5.4.2 Predictions of NV^- STEM-EELS Maps

Fig. 5.4 demonstrates the calculation of the STEM-EELS map of NV^- viewed along the [111] N-V axis. The transition charges in Figs. 5.4a and 5.4d are localized to within a few bond lengths of the defect. By comparison, the Møller potentials Figs. 5.4b and 5.4e extend over about 5 nm, attributable to the inelastic delocalization at this relatively low energy loss (~ 1 eV). The STEM-EELS maps exhibit a similar degree of delocalization. It would not be possible to observe the partial STEM-EELS maps in Fig. 5.4c and 5.4f, but only the total map shown in Fig. 5.4g (or Fig. 5.4h), which is the sum of the two partial maps. The total STEM-EELS maps show three symmetric lobes of higher intensity near the center of the defect, corresponding to

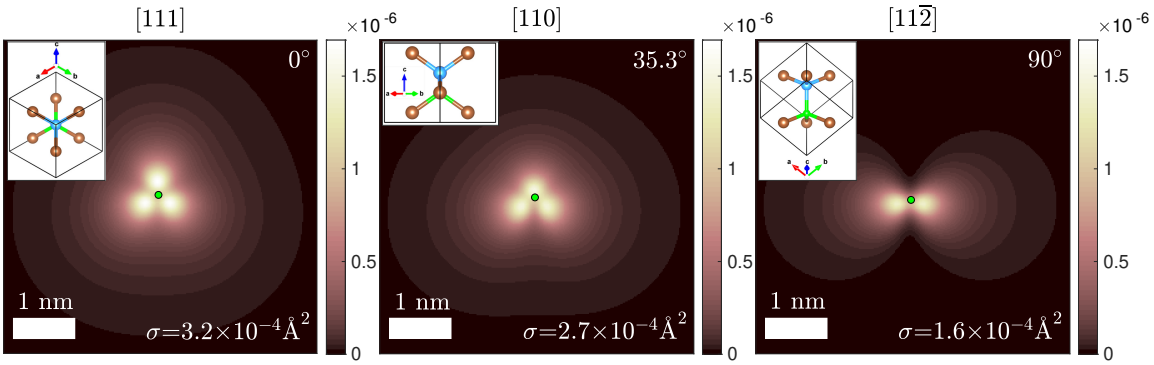

Figure 5.5: NV^- in three orientations. (a–c) STEM-EELS map for Gaussian probe with 1 Å FWHM, viewed along [111], [110], and [11$\bar{2}$], respectively. The green dot marks the position of the vacancy. The inclination with respect to the [111] N-V axis and the corresponding plane-wave cross-section are given in each case.

the positions of the three dangling bonds $c_{1,2,3}$ adjacent to the vacancy. The total map in Fig. 5.4h includes the blurring effect of a Gaussian-shaped STEM probe with 1 Å FWHM.

Fig. 5.5 shows STEM-EELS maps of the NV^- center viewed along three orientations. The three orientations represent a rotation about the $\hat{\mathbf{x}}$-axis defined above, which lies horizontal throughout Fig. 5.5. In each orientation, the lobes of higher intensity correspond to the dangling bonds $c_{1,2,3}$, as seen in projection. From the maps we conclude that it is possible in principle to locate the NV^- center transversely to within about 1 nm. An extremely important consideration is the intensity (scattering probability) obtained in the STEM-EELS maps. From Fig. 5.5 we see that the intensity is approximately 10^{-6} over an area approximately 1 nm^2. Fig. 5.5 also states the cross-sections σ for exciting the NV^- center using an incident electron plane wave. The cross-sections are the integrals of the map intensities and they provide convenient measures of the scattering probabilities. The map intensities and plane-wave cross sections are greatest when the defect is viewed along the [111] N-V axis. A detailed discussion of these points will be presented in Section 5.6.

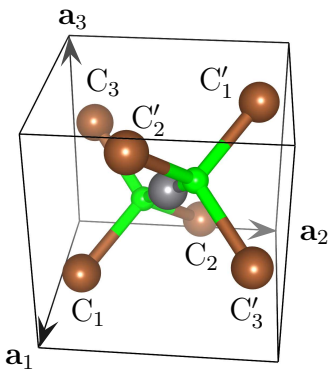

Figure 5.6: The atomic structure of the SiV center in diamond, with brown carbon atoms, a gray silicon atom, and green (split) vacancies. Carbon atoms $C_{1,2,3}$ are labeled as per the NV center, and carbon atoms $C'_{1,2,3}$ are related to $C_{1,2,3}$ by inversion about the silicon atom.

5.5 The SiV^0 Defect Center

This section contains our results for the SiV^0 defect center in diamond. The unusual structure of this defect, shown in Fig. 5.6, has been recognized since early DFT studies (Goss *et al.* (1996)). Unlike the NV center, where the nitrogen atom is only slightly shifted from a carbon position in the undefected structure, the SiV^0 adopts a split-vacancy configuration with the silicon atom midway between two vacant sites. This results in a point group symmetry of D_{3d}, which is related to C_{3v} by a direct product with inversion group: $D_{3d} = C_{3v} \times i$.

5.5.1 Electronic Structure of SiV^0

The electronic structure of the SiV^0 center has been treated via quantum chemistry (Moliver (2003)) and DFT approaches (Gali and Maze (2013); Hepp *et al.* (2014)). Experimental work in fully characterizing this defect is still ongoing, but recent measurements by Green *et al.* (2019)), correcting the earlier claim of D'Haenens-Johansson *et al.* (2011), support the identification of the 1.31 eV optically-allowed transition as a ${}^3A_{2g} \rightarrow {}^3E_u$ transition, matching the DFT prediction of Gali and Maze (2013).

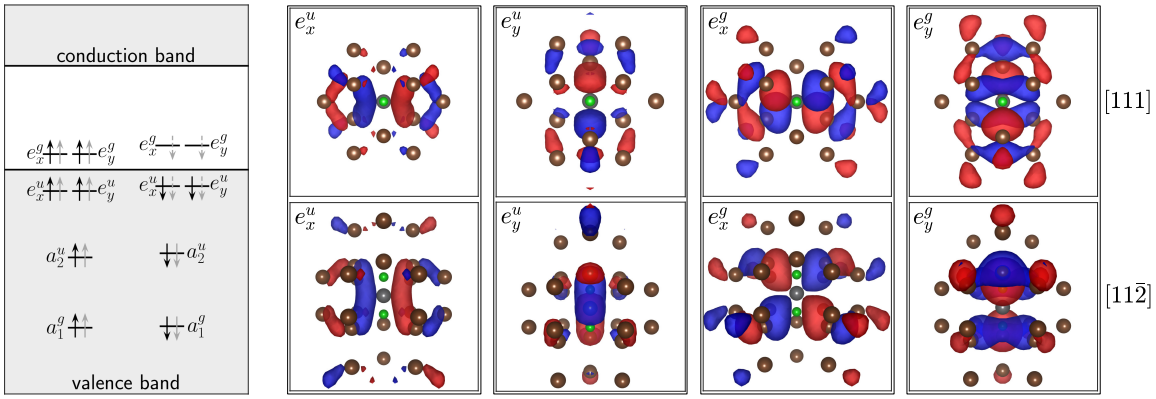

Figure 5.7: *Left:* The DFT (LSDA) predicted ground and excited states of the SiV^0 defect, with arrow conventions following those of Fig. 5.2. The ground state has a ${}^3A_{2g}$ (triplet) electronic configuration. The optically-allowed excited state is the ${}^3E_u$ (triplet) configuration resulting from a single-electron transition from e_x^u or e_y^u to e_x^g or e_y^g (Gali and Maze (2013)). *Right:* Isosurface plots of the wave functions involved in the excitation (red is negative, blue is positive), viewed both along and perpendicular to the SiV defect axis [111] (carbon atoms are brown, silicon atom is grey, vacancy is green).

