Advanced Applications in Nanophotonics

Hao Yang

Submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2019
ABSTRACT

Advanced Applications in Nanophotonics

Hao Yang

Nanophotonics is a fast-growing area of both scientific significance and practical value for applications. Nanophotonics studies the interaction between light and electronic systems in nanomaterials and nanostructures as well as the behavior of light in nanometer scales. It covers many hot topics such as plasmonics, two-dimensional materials, and silicon photonics. Increasing attention is given to the area and nanophotonics is expected to have significant impact on future technology advances.

This thesis work focuses on three aspects of nanophotonics. The first aspect is in exploring the nonlocal effect and surface correction for nanometer-length-scale plasmonic structures. Plasmonics is the study of the interaction between electromagnetic fields and free electrons in a metal. It exploits the unique optical properties of metallic nanostructures to enable routing and manipulation of light at the nanoscale, where nonlocal effect becomes important. Here we introduce a new surface hydrodynamic model for plasmon propagation at interfaces, which incorporates both nonlocality and surface contributions. This surface correction is calculated via a discontinuity in the normal component of the electric displacement in conjunction with Feibelman’s d-parameters, thus enabling rapid numerical calculation of nanostructures without requiring a full quantum calculation because of its large computational requirement. We examine numerical calculations of surface plasmon polaritons propagation at a single interface structure, and then for a more complex thin-film structures.
The second aspect is investigating the third-harmonic generation in thick multilayer graphene. Graphene is the first two-dimensional material to be discovered and has attracted much interest because of its remarkable two-dimensional electronic, optical, mechanical, and thermal properties. Multilayer graphene, can be seen as stacking of monolayer graphene, and it offers an array of properties that are of interest for optical physics and devices. We describe the layer-dependent for third-harmonic generation in thick multilayer graphene on quartz substrate. The third harmonic signal of multilayer graphene exhibits a complex dependence on its layer number showing that the optimal third harmonic signal at 24 layers, in good agreement with two theoretical models.

The third aspect is an exploration in silicon photonics of design and demonstration of a differential phase shift keying demodulator based on coherent perfect absorption effect. Silicon photonics is considered a potential future communication system mainly due to its compact footprint, dense integration, and compatibility with mature silicon integrated circuit manufacturing. Differential phase shift keying based system offers advantages, e.g., dispersion tolerance, improved sensitivity, and does not require coherent detection. Coherent perfect absorption uses a ring resonator works for the critical coupling condition at resonance frequency. This work shows a new compact demodulator circuit can be integrated in all optical-system.
## Contents

List of Figures ................................................................. iv

Acknowledgements ............................................................... xii

Dedication .............................................................................. xiv

Introduction ........................................................................... 1

0.1 Plasmonics ...................................................................... 3

0.2 Graphene ...................................................................... 5

0.3 Silicon Photonics ............................................................ 6

0.4 Scope of the Thesis .......................................................... 7

1 Experiment ......................................................................... 10

1.1 Numerical Methods ......................................................... 10

1.2 Theoretical Approach ...................................................... 36

1.3 Third Harmonic Generation Experiment .......................... 76

1.4 Fabrication Process ......................................................... 80

1.5 Other Tools Used ............................................................ 81
2 Numerical Surface-Corrected Nonlocal Electrodynamic Model for Nanophotonic Structure

2.1 Introduction .................................................. 84
2.2 Basic Electromagnetic Framework ............................. 88
2.3 Feibelman $d$ parameters ..................................... 90
2.4 Surface Hydrodynamic Model ............................... 91
2.5 Application to Planar Interfaces .............................. 93
2.6 Solution for a Thin Film Stack ............................... 99
2.7 Conclusion ..................................................... 103
2.8 APPENDIX A. Modified Reflection Coefficient for the Surface Hydrodynamic Model ........................................... 104
2.9 APPENDIX B. Numerical Implementation of the Surface Current ..................................................... 106

3 Layer Dependence of Third-Harmonic Generation in Thick Multilayer Graphene

3.1 Introduction .................................................. 109
3.2 Theoretical Considerations ...................................... 112
3.3 Experiment .................................................... 117
3.4 Results and Discussion ........................................ 118

4 Differential Phase-Shift Keying Demodulation by Coherent Perfect Absorption in Silicon Photonics

4.1 Introduction .................................................. 122
4.2 Device Design ................................................ 124
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.3 Device Performance</td>
<td>128</td>
</tr>
<tr>
<td>4.4 DPSK Demodulation</td>
<td>131</td>
</tr>
<tr>
<td>4.5 CPA-DPSK with Grating Coupler</td>
<td>134</td>
</tr>
<tr>
<td>4.6 Optsim Demonstration of CPA</td>
<td>136</td>
</tr>
<tr>
<td>5 Conclusion and Future Work</td>
<td>139</td>
</tr>
<tr>
<td>5.1 Conclusion</td>
<td>139</td>
</tr>
<tr>
<td>5.2 Future Works</td>
<td>141</td>
</tr>
<tr>
<td>Other Relevant Works</td>
<td>143</td>
</tr>
<tr>
<td>Bibliography</td>
<td>145</td>
</tr>
<tr>
<td>Appendix. Publication</td>
<td>159</td>
</tr>
</tbody>
</table>
List of Figures

0.1 Three aspects of nanophotonics ........................................ 3
0.2 Schematic diagram of surface plasmon polaritons. .................... 4
0.3 (a) Hexagonal honeycomb lattice of graphene with two atoms (A and B) per unit cell. (b) The 3D band structure of graphene.[45] .................. 5
0.4 Photograph of SOI photonic chip fabricated at IME A STAR.[58] ..... 7
0.5 A optical communication channel. ...................................... 7

1.1 User interface of COMSOL Multiphysics ................................. 18
1.2 Position of the electric and magnetic field vector components about a cubic unit cell of the Yee space lattice ................................. 24
1.3 Position of the electric and magnetic field vector components about a cubic unit cell of the Yee space lattice ................................. 26
1.4 User interface of Synopsys RSOFT ....................................... 36
1.5 Geometry relation of the slab model .................................... 38
1.6 Schematic of a slab system ................................................. 40
1.7 Geometry of harmonic fields arising from interface between two medium ........................................................ 42
1.8 Schematic of beam incident on an interface. ............................ 49
1.9 Diagram of a simple interface system .................................... 53
1.10 Schematic diagram of a ring resonator with two bus waveguides for coupled mode analysis. .................................................. 71
1.11 Schematic diagram of a ring resonator with two bus waveguides for matrix analysis. .................................................. 75
1.12 Energy-level description of THG process ................................. 77
1.13 Schematic diagram of the THG measurement setup .................... 78
1.14 Photograph of a custom nonlinear optical scanning microscope for SHG, THG, and photoluminescence imaging experiments. .............. 78
1.15 Labview program for nonlinear optical scanning microscope ......... 79
1.16 SEM images of (a) a delay line (b) a ring resonator ................... 81

2.1 Notional diagram of (a) the equilibrium and non-equilibrium near-surface electron density, including $d_\perp$ parameter; (b) HDM pressure-driven electrons. Panel (c) provides a schematic sketch indicating the polarization quantities used in this paper. The spill-over electron density $\delta n(r)$ via the $d$-parameter gives rise to a non-equilibrium discontinuity in the dielectric constant, which is used to describe the nonlocal response of the metal-dielectric system. .................................................. 87

2.2 (a) Plot showing the real $d_\perp'$ (solid red) and imaginary part $d_\perp''$ (dashed blue) of the Feibelman $d$-parameter $d_\perp$. (b) Calculated real $\Delta \epsilon'$ (solid red) and imaginary part $\Delta \epsilon''$ vs. wavelength (dashed blue) of the induced discontinuity in the dielectric function $\Delta \epsilon$ across a Ag/GaN interface vs. wavelength. .................................................. 94
2.3 (a) Comparison of the theoretical dispersion relations of SPPs at an Ag/GaN interface as predicted by a local dielectric model (gray line), an analytical result from [74] (green line), the hydrodynamic dielectric model (blue triangles), and the surface-hydrodynamic model (present work - red circle). For a SPP energy in the vicinity of the surface plasmon resonance frequency, e.g., 3.02eV, the surface-hydrodynamic model predicts that the effective wavelength of the SPP mode lies between that obtained with the hydrodynamic model and that from the standard local dielectric model.

(b) Plot of the simulated electric-field distribution along the interface in the presence of nonlocal polarization terms at a wavelength of 440nm using the surface-HDM model for a single Ag/GaN interface.

2.4 Comparison of the relation of SPPs at a Ag/GaN thin film stack with thickness (a) $w = 50$ nm and (b) $w = 20$ nm using the local Maxwell-equation formalism (gray), HDM (blue) and surface-HDM (red). The $x$-axis is the propagation constant $k'$ for the main figures and decay constant $k''$ for the inset figures. Note that for the bottom sketches the SPP propagation is from left to right.
3.1  (a) Linear transmission, reflection, and absorption coefficients for MLG between air and quartz substrate vs. layer number $N$ for the interface (blue curves) and slab (red curves) models. (Left inset) Interface model: MLG is treated as a single source of interface current via the surface conductivity $\sigma_{N,g}$. The TH reflected and transmitted waves, $E_{3\omega,R}$ and $E_{3\omega,T}$, are generated via the third-order conductivity $\sigma_{N,g}^{(3)}$. (Right inset) Slab model: the MLG is treated as a single slab with third-order nonlinearity $\chi^{(3)}$. Four TH waves are generated: $E_{3\omega,R}$, $E_{3\omega,T}$, $E_{3\omega,M}$ and $E_{3\omega,M'}$, which denote the reflected, transmitted, internal down- and up-going waves, respectively. (b) Reflected THG signal predicted by the interface and slab models for MLG. Inset: prediction by slab model for a large $N$.

3.2  Experimental setup. The fundamental beam from a Ti:sapphire laser (TS) passes through dichroic mirror (DC) and is focused on the MLG (G) sample via a microscope objective (MO), which generates a TH signal that is reflected by the DC into a Pellin-Broca prism (BP), reflected by mirror (M), and focused by a lens (L) into the monochromator (MC). The detection system consists of a photomultiplier tube (PMT) and a photon counter (PC). Left inset: cubic-power dependence of TH signal; right inset: TH spectrum.

3.3  (a) Optical microscopy of a typical region of a graphene sample with a large variation in layer thickness; (b) TH mapping of the same region; (c) AFM image of an $13 \times 20 \mu m^2$ region, and (d) comparison of TH and AFM line scans of AFM image indicated in (c).
3.4 Normalized layer-dependent TH generation in MLG. Circles: Experimental data. Blue and red curves: theoretical predictions using the interface and slab models, respectively. ........................................... 119

3.5 (a) MLG sample on a quartz substrate (b) Silicon/SiO$_2$ sample used for calibration. .......................................................... 120
4.1 (a) Experimental setup used to characterize the demodulator circuit. The input light from a tunable laser source (TLS) is edge coupled to the chip using a 2.5 μm spot lens-tapered fiber (LTF). A polarization rotator (PR) is used to maximize the current output. An inverted taper of 180 nm width is used to couple the light from the fiber to the SOI chip. The schematic of the photonic circuit shows the Y-branch where one signal branch is routed directly to the ring, while the other signal goes through a bit-rate-dependent delay line (DL). The two signals are coupled into the ring resonator. Four photodiodes (PDs) collect the photocurrent from the two signal buses and the two ring resonator ports. The photocurrents are measured with the help of a digital multimeter (DM) through a 150 μm pitch electrical probe. (b) Mode profile of the fundamental TE mode of the 500 nm × 220 nm Si waveguide at 1550 nm. (c) Optical microscope image of one 10 Gbps device. (d) Power transmission through the signal bus and the ring resonator as a function of phase shift (Δϕ). (e) Theoretical DPSK bit sequence and simulated resultant power absorption. The blue curve represents power through the signal bus while the red curve represents power inside the ring. When Δϕ = 0, the two inputs are identical and the ring is in resonance resulting in full power loading of the ring and extinguishing power in the signal bus. This situation is reversed for Δϕ = π.125
4.2 Measured photocurrent vs. wavelength for the signal and the RR1 port for a 10 Gbps device with a 434 nm bus and ring gap. The top right inset shows one resonant dip of the delayed signal port used to calculate the Q-factor of the device. The bottom right inset shows the photocurrent inside the ring used to calculate the wavelength change for a $2\pi$ phase shift. 130


4.4 (a) Eye diagrams for data rates of 9Gbps, 9.5Gbps, 10 Gbps, 10.5 Gbps, and 11Gbps. (b) Eye diagram with different duty ratios of 50%, 75%, and 87.5%. The minimum BER and maximum extinction ratio occurs at 10Gbps, i.e. their designed values. ........................................ 132

4.5 (a) Structure of CPA-DPSK with grating coupler (b) TE grating coupler (c) Transmission spectra of CPA-DPSK with grating coupler. ......... 135

4.6 Optsim setup for CPA-DPSK demonstration ................................. 137

4.7 (a) Original signal coming from a PRBS generator. (b) Demodulated DPSK voltage signal. (c) Eye diagram of the output DPSK demodulated voltage signal. .................................................. 138
5.1 (a) Normalized, time-averaged electric field intensity spatial distribution in response to two-color excitation for a 55 nm radius tungsten tip, and (b) the plasmonic mode intensity at the midpoint between the tip apex and the substrate as a function of the tip radius. The tip apex-substrate spacing is 2 nm in both (a) and (b).

5.2 (a) Time-resolved 2C-SPPX photoemission signal (open circles). Solid line: delay-time derivative of the 3D-FDTD-simulated 2C-SPPX signal \( \frac{d\gamma}{dt_d} \), \( \gamma \) is defined in the main text). (b) Measured STM current (circles), 2C-SPPX signal (squares) and \( \gamma \) (proportional to \( \frac{d\gamma}{dt_d} \) at fixed delay, triangles) versus tip-sample distance. Solid lines are single-exponential-decay fits.
Acknowledgements

Over the past few years at Columbia University, I always feel blessed because I have met so many nice people who have taught me, helped me, and influenced me. First of all, I would like to give my sincerest thanks to my advisor, Prof. Richard Osgood, for his continuous support of my Ph.D. study, and for his passion, enthusiasm and immense knowledge that has deeply inspired me. Then I would like to acknowledge all my group members. I would also like to thank Dr. Xiang Meng, who gave me a lot of suggestions in both research and life. I would like to thank Dr. Jerry Dadap, who gave me a lot of guidance in the research, e.g. laser operation, writing papers. He never hesitated to share any new ideas he may have that is relevant to our work. And he contributed a lot in the projects. I would like to thank Asif Ahmed, who worked together with me in one of the projects. We went to BNL and West Point for experiments together. I still remember a lot of the moments that we shared on the bus and in the labs. I would like to thank Dr. Brian Souhan, and LTC Kirk Ingold, who offered great help for organizing the experiment at West Point and taught us essential techniques like edge coupling. I would like to thank Christine Chen and Abram Nathan, who shared a lot of the experience in the setting up the high-speed measurement experiment. I would like to thank Prof. Keren Bergman for allowing us
to use her laboratory. I would also like to thank Yi Lin, Dr. Wencan Jin, Honghua Guan, Yawen Sun and all the people who worked with me.

I would also like to thank the rest of my thesis committee: Prof. Irving Herman, Prof. John Kymissis, Prof. Christine Hendon and Prof. Yuan Yang, for their time, interest and valuable comments on my thesis.

Finally, I give my deepest gratitude to my family, who gives me caring, love, and support through my life.

Last but not the least, there are many other people who have helped me a lot but did not appear in my papers. I would also thank a lot of friends in my life: Xiaoqi Lu, Zhong Zhuang, Qian Huang, Wei Huang, Fengqi Zhang, Yu Cheng, Siyan Wang and so many people that make my life colorful and memorable.
Dedication

To my beloved parents,

Fan Yang, Jiangchun Sui

for their wisdom, patience and open-mindedness.
Human made efforts to understand and manipulate light for many thousands of years. In ancient Greek of 2nd century AD, it was written that Archimedes has utilized mirrors acting collectively as a parabolic reflector to burn ships attacking Syracuse. Although Archimedes “death ray” might just be a legendary, the effort and progress in understanding and manipulating lights remains through the long history and keeps advancing from the ray optics, to wave description of light, then to quantum description of light.

Recent advancement in fabrication and characterization techniques allows optical structures to be scaled down to nanoscale size or even to the atomic level [1–7]. These nanostructures can have unique, controllable, and tunable optical properties and their interactions with quantum materials can have important near-field and far-field optical response [8, 9]. Undoubtedly, these optical properties can have many important applications, ranging from the efficient and tunable light sources [10–14], detectors [15–17], filters [18, 19], modulators [20–22], high-speed all-optical switches [23, 24]; to the next-generation classical and quantum computation, and biophotonic medical sensors [25–27]. Interest in nanoscience is a realization of a famous statement by Feynman that “There’s Plenty of Room at the Bottom” (Feynman, 1961) [28]. He
pointed out that if one takes a length and divides it into several pieces for many times, it will give you billions of segments to manipulate. This emerging research of nanoscience, known as nanophotonics, is a highly interdisciplinary field requiring expertise in materials science, physics, electrical engineering, and scientific computing, modeling and simulation. It has also become an important research field for investigating the science and engineering of light-matter interactions that take place on wavelength and subwavelength scales where the nature of the nanostructured matter controls the interactions.

Nanophotonics can conceptually be divided into three aspects: nanoscale confinement of radiation, nanoscale confinement of matter, and nanoscale photoprocesses (Shen et al., 2000) [29]. The first aspect, nanoscale confinement of radiation, is to confine light to nanoscale dimensions that are much smaller than the wavelength of light, which usually pays attention to the control of propagation properties and its interaction with matter. The second aspect, nanoscale confinement of matter, is to confine matter to nanoscale dimensions, thereby limiting interactions between light and matter to nanoscopic dimensions. This aspect usually pays attention to the control of the optical properties, excitation dynamics and energy transfer through the confinement of matter. The third aspect, nanoscale photoprocess, which usually pays attention to the control of spatial confinement of photochemical and photophysical processes, thus which can be used for nanolithography to fabricate nanostructures. Plasmonics, graphene, and silicon photonics are currently three of the most important areas in the field of nanophotonics, which correspond to nanoscale confinement of radiation, nanoscale confinement of matter, and nanoscale photoprocesses respec-
Figure 0.1: Three aspects of nanophotonics

tively.

0.1 Plasmonics

Plasmonics is a very important concept in the field of nanophotonics, in which we are concerned primarily with the manipulation of light at the nanoscale. Surface plasmons are essentially light waves that are trapped on a surface of a conductor due to their interaction with free electrons near the surface. This mode occurs at the interface of a material exhibiting positive real part of their relative permittivity, e.g. vacuum, air, glass and other dielectrics, and a material whose real part of permittivity is negative at the given frequency of light, typically a metal or heavily doped semiconductors. In this interaction, the free electrons respond collectively by oscillating in resonance with the light wave. The resonant interaction between the surface charge oscillation and the electromagnetic field of the light constitutes the surface plasmon (SP) and gives rise to its unique properties. The surface plasmon has a propagation vector parallel to the interface, while its amplitude decays exponentially in the direction orthogonal to the surface. Unlike pure electromagnetic waves, surface plasmons can be localized
to subwavelength dimensions in the plane perpendicular to the propagation direction, providing a viable route to nanoscale optics.

Much of today’s research is aimed at structures that provide additional localization, such as localized surface plasmons of nano metal particles [30–32], and guiding structures utilizing surface plasmon polaritons [33–35]. Localization of the electromagnetic fields at the nanoscale also yields a dramatic decrease in footprint and increases in the field intensity, thus suggesting the use of surface plasmons in application, such as waveguides [33–35], scanning near-field microscopy [36, 37], surface plasmon laser [38, 39], and photoluminescence enhancement of quantum emitters [40]. Nanophotonic applications based on the use of surface plasmons can be applied for a large variety of tasks, through the design and manipulation of the geometry of metallic structures, and consequently their specific plasmon-resonant or plasmon-propagating properties.
0.2 Graphene

The interest in two-dimensional (2D) materials and materials physics has also grown dramatically over the past decade. The family of two-dimensional materials, which includes graphene, transition metal dichalcogenides, phosphorene, hexagonal boron nitride, etc., can be fabricated into atomically thin films since the intralayer bonding arises from their strong covalent character, while the interlayer interaction is mediated by weak van der Waals forces. Graphene is the first two-dimensional material to be discovered [41] and has attracted much interest. Graphene can be described as a one-atomic-layer thick graphite [42]. It is also the basic structural element of other allotropes, including graphite, nanotubes and fullerenes. Because of its 2D structure, it has linear dispersion near K point in its electron band structure, and thus a massless Dirac Fermion [43, 44]. This state intrinsically induces its zero bandgap in between. Multilayer graphene (MLG) can be seen as a stacking of monolayer graphene. Thus

Figure 0.3: (a) Hexagonal honeycomb lattice of graphene with two atoms (A and B) per unit cell. (b) The 3D band structure of graphene.[45]

it is also a transition state between graphene and graphite. MLG also offers an
array of properties that are of interest for optical physics and applications, including strong optical absorption for visible and IR wavelengths [46], strong doping-dependent absorption edge and pronounced excitonic effects near M point [47, 48]. Besides the linear effect in graphene, there is interest in developing nonlinear optical effects in graphene of all degrees of thickness, and in optimization or enhancement of its nonlinear optical response.

### 0.3 Silicon Photonics

Silicon photonics is currently at the same early stage of expansion as electronics was in the 1970s, but with a major advantage for chip fabrication: existing silicon foundries that produce highly controlled wafers for microelectronics already exist [49]. The same foundries and processes that were developed to build transistors are being re-purposed to build chips that can generate [50–52], detect [53, 54], modulate [55–57], and otherwise manipulate light. There are a number of applications that are emerging for complex silicon photonic systems, the most common being data communication. A communication channel consists of a transmitter, transmission channel, and a receiver. The transmitter converts information into an energy form appropriate for the transmission channel. The channel carries the energy, but also distorts the signal and adds noise. Following detection, the receiver regenerates the information in a nearly identical form to the original signal. The application of silicon photonics in data communication includes high-bandwidth digital communications for short-reach applications, complex modulation schemes and coherent communications
for long-reach applications, and so on.

![Figure 0.5: A optical communication channel.](image)

### 0.4 Scope of the Thesis

The research fields in nanophotonics are vast and extensive; only few of them have been mentioned above. In this dissertation, three aspects of nanophotonics are analyzed. Chapter 1 discusses our numerical calculation tools, theoretical tools, and experimental tools used in the following chapter for exploring the application of nanophotonics. Chapter 2 explores nonlocal effect and surface correction for nanometer-length-scale plasmonic structures. Recent studies have shown that in plas-
monic guided-wave structures accurate simulation of polaritonic physics is limited in many cases by the importance of spatial dispersion or nonlocal polarization, particularly for frequencies near the surface plasmon resonance frequency. We introduce a new surface hydrodynamic model for plasmon propagation at interfaces, which incorporates both nonlocality and surface contributions. This surface correction is calculated via a discontinuity in the normal component of the electric displacement in conjunction with Feibelman’s d-parameters, thus enabling rapid numerical calculation of nanostructures without exploiting a full quantum calculation because of its large computational requirement. We examine numerical calculations of SPP propagation at a single interface structure, and then in more complex thin-film structures. Chapter 3 investigates the third-harmonic generation in thick multilayer graphene. Graphene has large third-order optical nonlinearities and therefore has potential applications in optoelectronics and optical devices. While other third-order nonlinear optical processes have been demonstrated in few layer graphene, there has been a lack of investigation with regards to much thicker graphene films. We investigate layer-dependent third-harmonic (TH) generation in thick multilayer graphene on quartz substrate, wherein the layer number N varies from 3 to 50. The results are also compared with two theoretical models. Chapter 4 is the design and demonstration of coherent perfect absorption (CPA) effect for differential phase shift keying (DPSK) demodulator circuit using a ring resonator. Coherent perfect absorption uses a ring resonator working at the critical coupling condition at resonance frequency. DPSK-based system offers advantages, e.g., dispersion tolerance, improved sensitivity, and no need of coherent detection. This work shows a new compact demodulator circuit
can be integrated in all optical-system.
Chapter 1

**Experiment**

1.1 Numerical Methods

Maxwell’s partial differential equations of electrodynamics, formulated around 1870, represent a fundamental unification of electric and magnetic fields predicting electromagnetic wave phenomena which Nobel Laureate Richard Feynman has recalled as the most outstanding achievement of 19th-century science. Contemporary engineers and scientists worldwide use computers ranging from simple desktop machines to massively parallel arrays of processors to obtain solutions of these equations for the purpose of investigating electromagnetic wave guiding, radiation, and scattering phenomena and technologies. Maxwell’s equation in time and frequency domain are both popular, and versatile enough to deal with a variety of problems. In this section, we give a review about the select numerical methods used in nanophotonics.

Due to the complex nature of the light wave interaction and the ultra-small scale of the photonic components, analytical solutions of Maxwell’s equations in most real-world cases may not exist. Thus experimental studies rely heavily on numerical analysis to provide guidance both for the design of the photonic components as well as for the interpretation of their performance prior to fabrication. In most cases, one
should first develop a quantitative theoretical description of the photonic systems using advanced computational techniques, which requires solving the corresponding partial differential equations numerically. In a broad sense, there are two categories of modeling methods: finite-element method (FEM) and finite-difference method (FDM), as well as two categories of equation solving techniques: frequency-domain solvers, and time-domain simulations. In this section, we briefly present an overview of the modeling methods and solving techniques.

**Finite Element Method**

The finite element method (FEM) is a numerical technique for obtaining approximate solution to boundary-value problems of mathematical physics. The method was originally developed for mechanical and structural analysis in the 1950s. It became popular in solving the vector electromagnetic problems after an important breakthrough occurred in the 1980s namely the development of edge-based vector element [59, 60]. In principle, FDM finds an approximation to the differential operators, and then use these difference equations to solve for the fields at each grid; while the FEM makes an approximation to the solution of the differential equation over the domain of the problem, and then tailors that approximation to minimize its difference with the exact solution. The FEM is a numerical procedure to convert partial differential equations into a set of linear algebraic equations to obtain approximate solutions to boundary-value problems. In particular, it divides the simulation space into small areas or volumes, which can be arbitrarily shaped and oriented; for this reason, the
FEM is well suited to problems with complex geometry. The solution to Maxwell’s equations over each subdomain is then approximated with some functional form, usually a low-order polynomial. The solutions in each subdomain are then made to be continuous across their boundaries, and the solution must be made to fit with the global boundary conditions. The primary reasons for using FEM for electromagnetic problems are its geometric flexibility and the ability to work in higher orders of accuracy. Geometric flexibility arises because the grid in FEM can use arbitrary polygons or polyhedral (in 2D or 3D, respectively), and these can be designed to match the shapes of objects in the simulation space.

In this section, we illustrate the basic principle of the FEM by briefly introducing the methods for solving the boundary-value problems in mathematical modeling. We then present the formulation procedure of the FEM to solve the electromagnetic problems in frequency domain. In the end, we introduce a FEM based software — COMSOL.

**Ritz Method and Galerkin’s Method**

The Ritz method is a variational method, in which the boundary-value problem is formulated into its variational functional. The minimum of this functional corresponds to the governing differential equation under the given boundary conditions. The approximate solution is then obtained by minimizing the functional with respect to variables that approximately expand into the solution.
To illustrate this procedure, we first define the inner product denoted by

$$\langle \phi, \psi \rangle = \int_{\Omega} \phi \psi^* d\Omega$$  \hspace{1cm} (1.1)

where the asterisk denotes the complex conjugate. A typical boundary-value problem can be defined by a governing differential equation in a domain \(\Omega\) as

$$L\phi = f$$  \hspace{1cm} (1.2)

then the solution to (1.2) can be obtained by minimizing the functional

$$F(\tilde{\phi}) = \frac{1}{2} \langle L\tilde{\phi}, \tilde{\phi} \rangle - \frac{1}{2} \langle \tilde{\phi}, f \rangle - \frac{1}{2} \langle f, \tilde{\phi} \rangle$$  \hspace{1cm} (1.3)

with respect to the trial function \(\tilde{\phi}\) [61]. Suppose the solution can be approximated by the expansion

$$\tilde{\phi} = \sum_{j=1}^{N} c_j v_j = \{c\}^T \{v\}$$  \hspace{1cm} (1.4)

where \(v_j\) are the expansion functions defined over the entire domain. To minimize \(F(\tilde{\phi})\), we force its partial derivatives with respect to \(c_i\) to vanish. This yields a set of linear algebraic equations

$$\frac{\partial F}{\partial c_i} = \frac{1}{2} \int_{\Omega} v_i L\{v\}^T d\Omega \{c\} + \frac{1}{2} \{c\}^T \int_{\Omega} \{v\} Lv_i d\Omega - \int_{\Omega} v_i f d\Omega = 0 \hspace{1cm} i = 1, 2, 3, ..., N$$  \hspace{1cm} (1.5, 1.6, 1.7)
which can be written as the matrix equation

\[ [S]{c} = {b} \]  \hspace{1cm} (1.8)

with the elements in \([S]\) given by

\[ S_{ij} = \frac{1}{2} \int_{\Omega} (v_iLv_j + v_jLv_i) d\Omega \]  \hspace{1cm} (1.9)

and the element in \([b]\)

\[ b_i = \int_{\Omega} v_i f d\Omega \]  \hspace{1cm} (1.10)

Galerkin’s Method is a weighted residual method, which is also called the weak form formulation. It seeks a solution by weighting the residual of the differential equation over the entire domain, thus enforcing the condition

\[ R_i = \int_{\Omega} w_i (L\tilde{\phi} - f) = 0 \]  \hspace{1cm} (1.11)

Usually the weighted functions are chosen to be the same as expansion functions of the approximate solution \(w_i = v_i\). So that

\[ R_i = \int_{\Omega} (v_i L\{v\}^T\{c\} - v_i f) d\Omega = 0 \]  \hspace{1cm} (1.12)

This again leads to the matrix system in (1.8).
The Basic Steps

Thus, a finite element analysis of a boundary-value problem should include the following basic steps:

1. Discretization of subdivision of the domain
2. Selection of the interpolation function
3. Formulation of the system of equations
4. Solution of the system of equations

FEM in Electrostatic Problems

The finite element formulation can be applied to all electrostatic problems, which amount to solving the Poisson equation

\[-\nabla \cdot (\epsilon_r \nabla \phi) = \rho/\epsilon_0\] (1.13)

Boundary conditions that are often used in electrostatic problems are Dirichlet condition \( \phi = p \), which specifies the potential on conductor surfaces and the homogeneous Neumann condition \( \frac{\partial \phi}{\partial n} = 0 \), which applies at the plane of symmetry. Once the potential \( \phi \) is found by the finite element method, the electric field can be obtained from

\[E = -\nabla \phi\] (1.14)
FEM in Time-Harmonic EM Field Problems

In the time-harmonic field problems, we are dealing with either the electric field formulation

$$\nabla \times \left( \frac{1}{\mu_r} \nabla \times \mathbf{E} \right) - k_0^2 \epsilon_r \mathbf{E} = -jk_0 Z_0 \mathbf{J} \quad (1.15)$$

or with the magnetic field formulation

$$\nabla \times \left( \frac{1}{\epsilon_r} \nabla \times \mathbf{H} \right) - k_0^2 \mu_r \mathbf{H} = \nabla \times \left( \frac{1}{\epsilon_r} \mathbf{J} \right) \quad (1.16)$$

Boundary conditions often encountered are electrical conducting surface condition

$$\hat{n} \times \mathbf{E} = 0 \quad (1.17)$$
$$\hat{n} \times \nabla \times \mathbf{H} = 0 \quad (1.18)$$

magnetically conducting surface condition

$$\hat{n} \times \nabla \times \mathbf{E} = 0 \quad (1.19)$$
$$\hat{n} \times \mathbf{H} = 0 \quad (1.20)$$

and the third kind condition

$$\frac{1}{\mu_r} \hat{n} \times (\nabla \times \mathbf{E}) + \gamma_e \hat{n} \times (\hat{n} \times \mathbf{E}) = \mathbf{U} \quad (1.21)$$
$$\frac{1}{\epsilon_r} \hat{n} \times (\nabla \times \mathbf{H}) + \gamma_h \hat{n} \times (\hat{n} \times \mathbf{H}) = \mathbf{V} \quad (1.22)$$
where \( \gamma_e, \gamma_h, U \) and \( V \) are known parameters. The third kind conditions are usually used to represent impedance boundary condition and Sommerfeld radiation conditions. The continuity conditions at the interface are

\[
\hat{n} \times \mathbf{E}^+ = \hat{n} \times \mathbf{E}^-	ag{1.23}
\]

\[
\hat{n} \times \mathbf{H}^+ = \hat{n} \times \mathbf{H}^-	ag{1.24}
\]

which may also be written as

\[
\frac{1}{\mu_r^+} \hat{n} \times (\nabla \times \mathbf{E}^+) = \frac{1}{\mu_r^-} \hat{n} \times (\nabla \times \mathbf{E}^-)	ag{1.25}
\]

\[
\frac{1}{\varepsilon_r^+} \hat{n} \times (\nabla \times \mathbf{H}^+) = \frac{1}{\varepsilon_r^-} \hat{n} \times (\nabla \times \mathbf{H}^-)	ag{1.26}
\]

In accordance with the variational principle, the electric field can be obtained by solving the stationary point of the functional

\[
F(\mathbf{E}) = \frac{1}{2} \iiint_V \left[ \frac{1}{\mu_r} (\nabla \times \mathbf{E}) \cdot (\nabla \times \mathbf{E})^* - k_0^2 \varepsilon_r \mathbf{E} \cdot \mathbf{E}^* \right] dV
\]

\[
+ \frac{1}{2} \iint_{S_2} [\gamma_e (\hat{n} \times \mathbf{E}) \cdot (\hat{n} \times \mathbf{E})^* + \mathbf{E}^* \cdot \mathbf{U} + \mathbf{E} \cdot \mathbf{U}^*] dS
\]

\[
+jk_0 Z_0 \frac{1}{2} \iiint_V (\mathbf{E}^* \cdot \mathbf{J} - \mathbf{E} \cdot \mathbf{J}^*)
\]

(1.27)
and

\[
F(H) = \frac{1}{2} \iiint_V \left[ \frac{1}{\epsilon_r} (\nabla \times H) \cdot (\nabla \times H)^* - k_0^2 \mu_r H \cdot H^* \right] dV \\
+ \frac{1}{2} \iint_{S_2} \left[ \gamma_h (\hat{n} \times H) \cdot (\hat{n} \times H)^* + H^* \cdot V + H \cdot V^* \right] dS \\
- \frac{1}{2} \iiint_V \left[ H^* \cdot \left( \nabla \times \frac{1}{\epsilon_r} J \right) + H \cdot \left( \nabla \times \frac{1}{\epsilon_r} J \right)^* \right]
\]  

(1.28)

COMSOL Multiphysics

We utilize COMSOL Multiphysics to fulfill the simulation based on finite element method. COMSOL Multiphysics software can be used to simulate designs, devices, and processes in all fields of engineering, manufacturing, and scientific research. It has several built-in modules, e.g. electromagnetic waves, ray optics, semiconductor optoelectronics, etc. It also offers the flexibility to build your own modules modified from a built-in one or start from any partial differential equations. The typical user interface of COMSOL is shown in Fig. (1.1). In order to simulate in a set of modified

![Figure 1.1: User interface of COMSOL Multiphysics](image-url)
differential equations, we need to derive and enter the modified weak expressions.