As before, we gain insight into these states from a tight-binding model. Once again letting $c_{1,2,3}$ denote the dangling bond orbitals from the three carbon atoms $C_{1,2,3}$, we must also include $c'_{1,2,3}$ from the three carbon atoms $C'_{1,2,3}$ for this geometry. The full tight-binding treatment would require mixing with the valence orbitals of silicon, but previous investigators Hepp *et al.* (2014) have found that the defect states of interest are well-described by the carbon orbitals alone:

$$\begin{aligned} a_{1/2}^{g/u} &= \frac{1}{\sqrt{6}}\left((c_1+c_2+c_3) \pm (c'_1+c'_2+c'_3)\right), \\ e_x^{g/u} &= \frac{1}{2}\left((c_1-c_2) \pm (c'_1-c'_2)\right), \\ e_y^{g/u} &= \frac{1}{\sqrt{12}}\left((c_1+c_2-2c_3) \pm (c'_1+c'_2-2c'_3)\right). \end{aligned} \tag{5.17}$$

In these expressions, the superscripts "g" and "u" denote the gerade and ungerade states, corresponding to the upper and lower signs, respectively. Ten electrons occupy these defect states: six from the carbon orbitals, plus four from the silicon valence), as illustrated by the leftmost panel of Fig. 5.7. The promotion of fully-occupied $e_x^{u\downarrow}$ and

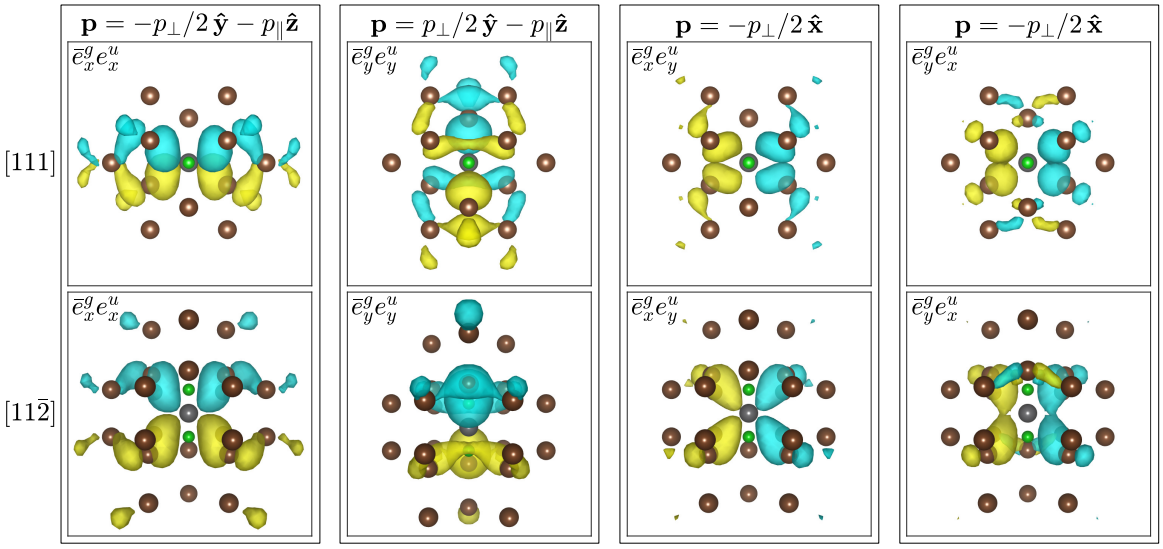

Figure 5.8: The four transition charges for the SiV^0 excitation shown in Fig. 5.7. Each transition charge is viewed both along and perpendicular to the SiV defect axis [111], and, as with the NV^- excitation, each has an obvious dipole character indicative of an optically-allowed excitation (cyan is negative, yellow is positive).

$e_y^{u\downarrow}$ states to half-occupied $e_x^{u\downarrow}/e_y^{u\downarrow}$ and $e_x^{g\downarrow}/e_y^{g\downarrow}$ states preserves the D_{3d} symmetry and the electron spin. The DFT prediction for the transition energy is 1.39 eV, to be compared with the experimental value of 1.31 eV. The relevant wave functions pictured on the right in Fig. 5.7 qualitatively resemble their tight-binding counterparts.

The transition described above leads to four separate transition charges from the orbitals of Eq. 5.17. But if we assume once again for simplicity that the bonds are separated in space, we find only three *independent* transition charges:

$$\begin{aligned}
\bar{e}^g_x e^u_x &\approx \frac{1}{4}\left(|c_1|^2 + |c_2|^2 - |c_1'|^2 - |c_2'|^2\right), \\
\bar{e}^g_y e^u_y &\approx \frac{1}{12}\left(|c_1|^2 + |c_2|^2 + 4|c_3|^2 - |c_1'|^2 - |c_2'|^2 - 4|c_3'|^2\right), \\
\bar{e}^g_x e^u_y &\approx \bar{e}^g_y e^u_x \approx \frac{\sqrt{3}}{12}\left(|c_1|^2 - |c_2|^2 - |c_1'|^2 + |c_2'|^2\right).
\end{aligned} \tag{5.18}$$

To find the transition dipole moments, we reuse the ansatz of Eq. 5.15 with the additional stipulation that the dipole moments $\mathbf{p}'_{1,2,3}$ of $|c'_{1,2,3}|^2$ are related to $\mathbf{p}_{1,2,3}$ by inversion, i.e., $\mathbf{p}'_{1,2,3} = -\mathbf{p}_{1,2,3}$. Inserting these expressions into Eq. 5.14, we find that

the dipole moments of $\bar{e}_x^g e_x^u$ and $\bar{e}_y^g e_y^u$ have components along both $\hat{\mathbf{y}}$ and the symmetry axis $\hat{\mathbf{z}}$, while the dipole moments of $\bar{e}_x^g e_y^u$ and $\bar{e}_y^g e_x^u$, equal in this approximation, lie strictly along $\hat{\mathbf{x}}$:

$$\begin{aligned}
\mathbf{p}(\bar{e}_x^g e_x^u) &\approx (2\mathbf{p}_1 + 2\mathbf{p}_2)/4 = -\frac{1}{2}p_\perp \hat{\mathbf{y}} - p_\parallel \hat{\mathbf{z}}, \\
\mathbf{p}(\bar{e}_y^g e_y^u) &\approx (2\mathbf{p}_1 + 2\mathbf{p}_2 + 8\mathbf{p}_3)/12 = \frac{1}{2}p_\perp \hat{\mathbf{y}} - p_\parallel \hat{\mathbf{z}}, \\
\mathbf{p}(\bar{e}_y^g e_x^u) &\approx \mathbf{p}(\bar{e}_x^g e_y^u) \approx \frac{\sqrt{3}}{12}(2\mathbf{p}_1 - 2\mathbf{p}_2) = -\frac{1}{2}p_\perp \hat{\mathbf{x}}.
\end{aligned} \tag{5.19}$$

Fig. 5.8 shows the transition charges formed from the DFT wave functions in Fig. 5.7. As with the NV^- transition charges, there is an obvious dipole nature to these transition charges when viewed along the [111] Si-V axis. Yet what is strikingly different for the SiV^0 case is that the dipole character of $\bar{e}_x^g e_x^u$ and $\bar{e}_y^g e_y^u$ are even more pronounced when they are viewed along $[11\bar{2}]$, perpendicular to the symmetry axis. Using the DFT transition charges, we find $p_\perp \approx 0.920$ Å, and $p_\parallel \approx 0.853$ Å, consistently among all the transitions. Again these dipole moments have the same order of magnitude as that for a $1s \leftrightarrow 2p$ transition in hydrogen.

As we discuss below, the fact that $p_\parallel > p_\perp/2$ is significant to the character of spectral maps obtained at different orientations. We also note that the $p_\perp$ calculated for the SiV^0 defect is inconsistent with the $p_\perp$ calculated from the NV^- defect, highlighting the quantitative limitations of the tight-binding analysis.

5.5.2 Predictions of SiV^0 STEM-EELS maps

Fig. 5.9 shows STEM-EELS maps for the SiV^0 defect along three orientations. As with the NV^- case in Fig. 5.5, the maps in Fig. 5.9 indicate that it is possible in principle to locate the defect to within about 1 nm. But the SiV^0 maps are distinct from those of the NV^- case. At the [111] orientation the SiV^0 map exhibits a six-fold

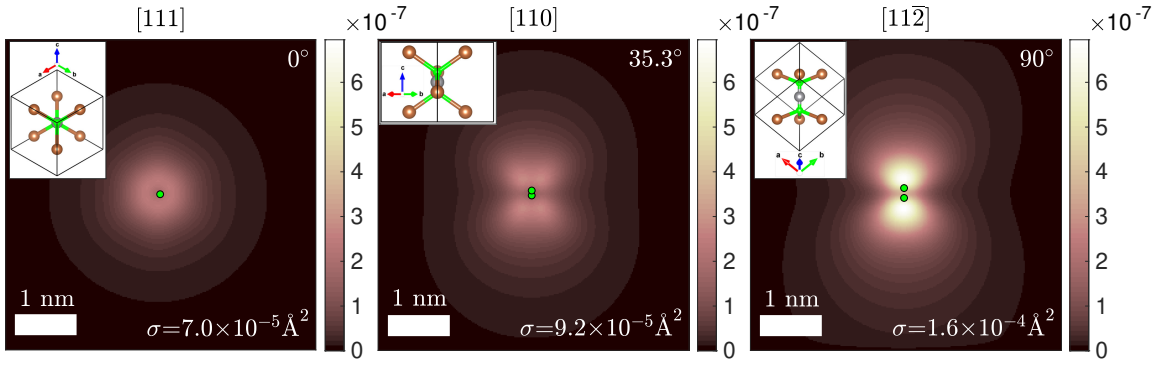

Figure 5.9: SiV^0 in three orientations. (a–c) STEM-EELS maps for Gaussian probe with 1 Å FWHM, viewed along [111], [110] and [11$\bar{2}$], respectively. The green dots mark the position of the split vacancy. The inclination with respect to the [111] defect axis and the corresponding plane-wave cross-sections are given for each case.

rotational symmetry consistent with its D_{3d} point group. At the [11$\bar{2}$] orientation, the SiV^0 map exhibits two vertical lobes along the defect axis arising from the dipole component along that axis, which is in contrast to the lobes along $\hat{\mathbf{x}}$ for the NV^- case. The SiV^0 map intensities and cross-sections tend to be smaller than those for NV^-, though the same order of magnitude. Moreover, as the SiV^0 is rotated away from the defect axis, the plane-wave cross-section increases, whereas under the equivalent rotation the plane-wave cross-section for NV^- decreases.

5.6 Discussion

A key conclusion from the results in Figs. 5.5 and 5.9 is that it is possible in principle to use STEM-EELS to locate individual defect centers transversely to within about 1 nm. Such a precision is well beyond that permitted by the diffraction limit in photodetection. One might compare this precision to the recent nanoscale characterization of the NV^- charge environment by microwave spectroscopy (Mittiga *et al.* (2018)), which claimed the localization of a positive charge to within 2 nm of a NV^- center by inference from orientation-specific measurements, without the ability to localize the NV^- defect itself. The STEM-EELS claim of 1 nm precision is consistent

with the inelastic delocalization v/ω that determines the upper bound on the impact parameter at which a passing beam electron can cause an excitation, as was pointed out by Dwyer (2014). A precision of 1 nm would also enable the crystallographic environment of the point defect to be characterized using STEM imaging.

Another important consideration is whether the scattering probabilities in Figs. 5.5 and 5.9 would allow the detection of single defect centers with a beam electron dose that the sample could tolerate. For the NV^- and SiV^0 excitations considered here, the cross-sections for 60 keV plane wave electrons are in the ranges 1.6–3.2×10^{-4} Å^2 and 0.7–1.6×10^{-4} Å^2, depending on orientation. By comparison, the cross-section for 60 keV plane wave electrons to excite the C-K shell of a single carbon atom to within a 10 eV energy window immediately above the ionization threshold (285 eV) is 2×10^{-5} Å^2. Compared with this common core-level excitation, the NV^- and SiV^0 cross-sections are 5–15 times larger. Given this comparison and the fact that EELS experiments have already detected single atoms via their core-level excitations (see Varela *et al.* (2009); Senga and Suenaga (2015); Ramasse *et al.* (2013)), our results indicate that detecting single defect centers with STEM-EELS is within reach.

The specific appearance and symmetries of the STEM-EELS maps in Figs. 5.5 and 5.9 are determined by the nature of the projected transition charges. Within the tight-binding approximations of Sections 5.4.1 and 5.5.1, the transition charges for NV^- and SiV^0 can be written in terms of the dangling carbon bonds, so the maps in Figs. 5.5 and 5.9 can be interpreted as projections of these dangling bonds.

These transition charges explain the dependence of the NV^- and SiV^0 cross-sections on orientation, as the cross-section is determined largely by the transition charges' dipole moments. More specifically, partial cross sections are mainly determined by the component of these dipole moments lying perpendicular to the beam. For NV^-, the cross section is largest at [111] because at that orientation the dipole

moments of $\bar{e}_x a_1(2)$ and $\bar{e}_y a_1(2)$ are both perpendicular to the beam. On rotating from [111] to [11$\bar{2}$], the relevant component of $\bar{e}_x a_1(2)$ is constant but that of $\bar{e}_y a_1(2)$ decreases to zero, and hence the NV^- cross section decreases accordingly. The SiV^0 case is a bit more involved: At the [111] orientation, the perpendicular components of the four transition dipoles have equal magnitudes of $p_\perp/2$. On rotating from [111] to [11$\bar{2}$], the relevant dipole components of $\bar{e}_x^g e_x^u$ and $\bar{e}_y^g e_y^u$ both increase since $p_\parallel > p_\perp/2$ for this defect. This accounts for the increase in the SiV^0 cross section on rotating from [111] to [11$\bar{2}$].

An important effect that has been neglected in this calculation is that of phonons on the excitation of defect centers. Phonons give rise to multiple peaks in excitation spectra, where each peak corresponds to the creation of a specific number of phonon quanta within the underlying electronic excitation. In a photoluminescence spectrum, the zero-phonon line sits at the upper edge of a sideband of lower-energy photons, representing the relatively shallower transitions from the zero-phonon electronic excited state to an n-phonon electronic ground state. In the EEL spectrum, the sideband will manifest at energies above the zero-phonon line, representing excitaitions from the zero-phonon electronic ground state to an n-phonon electronic excited state. But while luminescence and excitation are not entirely symmetric (phonon states differ for the excited and ground electronic states), the work of Alkauskas *et al.* (2014, 2016) suggests that in either case phonons will spread the spectral intensity acorss multiple peaks while having little effect on the total spectral intensity.

Still, detectability prospects improve when the signal is confined to a narrow energy range. This is one way that SiV^0 might compete alongside NV^- as an experimental candidate. While Alkauskas *et al.* (2014) reports that the luminescence line shape for NV^- is contained mainly within a $\sim$0.4 eV window, Rose *et al.* (2018) reports that SiV^0 emits primarily to its zero-phonon line at low temperatures. Unfortunately,

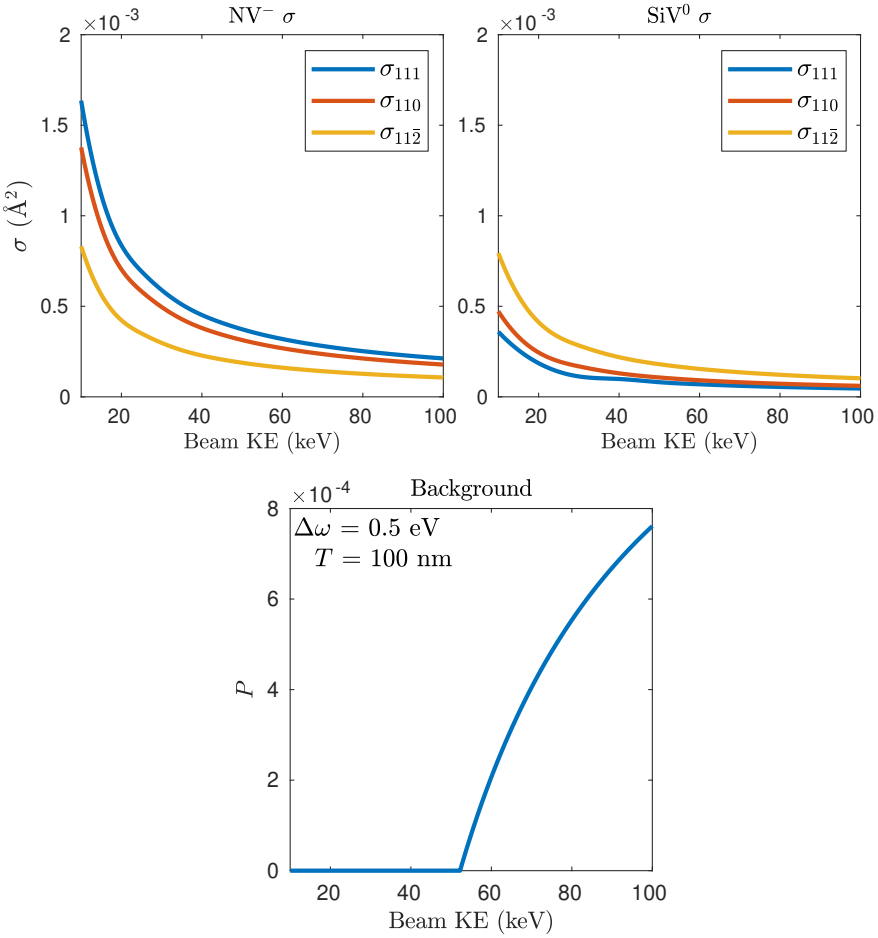

Figure 5.10: *Top:* Plane-wave cross sections as a function of beam energy for NV^- (left) and Si^0 (right), for the [111], [110] and [11$\bar{2}$] orientations. *Bottom:* The Cherenkov background, approximated using the Frank-Tamm formula (Eq. 3.28) for an integration window of 0.5 eV and a sample thickness of 100 nm.

these energy ranges may not translate directly to the EEL spectra. Electrons carry much more momentum than photons, so one might plausibly expect more phonons to be created by electron beams, and hence to see a slightly broader energy range in EEL spectra than in luminescence spectra.

For unambiguous detection, one must have both an interaction cross section that is acceptably large and a spread of spectral intensities over a range that is acceptably small. Recent improvements in electron detection enhance the experimental feasibility, but there are many factors to weigh. It will also be important to use a dose rate

of beam electrons that respects the excitation lifetimes of the defects ($\sim$ 10–20 ns for NV^- and $\sim$ 1 ns for SiV^0), so as not to cause significant multiple excitations which may lead to their ionization. It will also be important to minimize low-loss background processes, such as Cherenkov radiation (past its onset for beam energies above $\sim$ 52 keV), which could potentially mask spectral signatures.