Here we give an example of the case we consider in the Chapter 3, the hydrodynamic model in 2-dimensional simulation. This model is governed by two main equations, let us repeat it here

\[ \nabla \times \nabla \times E = \frac{\omega^2}{c^2} E + \omega^2 \mu_0 P \quad (1.29) \]

\[ \beta^2 \nabla (\nabla \cdot P_f) + \dot{P}_f + \gamma \ddot{P}_f = \epsilon_0 \omega_p^2 E \quad (1.30) \]

Since Eq. (1.29) is a Maxwell wave equation, we can modify it from a built-in electromagnetic wave modules. Then we need to modify the weak expression to be

\[ -\mu_0 \text{const} \ast (-i \mu_0 \text{curl} \text{dep} E_z + \epsilon_0 \omega_0 \text{test}(\text{emw}.Ex) + \epsilon_0 \omega_0 \text{test}(\text{emw}.Ey) + Jh_1 \text{test}(\text{emw}.Ex) + Jh_2 \text{test}(\text{emw}.Ey)) \]

Also we then need to build a self-defined modules using a “weak form PDE” module, and put in the weak expression of eq. (1.30). The dependent variables of current in x and y direction are named “Jh1” and “Jh2” respectively. The weak form are given by

\[ \beta^2 \text{test}(Jh_1 x) \ast (Jh_1 x + Jh_2 y) - \text{omega0} \ast (\text{omega0} - j \ast \text{gamma}) \ast Jh_1 \text{test}(Jh_1) - j \ast \text{omega0} \ast \text{epsilon0 const} \ast \omega_0^2 \ast \text{emw}.Ex \ast \text{test}(Jh1) \]

\[ \beta^2 \text{test}(Jh_2 y) \ast (Jh_1 x + Jh_2 y) - \text{omega0} \ast (\text{omega0} - j \ast \text{gamma}) \ast Jh_2 \text{test}(Jh_2) - j \ast \text{omega0} \ast \text{epsilon0 const} \ast \omega_0^2 \ast \text{emw}.Ey \ast \text{test}(Jh2) \]

The names of the variables in the expressions are accordingly to the name list of variables in the built-in modules and physics constants. These two modules are then implemented on the metal regions and another original built-in electromagnetic wave module is implement on the dielectric regions.
Finite Difference Time Domain Method

Finite-difference time-domain (FDTD) is a well-known numerical technique in electrodynamics to compute the Maxwell’s equations. It translates the differential form of Maxwell’s equations into difference equations that can be solved numerically by computer. However, before the 1990s, the FDTD method was limited by the need to discretize the simulation space on subwavelength scales, with relatively small time steps. Thus, a typical photonics modeling would require a large amount of computer memory that exceed the technology limits at that time. However, since the 1990s, the FDTD became more computationally affordable with the fast increases in computer memory and speed.

There are several advantages for using FDTD. First, the method is accurate and robust, such that approximations are minimized and detailed solutions are provided with accuracy determined by the grid resolution. Second, the method is naturally including the effects such as polarization, dispersion, and nonlinearities. Furthermore, FDTD is able to calculate the full-wave response, which includes the transient behavior of an electromagnetic system.

We first introduce Yee’s unique, a highly powerful, FDTD scheme for solving Maxwell’s equations. Taking the advantage of the simplicity of one dimensional problems, we demonstrate the basic principle and formulation of the FDTD method for analysis of electrodynamic problems. We then discuss stability analysis, boundary conditions and extensions to the analysis of 2D/3D problems.

Consider a region of space that has no electric or magnetic current sources, but may
have materials that absorb electric or magnetic field energy. The time-dependent Maxwell’s equations are given in differential form by

\[
\frac{\partial B}{\partial t} = \nabla \times E - M \tag{1.31}
\]
\[
\frac{\partial D}{\partial t} = \nabla \times H - J \tag{1.32}
\]
\[
\nabla \cdot D = 0 \tag{1.33}
\]
\[
\nabla \cdot B = 0 \tag{1.34}
\]

the following symbols are defined:

**E**: electric field (volts / meter)

**D**: electric flux density (coulombs / meter\(^2\))

**H**: magnetic field (amperes / meter)

**B**: magnetic flux density (webers / meter\(^2\))

**J**: electric current density (amperes / meter\(^2\))

**M**: equivalent magnetic current density (volts / meter\(^2\))

In linear, isotropic, nondispersive materials (i.e., materials having field-independent, direction-independent, and frequency-independent electric and magnetic properties), we can relate \(D\) to \(E\) and \(B\) to \(H\) using simple proportions:

\[
D = \varepsilon E = \varepsilon_r \varepsilon_0 E \tag{1.35}
\]
\[
B = \mu H = \mu_r \mu_0 H \tag{1.36}
\]
where

\( \epsilon \) : electric permittivity (farads / meter)

\( \epsilon_r \) : relative permittivity (dimensionless scalar)

\( \epsilon_0 \) : free-space permittivity \( (8.854 \times 10^{-12} \) farads / meter\)

\( \mu \) : magnetic permeability (henrys / meter)

\( \mu_r \) : relative permeability (dimensionless scaler)

\( \mu_0 \) : free-space permeability \( (4\pi \times 10^{-7} \) henrys / meter\)

Note that \( J \) and \( M \) can act as independent sources of E- and H-field energy, \( J_{\text{source}} \) and \( M_{\text{source}} \). These proportions yield:

\[
J = J_{\text{source}} + \sigma E
\]

\[
M = M_{\text{source}} + \sigma^* H
\]

where

\( \sigma \) : electric conductivity (siemens / meter)

\( \sigma^* \) : equivalent magnetic loss (ohms / meter)

Finally, we get the Maxwell’s curl equations in linear, isotropic, nondispersive, lossy materials:

\[
\frac{\partial H}{\partial t} = -\frac{1}{\mu} \nabla \times E - \frac{1}{\mu}(M_{\text{source}} + \sigma^* H)
\]

\[
\frac{\partial E}{\partial t} = \frac{1}{\epsilon} \nabla \times H - \frac{1}{\epsilon}(J_{\text{source}} + \sigma E)
\]
We then write out the vector components of the curl operators in the Cartesian coordinate system.

\[
\frac{\partial H_x}{\partial t} = 1 \left[ \frac{\partial E_y}{\partial y} - \frac{\partial E_z}{\partial z} - (M_{\text{source}_x} + \sigma H_x) \right] \quad (1.41) \\
\frac{\partial H_y}{\partial t} = 1 \left[ \frac{\partial E_z}{\partial z} - \frac{\partial E_x}{\partial x} - (M_{\text{source}_y} + \sigma H_y) \right] \quad (1.42) \\
\frac{\partial H_z}{\partial t} = 1 \left[ \frac{\partial E_x}{\partial x} - \frac{\partial E_y}{\partial y} - (M_{\text{source}_z} + \sigma H_z) \right] \quad (1.43) \\
\frac{\partial E_x}{\partial t} = 1 \left[ \frac{\partial H_z}{\partial z} - \frac{\partial H_y}{\partial y} - (J_{\text{source}_x} + \sigma E_x) \right] \quad (1.44) \\
\frac{\partial E_y}{\partial t} = 1 \left[ \frac{\partial H_x}{\partial x} - \frac{\partial H_z}{\partial z} - (J_{\text{source}_y} + \sigma E_y) \right] \quad (1.45) \\
\frac{\partial E_z}{\partial t} = 1 \left[ \frac{\partial H_y}{\partial y} - \frac{\partial H_x}{\partial x} - (J_{\text{source}_z} + \sigma E_z) \right] \quad (1.46)
\]

The six coupled partial differential equations form the basis of the FDTD numerical algorithm for electromagnetic wave interactions with general three-dimensional objects. The FDTD algorithm need not explicitly enforce the Gauss’s law relations indicating zero free electric and magnetic charge. This is because these relations are theoretically a direct consequence of the curl equations. However, the FDTD space grid must be structured so that the Gauss’ Law relations are implicit in the positions of the E- and H-field vector components in the grid, e.g. Yee’s grid.

**The Yee Algorithm**

Now, what is the Yee algorithm. In 1966, Kane Yee originated a set of finite-difference equations for the time-dependent Maxwell’s curl equations system for the lossless materials case.

As illustrated in Fig. 1.2, the Yee algorithm centers its \( \mathbf{E} \) and \( \mathbf{H} \) components in
three-dimensional space so that every \( E \) component is surrounded by four circulating \( H \) components, and every \( H \) component is surrounded by four circulating \( E \) components. Continuity of tangential \( E \) and \( H \) is naturally maintained across an interface of dissimilar materials if the interface is parallel to one of the lattice coordinate axes. For this case, there is no need to enforce field boundary conditions at the interface. The location of the \( E \) and \( H \) components in the Yee space lattice and the central-difference operations on these components implicitly enforce the two Gauss’ Law relations. Thus, the Yee mesh is divergence-free with respect to its \( E \) and \( H \) fields in the absence of free and magnetic charge.

The Yee algorithm also centers its \( E \) and \( H \) components in time in a leapfrog arrangement. All of the \( E \) computations in the modeled space are completed and stored in memory for a particular time point using previously stored \( H \) data. All of the \( H \) computations in the space are then completed and stored in memory using the \( E \) data just computed. The cycle begins again with the recomputation of the \( E \)
components based on the newly obtained $H$.

**Finite Difference and Notations**

We denote a space point in a uniform, rectangular lattice as

$$(i, j, k) = (i\Delta x, j\Delta y, j\Delta z) \quad (1.47)$$

Further, we denote any function, $u$, of space and time evaluated at a discrete point in the grid and at a discrete point in time as

$$u(i\Delta x, j\Delta y, k\Delta z, n\Delta t) = u_{i,j,k}^n \quad (1.48)$$

We use center finite-difference expressions for the space and time derivatives. For example, we consider this expression for the first partial space derivative of $u$ in the $x$-direction, evaluated at the fixed time $t_n = \Delta t$:

$$\frac{\partial u}{\partial x}(i\Delta x, j\Delta y, k\Delta z, n\Delta t) = \frac{u_{i+1/2,j,k}^n - u_{i-1/2,j,k}^n}{\Delta x} + O[(\Delta x)^2] \quad (1.49)$$

and first time partial derivative of $u$:

$$\frac{\partial u}{\partial t}(i\Delta x, j\Delta y, k\Delta z, n\Delta t) = \frac{u_{i,j,k}^{n+1/2} - u_{i,j,k}^{n-1/2}}{\Delta t} + O[(\Delta t)^2] \quad (1.50)$$
Finite-Difference Expression for Maxwell’s Equations in Three Dimensions

We now apply the above ideas and notations to achieve a numerical approximation of the Maxwell’s curl equations in three dimensions. We begin by considering the $E_x$ field-component equation (1.44)

$$
\frac{E_x|_{i,j+1/2,k+1/2}^{n+1/2} - E_x|_{i,j+1/2,k+1/2}^{n-1/2}}{\Delta t} = \frac{1}{\epsilon_{i,j+1/2,k+1/2}}.
$$

$$
\left(\frac{H_z|_{i,j+1/2,k+1/2}^{n} - H_z|_{i,j,k+1/2}^{n}}{\Delta t} - \frac{H_y|_{i,j+1/2,k+1/2}^{n} - H_y|_{i,j+1/2,k}^{n}}{\Delta t} - J_{\text{source}}|_{i,j+1/2,k+1/2}^{n} - \sigma_{i,j+1/2,k+1/2} E_x|_{i,j+1/2,k+1/2}^{n}\right)
$$

(1.51)

Figure 1.3: Position of the electric and magnetic field vector components about a cubic unit cell of the Yee space lattice
Since we only know the previous values of \( E_x \) at time-step \( n - 1 \), we replace \( E_x \) with

\[
E_x |_{i,j+1/2,k+1/2}^{n+1/2} = \frac{E_x |_{i,j+1/2,k+1/2}^{n+1/2} + E_x |_{i,j+1/2,k+1/2}^{n-1/2}}{2}.
\]

(1.52)

Substituting (1.52) into (1.51), finally we obtain

\[
E_x |_{i,j+1/2,k+1/2}^{n+1/2} = \left( 1 - \frac{\sigma_{i,j+1/2,k+1/2}^{\Delta t}}{2\epsilon_{i,j+1/2,k+1/2}} \right) E_x |_{i,j+1/2,k+1/2}^{n-1/2} + \left( \frac{\Delta t}{\epsilon_{i,j+1/2,k+1/2}^{\Delta t}} \right) \begin{pmatrix}
\frac{H_x |_{i,j+1/2,k+1/2}^{n} - H_x |_{i,j+1/2,k+1/2}^{n-1}}{\Delta y} \\
\frac{H_y |_{i,j+1/2,k+1}^{n} - H_y |_{i,j+1/2,k}^{n}}{\Delta z} \\
- J_{\text{source}} |_{i,j+1/2,k+1/2}^{n}
\end{pmatrix}.
\]

(1.53)

In the same way, we can obtain the expression for \( E_y, E_z \) and also \( H_x, H_y, H_z \). We give the expression for \( H_x \) for instance

\[
H_x |_{i-1/2,j+1,k}^{n+1} = \left( 1 - \frac{\sigma_{i-1/2,j+1,k+1}^{\Delta t}}{2\mu_{i-1/2,j+1,k+1}} \right) H_x |_{i-1/2,j+1,k}^{n} + \left( \frac{\Delta t}{\mu_{i-1/2,j+1,k+1}^{\Delta t}} \right) \begin{pmatrix}
\frac{E_x |_{i-1/2,j+1,k+1}^{n+1/2} - E_x |_{i-1/2,j+1,k+1/2}^{n+1/2}}{\Delta z} \\
\frac{E_y |_{i-1/2,j+3/2,k}^{n+1/2} - E_y |_{i-1/2,j+1/2,k+1}^{n+1/2}}{\Delta y} \\
- M_{\text{source}} |_{i-1/2,j+1,k+1}^{n+1/2}
\end{pmatrix}.
\]

(1.54)

**Divergence-Free Nature**

As mentioned earlier, it is crucial for any grid-based solution of Maxwell’s curl equations to implicitly force (1.33) and (1.34), the Gauss’ Law relation for the electric and
magnetic fields in the source-free space being modeled.

Now we assume lossless free space and consider forming the time derivative of the total electric flux over the surface of a single Yee cell:

\[
\frac{\partial}{\partial t} \int_Y D \cdot dS = \epsilon_0 \frac{\partial}{\partial t} \left( E_x |_{i,j+1/2,k+1/2} - E_x |_{i-1,j+1/2,k+1/2} \Delta y \Delta z \right) \\
+ \epsilon_0 \frac{\partial}{\partial t} \left( E_y |_{i-1/2,j+1,k+1/2} - E_y |_{i-1/2,j,k+1/2} \Delta x \Delta z \right) \\
+ \epsilon_0 \frac{\partial}{\partial t} \left( E_z |_{i-1/2,j+1/2,k+1} - E_z |_{i-1/2,j+1/2,k} \Delta x \Delta y \right)
\]

(1.55)

We substitute appropriate \( H \)-field spatial finite differences for the \( E \)-field time derivatives in each term:

**Term 1**

\[
\text{Term 1} = \left( \frac{H_z |_{i,j+1,k+1/2} - H_z |_{i,j,k+1/2}}{\Delta y} - \frac{H_y |_{i,j+1/2,k+1} - H_y |_{i,j+1/2,k}}{\Delta z} \right) - \\
\left( \frac{H_z |_{i-1,j+1,k+1/2} - H_z |_{i-1,j,k+1/2}}{\Delta y} - \frac{H_y |_{i-1,j+1/2,k+1} - H_y |_{i-1,j+1/2,k}}{\Delta z} \right)
\]

(1.56)

**Term 2**

\[
\text{Term 2} = \left( \frac{H_z |_{i-1/2,j+1,k+1} - H_z |_{i-1/2,j+1,k}}{\Delta z} - \frac{H_z |_{i,j+1,k+1/2} - H_z |_{i-1,j+1,k+1/2}}{\Delta x} \right) \\
- \left( \frac{H_z |_{i-1/2,j,k+1} - H_z |_{i-1/2,j,k}}{\Delta z} - \frac{H_z |_{i,j+k+1/2} - H_z |_{i-1,j+k+1/2}}{\Delta x} \right)
\]

(1.57)

**Term 3**

\[
\text{Term 3} = \left( \frac{H_y |_{i,j+1/2,k+1} - H_y |_{i-1,j+1/2,k+1}}{\Delta x} - \frac{H_x |_{i-1/2,j+1,k+1} - H_x |_{i-1/2,j,k+1}}{\Delta y} \right) \\
- \left( \frac{H_y |_{i,j+1/2,k} - H_y |_{i-1,j+1/2,k}}{\Delta x} - \frac{H_x |_{i-1/2,j+1,k} - H_x |_{i-1/2,j,k}}{\Delta y} \right)
\]

(1.58)
For all time steps, this results in

\[
\frac{\partial}{\partial t} \oint D \cdot \, dS = (\text{Term 1})\Delta y \Delta z + (\text{Term 2})\Delta x \Delta z + (\text{Term 3})\Delta x \Delta y
\]

\[= 0 \]  

(1.59)

Assuming zero initial conditions, it means this flux is always zero

\[
\oint D \cdot \, dS = 0
\]

(1.60)

Therefore, the Yee cell satisfies Gauss’ Law for the \(E\)-field in charge-free space. In the same way, we can proof it is also satisfied for \(H\)-field.

**Numerical Dispersion**

We take a two-dimensional TM\(_z\) as an example. We assume the following plane, monochromatic, traveling-wave trial solution for the TM\(_z\) mode:

\[
E_{z|I,J}^{n} = E_{z0}e^{i(\omega n\Delta t - \tilde{k}_x I \Delta x - \tilde{k}_y J \Delta y)}
\]

(1.61)

\[
H_{x|I,J}^{n} = H_{x0}e^{i(\omega n\Delta t - \tilde{k}_x I \Delta x - \tilde{k}_y J \Delta y)}
\]

(1.62)

\[
H_{y|I,J}^{n} = H_{y0}e^{i(\omega n\Delta t - \tilde{k}_x I \Delta x - \tilde{k}_y J \Delta y)}
\]

(1.63)

where \(\tilde{k}_x\) and \(\tilde{k}_y\) are the \(x\)- and \(y\)-components of the numerical wavevector and \(\omega\) is the wave angular frequency. Substituting the traveling-wave expressions into the
finite-difference equations yields

\[
H_{x_0} = \frac{\Delta t E_{z_0}}{\mu \Delta y} \frac{\sin(\tilde{k}_y \Delta y/2)}{\sin(\omega \Delta t/2)}
\]

(1.64)

\[
H_{y_0} = \frac{\Delta t E_{z_0}}{\mu \Delta x} \frac{\sin(\tilde{k}_x \Delta x/2)}{\sin(\omega \Delta t/2)}
\]

(1.65)

\[
E_{z_0} \sin \left( \frac{\omega \Delta t}{2} \right) = \frac{\Delta t}{\epsilon} \left[ \frac{H_{x_0}}{\Delta y} \sin \left( \frac{\tilde{k}_y \Delta y}{2} \right) - \frac{H_{y_0}}{\Delta x} \sin \left( \frac{\tilde{k}_x \Delta x}{2} \right) \right]
\]

(1.66)

Upon substituting (1.64) and (1.65) into (1.66), we obtain

\[
\left[ \frac{1}{c \Delta t} \sin \left( \frac{\omega \Delta t}{2} \right) \right]^2 = \left[ \frac{1}{\Delta x} \sin \left( \frac{\tilde{k}_x \Delta x}{2} \right) \right]^2 + \left[ \frac{1}{\Delta y} \sin \left( \frac{\tilde{k}_y \Delta y}{2} \right) \right]^2
\]

(1.67)

where \( c = 1/\sqrt{\mu \epsilon} \) is the speed of light in the material. And it is easy to extend to three dimensional case

\[
\left[ \frac{1}{c \Delta t} \sin \left( \frac{\omega \Delta t}{2} \right) \right]^2 = \left[ \frac{1}{\Delta x} \sin \left( \frac{\tilde{k}_x \Delta x}{2} \right) \right]^2 + \left[ \frac{1}{\Delta y} \sin \left( \frac{\tilde{k}_y \Delta y}{2} \right) \right]^2 + \left[ \frac{1}{\Delta z} \sin \left( \frac{\tilde{k}_z \Delta z}{2} \right) \right]^2
\]

(1.68)

In contrast to the numerical dispersion relation (1.68), the analytical (ideal) dispersion relation for a physical plane wave is

\[
\left( \frac{\omega}{c} \right)^2 = (k_x)^2 + (k_y)^2 + (k_z)^2
\]

(1.69)

The two dispersion relations are identical in the limit as \( \Delta x, \Delta y, \Delta z, \) and \( \Delta t \) approach zero. Further, consider the special case of a cubic-cell grid having \( \Delta x = \Delta y = \Delta, \)
and define the Courant stability factor \( S = c \Delta t / \Delta \) and the grid sampling density \( N_\lambda = \lambda_0 / \Delta \), we can rewrite (1.67) as

\[
\frac{1}{S^2} \sin^2 \left( \frac{\pi S}{N_\lambda} \right) = \sin^2 \left( \frac{\Delta \cdot \ddot{k} \cos \phi}{2} \right) + \sin^2 \left( \frac{\Delta \cdot \ddot{k} \sin \phi}{2} \right)
\]  

(1.70)

**Numerical Stability**

We can now solve \( \omega \) from (1.68). This operation yields

\[
\omega = \frac{2}{\Delta t} \sin^{-1}(\xi)
\]  

(1.71)

where

\[
\xi = c \Delta t \sqrt{\sum_{\eta=x,y,z} \frac{1}{(\Delta \eta)^2} \sin^2 \left( \frac{\ddot{k}_\eta \Delta \eta}{2} \right)}
\]  

(1.72)

thus, we observe that

\[
0 \leq \xi \leq c \Delta t \sqrt{\frac{1}{(\Delta x)^2} + \frac{1}{(\Delta y)^2} + \frac{1}{(\Delta z)^2}} = \xi_{\text{upper bound}}
\]  

(1.73)

for all possible real value of \( \ddot{k} \). \( \xi_{\text{upper bound}} \) is obtained when

\[
\ddot{k}_x = \pm \frac{\pi}{\Delta x}; \quad \ddot{k}_y = \pm \frac{\pi}{\Delta y}; \quad \ddot{k}_z = \pm \frac{\pi}{\Delta z}
\]  

(1.74)

\( \xi_{\text{upper bound}} \) can exceed 1 depending upon the choice of \( \Delta t \). This can yield complex values of \( \sin^{-1}(\xi) \), and therefore complex values for \( \omega \) which give rise to numerical
instability. An instability exists only if

\[ \xi_{\text{upper bound}} > 1 \quad (1.75) \]

thus

\[ \Delta t > \frac{1}{c \sqrt{\frac{1}{(\Delta x)^2} + \frac{1}{(\Delta y)^2} + \frac{1}{(\Delta z)^2}}} \quad (1.76) \]

**Perfectly Matched Layer Absorbing Boundary Conditions**

One of the greatest challenges of FDTD method has been the efficient and accurate solution of electromagnetic wave interaction problems in unbounded regions. For such problems, an absorbing boundary condition (ABC) must be introduced at the outer lattice boundary to simulate the extension of the lattice to infinity. In 1994, J.-P Berenger introduced a highly effective absorbing-material ABC, designated as a perfectly match layer (PML), which makes plane waves of arbitrary incidence, polarization, and frequency matched at the boundary.

Consider again a $\text{TE}_z$ plane wave incident on the planar interface $x = 0$ of the material half-space region 2 and normal material region 1, where Maxwell’s equations
is modified as expressed

$$\epsilon_2 \frac{\partial E_x}{\partial t} + \sigma_y E_x = \frac{\partial H_z}{\partial y}$$  \hspace{1cm} (1.77)$$

$$\epsilon_2 \frac{\partial E_y}{\partial t} + \sigma_x E_y = - \frac{\partial H_z}{\partial x}$$  \hspace{1cm} (1.78)$$

$$\mu_2 \frac{\partial H_{zx}}{\partial t} + \sigma_x H_{zx} = - \frac{\partial E_y}{\partial x}$$  \hspace{1cm} (1.79)$$

$$\mu_2 \frac{\partial H_{zy}}{\partial t} + \sigma_y H_{zy} = \frac{\partial E_x}{\partial y}$$  \hspace{1cm} (1.80)$$

Here, $H_z$ is assumed to be split into two additive subcomponents

$$H_z = H_{zx} + H_{zy}$$  \hspace{1cm} (1.81)$$

Now we consider fields are expressed in phasor quantity

$$j \omega \epsilon_2 \left(1 + \frac{\sigma_y}{j \omega \epsilon_2}\right) E_x = \frac{\partial}{\partial y}(H_{zx} + H_{zy})$$  \hspace{1cm} (1.82)$$

$$j \omega \epsilon_2 \left(1 + \frac{\sigma_x}{j \omega \epsilon_2}\right) E_y = - \frac{\partial}{\partial x}(H_{zx} + H_{zy})$$  \hspace{1cm} (1.83)$$

$$j \omega \mu_2 \left(1 + \frac{\sigma_x^*}{j \omega \mu_2}\right) H_{zx} = \frac{\partial E_y}{\partial x}$$  \hspace{1cm} (1.84)$$

$$j \omega \mu_2 \left(1 + \frac{\sigma_y^*}{j \omega \mu_2}\right) H_{zy} = \frac{\partial E_x}{\partial y}$$  \hspace{1cm} (1.85)$$

we denote

$$s_w = \left(1 + \frac{\sigma_w}{j \omega \epsilon_2}\right); \quad s_w^* = \left(1 + \frac{\sigma_w^*}{j \omega \mu_2}\right)$$  \hspace{1cm} (1.86)$$
thus we can have

\[-\omega^2 \mu_2 \varepsilon_2 H_{xx} = -\frac{1}{s_x^*} \frac{\partial}{\partial x} \frac{1}{s_x} \frac{\partial}{\partial x} (H_{xx} + H_{zy}) \]  
\[-\omega^2 \mu_2 \varepsilon_2 H_{zy} = -\frac{1}{s_y^*} \frac{\partial}{\partial y} \frac{1}{s_y} \frac{\partial}{\partial y} (H_{xx} + H_{zy}) \]  

Adding these together leads to the representative wave equation

\[\frac{1}{s_x^*} \frac{\partial}{\partial x} \frac{1}{s_x} \frac{\partial}{\partial x} H_x + \frac{1}{s_y^*} \frac{\partial}{\partial y} \frac{1}{s_y} \frac{\partial}{\partial y} H_x + \omega^2 \mu_2 \varepsilon_2 H_z = 0 \]  

This wave equation supports the solutions

\[H_z = H_0 \tau e^{-j\sqrt{s_x s_x^*} \beta_{2x} x - j\sqrt{s_y s_y^*} \beta_{2y} y} \]  
\[E_x = -H_0 \frac{\beta_{2y}}{\omega \varepsilon_2} \sqrt{\frac{s_y^*}{s_y}} e^{-j\sqrt{s_x s_x^*} \beta_{2x} x - j\sqrt{s_y s_y^*} \beta_{2y} y} \]  
\[E_y = -H_0 \frac{\beta_{2x}}{\omega \varepsilon_2} \sqrt{\frac{s_x^*}{s_x}} e^{-j\sqrt{s_x s_x^*} \beta_{2x} x - j\sqrt{s_y s_y^*} \beta_{2y} y} \]

with the dispersion relationship

\[(\beta_{2x})^2 + (\beta_{2y})^2 = (k_2)^2 \]

To enforce this field continuity, we have \(s_y = s_y^* = 1\), or equivalently \(\sigma_y = \sigma_y^* = 0\). This yields the phase-matching condition \(\beta_{2y} = \beta_{1y} = k_1 \sin \theta\). Further, we can obtain
the $H$-field reflection and transmission coefficients

$$\Gamma = \left( \frac{\beta_{1x}}{\omega \epsilon_1} - \frac{\beta_{2x}}{\omega \epsilon_2} \sqrt{s_x^s} \right) \cdot \left( \frac{\beta_{1x}}{\omega \epsilon_1} + \frac{\beta_{2x}}{\omega \epsilon_2} \sqrt{s_x^s} \right)^{-1}$$  

(1.94)

$$\tau = 1 + \Gamma$$  

(1.95)

Now assume $\epsilon_1 = \epsilon_2$, $\mu_1 = \mu_2$, and $s_x = s_x^s$, which is equivalent to $k_1 = k_2$, $\mu_1 = \sqrt{\mu_1/\epsilon_1} = \sqrt{\mu_2/\epsilon_2}$, and $\sigma_x/\epsilon_1 = \sigma_x^s/\mu_1$. With $\beta_{2y} = \beta_{1y}$, the dispersion relation yields $\beta_{2x} = \beta_{1x}$. Substituting into (1.94) gives $\Gamma = 0$ for all incident angles $\theta$. Then we obtain the transmitted fields with PML in the region $x > 0$

$$H_z = H_0 e^{-j\beta_{1x}x - j\beta_{1y}y} e^{-\sigma_x \eta_1 \cos \theta}$$  

(1.96)

$$E_x = -H_0 \eta_1 \sin \theta e^{-j\beta_{1x}x - j\beta_{1y}y} e^{-\sigma_x \eta_1 \cos \theta}$$  

(1.97)

$$E_y = H_0 \eta_1 \cos \theta e^{-j\beta_{1x}x - j\beta_{1y}y} e^{-\sigma_x \eta_1 \cos \theta}$$  

(1.98)

Within the PML, the transmitted wave propagates with the same speed and direction as the impinging wave while simultaneously undergoing exponential decay along the x-axis normal to the interface. The attenuation factor $\sigma_x \mu_1 \cos \theta$ is independent of frequency.

**Synopsys RSOFT**

We utilize Synopsys RSOFT to fulfill the simulation based on finite difference time domain method. Fig. (1.4) shows an example of a 3D FDTD simulation of a ring-bus waveguide structure in RSOFT.
1.2 Theoretical Approach

Nonlinear Slab Model

A slab model is introduced in order to describe the effects at the boundary of a nonlinear medium by N. Bloembergen and P.S. Pershan [62]. When a light incidents on a plane boundary of a nonlinear medium, the nonlinear susceptibility will give rise to a polarization at harmonic frequencies, which in turn will radiate energy at these frequencies. For example, for the 2nd harmonic generation, the effective nonlinear source term is given by

$$P^{NLS}(2\omega_1) = \chi(\omega_2 = \omega_1) : \mathbf{E}_1^T \mathbf{E}_1 e^{i(k^s \cdot r - 2\omega_1 t)}$$

(1.99)
The effective nonlinear source term can be incorporated into Maxwell’s equations for the nonlinear medium

\[
\nabla \times \mathbf{E} = -\frac{1}{c} \frac{\partial \mu \mathbf{H}}{\partial t} \tag{1.100}
\]

\[
\nabla \times \mathbf{H} = \frac{1}{c} \frac{\partial (\epsilon \mathbf{E})}{\partial t} + \frac{4\pi}{c} \frac{\partial \mathbf{P}^{NLS}}{\partial t} \tag{1.101}
\]

usually we treat \( \epsilon \) as a scalar and \( \mu = 1 \) for most material. Thus we can obtain the nonlinear wave equation for the \( 2\omega \) frequency

\[
\nabla \times \nabla \times \mathbf{E}_2 + \frac{\epsilon(\Omega)}{c} \frac{\partial^2 \mathbf{E}_2}{\partial t^2} = -\frac{4\pi}{c^2} \frac{\partial^2 \mathbf{P}^{NLS}(2\omega)}{\partial t^2} \tag{1.102}
\]

The general solution to the above equation consists of the solution of the homogeneous equation plus one particular solution of the inhomogeneous equation

\[
\mathbf{E}_2^T = \mathbf{\hat{e}}_T \mathbf{E}_2^T e^{i(\mathbf{k}^T_2 \cdot \mathbf{r} - \omega t)} - \frac{4\pi P_2^{NLS}(\Omega^2/c^2)}{(k^T_2)^2 - (k^S)^2} \times \left[ \mathbf{\hat{p}} - \frac{k^S (k^S \cdot \mathbf{\hat{p}})}{(k^T)^2} \right] e^{i(k^T_2 \cdot \mathbf{r} - \omega t)} \tag{1.103}
\]

\[
\mathbf{H}_2^T = \frac{c}{\Omega} (\mathbf{k}_2^T \times \mathbf{\hat{e}}_T) \mathbf{E}_2^T e^{i(\mathbf{k}^T_2 \cdot \mathbf{r} - \omega t)} - \frac{4\pi P_2^{NLS}(\Omega^2/c^2)}{(k^T_2)^2 - (k^S)^2} \times \frac{c}{\Omega} (k^S \times \mathbf{\hat{p}}) e^{i(k^S \cdot \mathbf{r} - \Omega t)} \tag{1.104}
\]

In vacuum the usual plane wave solutions are

\[
\mathbf{E}_2^R = \mathbf{\hat{e}}_R \mathbf{E}_2^R e^{i(k^R_2 \cdot \mathbf{r} - 2\Omega t)} \tag{1.105}
\]

\[
\mathbf{H}_2^R = \frac{c}{\Omega} (\mathbf{k}_2^R \times \mathbf{\hat{e}}_R) \mathbf{E}_2^R e^{i(k^R_2 \cdot \mathbf{r} - \Omega t)} \tag{1.106}
\]
The direction of the wave vectors $k_2^R$ and $k_2^T$ as well as the polarization vectors $\hat{e}_T$ and $\hat{e}_R$ and the magnitude of the reflected and transmitted amplitudes $E_2^R$ and $E_2^T$ is determined by the boundary condition. Please note that here we assume the medium is isotropic. However, it can be derived in a similar way for anisotropic medium.