Fig. 5.10 illustrates how plane-wave cross sections and Cherenkov backgrounds vary with beam energy, and provides an example of how the benefits and costs must be weighed. Interaction probabilities are larger at lower beam energies, but for the inelastically-scattered electron to exit without significant spatial dispersion, high beam energies are preferred. But at high beam energies, Cherenkov radiation will contribute an increasing background. These considerations and more must be weighed by experimentalists hoping to observe individual point defects using EELS.

5.7 Summary

Density function theory and a Møller potential scattering formalism were used to calculate STEM-EELS spectral maps and plane-wave cross sections for two important defect centers in bulk diamond: the NV^- center and the SiV^0 center. These calculations indicate that STEM-EELS should in principle be able to locate individual occurrences of these defects with a precision of around 1 nm, with plane-wave scattering cross sections of the order of 10^{-4} Å^2. Such cross-sections are similar to or greater than core-level EELS cross sections that have already enabled individual atoms to be individually located. Similar conclusions are likely to apply to optically-allowed transitions of other point defects, which warrants optimism that individual point defects in bulk materials can be successfully probed using STEM-EELS.

Many advanced materials applications utilize such point defects, so these results indicate that STEM-EELS could become an extremely valuable tool for the physicists,

chemists, and materials scientists who need a detailed understanding of point defect properties. On the experimental side, continuing improvements in detector efficiencies will allow for an increasing proportion collected electron to be directly interpretable. And on the theory side, calculations of spectra that include line broadening from localized phonon modes will aid both in interpretation and in understanding the expected signal-to-background for practical detection.

Chapter 6

FINAL SUMMARY

This dissertation has dealt with two basic types of questions in the theory of STEM-EELS, both motivated by the recent advent of meV EELS energy resolution in the STEM. First, it has asked what sorts of mesoscopic STEM-EELS effects can be observed in simple nanoscale geometries, as captured by the physical description of the local dielectric response. The first type of question occupied Ch. 2, which explored the electrostatic approach to spectral modeling and nanoparticle shape effects, and Ch. 3, which explored the spatial imaging of waveguide modes and electrodynamic effects via spectral mapping. Such questions link this work with contemporary nanophotonics and materials science. Second, this dissertation has asked what sorts of new microscopic STEM-EELS effects can probed in the low-loss regime, given the resolution capabilities of currently available microscopes. This second type of question occupied Ch. 4, which developed a model for how the vibrational signals of adsorbate molecules can be enhanced by nanoparticle surface plasmons, and Ch. 5, which analyzed the prospects for spatial localization of selected point defects in diamond using density functional theory. Such questions link this work with contemporary surface science and quantum materials.

The treatment of the electrostatic approach in Ch. 2 introduced two complimentary approaches to simulating low-loss STEM-EEL spectra. In one approach, the "classical" approach, the electron beam's energy loss arises as the result of the force applied to the passing electron by the potential arising from the material response to it. In the other approach, the "quantum" approach, the dielectric material's normal modes come first, and EEL spectra are built up from the $|0\rangle \rightarrow |1_n\rangle$ mode transition

probabilities, for modes labeled by quantum numbers n. In the classical approach, energy loss seems to depend on the presence of an imaginary part to the dielectric function $\epsilon(\omega)$, as forces acting on the beam electron upon its approach and departure from the nanoparticle are otherwise matched. However, it was shown for a particular case, aloof scattering from a dielectric sphere, that the "quantum" result arises naturally from the "classical" result in the limit as the damping vanishes. Presumably, this quantum-classical corresponence could be extended to include damping using a more flexible density operator formalism for open systems.

In the electrostatic analysis of the Born-Huang model for dielectric response, the "quantum" analysis revealed three types of modes: transverse modes at ω_0, the resonance frequency of the material dipoles, which have no associated macroscopic electric field; longitudinal modes at $\omega_\ell = (\epsilon_0/\epsilon_\infty)^{1/2}\omega_0$, whose macroscopic electric field is confined within the material, and zero outside; and the harmonic modes at ω_h, such that $\omega_0 < \omega_h < \omega_\ell$, which only exist when $\mathrm{Re}[\epsilon(\omega)] < 0$. The harmonic modes of the sphere cover a minimal frequency range for surface states, from $\epsilon(\omega) = -2$ to $\epsilon(\omega) = -1$. With increasing aspect ratio, the prolate spheroid stretches these mode frequencies downward toward ω_0, evolving toward the limiting case of the semi-infinite cylinder. With increasing aspect ratio, oblate spheroid stretches these mode frequencies both downward toward ω_0, for modes having an even top/bottom symmetry in surface charge, and upward toward ω_ℓ, for modes having a top/bottom antisymmetry in surface charge. The limiting frequency $\epsilon(\omega) = -1$, lying between ω_0 and ω_ℓ, holds for surface modes of high spatial frequencies in all geometries, as it indicates that the potential near the surface depends only on local surface charge. At high enough spatial frequencies, the relevant description for EELS shifts from dipole to impact scattering, and an explicit k-space cutoff may be imposed to respect this limit.

The work in Ch. 3 revealed that the "classical" and "quantum" approaches to

spectrum simulation can be extended to a beam description or mode description incorporating electrodynamic effects. While in the electrostatic approach to dielectric losses significant signal is expected only where $\epsilon(\omega) \leq 0$, in the electrodynamic approach modes can exist where $\epsilon(\omega) > 0$, modes that trap light via total internal reflection. In the electrostatic approximation, these modes all exist at the transverse resonance frequencies of the dielectric ω_0, and do not contribute to EEL spectra. In the electrodynamic approach, these modes have lowered frequencies and nonzero field intensities. For electron beams piercing thick particles, momentum conservation constrains which of these modes can be excited by a beam of a particular energy, leading (counter-intuitively) to lower-energy modes being excited by higher-energy beams.

For dielectric ribbons, the "classical" approach to numerical computations showed how the spatial distributions of such modes should be most easily detectible for ribbon widths a few times the wavelength of light in the material: thinner, and the spatial variation will not be apparent; thicker, and the standing waves will become so closely spaced in frequency as to be indistinguishable from bulk Cherenkov loss. The transverse electric (TE) modes appear in these spectra as sharp, narrow spikes, and the transverse magnetic (TM) modes appear as shallow, broad bands. Together, they combine to form standing-wave patterns in the spectral maps across the ribbon, with modes with more nodes appearing at higher energies. These solutions for STEM-EELS were worked out long ago by Bolton and Chen (1995), but the spatial variations in waveguide modes had not been explicitly shown in published work.

The "quantum" approach was applied to modeling waveguide modes in amorphous silicon cylinders to explain the experimental results of Flauraud and Alexander (2019) on discs of varying diameters. The waveguide modes for an unbounded cylinder split into TE and TM modes only when the modes have azimuthal symmetry. Otherwise, they split into hybridized EH and HE modes, which resemble TE and TM modes,

respectively. TM and HE modes are most strongly imaged in these models, for a beam running parallel to the cylindrical axis. Modes with the expected spatial behavior were experimentally tracked across several discs, and demonstrated the expected qualitative behavior, but appeared at energies higher than those of predicted by the unbounded cylinder model. Upper bounds on modal energies were determined using a "thin disc" approximation, supplementing the waveguide model's lower bounds. The ability to interpret spatial variations in low-loss STEM-EELS signals could prove useful for designing and characterizing engineered materials, especially when the particular causes for effects are unclear from optical measurements.

The microscopic modelling of Ch. 4 and Ch. 5 returned to the electrostatic approach. In the model presented in Ch. 4, an adsorbate molecule (modeled as a point dipole) on the surface of a silver nanoparticle (modeled as a prolate spheroid) was shown to have an enhanced EEL signal. This signal enhancement is caused first by the molecule's experience of a stronger field due to the presence of the particle, then by the particle's altered field that acts back on the electron beam. The molecular signal scales with the square of the electric field at the molecular adsorption site, and sits atop the EEL signal of the "bare" nanoparticle. The spectral line shape depends critically on the phase shift of the molecularly-induced electric field relative to the beam's Coulomb field. Nanoparticles of an appropriately small thickness ($\sim$10 nm) and an appropriately large aspect ratio ($\sim$10) were shown to produce signal enhancements from 10^2-10^3, and nanoparticles tuned to the vibrational frequency of the adsorbate could do even better, albeit with higher background signals and more pronounced asymmetric Fano-type signatures. The ability to interpret such signals could lead to significantly increased spatial control over molecular vibrational EELS, and could aid in the characterization of nanoparticles.