The tangential components of $E$ and $H$ should be continuous on the boundary, which leads to

$$2k_{1x}^T = k_{2x}^S = k_{2x}^R = k_{1x}^T$$  \hspace{1cm} (1.107)

thus we can obtain the angles of reflection and refraction of the nonlinear wave

$$\sin \theta_2^R = \sin \theta^i$$  \hspace{1cm} (1.108)

$$\sin \theta_2^T = \frac{1}{\sqrt{\epsilon(\omega)}} \sin \theta^i$$  \hspace{1cm} (1.109)

$$\sin \theta_2^S = \frac{1}{\sqrt{\epsilon(\omega)}} \sin \theta^i$$  \hspace{1cm} (1.110)
The geometrical relationships are sketched in Fig. 1.5. The continuity of the tangential components $E_y$ at the boundary $z = 0$ leads to the condition

$$E_2^R = E_2^T - 4\pi P^{NLS}/[\epsilon(\Omega) - \epsilon(\omega)] \quad (1.111)$$

The continuity of the x components of the magnetic fields yields

$$-E_2^R \cos \theta^R = \sqrt{\epsilon(\Omega)} E_2^T \cos \theta^T - \sqrt{\epsilon(\omega)} \cos \theta^S A_{\Omega} \epsilon S^{NLS} \epsilon(\Omega) - \epsilon(\omega) \quad (1.112)$$

As a specific case we consider is the third-harmonic generation in a slab system, which is used in the one of Chapter 4. We consider an infinite slab of a nonlinear dielectric medium with boundaries at $z = 0$ and $z = d$, embedded between two linear media.

The linear wave at $\omega$ is incident from the medium R for $z < 0$, as shown in Fig. 1.6. First, we need to calculate the fields of fundamental light for this slab system. Using the boundary condition as used in deriving the Fresnel equation,

$$E_{1,y}(z = 0) = E_{1,i} + E_{1,R} = E_{1,M} + E_{1,M'} \quad (1.113)$$

$$E_{1,y}(z = d) = E_{1,M} e^{i\psi} + E_{1,M'} e^{-i\psi} = E_{1,T} \quad (1.114)$$

$$H_{1,z}(z = 0) = \sqrt{\epsilon_R} E_{1,i} - \sqrt{\epsilon_R} E_{1,R} = \sqrt{\epsilon_s} E_{1,M} - \sqrt{\epsilon_s} E_{1,M'} \quad (1.115)$$

$$H_{1,z}(z = d) = \sqrt{\epsilon_s} E_{1,M} e^{i\psi} - \sqrt{\epsilon_s} E_{1,M} e^{-i\psi} = \sqrt{\epsilon_T} E_{1,T} \quad (1.116)$$

where $\psi = \frac{\sqrt{\epsilon_s(\omega)} d\omega}{c}$. Please note that for equations describing the fundamental light above, $\epsilon_s$ denotes the relative permittivity of the nonlinear medium at $\omega$. We can
also write this relation in a matrix system

\[
\begin{pmatrix}
1 & -1 & -1 & 0 \\
0 & e^{i\psi} & e^{-i\psi} & -1 \\
\sqrt{\epsilon_R} & \sqrt{\epsilon_s} & -\sqrt{\epsilon_s} & 0 \\
0 & \sqrt{\epsilon_s}e^{i\psi} & i\sqrt{\epsilon_s}e^{-i\psi} & -\sqrt{\epsilon_T}
\end{pmatrix}
\begin{pmatrix}
E_{1,R} \\
E_{1,M} \\
E_{1,M'} \\
E_{1,T}
\end{pmatrix} =
\begin{pmatrix}
-1 \\
0 \\
\sqrt{\epsilon_R} \\
0
\end{pmatrix}
\]

(1.117)

The created forward moving wave \(E_{1,M}\) and backward moving wave \(E_{1,M'}\) in the non-linear medium, produce a nonlinear polarization \(P^{NLS}(\Omega)\) at 3rd harmonic frequency \(\Omega = 3\omega\).

\[
P^{NLS}(\Omega) = \chi \exp \left[i(3k_{1,x}x + 3k_{1,y}y)\right] \times \left\{ E_{1,M}E_{1,M}E_{1,M}e^{i3k_{1,z}z} + 3E_{1,M}E_{1,M}E_{1,M'}e^{i3k_{1,z}z} \right\}
\]

(1.118)
Because of the absorption effect in the nonlinear media, we assume $E_{1,M'} \ll E_{1,M}$, thus the nonlinear polarization term becomes

$$P_{NLS}(\Omega) = \chi \exp \left[i(3k_{1,x} x + 3k_{1,y} y)\right] \times \left\{ E_{1,M} E_{1,M} e^{i3k_{1,z} z} \right\}$$ (1.119)

The boundary conditions in the case of perpendicular polarization can be written as

$$E_y(z = 0) = E_R = E_M + E_{M'} + \frac{P_{NLS}}{\epsilon_s - \epsilon_M}$$ (1.120)

$$E_y(z = d) = E_T = E_M e^{i\phi_M} + E_{M'} e^{-i\phi_M} + \frac{P_{NLS}}{\epsilon_s - \epsilon_M} e^{i\phi_s}$$ (1.121)

$$H_z(z = 0) = -\sqrt{\epsilon_R} E_R = \sqrt{\epsilon_M}(E_M - E_{M'}) + \sqrt{\epsilon_s} P_{NLS} e^{i\phi_s}$$ (1.122)

$$H_z(z = d) = \sqrt{\epsilon_T} E_T = \sqrt{\epsilon_M} [E_M e^{i\phi_M} - E_{M'} e^{-i\phi_M}] + \sqrt{\epsilon_s} P_{NLS} e^{i\phi_s}$$ (1.123)

where $\phi_s = \frac{\sqrt{\epsilon_s} d\Omega}{c}$, $\phi_M = \frac{\sqrt{\epsilon_M} d\Omega}{c}$, and $\epsilon_s = \epsilon(\omega)$, $\epsilon_M = \epsilon(\Omega)$. We can also write this relation in a matrix system

$$
\begin{bmatrix}
1 & -1 & -1 & 0 \\
\sqrt{\epsilon_R} & \sqrt{\epsilon_M} & -\sqrt{\epsilon_M} & 0 \\
0 & e^{i\phi_M} & e^{-i\phi_M} & -1 \\
0 & \sqrt{\epsilon_M} e^{i\phi_M} & -\sqrt{\epsilon_M} e^{-i\phi_M} & -\sqrt{\epsilon_T}
\end{bmatrix}
\begin{bmatrix}
E_R \\
E_M \\
E_{M'} \\
E_T
\end{bmatrix}
= 
\begin{bmatrix}
\frac{P_{NLS}}{\epsilon_s - \epsilon_M} \\
-\sqrt{\epsilon_s} P_{NLS} e^{i\phi_s} \\
-\sqrt{\epsilon_s} P_{NLS} e^{i\phi_s} \\
-\sqrt{\epsilon_s} e^{i\phi_s}
\end{bmatrix}
$$ (1.124)

41
Reflection and Transmission under Sheet Conductivity of the Interface Model

To start, we consider a light incidents the interface of two medium (1 and 2) with the angle of $\theta$ in the x-z plane. To get the amplitudes of the reflected and transmitted waves, we apply the boundary conditions on two cases.

(1): S polarization, or TE waves, $E_{\perp}$ plane of incident, possesses the electromagnetic field components $\mathbf{E} = \{0, E_y, 0\}$ and $\mathbf{H} = \{H_x, 0, H_z\}$.

(2): P polarization, or TM waves, $E_{\parallel}$ plane of incident, possesses the electromagnetic field components $\mathbf{E} = \{E_x, 0, E_z\}$ and $\mathbf{H} = \{0, H_y, 0\}$. The S-polarization and P-polarization are illustrated in Fig. 1.7. Boundary conditions in Gaussian Units are

\[
\hat{n} \times (\mathbf{E}_2 - \mathbf{E}_1) = 0 \quad (1.125)
\]

\[
\hat{n} \times (\mathbf{H}_2 - \mathbf{H}_1) = \mathbf{K} \quad (1.126)
\]
Here \( \hat{n} \) is the unit normal vector pointing from 1 to 2 at the interface, so \( \hat{n} = \{0, 0, 1\} \). And \( \mathbf{K} \) is the surface current density. Now we substitute \( \hat{n} \) into the equations, we can get the boundary condition at \( z = 0 \)

\[
E_x^{(2)}(x, 0) = E_x^{(1)}(x, 0) \tag{1.127}
\]
\[
E_y^{(2)}(x, 0) = E_y^{(1)}(x, 0) \tag{1.128}
\]
\[
H_x^{(2)}(x, 0) - H_x^{(1)}(x, 0) = K_y \tag{1.129}
\]
\[
H_y^{(2)}(x, 0) - H_y^{(1)}(x, 0) = -K_x \tag{1.130}
\]

Here surface density \( \mathbf{K} \) connects the electric field and magnetic field through the following relation

\[
\mathbf{K} = \sigma \mathbf{E} \tag{1.131}
\]

where \( \sigma \) is the electric conductivity of the interface material. In the next step, we express \( E_x^{(1)}(x, z), E_x^{(2)}(x, z), E_y^{(1)}(x, z), E_y^{(2)}(x, z) \) and \( H_x^{(1)}(x, z), H_x^{(2)}(x, z), H_y^{(1)}, H_y^{(2)}(x, z) \) using \( H_r^\alpha, H_i^\alpha, H_t^\alpha \) and \( E_r^\alpha, E_i^\alpha, E_t^\alpha \). Here \( \alpha = s, p \), represents different polarization. After that, we can eliminate \( \mathbf{E} \) and \( \mathbf{H} \) fields to obtain the relation between reflection (transmission) and incidence.

Now we discuss the wave vector \( k \), which is the same for S polarization and P polar-
In medium 1, the wave vector can be written as

\[
\sqrt{k_{x}^{(1)}{}^2 + k_{y}^{(1)}{}^2 + k_{z}^{(1)}{}^2} = |\mathbf{k}^{(1)}| = \frac{\omega}{c} \sqrt{\varepsilon_1}
\]

\[
k_{x}^{(1)} = \frac{\omega}{c} \sqrt{\varepsilon_1} \sin \theta_i
\]

(1.132)

\[
k_{y}^{(1)} = 0
\]

\[
k_{z}^{(1)} = \frac{\omega}{c} \sqrt{\varepsilon_1} \cos \theta_i
\]

In medium 2, according to Snell’s law, the wave vector can be written as

\[
k_{x}^{(2)}{}^2 + k_{y}^{(2)}{}^2 + k_{z}^{(2)}{}^2 = |\mathbf{k}^{(2)}| = \frac{\omega}{c} \sqrt{\varepsilon_2}
\]

\[
k_{x}^{(2)} = \frac{\omega}{c} \sqrt{\varepsilon_2} \sin \theta_i = \frac{\omega}{c} \sqrt{\varepsilon_1} \sin \theta_i = k_{x}^{(1)}
\]

(1.133)

\[
k_{y}^{(2)} = 0
\]

\[
k_{z}^{(2)} = \frac{\omega}{c} \sqrt{\varepsilon_2 - \varepsilon_1 \sin^2 \theta_i}
\]

After preparation, we can expand boundary condition (1.127) to (1.130) as follows.

(1) S polarization.

In this case, \( \mathbf{E} = \{0, E_y, 0\} \) and \( \mathbf{H} = \{H_x, 0, H_z\} \). The electric field has only a y-component. In medium 1, we can write \( E_y \) as

\[
E_{y}^{(1)}(x, z) = [E_{t}^{s} \exp(i k_{x}^{(1)} x + i k_{z}^{(1)} z) + E_{r}^{s} \exp(i k_{x}^{(1)} x - i k_{z}^{(1)} z)]
\]

(1.134)

We add “-” in the reflection part because its direction lies in opposite to z-axis.

In medium 2

\[
E_{y}^{(2)}(x, z) = [E_{t}^{s} \exp(i k_{x}^{(2)} x + i k_{z}^{(2)} z)]
\]

(1.135)
At the same time, we can express magnetic field as \( \mathbf{H} = -\frac{\mathbf{E}}{\omega} \times \mathbf{k} \)

For the incident component:

\[
\mathbf{H}_i = -\frac{c}{\omega} (\mathbf{E}_i \times \mathbf{k}_i) = -\frac{c}{\omega} \begin{vmatrix} \bar{x} & \bar{y} & \bar{z} \\ 0 & E_{i,y} & 0 \\ k_x & 0 & k_z \end{vmatrix} = \bar{x}(-\frac{c}{\omega} k_z E_{i,y}) + \bar{z}(\frac{c}{\omega} k_x E_{i,y}) \quad (1.136)
\]

For the reflective component:

\[
\mathbf{H}_r = -\frac{c}{\omega} (\mathbf{E}_r \times \mathbf{k}_r) = -\frac{c}{\omega} \begin{vmatrix} \bar{x} & \bar{y} & \bar{z} \\ 0 & E_{r,y} & 0 \\ k_x & 0 & -k_z \end{vmatrix} = \bar{x}(\frac{c}{\omega} k_z E_{r,y}) + \bar{z}(\frac{c}{\omega} k_x E_{r,y}) \quad (1.137)
\]

According to boundary conditions (1.127) to (1.130), only the x and y component of fields are significant. Thus in medium 1, the magnetic field can be written as

\[
H^{(1)}_x(x, z) = -\frac{ck_z^{(1)}}{\omega} \left[ E^{s}_i \exp(ik_x^{(1)}x + ik_z^{(1)}z) - E^{s}_r \exp(ik_x^{(1)}x - ik_z^{(1)}z) \right] \quad (1.138)
\]

In medium 2,

\[
H^{(2)}_x(x, z) = -\frac{ck_z^{(2)}}{\omega} \left[ E^{s}_i \exp(ik_x^{(2)}x + ik_z^{(2)}z) \right] \quad (1.139)
\]

Based on the boundary conditions (1.128), (1.129) and (1.131), we can get

\[
E^{s}_i + E^{s}_r = E^{s}_i \quad (1.140)
\]

\[
-\frac{ch_z^{(2)}}{\omega} E^{s}_i + \frac{ch_z^{(1)}}{\omega} (E^{s}_i - E^{s}_r) = \frac{4\pi\sigma}{c} H_y = \frac{4\pi\sigma}{c} (E^{s}_i + E^{s}_r) \quad (1.141)
\]
Finally, we can obtain the amplitudes of the reflected and transmitted waves in the case of S-polarization,

\[ E_r^s = -\frac{k_2^{(2)} - k_1^{(1)}}{k_2^{(2)} + k_1^{(1)} + \frac{4\sigma\omega}{c^2}} E_i^s \]  

\[ E_t^s = \frac{2k_1^{(1)}}{k_2^{(2)} + k_1^{(1)} + \frac{4\sigma\omega}{c^2}} E_i^s \]  

(1.142)

(1.143)

(2) P polarization.

In this case, \( \mathbf{H} = \{0, H_y, 0\} \) and \( \mathbf{E} = \{E_x, 0, E_z\} \).

Similarly, in medium 1, we can write \( H_y \) as

\[ H_y^{(1)}(x, z) = [H_i^p \exp(ik_x^{(1)}x + ik_z^{(1)}z) + H_t^p \exp(ik_x^{(1)}x - ik_z^{(1)}z)] \]  

(1.144)

In medium 2,

\[ H_y^{(2)}(x, z) = [H_i^p \exp(ik_x^{(2)}x + ik_z^{(2)}z)] \]  

(1.145)

We can also express electrical field.