In Ch. 5, the the output of density functional theory code was used to model two

point defects in diamond: the negatively-charged nitrogen-vacancy center (NV^-), and the neutral silicon-vacancy center (SiV^0). These two defects have different symmetries, but the outcomes of each could be interpreted using a tight binding model, and simulated spectral maps indicated a possible transverse localization to around 1 nm. The orientation dependence of the excitation probabilities was found to scale with the component of the dipole monment of the the transition charge in the sample plane. The calculated plane-wave scattering cross sections ($\sim 10^{-4}$ Å^2) are similar to the cross-sections of core-level excitations that have already been observed, leading to optimism that observation of such individual defects can be experimentally realized, and that STEM-EELS could become increasingly useful as a tool for characterizing advanced quantum materials in the coming decade.

In closing, it should be emphasized that each topic explored here contains numerous opportunities for continued research. Thumbing backward through the dissertation, I notice that the calculations for point defects (Ch. 5) could incorporate vibrational broadening into the spectral signatures of the electronic transitions. Likewise, the work on enhanced signals from adsorbates (Ch. 4) could be extended to treat adsorbates near a plasmonic dimer, which would lead to greater enhancement effects, at the cost of interpretational difficulties. The treatment of electrodynamic modes (Ch. 3) could be extended to treat exotic geometries and particle arrays, including studies of metamaterials. And the "classical" to "quantum" transition (Ch. 2) could be extended to include damping effects in the "quantum" treatement, which might yield insights into the nature of the classical/quantum correspondance.

This is, of course, only a partial tally. STEM-EELS is just one way of tickling the universal elephant (Ch. 1), but its possibilities are far from exhausted.

Adachi, S., "Model dielectric constants of Si and Ge", Phys. Rev. B **38**, 12966–12976 (1988).

Alkauskas, A., B. B. Buckley, D. D. Awschalom and C. G. V. de Walle, "First-principles theory of the luminescence lineshape for the triplet transition in diamond NV centres", New Journal of Physics **16**, 7, 073026 (2014).

Alkauskas, A., M. D. McCluskey and C. G. Van de Walle, "Tutorial: Defects in semiconductors—combining experiment and theory", Journal of Applied Physics **119**, 18, 181101 (2016).

Allen, L., A. D'Alfonso and S. Findlay, "Modelling the inelastic scattering of fast electrons", Ultramicroscopy **151**, 11–22 (2015).

Arnbak, J., "Leaky waves on a dielectric rod", Electronics Letters **5**, 41–42(1) (1969).

Aspnes, D. E. and A. A. Studna, "Dielectric functions and optical parameters of Si, Ge, GaP, GaAs, GaSb, InP, InAs, and InSb from 1.5 to 6.0 eV", Phys. Rev. B **27**, 985–1009 (1983).

Batson, P., N. Dellby and O. L. Krivanek, "Sub-ångstrom resolution using aberration corrected electron optics.", Nature **418**, 6898, 617–620 (2002).

Big, E. J., "A short history of the electron microscope", Bios **27**, 1, 33–37 (1956).

Blöchl, P. E., "Projector augmented-wave method", Phys. Rev. B **50**, 17953–17979 (1994).

Bohm, D. and D. Pines, "A Collective Description of Electron Interactions: III. Coulomb Interactions in a Degenerate Electron Gas", Phys. Rev. **92**, 609–625 (1953).

Bohren, C. F., "What did Kramers and Kronig do and how did they do it?", European Journal of Physics **31**, 3, 573–577 (2010).

Bolton, J. P. R. and M. Chen, "Electron energy loss in multilayered slabs. II. parallel incidence", Journal of Physics: Condensed Matter **7**, 18, 3389–3403 (1995).

Born, M. and K. Huang, *Dynamical Theory of Crystal Lattices*, International series of monographs on physics (Clarendon Press, 1954).

Chang, S. L. Y., A. S. Barnard, C. Dwyer, C. B. Boothroyd, R. K. Hocking, E. Ōsawa and R. J. Nicholls, "Counting vacancies and nitrogen-vacancy centers in detonation nanodiamond", Nanoscale **8**, 10548–10552 (2016).

Cohen, M. L. and S. G. Louie, *Fundamentals of Condensed Matter Physics* (Cambridge UP, 2016).

Cohen de Lara, E., "Electric field effect on molecules: Relation between the orientation of the molecule with respect to the field and the vibrational frequency shift observed in IR spectra of molecules adsorbed in zeolites", Phys. Chem. Chem. Phys. **1**, 501–505 (1999).

Couillard, M., A. Yurtsever and D. A. Muller, "Competition between bulk and interface plasmonic modes in valence electron energy-loss spectroscopy of ultrathin Sio_2 gate stacks", Phys. Rev. B **77**, 085318 (2008).

Crewe, A. V., J. Wall and L. M. Welter, "A high-resolution scanning transmission electron microscope", Journal of Applied Physics **39**, 13, 5861–5868 (1968).

Cruz y Cruz, S. and O. Rosas-Ortiz, "Leaky modes of waveguides as a classical optics analogy of quantum resonances", Advances in Mathematical Physics **2015** (2015).

de Solla Price, D., *Science Since Babylon*, A Yale Paperbound (Yale University Press, 1975).

Deinega, A. and S. John, "Effective optical response of silicon to sunlight in the finite-difference time-domain method", Opt. Lett. **37**, 1, 112–114 (2012).

Deng, B., R. Q. Zhang and X. Q. Shi, "New insight into the spin-conserving excitation of the negatively charged nitrogen-vacancy center in diamond", Scientific Reports **4** (2014).

Dethe, T., H. Gill, D. Green, A. Greensweight, L. Gutierrez, M. He, T. Tajima and K. Yang, "Causality and dispersion relations", American Journal of Physics **87**, 4, 279–290 (2019).

D'Haenens-Johansson, U. F. S., A. M. Edmonds, B. L. Green, M. E. Newton, G. Davies, P. M. Martineau, R. U. A. Khan and D. J. Twitchen, "Optical properties of the neutral silicon split-vacancy center in diamond", Phys. Rev. B **84**, 245208 (2011).

Dhomkar, S., P. R. Zangara, J. Henshaw and C. A. Meriles, "On-demand generation of neutral and negatively charged silicon-vacancy centers in diamond", Phys. Rev. Lett. **120**, 117401 (2018).

Doherty, M. W., N. B. Manson, P. Delaney, F. Jelezko, J. Wrachtrup and L. C. Hollenberg, "The nitrogen-vacancy colour centre in diamond", Physics Reports **528**, 1, 1 – 45 (2013).

Dwyer, C., "Multislice theory of fast electron scattering incorporating atomic inner-shell ionization", Ultramicroscopy **104**, 141–151 (2005).

Dwyer, C., "Chapter Three - Atomic-Resolution Core-Level Spectroscopy in the Scanning Transmission Electron Microscope", in "Advances in Imaging and Electron Physics", edited by P. W. Hawkes, vol. 175, pp. 145 – 199 (Elsevier, 2013).

Dwyer, C., "Localization of high-energy electron scattering from atomic vibrations", Phys. Rev. B **89**, 054103 (2014).

Dwyer, C., "The role of symmetry in the theory of inelastic high-energy electron scattering and its application to atomic-resolution core-loss imaging", Ultramicroscopy **151**, 68–77 (2015).

Dwyer, C., "Prospects of spatial resolution in vibrational electron energy loss spectroscopy: Implications of dipolar scattering", Phys. Rev. B **96**, 224102 (2017).

Dwyer, C., T. Aoki, P. Rez, S. L. Y. Chang, T. C. Lovejoy and O. L. Krivanek, "Electron-beam mapping of vibrational modes with nanometer spatial resolution", Phys. Rev. Lett. **117**, 256101 (2016).

Dwyer, C. and J. S. Barnard, "Relativistic effects in core-loss electron diffraction", Phys. Rev. B **74**, 064106 (2006).

Echenique, P. M. and J. B. Pendry, "Absorption profile at surfaces", Journal of Physics C: Solid State Physics **8**, 18, 2936–2942 (1975).

Egerton, R., *Electron Energy-Loss Spectroscopy in the Electron Microscope* (Springer US, 2011).

Egerton, R., "TEM-EELS: A personal perspective", Ultramicroscopy **119**, 24 – 32 (2012).

Egerton, R. F., "Electron energy-loss spectroscopy in the TEM", Reports on Progress in Physics **72**, 1, 016502 (2008).

Eljarrat, A., S. Estradé and F. Peiró, "Chapter One - Introduction to EELS", in "Advances in Imaging and Electron Physics", edited by P. W. Hawkes and M. Hÿtch, vol. 209 of *Advances in Imaging and Electron Physics*, pp. 1 – 47 (Elsevier, 2019).

Englman, R. and R. Ruppin, "Optical phonons in finite crystals", Phys. Rev. Lett. **20**, 898 (1966).

Eremets, M. I., M. Gauthier, A. Polian, J. C. Chervin, J. M. Besson, G. A. Dubitskii and Y. Y. Semenova, "Optical properties of cubic boron nitride", Phys. Rev. B **52**, 8854–8863 (1995).