For the incident component:

\[ \mathbf{E}_i = \frac{c}{\epsilon \omega} (\mathbf{H}_i \times k_i) = \frac{c}{\epsilon \omega} \begin{vmatrix} \bar{x} & \bar{y} & \bar{z} \\ 0 & H_{i,y} & 0 \\ k_x & 0 & k_z \end{vmatrix} = \bar{x}(\frac{c}{\epsilon \omega} k_z H_{i,y}) + \bar{z}(\frac{c}{\epsilon \omega} k_x H_{i,y}) \]

(1.146)
For the reflective component:

\[
E_r = \frac{c}{\epsilon \omega} (\mathbf{H}_r \times k_r) = \begin{vmatrix}
\vec{x} & \vec{y} & \vec{z} \\
0 & H_{r,y} & 0 \\
k_x & 0 & -k_z
\end{vmatrix} = \vec{x}(-\frac{c}{\epsilon \omega} k_z H_{r,y}) + \vec{z}(-\frac{c}{\epsilon \omega} k_x H_{r,y}) \tag{1.147}
\]

Thus in medium 1, the electrical field can be written as:

\[
E_x^{(1)}(x, z) = \frac{ck_z^{(1)}}{\epsilon_1 \omega} \left[H_t^p \exp(ik_z^{(1)} x + ik_z^{(1)} z) - H_r^p \exp(ik_z^{(1)} x - ik_z^{(1)} z)\right] \tag{1.148}
\]

In medium 2,

\[
E_x^{(2)}(x, z) = \frac{ck_z^{(2)}}{\epsilon_2 \omega} \left[H_t^p \exp(ik_z^{(2)} x + ik_z^{(2)} z)\right] \tag{1.149}
\]

Based on the boundary conditions (1.127), (1.130) and (1.131), we can get

\[
\frac{ck_z^{(1)}}{\epsilon_1 \omega} (H_t^p - H_r^p) = \frac{ck_z^{(2)}}{\epsilon_2 \omega} H_t^p \tag{1.150}
\]

\[
H_t - (H_t^p + H_r^p) = -\frac{4\pi \sigma}{c} E_x = -\frac{4\pi \sigma}{\epsilon_2 \omega} c k_z^{(2)} H_t^p \tag{1.151}
\]

We can obtain the amplitudes of the reflected and transmitted waves in the case of P-polarization,

\[
H_r^p = \frac{\epsilon_2 k_z^{(1)} - \epsilon_1 k_z^{(2)} + \frac{4\pi \sigma}{\epsilon_2} k_z^{(1)} k_z^{(2)}}{\epsilon_2 k_z^{(1)} + \epsilon_1 k_z^{(2)} + \frac{4\pi \sigma}{\epsilon_2} k_z^{(1)} k_z^{(2)}} H_t^p \tag{1.152}
\]

\[
H_t^p = \frac{\frac{2\epsilon_2 k_z^{(1)}}{\epsilon_2 k_z^{(1)} + \epsilon_1 k_z^{(2)} + \frac{4\pi \sigma}{\epsilon_2} k_z^{(1)} k_z^{(2)}} H_t}{1.153}
\]
**Power Coefficients at the Boundary**

Now we move to reflection and transmission efficiency. The Poynting vector is used to measure the power ratio

$$ S = E \times H $$

(1.154)

Its intensity,

$$ I = \langle \vec{S} \rangle $$

(1.155)

can be simplified as

$$ I = \frac{\sqrt{\varepsilon}}{2} |E_0|^2 = \frac{1}{2\sqrt{\varepsilon}} |H_0|^2 $$

(1.156)

The power coefficients cannot be determined by simply measuring the electric field $E$ or $H$ because no detector can measure the field value directly. At the same time, simply squaring the electric fields to obtain the power densities and taking the ratio typically leads to incorrect results. Due to diffraction, the beam size of transmitted wave is actually not the same as that of the incident wave, which is shown in Fig. 1.8

By defining the beam area at boundary as $A_0$, we have

$$ A_i \cos \theta_i = A_r \cos \theta_r = A_0 \cos \theta_i $$

$$ A_t = A_0 \cos \theta_i $$

Thus the power $W$ carrying by a light wave with an area can be obtained by

$$ W_\alpha = I_a \cdot A_a $$

48
Here $\alpha = i, r, t$.

For the incident wave,

$$W_i = \frac{\sqrt{\varepsilon_1}}{2} |E_i|^2 A_0 \cos \theta_i = \frac{1}{2\sqrt{\varepsilon_1}} |H_i|^2 A_0 \cos \theta_i$$  \hspace{1cm} (1.157)

For the reflected wave,

$$W_r = \frac{\sqrt{\varepsilon_2}}{2} |E_r|^2 A_0 \cos \theta_i = \frac{1}{2\sqrt{\varepsilon_1}} |H_r|^2 A_0 \cos \theta_i$$  \hspace{1cm} (1.158)

For the transmitted wave,

$$W_t = \frac{\sqrt{\varepsilon_2}}{2} |E_t|^2 A_0 \cos \theta_t = \frac{1}{2\sqrt{\varepsilon_2}} |H_t|^2 A_0 \cos \theta_t$$  \hspace{1cm} (1.159)

Here

$$\cos \theta_t = \sqrt{1 - \frac{\varepsilon_1}{\varepsilon_2}} \sin^2 \theta_i$$  \hspace{1cm} (1.160)
For S-polarization (1.142) and (1.143),

\[
R_s = \frac{W_s}{W_t} = \frac{|E_i^s|^2}{E_i^s} \quad (1.161)
\]

\[
T_s = \frac{W_s}{W_t} = \frac{\sqrt{\varepsilon_2 - \varepsilon_1 \sin^2 \theta_i}}{\varepsilon_1 \cos \theta_i} |E_i^s|^2 \quad (1.162)
\]

For P-polarization (1.152) and (1.153),

\[
R_p = \frac{W_p}{W_t} = \frac{|H_i^p|^2}{H_i^p} \quad (1.163)
\]

\[
T_p = \frac{W_p}{W_t} = \frac{\sqrt{\varepsilon_1 \varepsilon_2 - \varepsilon_1^2 \sin^2 \theta_i}}{\varepsilon_2 \cos \theta_i} |H_i^p|^2 \quad (1.164)
\]

Substituting all related items into (1.161) - (1.164), we can obtain

\[
R_s = \frac{\left| \frac{\sqrt{\varepsilon_2 - \varepsilon_1 \sin^2 \theta_i} - \sqrt{\varepsilon_1 \cos \theta_i + \frac{4 \varepsilon \sigma}{c}}}{\sqrt{\varepsilon_2 - \varepsilon_1 \sin^2 \theta_i} + \sqrt{\varepsilon_1 \cos \theta_i + \frac{4 \varepsilon \sigma}{c}}} \right|^2}{\varepsilon_2 - \varepsilon_1 \sin^2 \theta_i - \sqrt{\varepsilon_1 \cos \theta_i + \frac{4 \varepsilon \sigma}{c}}} \quad (1.165)
\]

\[
R_p = \frac{\left| \frac{\varepsilon_2 \sqrt{\varepsilon_1 \cos \theta_i} - \varepsilon_1 \sqrt{\varepsilon_2 - \varepsilon_1 \sin^2 \theta_i + \frac{4 \varepsilon \sigma}{c}} \sqrt{\varepsilon_1 \varepsilon_2 - \varepsilon_1^2 \sin^2 \theta_i - \cos \theta_i}}{\varepsilon_2 \sqrt{\varepsilon_1 \cos \theta_i} + \varepsilon_1 \sqrt{\varepsilon_2 - \varepsilon_1 \sin^2 \theta_i + \frac{4 \varepsilon \sigma}{c}} \sqrt{\varepsilon_1 \varepsilon_2 - \varepsilon_1^2 \sin^2 \theta_i - \cos \theta_i}} \right|^2}{\varepsilon_2 - \varepsilon_1 \sin^2 \theta_i} \quad (1.166)
\]

\[
T_s = \frac{\left| \frac{\varepsilon_2 \sqrt{\varepsilon_1 \cos \theta_i} \sqrt{\varepsilon_2 - \varepsilon_1 \sin^2 \theta_i} + \sqrt{\varepsilon_1 \varepsilon_2 - \varepsilon_1^2 \sin^2 \theta_i + \frac{4 \varepsilon \sigma}{c}} \sqrt{\varepsilon_1 \varepsilon_2 - \varepsilon_1^2 \sin^2 \theta_i - \cos \theta_i}}{\varepsilon_2 \sqrt{\varepsilon_1 \cos \theta_i} + \varepsilon_1 \sqrt{\varepsilon_2 - \varepsilon_1 \sin^2 \theta_i + \frac{4 \varepsilon \sigma}{c}} \sqrt{\varepsilon_1 \varepsilon_2 - \varepsilon_1^2 \sin^2 \theta_i - \cos \theta_i}} \right|^2}{\varepsilon_2 \sqrt{\varepsilon_1 \cos \theta_i} + \varepsilon_1 \sqrt{\varepsilon_2 - \varepsilon_1 \sin^2 \theta_i + \frac{4 \varepsilon \sigma}{c}} \sqrt{\varepsilon_1 \varepsilon_2 - \varepsilon_1^2 \sin^2 \theta_i - \cos \theta_i}} \quad (1.167)
\]

\[
T_p = \frac{\left| \frac{\varepsilon_2 \sqrt{\varepsilon_1 \cos \theta_i} \sqrt{\varepsilon_2 - \varepsilon_1 \sin^2 \theta_i} + \sqrt{\varepsilon_1 \varepsilon_2 - \varepsilon_1^2 \sin^2 \theta_i + \frac{4 \varepsilon \sigma}{c}} \sqrt{\varepsilon_1 \varepsilon_2 - \varepsilon_1^2 \sin^2 \theta_i - \cos \theta_i}}{\varepsilon_2 \sqrt{\varepsilon_1 \cos \theta_i} + \varepsilon_1 \sqrt{\varepsilon_2 - \varepsilon_1 \sin^2 \theta_i + \frac{4 \varepsilon \sigma}{c}} \sqrt{\varepsilon_1 \varepsilon_2 - \varepsilon_1^2 \sin^2 \theta_i - \cos \theta_i}} \right|^2}{\varepsilon_2 \sqrt{\varepsilon_1 \cos \theta_i} + \varepsilon_1 \sqrt{\varepsilon_2 - \varepsilon_1 \sin^2 \theta_i + \frac{4 \varepsilon \sigma}{c}} \sqrt{\varepsilon_1 \varepsilon_2 - \varepsilon_1^2 \sin^2 \theta_i - \cos \theta_i}} \quad (1.168)
\]

Here (1.165) to (1.168) are the general solutions for power coefficients in the interface system.

**Multilayer Graphene as an Interface**

Now we consider the presence of graphene at the boundary of two media. In this case, graphene will introduce so called “optical conductivity”. There are several prior
publications [47, 63, 64] are focusing on this point. Here we use some conclusions directly.

For our experiments, the wavelength of our incident laser pulse is ~792nm, and the optical conductivity of graphene in this range can be simplified as:

\[ \sigma_g = \sigma_{\text{unit}} = \frac{\pi e^2}{2h} \]  

(1.169)

As for multilayer graphene, optical conductivity can be written as

\[ \sigma_{N,g} = N \frac{\pi e^2}{2h} \]  

(1.170)

For convenience, we use the following expression

\[ \frac{4\pi \sigma_g}{c} = \pi \alpha \]  

(1.171)

Here \( \alpha \) is the “fine structure constant”, \( \alpha \approx 1/137 \). For multilayer graphene,

\[ \frac{4\pi \sigma_g}{c} = N \pi \alpha \]  

(1.172)
Substituting (1.172) into (1.165) to (1.168), we can deduce the power coefficients at multilayer graphene boundary

\[ R_s = \frac{\sqrt{\varepsilon_2 - \varepsilon_1 \sin^2 \theta_i - \sqrt{\varepsilon_1 \cos \theta_i + N \pi \alpha}}}{\sqrt{\varepsilon_2 - \varepsilon_1 \sin^2 \theta_i + \sqrt{\varepsilon_1 \cos \theta_i + N \pi \alpha}}}^2 \quad (1.173) \]

\[ R_p = \frac{\varepsilon_2 \sqrt{\varepsilon_1 \cos \theta_i + \varepsilon_1 \sin^2 \theta_i + N \pi \alpha} \sqrt{\varepsilon_2 - \varepsilon_1 \sin^2 \theta_i - \cos \theta_i}}{\varepsilon_2 \sqrt{\varepsilon_1 \cos \theta_i + \varepsilon_1 \sin^2 \theta_i + N \pi \alpha} \sqrt{\varepsilon_2 - \varepsilon_1 \sin^2 \theta_i - \cos \theta_i}}^2 \quad (1.174) \]

\[ T_s = \frac{\sqrt{\varepsilon_2 - \varepsilon_1 \sin^2 \theta_i}}{\sqrt{\varepsilon_1 \cos \theta_i + \varepsilon_1 \sin^2 \theta_i + N \pi \alpha}} \frac{2 \sqrt{\varepsilon_1 \cos \theta_i}}{\sqrt{\varepsilon_2 - \varepsilon_1 \sin^2 \theta_i + \sqrt{\varepsilon_1 \cos \theta_i + N \pi \alpha}}}^2 \quad (1.175) \]

\[ T_p = \frac{\sqrt{\varepsilon_1 \varepsilon_2 - \varepsilon_1 \sin^2 \theta_i}}{\varepsilon_2 \cos \theta_i} \frac{2 \varepsilon_2 \sqrt{\varepsilon_1 \cos \theta_i}}{\varepsilon_2 \sqrt{\varepsilon_1 \cos \theta_i + \varepsilon_1 \sin^2 \theta_i + N \pi \alpha} \sqrt{\varepsilon_2 - \varepsilon_1 \sin^2 \theta_i - \cos \theta_i}}^2 \quad (1.176) \]

Now we consider a special case:

1. Normal incident: \( \theta_i = 0 \)
2. Both medium 1 and 2 are air: \( \varepsilon_1 = \varepsilon_2 = 1 \)

Under such conditions, we can simplify the power coefficients as

\[ R_s = R_p = \left| \frac{N \pi \alpha}{2 + N \pi \alpha} \right|^2 \quad (1.177) \]

\[ T_s = T_p = \left| \frac{2}{2 + N \pi \alpha} \right|^2 \quad (1.178) \]

\[ A = 1 - R - T = \frac{4 N \pi \alpha}{\left| 2 + N \pi \alpha \right|^2} \quad (1.179) \]

Here \( A \) represents the absorption coefficient of multilayer graphene. These solutions keep with deductions in [65, 66].

For other cases, such as tilting of the incident pulse or other substrate, we can refer to general solutions from (1.165) to (1.168). The reflectance and transmittance of the structure are key components for studying optical nonlinear intensities of multilayer
Surface Hydrodynamic Model

We consider the case of a simple interface of two media with relative permittivity $\epsilon_1$ and $\epsilon_2$, as shown in the Fig 1.2. From [67], the total polarization can be written

$$P = P_b + P_f,$$

(1.180)

where $P_b = \epsilon_0 \chi_b E$. So $D = \epsilon_0 E + P_b + P_f$, satisfies $\nabla \cdot D = 0$. Then we have

$$\nabla \cdot P_f = -\epsilon_0 (1 + \chi_b) \nabla \cdot E,$$

(1.181)

In addition, the second equation for hydrodynamic model is [67]

$$-\beta^2 \nabla (\nabla \cdot P_f) + \frac{\partial^2}{\partial t^2} P_f + \gamma \frac{\partial}{\partial t} P_f = \epsilon_0 \omega_p^2 E,$$

(1.182)
Assuming a harmonic solution of the form $e^{-i\omega t}$, the polarization $P_f$ can finally be written as

$$P_f = -\epsilon_0 \frac{\omega_p^2}{\omega^2 + i\gamma\omega} \left( E - (1 + \chi_b) \frac{\beta^2}{\omega_p^2} \nabla(\nabla \cdot E) \right)$$  \hspace{1cm} (1.183)$$

where

$$\chi_f = -\frac{\omega_p^2}{\omega^2 + i\gamma\omega}$$  \hspace{1cm} (1.184)$$

The Transverse and Longitudinal Modes in Metal Material

In a metal

$$\nabla \times E = i\omega \mu_0 H$$  \hspace{1cm} (1.185)$$

$$\nabla \times H = -i\omega (\epsilon_0 (1 + \chi_b) E + P_f)$$  \hspace{1cm} (1.186)$$

$$= -\omega \epsilon_0 \epsilon (E - \alpha \nabla(\nabla \cdot E))$$  \hspace{1cm} (1.187)$$

where $\epsilon = 1 + \chi_b + \chi_f$ is the local relative permittivity of the metal, and

$$\alpha = \frac{\chi_f (1 + \chi_b) \beta^2}{\epsilon \frac{\omega_p^2}{\omega^2}}$$  \hspace{1cm} (1.188)$$

$$= \frac{\beta^2}{\frac{\omega_p^2}{1 + \chi_b} - \omega^2 - i\gamma\omega}$$  \hspace{1cm} (1.189)$$
There are two different solutions to these equations corresponding to two different kinds of waves. The first solution satisfies $\nabla \cdot \mathbf{E} = 0$, then these become

\begin{align*}
\nabla \times \mathbf{E} &= i\omega \mu_0 \mathbf{H} \\
\nabla \times \mathbf{H} &= -i\omega \epsilon_0 c \mathbf{E}
\end{align*}

Since the divergence of the electric field is zero, the electric field is orthogonal to the wave vector when the wave is a propagating wave, which means it is transverse. The second kind of solution is curl free, which means it satisfies $\nabla \times \mathbf{E} = 0$ and there is no accompanying magnetic field. This longitudinal wave yields the wave equation for a bulk plasmon $\nabla \times \mathbf{H} = 0$

\begin{equation}
\nabla (\nabla \cdot \mathbf{E}) - \frac{1}{\alpha} \mathbf{E} = \nabla^2 \mathbf{E} - \frac{1}{\alpha} \mathbf{E} = 0
\end{equation}

and the corresponding dispersion relation

\begin{equation}
k^2 = -\frac{1}{\alpha} = \frac{1}{\beta^2} \left( \omega^2 - \frac{\omega_p^2}{1 + \chi_b} + i\gamma \omega \right)
\end{equation}