Essex, E. A., "Hertz vector potentials of electromagnetic theory", American Journal of Physics **45**, 11, 1099–1101 (1977).

Etchegoin, P. G. and E. C. Le Ru, *Basic Electromagnetic Theory of SERS*, chap. 1, pp. 1–37 (John Wiley & Sons, Ltd, 2010).

Feenstra, R. M., N. Srivastava, Q. Gao, M. Widom, B. Diaconescu, T. Ohta, G. L. Kellogg, J. T. Robinson and I. V. Vlassiouk, "Low-energy electron reflectivity from graphene", Phys. Rev. B **87**, 041406 (2013).

Fermi, E., "The Ionization Loss of Energy in Gases and in Condensed Materials", Phys. Rev. **57**, 485–493 (1940).

Ferrell, T. L., R. J. Warmack, V. E. Anderson and P. M. Echenique, "Analytical calculation of stopping power for isolated small spheres", Phys. Rev. B **35**, 7365–7371 (1987).

Flauraud, V. and D. T. Alexander, "STEM-EELS Imaging of Resonant Modes in Dielectric Silicon Nanostructures", Microscopy and Microanalysis **25**, S2, 634–635 (2019).

Ford, G. and W. Weber, "Shift and broadening of the vibrational mode for a diatomic molecule adsorbed on a metal surface", Surface Science **129**, 1, 123 – 136 (1983).

Forouhi, A. R. and I. Bloomer, "Optical dispersion relations for amorphous semiconductors and amorphous dielectrics", Phys. Rev. B **34**, 7018–7026 (1986).

Frame, J. W., "On the mathematical equivalence of two ways of regarding the excitation of an atom by a fast moving α-particle", Mathematical Proceedings of the Cambridge Philosophical Society **27**, 4, 511–517 (1931).

Fuchs, R. and K. L. Kliewer, "Optical modes of vibration in an ionic crystal slab", Phys. Rev. A **140**, 2076–2088 (1965).

Gali, A., M. Fyta and E. Kaxiras, "Ab initio supercell calculations on nitrogen-vacancy center in diamond: Electronic structure and hyperfine tensors", Phys. Rev. B **77**, 155206 (2008).

Gali, A., E. Janzén, P. Deák, G. Kresse and E. Kaxiras, "Theory of spin-conserving excitation of the $N-V^-$ center in diamond", Phys. Rev. Lett. **103**, 186404 (2009).

Gali, A. and J. R. Maze, "Ab initio study of the split silicon-vacancy defect in diamond: Electronic structure and related properties", Phys. Rev. B **88**, 235205 (2013).

García de Abajo, F. J., "Optical excitations in electron microscopy", Rev. Mod. Phys. **82**, 209–275 (2010).

Garcia-Molina, R., A. Gras-Marti, A. Howie and R. H. Ritchie, "Retardation effects in the interaction of charged particle beams with bounded condensed media", Journal of Physics C: Solid State Physics **18**, 27, 5335–5345 (1985).

Gil, A. and J. Segura, "A code to evaluate prolate and oblate spheroidal harmonics", Computer Physics Communications **108**, 2, 267 – 278 (1998).

Goss, J. P., R. Jones, S. J. Breuer, P. R. Briddon and S. Öberg, "The twelve-line 1.682 ev luminescence center in diamond and the vacancy-silicon complex", Phys. Rev. Lett. **77**, 3041–3044 (1996).

Green, B. L., M. W. Doherty, E. Nako, N. B. Manson, U. F. S. D'Haenens-Johansson, S. D. Williams, D. J. Twitchen and M. E. Newton, "Electronic structure of the neutral silicon-vacancy center in diamond", Phys. Rev. B **99**, 161112 (2019).

Green, B. L., S. Mottishaw, B. G. Breeze, A. M. Edmonds, U. F. S. D'Haenens-Johansson, M. W. Doherty, S. D. Williams, D. J. Twitchen and M. E. Newton, "Neutral silicon-vacancy center in diamond: Spin polarization and lifetimes", Phys. Rev. Lett. **119**, 096402 (2017).

Gumerov, N. A. and R. Duraiswami, *Recursive Computation of Spherical Harmonic Rotation Coefficients of Large Degree*, pp. 105–141 (Springer International Publishing, Cham, 2015).

Gunawan, A. A., K. A. Mkhoyan, A. W. Wills, M. G. Thomas and D. J. Norris, "Imaging 'invisible' dopant atoms in semiconductor nanocrystals", Nano Letters **11**, 12, 5553–5557 (2011).

Hage, F. S., D. M. Kepaptsoglou, Q. M. Ramasse and L. J. Allen, "Phonon Spectroscopy at Atomic Resolution", Phys. Rev. Lett. **122**, 016103 (2019).

Hage, F. S., G. Radtke, D. M. Kepaptsoglou, M. Lazzeri and Q. M. Ramasse, "Single-atom vibrational spectroscopy in the scanning transmission electron microscope", Science **367**, 6482, 1124–1127 (2020).

Haider, M., S. Uhlemann, E. Schwan, H. Rose, B. Kabius and K. Urban, "Electron microscopy image enhanced", Nature **392**, 6678, 768–769 (1998).

Hasnip, P. J., K. Refson, M. I. J. Probert, J. R. Yates, S. J. Clark and C. J. Pickard, "Density functional theory in the solid state", Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **372**, 2011, 20130270 (2014).

Hawkes, P. W. and J. C. H. Spence, eds., *Science of Microscopy* (Springer New York, New York, NY, 2007).

Hepp, C., T. Müller, V. Waselowski, J. N. Becker, B. Pingault, H. Sternschulte, D. Steinmüller-Nethl, A. Gali, J. R. Maze, M. Atatüre and C. Becher, "Electronic structure of the silicon vacancy color center in diamond", Phys. Rev. Lett. **112**, 036405 (2014).

Hertkorn, J., J. Wrachtrup and M. Fyta, "Vacancy defect centers in diamond: influence of surface termination", The European Physical Journal Special Topics **227**, 14, 1591–1601 (2019).

Hillier, J. and R. F. Baker, "Microanalysis by Means of Electrons", Journal of Applied Physics **15**, 9, 663–675 (1944).

Hobson, E., *The Theory of Spherical and Ellipsoidal Harmonics* (Cambridge University Press, 1931).

Hohenberg, P. and W. Kohn, "Inhomogeneous Electron Gas", Phys. Rev. **136**, B864–B871 (1964).

Hu, B. Y., "Kramers–Kronig in two lines", American Journal of Physics **57**, 9, 821–821 (1989).

Hu, J. and C. R. Menyuk, "Understanding leaky modes: slab waveguide revisited", Adv. Opt. Photon. **1**, 1, 58–106 (2009).

Hubbard, J., "The dielectric theory of electronic interactions in solids", Proceedings of the Physical Society. Section A **68**, 11, 976–986 (1955).

Illman, B. L., V. E. Anderson, R. J. Warmack and T. L. Ferrell, "Spectrum of surface-mode contributions to the differential energy-loss probability for electrons passing by a spheroid", Phys. Rev. B **38**, 3045–3049 (1988).

Jackson, J., *Classical Electrodynamics* (Wiley, 1999).

Jelley, J. V., "Cerenkov radiation and its applications", British Journal of Applied Physics **6**, 7, 227–232 (1955).

Jellison, G., "Optical functions of silicon determined by two-channel polarization modulation ellipsometry", Optical Materials **1**, 1, 41 – 47 (1992).

Johnson, J. M., S. Im, W. Windl and J. Hwang, "Three-dimensional imaging of individual point defects using selective detection angles in annular dark field scanning transmission electron microscopy", Ultramicroscopy **172**, 17 – 29 (2017).

Kho, K. W., U. S. Dinish, A. Kumar and M. Olivo, "Frequency shifts in sers for biosensing", ACS nano **6**, 6, 4892–4902 (2012).

Khrennikov, A., B. Nilsson, S. Nordebo and I. Volovich, "On the quantization of the electromagnetic field of a layered dielectric waveguide", AIP Conference Proceedings **1508**, 1, 285–299 (2012).

Kimoto, K., "Practical aspects of monochromators developed for transmission electron microscopy", Microscopy **63**, 5, 337–344 (2014).

Kliewer, K. L. and R. Fuchs, "Optical modes of vibration in an ionic crystal slab including retardation. I. Nonradiative region", Phys. Rev. **144**, 495–503 (1966a).

Kliewer, K. L. and R. Fuchs, "Optical modes of vibration in an ionic crystal slab including retardation. II. Radiative region", Phys. Rev. **150**, 573–588 (1966b).

Kohl, H. and L. Reimer, *Transmission Electron Microscopy: Physics of Image Formation* (Springer, New York, NY, 2008).

Kohn, W. and L. J. Sham, "Self-Consistent Equations Including Exchange and Correlation Effects", Phys. Rev. **140**, A1133–A1138 (1965).

Konečná, A., T. Neuman, J. Aizpurua and R. Hillenbrand, "Surface-enhanced molecular electron energy loss spectroscopy", ACS Nano **12**, 5, 4775–4786 (2018).

Kordahl, D., "Pop Goes the Physics", The New Atlantis **1**, 52, 87–99 (2017).

Kordahl, D., "Did Thomas Kuhn Kill Truth?", The New Atlantis **1**, 55, 109–122 (2018).

Kordahl, D., "Steven Weinberg Glimpses the Promised Land", The New Atlantis **1**, 57, 120–128 (2019).

Kordahl, D. and C. Dwyer, "Enhanced vibrational electron energy-loss spectroscopy of adsorbate molecules", Phys. Rev. B **99**, 104110, URL `https://link.aps.org/doi/10.1103/PhysRevB.99.104110` (2019).

Kordahl, D., L. W. Q. Xu, S. L. Y. Chang and C. Dwyer, "Prospects for detecting individual defect centers using spatially resolved electron energy loss spectroscopy", Phys. Rev. B **100**, 134103 (2019).

Kresse, G. and J. Furthmüller, "Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set", Computational Materials Science **6**, 1, 15 – 50 (1996).

Kresse, G. and J. Furthmüller, "Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set", Phys. Rev. B **54**, 11169–11186 (1996).

Kresse, G. and J. Hafner, "Ab initio molecular dynamics for liquid metals", Phys. Rev. B **47**, 558–561 (1993).

Kresse, G. and J. Hafner, "Ab initio molecular-dynamics simulation of the liquid-metal–amorphous-semiconductor transition in germanium", Phys. Rev. B **49**, 14251–14269 (1994).

Kresse, G. and D. Joubert, "From ultrasoft pseudopotentials to the projector augmented-wave method", Phys. Rev. B **59**, 1758–1775 (1999).

Krivanek, O. L., T. C. Lovejoy, N. Dellby, T. Aoki, R. W. Carpenter, P. Rez, E. Soignard, J. Zhu, P. E. Batson, M. Lagos, R. F. Egerton and P. A. Crozier, "Vibrational spectroscopy in the electron microscope", Nature **514**, 209–212 (2014).

Krivanek, O. L., J. P. Ursin, N. J. Bacon, G. J. Corbin, N. Dellby, P. Hrncirik, M. F. Murfitt, C. S. Own and Z. S. Szilagyi, "High-energy-resolution monochromator for aberration-corrected scanning transmission electron microscopy/electron energy-loss spectroscopy", Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **367**, 1903, 3683–3697 (2009).

Lagos, M. J., A. Trügler, V. Amarasinghe, L. C. Feldman, U. Hohenester and P. E. Batson, "Excitation of long-wavelength surface optical vibrational modes in films, cubes and film/cube composite system using an atom-sized electron beam", Microscopy **67**, 1 (2018).

Lagos, M. J., A. Trügler, U. Hohenester and P. E. Batson, "Mapping vibrational surface and bulk modes in a single nanocube", Nature **543**, 529–532 (2017).

Landau, L. and E. Lifshitz, *Electrodynamics of continuous media* (Pergamon Press, 1960).

Larsson, J. A. and P. Delaney, "Electronic structure of the nitrogen-vacancy center in diamond from first-principles theory", Phys. Rev. B **77**, 165201 (2008).

Lenef, A. and S. C. Rand, "Electronic structure of the N-V center in diamond: Theory", Phys. Rev. B **53**, 13441–13455 (1996).

Leng, J., J. Opsal, H. Chu, M. Senko and D. E. Aspnes, "Analytic representations of the dielectric functions of crystalline and amorphous Si and crystalline Ge for very large scale integrated device and structural modeling", Journal of Vacuum Science & Technology A **16**, 3, 1654–1657 (1998).

Limonov, M. F., M. V. Rybin, A. N. Poddubny and Y. S. Kivshar, "Fano resonances in photonics", Nature Photonics **11**, 543–554 (2017).

Liu, K., "Relation between charge density and curvature of surface of charged conductor", American Journal of Physics **55**, 9, 849–852 (1987).

Liu, Z., K. Suenaga, Z. Wang, Z. Shi, E. Okunishi and S. Iijima, "Identification of active atomic defects in a monolayer of tungsten disulfide nanoribbon", Nature Communications **2** (2011).

Lourenço-Martins, H. and M. Kociak, "Vibrational surface electron-energy-loss spectroscopy probes confined surface-phonon modes", Phys. Rev. X **7**, 041059 (2017).

Lucas, A. and M. Šunjić, "Fast-electron spectroscopy of collective excitations in solids", Progress in Surface Science **2**, 75 – 137 (1972).

Lucas, A. A. and E. Kartheuser, "Energy-loss spectrum of fast electrons in a dielectric slab. I. Nonretarded losses and Cherenkov bulk loss", Phys. Rev. B **1**, 3588–3598 (1970).

Lucas, A. A., E. Kartheuser and R. G. Badro, "Electron-phonon interaction in dielectric films. Application to electron energy loss and gain spectra", Phys. Rev. B **2**, 2488–2499 (1970).

Manson, N. B., J. P. Harrison and M. J. Sellars, "Nitrogen-vacancy center in diamond: Model of the electronic structure and associated dynamics", Phys. Rev. B **74**, 104303 (2006).

Mansuripur, M., M. Kolesik and P. Jakobsen, "Leaky modes of dielectric cavities", in "Spintronics IX", edited by H.-J. Drouhin, J.-E. Wegrowe and M. Razeghi, vol. 9931, pp. 14 – 33, International Society for Optics and Photonics (SPIE, 2016).

Marcuse, D., *Theory of Dielectric Optical Waveguides*, Quantum Electronics–Principles & Applications (Academic Press, 1991).

Maze, J. R., A. Gali, E. Togan, Y. Chu, A. Trifonov, E. Kaxiras and M. D. Lukin, "Properties of nitrogen-vacancy centers in diamond: the group theoretic approach", New Journal of Physics **13**, 2, 025025 (2011).

Mittal, A. and K. A. Mkhoyan, "Limits in detecting an individual dopant atom embedded in a crystal", Ultramicroscopy **111**, 8, 1101 – 1110 (2011).

Mittiga, T., S. Hsieh, C. Zu, B. Kobrin, F. Machado, P. Bhattacharyya, N. Z. Rui, A. Jarmola, S. Choi, D. Budker and N. Y. Yao, "Imaging the local charge environment of nitrogen-vacancy centers in diamond", Phys. Rev. Lett. **121**, 246402 (2018).

Miyata, T., M. Fukuyama, A. Hibara, E. Okunishi, M. Mukai and T. Mizoguchi, "Measurement of vibrational spectrum of liquid using monochromated scanning transmission electron microscopy–electron energy loss spectroscopy", Microscopy **63**, 377–382 (2014).

Moliver, S. S., "Electronic structure of neutral silicon-vacancy complex in diamond", Technical Physics **48**, 11, 1449–1453 (2003).

Momma, K. and F. Izumi, "*VESTA3* for three-dimensional visualization of crystal, volumetric and morphology data", Journal of Applied Crystallography **44**, 6, 1272–1276 (2011).

Mozrzymas, M., M. Studziński and M. Horodecki, "Explicit constructions of unitary transformations between equivalent irreducible representations", Journal of Physics A: Mathematical and Theoretical **47**, 50, 505203 (2014).

Mulvey, T., "Origins and historical development of the electron microscope", British Journal of Applied Physics **13**, 5, 197–207 (1962).

Neumann, P., N. Mizuochi, F. Rempp, P. Hemmer, H. Watanabe, S. Yamasaki, V. Jacques, T. Gaebel, F. Jelezko and J. Wrachtrup, "Multipartite entanglement among single spins in diamond", Science **320**, 5881, 1326–1329 (2008).

Nicholls, R., F. S. Hage, J. Yates, D. McCulloch, D. M. Kepaptsoglou, T. C. Lovejoy, N. Dellby, O. L. Krivanek, K. Refson and Q. Ramasse, "Vibrational phonon spectroscopy of boron nitride polymorphs: a comparison between theory and experiment", Microsc. Microanal. **21 (Suppl. 3)**, 1469 (2015).

Perdew, J. P., K. Burke and M. Ernzerhof, "Generalized gradient approximation made simple", Phys. Rev. Lett. **77**, 3865–3868 (1996).

Phillips, P., M. D. Graef, L. Kovarik, A. Agrawal, W. Windl and M. Mills, "Atomic-resolution defect contrast in low angle annular dark-field STEM", Ultramicroscopy **116**, 47 – 55 (2012).

Plakhotnik, T., M. W. Doherty and N. B. Manson, "Electron-phonon processes of the nitrogen-vacancy center in diamond", Phys. Rev. B **92**, 081203 (2015).

Poladian, L., "A rigorous approach to normalization, perturbation and orthogonality of leaky modes in microstructured fibre algorithms", 2005 Pacific Rim Conference on Lasers & Electro-Optics pp. 592–593 (2005).