An alternative way to this dispersion relation is

\begin{equation}
\epsilon_\parallel = 1 + \chi_b - \frac{\omega_p^2}{\omega^2 + i\gamma \omega - \beta^2 k^2} = 0
\end{equation}
For the s-polarization, $\nabla \times \mathbf{E} = 0$ yields $E_y = 0$, so that no bulk plasmon can be excited. Nonlocality has no impact on a s polarized wave. For a p-polarized wave

$$\partial^2_z H_y = -(\epsilon k_0^2 - k_x^2) H_y$$

so that the magnetic field can then be written as

$$H_y = (A e^{i k z t} + B e^{-i k z t}) e^{i k_x x - i \omega t}$$

with $k_z t = \sqrt{\epsilon k_0^2 - k_x^2}$, where $k_0 = \frac{\omega}{c}$. The $E_x$ and $E_z$ accompanying fields can be found

$$E_x = \frac{1}{i \omega \epsilon_0 \epsilon} \partial_z H_y$$
$$E_z = -\frac{1}{i \omega \epsilon_0 \epsilon} \partial_z H_y$$

And for longitudinal waves, we have

$$\partial^2_x E_x = (k_x^2 + \frac{1}{\alpha}) E_x$$

so that

$$E_x = (C e^{i k_z z} + D e^{-i k_z z}) e^{i k_x x - i \omega t}$$
with

\[ k_i^2 = \frac{1}{\beta^2} \left( \frac{\omega_p^2}{1 + \chi_b} - \omega^2 - i\gamma\omega \right) + k_x^2 \]  
(1.201)

\[ = k_x^2 + \frac{\omega_p^2}{\beta^2} \left( \frac{1}{\chi_f} + \frac{1}{1 + \chi_b} \right) \]  
(1.202)

and \( E_z \) can then be obtained

\[ E_z = \frac{1}{ik_x} \partial_z E_x \]  
(1.203)

**Reflection from a metal surface**

For a p-polarized wave, we have the field in dielectric

\[ H_y = (e^{-ik_zz} + \rho e^{ik_zz}) e^{ik_x x - i\omega t} \]  
(1.204)

\[ E_x = \frac{ik_z}{i\omega\epsilon_0\epsilon_d} (\rho e^{ik_zz} - e^{-ik_zz}) e^{ik_x x - i\omega t} \]  
(1.205)

\[ E_z = -\frac{k_x}{\omega\epsilon_0\epsilon_d} (1 + \rho) \]  
(1.206)

where \( k_z = \sqrt{\epsilon_d k_0^2 - k_x^2} \).

In metal,

\[ H_y = Ae^{ik_zz} e^{ik_x x - i\omega t} \]  
(1.207)

\[ E_x = \left( \frac{k_i}{i\omega\epsilon_0\epsilon} Ae^{ik_zz} + Be^{ik_zz} \right) e^{ik_x x - i\omega t} \]  
(1.208)

\[ E_z = \left( -\frac{ik_z}{i\omega\epsilon_0\epsilon} Ae^{ik_zz} + \frac{k_i}{ik_x} Be^{ik_zz} \right) e^{ik_x x - i\omega t} \]  
(1.209)
consider the continuity of $H_y$ and $E_x$

\[ 1 + \rho = A \]  
\[ (\rho - 1) \frac{ik_z}{\epsilon_d} = \frac{k_t}{\epsilon} A + i\omega\epsilon_0 B \]

We can also obtain the polarization induced by free electrons as

\[ P_{f_z} = -\frac{1}{i\omega} \frac{\partial}{\partial x} H_y - \epsilon_0 (1 + \chi_b) E_z \]  
\[ = A \frac{k_x}{\omega} (-1 + \frac{1 + \chi_b}{\epsilon}) - \epsilon_0 (1 + \chi_b) \frac{k_t}{ik_x} B \]

**Reflection for pure hydrodynamic model $P_{f_z} = 0$**

If we consider the continuity of $D_z$, which is the same as $P_{f_z} = 0$, then we have

\[ \rho_p = \frac{ik_z}{\epsilon_d} + \frac{k_t}{\epsilon} - \Omega \]  
\[ = \frac{ik_z}{\epsilon_d} - \frac{k_t}{\epsilon} + \Omega \]

where $\Omega = \frac{k^2}{k_t} (\frac{1}{\epsilon} - \frac{1}{1 + \chi_b})$, $k_t = \sqrt{k^2_x - \epsilon k^2_0}$, and

\[ k^2_t = \frac{1}{\beta^2} \left( \frac{\omega^2_p}{1 + \chi_b} - \omega^2 - i\gamma \omega \right) + k^2_x \]  
\[ = k^2_x + \frac{\omega^2_p}{\beta^2} \left( \frac{1}{\chi_f} + \frac{1}{1 + \chi_b} \right) \]
Reflection for surface-hydrodynamic model $\Delta \epsilon$

If we consider discontinuity as $\Delta \epsilon$, and apply $D_{mz} = D_{dz}$

\[
D_{mz} = P_{fz} + \epsilon_0 (1 + \chi_b) E_z
\]
\[
D_{dz} = \epsilon_0 \epsilon_d E_z
\]

then we put $\Delta \epsilon$ in the way $P_{fz} = \epsilon_0 \Delta \epsilon E_{mz}$

\[
A \frac{k_x}{\omega} (-1 + \frac{1 + \chi_b}{\epsilon} - \epsilon_0 (1 + \chi_b) \frac{k_l}{i k_x} B) = -\epsilon_0 (1 + \chi_b) (-\frac{i k_0}{i \omega \epsilon_0} A e^{k_l z} + \frac{k_l}{i k_x} B e^{k_l z}) \Delta \epsilon
\]

thus

\[
A \left\{ k_x \left(-1 + \frac{1 + \chi_b}{\epsilon} (1 + \Delta \epsilon)\right) \right\} = B \epsilon_0 \omega \left\{ (1 + \chi_b) \frac{k_l}{i k_x} \right\} (1 + \Delta \epsilon)
\]

or we can write it as

\[
i \alpha A = \epsilon_0 \omega \beta B
\]

equivalently

\[
i \omega \epsilon_0 B = -\frac{\alpha}{\beta} A
\]
where

\[
\alpha = k_x \left[ 1 + \frac{1 + \chi_b}{\epsilon} (1 + \Delta \epsilon) \right]
\]

(1.224)

\[
\beta = (1 + \chi_b) k_l k_x (1 + \Delta \epsilon)
\]

(1.225)

\[
\Omega^* = \frac{\alpha}{\beta} = \frac{k^2_x}{k_l} \left( \frac{1}{\epsilon} - \frac{1}{(1 + \chi_b)(1 + \Delta \epsilon)} \right)
\]

(1.226)

substitutedly,

\[
(\rho - 1) \frac{i k_z}{\epsilon_d} = \left( \frac{k_l}{\epsilon} - \frac{\alpha}{\beta} \right) A
\]

(1.227)

then

\[
(1 + \rho) \left( \frac{k_l}{\epsilon} - \frac{\alpha}{\beta} \right) = (\rho - 1) \frac{i k_z}{\epsilon_d}
\]

(1.228)

finally, we obtain

\[
\rho = \frac{i k_z}{\epsilon_d} + \left( \frac{k_l}{\epsilon} - \frac{\alpha}{\beta} \right)
\]

(1.229)

\[
= \frac{i k_z}{\epsilon_d} - \frac{k_l}{\epsilon} + \Omega^*
\]

(1.230)

The \textit{d} function method

In the \textit{d} function method [68], the Fresnel equation for reflection considers the fluctuation of the electron density near a metal/dielectric interface. It always exhibits the limitation of the “long wavelength approximation” (LWA), which means the width of
surface region is much smaller than the incident wavelength ($\xi/\lambda \ll 1$).

In detail, the origin ($z = 0$) is put at the edge of the positive background, and the metal extends in the positive $z$-direction, and considers a $p$-polarized incidence. The equations need to satisfy

\begin{align}
\nabla \times \mathbf{E} &= i\omega \mu_0 \mathbf{H} \\
\nabla \times \mathbf{H} &= -i\omega \epsilon \epsilon_0 \mathbf{E}
\end{align}

where $\xi_1$ is such that for $z \leq \xi_1$, $E(k, z, \omega)$ is of the form of an incoming and a reflected wave

\begin{align}
E_x(k, z, \omega) &= -ik_z(e^{ik_z z} - \rho_p e^{-ik_z z}) \cdot e^{ik_x x} \\
E_z(k, z, \omega) &= ik_x(e^{ik_z z} + \rho_p e^{-ik_z z}) \cdot e^{ik_z z} e^{ik_x x}
\end{align}

and corresponding magnetic field satisfies $i\omega \mu_0 H_y(z) = (\nabla \times \mathbf{E})_y$

\begin{equation}
H_y(k, z, \omega) = -i\epsilon \epsilon_0 \omega(e^{ik_z z} + \rho_p e^{-ik_z z}) \cdot \epsilon_1 \cdot e^{ik_z z} e^{ik_x x}
\end{equation}

The other length $\xi_2$ is such that for $z \geq \xi_2$, $E(k, z, \omega)$ has the form of a transverse refracted wave

\begin{align}
E_x(k, z, \omega) &= E_x^t e^{ipz} \cdot e^{ik_x x} \\
E_z(k, z, \omega) &= \frac{-k_x}{p} E_x^t e^{ipz} \cdot e^{ik_x x}
\end{align}
and

\[ H_y(k, z, \omega) = \frac{\epsilon_2\epsilon_0\omega}{p_t} E_x^t e^{ip_t z} \]  \hspace{1cm} (1.238)

where \( p_t^2 = k_0^2 - k_z^2 \) (provided we are below the plasmon frequency). Integrating

\[ \frac{\partial}{\partial z} H_y = i\omega D_x \] at \( z \)-direction from \( \xi_1 \) to \( \xi_2 \), where

\[ D_x(k, z, \omega) = \int_{-\infty}^{\xi_2} dz' \epsilon(k, z, z', \omega) E_x(k, z', \omega) \]  \hspace{1cm} (1.239)

we get

\[ H_y(\xi_2) - H_y(\xi_1) = \left( i\omega \int_{\xi_1}^{\xi_2} \Delta D_x(z) dz \right) E_x^t + i\omega E_x^t (\xi_2 \epsilon_2 - \xi_1 \epsilon_1) \]  \hspace{1cm} (1.240)

where we introduce

\[ \Delta D_x(z) = (D_x(z) - D^0_x(z))/E_x^t \]  \hspace{1cm} (1.241)

and

\[ D^0_x(z) = E_x^t (\epsilon_2 \theta(z) + \epsilon_1 \theta(-z)) \]  \hspace{1cm} (1.242)
When \( k_z \xi \) and \( p_t \xi \ll 1 \), we can expand the exponential in field expression to the first order

\[
H_y(k, z, \omega) = -i\epsilon_0\epsilon_1\omega(1 + \rho_p + ik_z z - ik_z z\rho_p) + O((k_z z)^2) \tag{1.243}
\]

\[
H_y(k, z, \omega) = \frac{\epsilon_2\epsilon_0\omega}{p_t} E_t^t(1 + ip_t z) + O((p_t z)^2) \tag{1.244}
\]

thus

\[
H_y(\xi_2) - H_y(\xi_1) = \frac{\epsilon_0\omega}{p_t} E_x^t + i\epsilon_0\omega(1 + \rho_p) + i\epsilon_0\omega E_x^t \xi_2
\]

\[
-\epsilon_0\omega k_z(1 - \rho_p)\xi_1 + O(k_z z)^2 \tag{1.245}
\]

Then we get

\[
E_x^t \left( \frac{\epsilon_0\omega\epsilon_2}{p_t} - i\epsilon_0\omega \int_{\xi_1}^{\xi_2} dz \Delta \bar{D}_x(z) \right) + i\epsilon_1\epsilon_0\omega(1 + \rho_p) = 0 \tag{1.246}
\]

We can also duplicate this process for \( E_x \) by integrating \( i\mu_0\omega H_y = \frac{\partial}{\partial z} E_x - ik_x E_z \)

\[
E_x(\xi_2) - E_x(\xi_1) - ik \left( \int_{\xi_1}^{\xi_2} dz \Delta \bar{E}_z(z) \right) E_x^t
\]

\[
-ik E_z^t(-\frac{\epsilon_2}{\epsilon_1} \xi_1 + \xi_2) = i\xi \int_{\xi_1}^{\xi_2} H_y(z)dz \tag{1.247}
\]
where we define

$$\Delta E_z(z) = \frac{(E_z(z) - E^0_z(z))}{E^t_z}$$

$$E^0_z(z) = E^t_z(\theta(z) + \frac{\epsilon_2}{\epsilon_1}(\theta(-z)))$$

Then we obtain

$$E^t_x \left(1 + \frac{ik^2_x}{pt} \int_{\xi_1}^{\xi_2} dz \Delta E_z(z)\right) + ik_x(1 - \rho_p) = 0$$

### Modified Reflectance

So that we obtain the reflection of a p-polarized incidence

$$\rho_p = \frac{\epsilon_2 k_z \left(1 - \frac{ip}{\epsilon_2} \int_{\xi_1}^{\xi_2} \Delta \tilde{D}_z dz\right) - \epsilon_1 p_t \left(1 + \frac{ik^2_x}{pt} \int_{\xi_1}^{\xi_2} \Delta \tilde{E}_z dz\right)}{\epsilon_2 k_z \left(1 - \frac{ip}{\epsilon_2} \int_{\xi_1}^{\xi_2} \Delta \tilde{D}_z dz\right) + \epsilon_1 p_t \left(1 + \frac{ik^2_x}{pt} \int_{\xi_1}^{\xi_2} \Delta \tilde{E}_z dz\right)}$$

$$= \frac{\epsilon_2 \tilde{p} - \epsilon_1 \tilde{p}_t}{\epsilon_2 \tilde{p} + \epsilon_1 \tilde{p}_t}$$

$$= \frac{\epsilon_2 k_z \left(1 + \frac{ip}{\epsilon_2} (\epsilon_2 - \epsilon_1) d_{||}\right) - \epsilon_1 p_t \left(1 + \frac{ik^2_x}{pt} (\epsilon_2 - \epsilon_1) d_{\perp}\right)}{\epsilon_2 k_z \left(1 + \frac{ip}{\epsilon_2} (\epsilon_2 - \epsilon_1) d_{||}\right) + \epsilon_1 p_t \left(1 + \frac{ik^2_x}{pt} (\epsilon_2 - \epsilon_1) d_{\perp}\right)}$$

where

$$\tilde{p} = k_z(1 + ip_t(\epsilon_2 - \epsilon_1)d_{||}(\omega)/\epsilon_2)$$

$$\tilde{p}_t = p_t(1 + i k^2_z(\epsilon_2 - \epsilon_1)d_{\perp}(\omega)/p_t)$$
where

\[ d_{\parallel}(\omega) = \frac{1}{\epsilon_1 - \epsilon_2} \int_{\xi_1}^{\xi_2} \Delta \hat{D}_x(z) \, dz \] (1.254)

\[ d_{\perp}(\omega) = \frac{1}{\epsilon_2 - \epsilon_1} \int_{\xi_1}^{\xi_2} \Delta \hat{E}_x(z) \, dz \] (1.255)

**d_{\parallel} and d_{\perp}**

\[ D_x(z) = \int_{\infty}^{z} \epsilon_{xx}(0, z, z', \omega) E_x(z', \omega) \, dz' \] (1.256)

Since suppose \( E_x \) is a weakly varying near the surface of metal, thus

\[
d_{\parallel}(\omega) = \frac{1}{\epsilon_1 - \epsilon_2} \int_{\xi_1}^{\xi_2} dz \left( \int_{\infty}^{z} \epsilon_{xx}(0, z, z', \omega) dz' - (\epsilon_2 \theta(z) + \epsilon_1 \theta(-z)) \right)
= \int_{\xi_1}^{\xi_2} dz \left\{ \theta(z) - \left( \int dz' \sigma_{xx}(z, z') \right) \right\}^{\frac{1}{\sigma_{xx}(\infty)}}
\] (1.257)

We first write \( \epsilon_{xx} = \epsilon_1 \delta(z - z') + (4\pi i/\omega) \sigma_{xx}(z, z') \) where \( \sigma_{xx}(\infty) = (\omega/4\pi i)(\epsilon_2 - \epsilon_1) \) is the limiting value of \( \sigma_{xx}(z, z') \) far into the bulk. Making use of the RPA allows us to estimate \( d_{||} \) as

\[
\frac{\sigma_{xx}(z, z')}{\sigma_{xx}(\infty)} = \delta(z - z') \left[ n_0(z) + \frac{1}{\omega^2} \hat{\sigma} + O\left(\frac{1}{\omega^4}\right) \right]
\] (1.258)

where

\[
\hat{\sigma} = -\frac{12}{5} \left[ \frac{1}{3} \frac{d^2}{dz^2} n_0^{5/3} + \frac{4}{3} \frac{d}{dz} n_0^{5/3} \frac{d}{dz} + \frac{d^2}{dz^2} \right]
\] (1.259)
and \( n_0(z) \) is the normalized self-consistent ground-state electron-density profile of the metal. So that to the first order

\[
d_{||} = \int_{\xi_1}^{\xi_2} dz (\theta(z) - n_0(z)) = \frac{\int_{\xi_1}^{\xi_2} dzzn'_0}{\int_{\xi_1}^{\xi_2} dzn'_0}
\]  

(1.260)

Usually we consider charge neutrality in the system, thus \( d_{||} \) is identical to zero in this level of approximation.