Potapov, P., H. Engelmann, E. Zschech and M. Stöger-Pollach, "Measuring the dielectric constant of materials from valence EELS.", Micron **402**, 262–8 (2009).

Price, R. H. and R. J. Crowley, "The lightning-rod fallacy", American Journal of Physics **53**, 9, 843–848 (1985).

Ramasse, Q. M., C. R. Seabourne, D.-M. Kepaptsoglou, R. Zan, U. Bangert and A. J. Scott, "Probing the bonding and electronic structure of single atom dopants in graphene with electron energy loss spectroscopy", Nano Letters **13**, 10, 4989–4995 (2013).

Rez, P., T. Aoki, K. March, D. Gur, O. L. Krivanek, N. Dellby, T. C. Lovejoy, S. G. Wolf and H. Cohen, "Damage-free vibrational spectroscopy of biological materials in the electron microscope", Nature Comm. **7**, 10945 (2016).

Ritchie, R. and A. Howie, "Inelastic scattering probabilities in scanning transmission electron microscopy", Philosophical Magazine A **58**, 5, 753–767 (1988).

Ritchie, R. H., "Plasma losses by fast electrons in thin films", Phys. Rev. **106**, 874–881 (1957).

Ritchie, R. H., "Quantal aspects of the spatial resolution of energy-loss measurements in electron microscopy I. Broad-beam geometry", Philosophical Magazine A **44**, 4, 931–942 (1981).

Rogers, L. J., K. D. Jahnke, M. H. Metsch, A. Sipahigil, J. M. Binder, T. Teraji, H. Sumiya, J. Isoya, M. D. Lukin, P. Hemmer and F. Jelezko, "All-optical initialization, readout, and coherent preparation of single silicon-vacancy spins in diamond", Phys. Rev. Lett. **113**, 263602 (2014).

Rose, B. C., D. Huang, Z.-H. Zhang, P. Stevenson, A. M. Tyryshkin, S. Sangtawesin, S. Srinivasan, L. Loudin, M. L. Markham, A. M. Edmonds, D. J. Twitchen, S. A. Lyon and N. P. de Leon, "Observation of an environmentally insensitive solid-state spin defect in diamond", Science **361**, 6397, 60–63 (2018).

Rose, H., "Zur systematischen berechnung elektronenoptischer bildfehler", Optik; Zeitschrift fur das gesamte Gebiet der Licht- und Elektronenoptik. **33**, 151, 25 (1971).

Rose, H. H., "Optics of high-performance electron microscopes", Science and Technology of Advanced Materials **9**, 1, 014107 (2008).

Ru, E. L. and P. Etchegoin, "Rigorous justification of the $|E|^4$ enhancement factor in Surface Enhanced Raman Spectroscopy", Chemical Physics Letters **423**, 1, 63 – 66 (2006).

Ruska, E., "The Development of the Electron Microscope and of Electron Microscopy", Nobel Lecture (1986).

Sammut, R. and A. W. Snyder, "Leaky modes on circular optical waveguides", Appl. Opt. **15**, 2, 477–482 (1976).

Schattschneider, P., C. Hébert, H. Franco and B. Jouffrey, "Anisotropic relativistic cross sections for inelastic electron scattering, and the magic angle", Phys. Rev. B **72**, 045142 (2005).

Scherzer, O., "The theoretical resolution limit of the electron microscope", Journal of Applied Physics **20**, 1, 20–29 (1949).

Schmeits, M., "Inelastic scattering of fast electrons by spherical surfaces", Journal of Physics C: Solid State Physics **14**, 8, 1203 (1981).

Schönleber, M., D. Klotz and E. Ivers-Tiffée, "A method for improving the robustness of linear kramers-kronig validity tests", Electrochimica Acta **131**, 20 – 27, electrochemical Impedance Spectroscopy (2014).

Senga, R. and K. Suenaga, "Single-atom electron energy loss spectroscopy of light elements", Nature Communications **6**, 7943 (2015).

Senga, R., K. Suenaga, P. Barone, S. Morishita, F. Mauri and T. Pichler, "Position and momentum mapping of vibrations in graphene nanostructures", Nature **573**, 7773, 247–250 (2019).

Silcox, J., "Core-loss EELS", Current Opinion in Solid State and Materials Science **3**, 4, 336 – 342 (1998).

Smith, K. C., A. Olafsson, X. Hu, S. C. Quillin, J. C. Idrobo, R. Collette, P. D. Rack, J. P. Camden and D. J. Masiello, "Direct observation of infrared plasmonic fano antiresonances by a nanoscale electron probe", Phys. Rev. Lett. **123**, 177401 (2019).

Smythe, W., *Static and Dynamic Electricity, 2nd Ed.* (McGraw-Hill, 1950).

Snitzer, E., "Cylindrical dielectric waveguide modes", J. Opt. Soc. Am. **51**, 5, 491–498 (1961).

Snyder, A. W. and D. J. Mitchell, "Leaky rays on circular optical fibers", J. Opt. Soc. Am. **64**, 5, 599–607 (1974).

Sorini, A., *The passage of fast electrons through matter*, Ph.D. thesis, University of Washington (2008).

Spence, J., *High-Resolution Electron Microscopy*, Monographs on the Physics and Chemistry of Materials (OUP Oxford, 2009).

Stratton, J. A., *Electromagnetic Theory* (McGraw-Hill, New York, 1941), 1st edn.

Talebi, N., "Electron-light interactions beyond the adiabatic approximation: recoil engineering and spectral interferometry", Advances in Physics: X **3**, 1, 1499438 (2018).

Tizei, L. H. G., V. Mkhitaryan, H. Lourenço-Martins, L. Scarabelli, K. Watanabe, T. Taniguchi, M. Tencé, J.-D. Blazit, X. Li, A. Gloter, A. Zobelli, F.-P. Schmidt, L. M. Liz-Marzán, F. J. García de Abajo, O. Stéphan and M. Kociak, "Tailored nanoscale plasmon-enhanced vibrational electron spectroscopy", Nano Letters (2020).

Van Bladel, J., "On Helmholtz's theorem in finite regions", Midwestern Universities Research Association Report **440** (1958).

Varela, M., M. P. Oxley, W. Luo, J. Tao, M. Watanabe, A. R. Lupini, S. T. Pantelides and S. J. Pennycook, "Atomic-resolution imaging of oxidation states in manganites", Phys. Rev. B **79**, 085117 (2009).

Venkatraman, K., B. D. A. Levin, K. March, P. Rez and P. A. Crozier, "Vibrational spectroscopy at atomic resolution with electron impact scattering", Nature Physics **15**, 12, 1237–1241 (2019).

Verma, P., "Tip-Enhanced Raman Spectroscopy: Technique and Recent Advances", Chemical Reviews **117**, 9, 6447–6466, pMID: 28459149 (2017).

Wang, S., K. March, F. A. Ponce and P. Rez, "Identification of point defects using high-resolution electron energy loss spectroscopy", Phys. Rev. B **99**, 115312 (2019).

Wang, Z., "Valence electron excitations and plasmon oscillations in thin films, surfaces, interfaces and small particles", Micron **27**, 3, 265 – 299 (1996).

Weber, J. R., W. F. Koehl, J. B. Varley, A. Janotti, B. B. Buckley, C. G. Van de Walle and D. D. Awschalom, "Quantum computing with defects", Proceedings of the National Academy of Sciences **107**, 19, 8513–8518 (2010).

Yang, H. U., J. D'Archangel, M. L. Sundheimer, E. Tucker, G. D. Boreman and M. B. Raschke, "Optical dielectric function of silver", Phys. Rev. B **91**, 235137 (2015).

Yang, S. and J.-M. Song, "Analysis of Guided and Leaky TM0n and TE0n modes in circular dielectric waveguide", Progress in Electromagnetics Research B **66**, 143–156 (2016).

Yeh, C. and F. I. Shimabukuro, *The Essence of Dielectric Waveguides* (Springer US, Boston, MA, 2008).

Zabala, N., E. Ogando, A. Rivacoba and F. J. García de Abajo, "Inelastic scattering of fast electrons in nanowires: A dielectric formalism approach", Phys. Rev. B **64**, 205410 (2001).

Zaretskaya, G., A. V. Drozdovskii, A. B. Ustinov and B. A. Kalinikos, "Eigenfrequencies of whispering gallery modes of disk dielectric resonators: a dimensional quantization method", Journal of Physics: Conference Series **1038**, 012099 (2018).

Zhang, L., R. Erni, J. Verbeeck and G. Van Tendeloo, "Retrieving the dielectric function of diamond from valence electron energy-loss spectroscopy", Phys. Rev. B **77**, 195119 (2008).

Zheng, F., J. Tao and A. M. Rappe, "Frequency-dependent dielectric function of semiconductors with application to physisorption", Phys. Rev. B **95**, 035203 (2017).