Using the Poisson’s equation it can shown that \( d_{\perp}(\omega) \) can be expressed in terms of the induced electron density \( \delta \rho(z) \) due to the incident radiation

\[
d_{\perp}(\omega) = \frac{\int_{\xi_1}^{\xi_2} dzz\delta \rho(z)}{\int_{\xi_1}^{\xi_2} d\delta \rho(z)}
\]  

(1.261)

From previously mentioned results, we obtain

\[
\rho = \frac{ek_z - \epsilon_d k_t + \epsilon_d \epsilon^2}{ek_z + \epsilon_d k_t - \epsilon_d \epsilon^2} \rho
\]  

(1.262)

\[
\rho = \frac{ek_z}{ek_z + \epsilon_d k_t} (1 + \frac{i \mu \epsilon}{e}(\epsilon - \epsilon_d)d_{||}) - \epsilon_d p_t \left( 1 + \frac{i k^2}{p_t}(\epsilon - \epsilon_d)d_{\perp} \right)
\]  

(1.263)

Using the definition \( p_t^2 = k_0^2 \epsilon - k_z^2 \) and set \( d_{||} = 0 \), then we get

\[
\rho = \frac{ek_z - \epsilon_d p_t \left( 1 + \frac{i k^2}{p_t}(\epsilon - \epsilon_d)d_{\perp} \right)}{ek_z + \epsilon_d p_t \left( 1 + \frac{i k^2}{p_t}(\epsilon - \epsilon_d)d_{\perp} \right)}
\]  

(1.264)

\[
= 1 - 2 \epsilon_d \frac{i k^2(\epsilon - \epsilon_d)d_{\perp}}{ek_z + \epsilon_d p_t \left( 1 + \frac{i k^2}{p_t}(\epsilon - \epsilon_d)d_{\perp} \right)}
\]  

(1.265)
Using the two \( \rho \) of the same value, then

\[
-ik^2(\epsilon - 1)d_\perp = i\epsilon_\epsilon d\Omega^* \tag{1.266}
\]

where

\[
\Omega^* = \frac{k_x^2}{k_i} \left( \frac{1}{\epsilon} - \frac{1}{(1 + \chi_b)(1 + \Delta \epsilon)} \right) \tag{1.267}
\]

Finally, we obtain

\[
\Delta \epsilon = \frac{\epsilon}{1 + \frac{\epsilon - \epsilon_d}{\epsilon_d} k_i d_\perp} \cdot \frac{1}{1 + \chi_b} - 1 \tag{1.268}
\]

**SHDM in the case of two interfaces**

So far we have derived the method for the single interface case. Now we are considering the case of a metal/dielectric metal slab with two interfaces here. For the top dielectric region, the electric field \( E \) and magnetic field \( H \) can be written as

\[
H_y = (e^{-ik_z} + re^{ik_z}) e^{ik_{xz} - i\omega t} \tag{1.269}
\]

\[
E_x = \frac{ik_z}{i\omega \epsilon_0 \epsilon_d} (re^{ik_z} - e^{-ik_z}) e^{ik_{xz} - i\omega t} \tag{1.270}
\]

\[
E_z = -\frac{k_x}{\omega \epsilon_0 \epsilon_d} (e^{-ik_z} + re^{ik_z}) e^{ik_{xz} - i\omega t} \tag{1.271}
\]
The fields in metal can be written as

\[ \begin{align*}
H_y &= (A e^{k_{t}z} + C e^{-k_{t}z}) e^{i k_{x}x - i \omega t} \quad (1.272) \\
E_x &= \left( \frac{k_{t}}{i \omega \epsilon_0 \epsilon} A e^{k_{t}z} + B e^{k_{t}z} + \frac{-k_{t}}{i \omega \epsilon_0 \epsilon} C e^{-i k_{t}z} + D e^{-k_{t}z} \right) e^{i k_{x}x - i \omega t} \quad (1.273) \\
E_z &= \left( - \frac{i k_{x}}{i \omega \epsilon_0 \epsilon} A e^{k_{t}z} + \frac{k_{t}}{i k_{x}} B e^{k_{t}z} - \frac{i k_{x}}{i \omega \epsilon_0 \epsilon} C e^{-k_{t}z} + \frac{-k_{t}}{i k_{x}} D e^{-k_{t}z} \right) \quad (1.274)
\end{align*} \]

The fields in the button dielectric can be written as

\[ \begin{align*}
H_y &= (t e^{-i k_{z}z}) e^{i k_{x}x - i \omega t} \quad (1.275) \\
E_x &= \frac{i k_{z}}{i \omega \epsilon_0 \epsilon_d} (t e^{-i k_{z}z}) e^{i k_{x}x - i \omega t} \quad (1.276) \\
E_z &= - \frac{k_{x}}{\omega \epsilon_0 \epsilon_d} (t e^{-i k_{z}z}) e^{i k_{x}x - i \omega t} \quad (1.277)
\end{align*} \]

Thus the continuity conditions for \( H_y \) and \( E_x \) at \( z = 0 \) are given by

\[ \begin{align*}
1 + r &= A + C \quad (1.278) \\
(r - 1) \frac{i k_{z}}{\epsilon_d} &= \frac{k_{t}}{\epsilon} (A - C) + i \omega \epsilon_0 (B + D) \quad (1.279)
\end{align*} \]

Considering the additional boundary condition as a discontinuity of dielectric constant as \( \Delta \epsilon \), we can write the normal component of the polarization as \( P_{fz} = \epsilon_0 \Delta \epsilon E_{mz} \), we obatin

\[ i \omega \epsilon_0 (B - D) = -\Omega^*(A - C) \quad (1.280) \]
where $\Omega^* = \frac{k^2}{k_t} \left( \frac{1}{\epsilon} - \frac{1}{\epsilon_b (1+\Delta\epsilon)} \right)$, $k_t = \sqrt{\frac{k^2_x}{\epsilon} - \epsilon k_0^2}$, and $k_t^2 = \frac{1}{\beta^2} \left( \frac{\omega_p^2}{1+\chi_b} - \omega^2 - i \gamma \omega \right)$, as also defined in the original manuscript. Similarly, the continuity condition for $H_y$ and $E_z$ of another interface located at $z = -h$ are given by

$$
te^{ik_h z} = Ae^{-k_h h} + Ce^{k_h h}$$

$$-te^{ik_h z} \frac{i k_z}{\epsilon_d} = \frac{k_t}{\epsilon} (Ae^{-k_h h} - Ce^{k_h h}) + i \omega \epsilon_0 (Be^{-k_h h} + De^{k_h h})$$

And from the discontinuity condition, we can obtain

$$i \omega \epsilon_0 (Be^{-k_h h} - De^{k_h h}) = -\Omega^* (Ae^{-k_h h} - Ce^{k_h h})$$

By combining eq. (10-15), the reflection coefficient of the two-interface case can be obtained through the following matrix

$$
\begin{bmatrix}
1 & -1 & 0 & -1 & 0 & 0 \\
\frac{i k_z}{\epsilon_d} & -\frac{k_t}{\epsilon} & -i \omega \epsilon_0 & \frac{k_t}{\epsilon} & -i \omega \epsilon_0 & 0 \\
0 & \Omega^* & i \omega \epsilon_0 & -\Omega^* & -i \omega \epsilon_0 & 0 \\
0 & e^{-k_h h} & 0 & e^{k_h h} & 0 & -e^{ik_h z} \\
0 & \frac{k_t}{\epsilon} e^{-k_h h} & i \omega \epsilon_0 e^{-k_h h} & -\frac{k_t}{\epsilon} e^{-k_h h} & i \omega \epsilon_0 e^{k_h h} & \frac{i k_z}{\epsilon_d} e^{ik_h z} \\
0 & \Omega^* e^{-k_h h} & i \omega \epsilon_0 e^{-k_h h} & -\Omega^* e^{k_h h} & -i \omega \epsilon_0 e^{k_h h} & 0 \\
\end{bmatrix}
= \begin{bmatrix} r \\ A \\ B \\ C \\ D \\ t \end{bmatrix}
$$

On the other hand, we can also derive the reflection coefficient for the two-interface case from d-parameter approach. The reflection coefficient from dielectric to metal
has been defined in the original manuscript as

\[
 r_{D}^{12} = \frac{\epsilon k_{x} - \epsilon_{d} k_{t} - i k_{x}^{2} (\epsilon - \epsilon_{d}) d_{\perp}}{\epsilon k_{x} + \epsilon_{d} k_{t} + i k_{x}^{2} (\epsilon - \epsilon_{d}) d_{\perp}}
\]  

(1.285)

similarly, the reflection coefficient from metal to dielectric can also be written as

\[
 r_{D}^{21} = \frac{\epsilon_{d} k_{x} - \epsilon k_{t} - i k_{x}^{2} (\epsilon_{d} - \epsilon) d_{\perp}}{\epsilon_{d} k_{x} + \epsilon k_{t} + i k_{x}^{2} (\epsilon_{d} - \epsilon) d_{\perp}}
\]  

(1.286)

using the boundary condition of \( E_{x} \), the relation between the transmitted coefficient \( t \) and the reflection coefficient can be written as \( \frac{1+r_{D}^{12}}{\sqrt{r_{D}^{12}}} = \frac{t_{D}^{12}}{\sqrt{r_{D}^{12}}} \), the transmitted coefficient can also be approximately written as \( t_{D}^{12} = \sqrt{\frac{\epsilon_{d}}{\epsilon_{1}}}(1 + r_{D}^{12}) \) Thus the two-interface system can be described by a cascade of transfer matrix

\[
 \begin{bmatrix}
 E_{in}^+ \\
 E_r^+
 \end{bmatrix} = \begin{bmatrix}
 \frac{1}{t_{12}} & -\frac{r_{21}}{t_{12}} \\
 \frac{r_{12}}{t_{12}} & t_{21} - \frac{r_{12} r_{21}}{t_{12}}
 \end{bmatrix} \begin{bmatrix}
 E_{down}^+ \\
 E_{up}^+
 \end{bmatrix}
\]  

(1.287)

\[
 \begin{bmatrix}
 E_{down}^- \\
 E_{up}^-
 \end{bmatrix} = \begin{bmatrix}
 e^{ik_{z} h} & 0 \\
 0 & e^{-ik_{z} h}
 \end{bmatrix} \begin{bmatrix}
 E_{down}^- \\
 E_{up}^-
 \end{bmatrix}
\]  

(1.288)

\[
 \begin{bmatrix}
 E_{down}^- \\
 E_{up}^-
 \end{bmatrix} = \begin{bmatrix}
 \frac{1}{t_{21}} & -\frac{r_{12}}{t_{21}} \\
 \frac{r_{21}}{t_{21}} & t_{12} - \frac{r_{21} r_{12}}{t_{21}}
 \end{bmatrix} \begin{bmatrix}
 E_t \\
 0
 \end{bmatrix}
\]  

(1.289)

Using this transfer matrix method, we can calculate the reflection coefficient for the two-interface case from the d-parameter approach.

However, this consideration is beyond the scope of this thesis. It does however de-
serve further careful considerations when the thickness becomes smaller. We should also note that the skin depth of the metal is still much smaller than the thickness considering in this paper, so that $C$, $D$ is much smaller than $A$, $B$ for the SHDM, and $E_{down}^+$ is much larger than the $E_{up}^+$ for the d-parameter approach. Thus the reflection coefficient of the two-interface case for both SHDM and d-parameter approach both reduce to that of a single interface case, which validate our assumption in Chapter 3.

**Theory of Coherent Perfect Absorption (CPA)**

![Schematic diagram of a ring resonator with two bus waveguides for coupled mode analysis.](image)

Coherent perfect absorption (CPA) is a recently discovered phenomenon, which allows for phase controllable modulation of absorption in a resonator cavity [69, 70]. CPA rises from the interference between two degenerate resonator modes, which are excited by launching two sources into the same cavity. Perfect absorption can be
achieved at critical coupling condition by appropriately designing the coupling coefficients and intrinsic absorption coefficient. CPA has been proposed for a number of on-chip applications [71] and has been utilized by our group previously an integrated optical modulator using a ring resonator (RR) based device [72].

**Coupled Mode Analysis of CPA**

In the coupled modes analysis, the equation for the evolution of the resonator mode in time domain is given by

$$\frac{da}{dt} = \left( j\omega_0 - \frac{1}{\tau_o} - \frac{1}{\tau_e} - \frac{1}{\tau'_e} \right) a + \kappa_1 s_{+1} + \kappa_2 s_{+2} + \kappa_3 s_{+3} + \kappa_4 s_{+4}$$  \hspace{1cm} (1.290)

By power conservation, the out-going waves are

$$s_{-1} = e^{-j\beta d} (s_{+2} - \kappa_2^* a)$$  \hspace{1cm} (1.291)

$$s_{-2} = e^{-j\beta d} (s_{+1} - \kappa_1^* a)$$  \hspace{1cm} (1.292)

$$s_{-3} = e^{-j\beta d} (s_{+4} - \kappa_3^* a)$$  \hspace{1cm} (1.293)

$$s_{-4} = e^{-j\beta d} (s_{+3} - \kappa_4^* a)$$  \hspace{1cm} (1.294)

where $\beta$ and $\beta'$ are the propagation constants in the bus and the receiver waveguides.

We can write the coupling coefficients as

$$\kappa_i = \sqrt{\frac{1}{\tau_{ei}}} e^{j\theta_i}, \quad i = 1, \ldots, 4$$  \hspace{1cm} (1.295)
where \(1/\tau_{e1,3}\) and \(1/\tau_{2,4}\) are defined as the decay rates in the forward and backward direction respectively, satisfying

\[
\frac{1}{\tau_{e1}} + \frac{1}{\tau_{e2}} = \frac{2}{\tau_e} \quad (1.296)
\]
\[
\frac{1}{\tau_{e3}} + \frac{1}{\tau_{e4}} = \frac{2}{\tau_{e'}} \quad (1.297)
\]

When we only have one inputs \(s_{+1}\), then

\[
a = \frac{\sqrt{\frac{1}{\tau_{e1}} e^{j\theta_1 s_{+1}}}}{j(\omega - \omega_0) + \frac{1}{\tau_0} + \frac{1}{\tau_e} + \frac{1}{\tau_e'}} \quad (1.298)
\]

When we have two input \(s_{+1}\) and \(s_{+4}\), then

\[
a = \frac{\sqrt{\frac{1}{\tau_{e1}} e^{j\theta_1 s_{+1}}} + \sqrt{\frac{1}{\tau_{e4}} e^{j\theta_4 s_{+4}}}}{j(\omega - \omega_0) + \frac{1}{\tau_0} + \frac{1}{\tau_e} + \frac{1}{\tau_e'}} \quad (1.299)
\]

Assuming \(\theta_1 = \theta_4 = 0\), \(1/\tau_{e2} = 1/\tau_{e3} = 0\), \(d = 0\) and define \(\frac{2}{\tau_e} = \frac{2}{\tau_{e'}} = \gamma_c\), then we can get

\[
a = \frac{\sqrt{\gamma_c} (s_{+1} + s_{+4})}{j\Delta\omega + \frac{\eta_0}{2}} s_0 \quad (1.300)
\]
\[
s_{-2} = s_{+1} - \sqrt{\gamma_c} a \quad (1.301)
\]
\[
s_{-3} = s_{+4} - \sqrt{\gamma_c} a \quad (1.302)
\]
Thus we can obtain

\[
\Gamma_A = \frac{s_A}{s_0} = \frac{1}{\sqrt{2}} \left( \frac{\gamma_i}{2} - \gamma_c e^{j \Delta \phi} \right) + j \Delta \omega \frac{1}{2 \gamma_{\text{tot}} + j \Delta \omega}
\]  

(1.303)

When the ring is on resonance frequency, \( \Delta \omega = 0 \). The transmission will be zero, if it satisfies the critical coupling condition

\[
\gamma_i = 2 \gamma_c
\]  

(1.304)

**Matrix Analysis of CPA**

Typically CPAs have used waveguides and ring resonators. Here we use a transform matrix for analyzing the performance of this ring resonator. We consider only clockwise direction first. The top half of the ring can be treated as a 2×2 directional coupler, which is related with the transform matrix as

\[
s_1 I_1 + k_1 A = O_3
\]  

(1.305)

\[
k_1 I_1 + s_1 A = B
\]  

(1.306)

For the bottom half of the ring, similarly, we can also obtain

\[
s_2 I_4 + kB' = O_2
\]  

(1.307)

\[
k_2 I_4 + s_2 B' = A'
\]  

(1.308)
Figure 1.11: Schematic diagram of a ring resonator with two bus waveguides for matrix analysis.

where A, A’ and B, B’ are related by $A' = A e^{\gamma L e^{-j n_e f f L}}$ and $B' = B e^{-\gamma L e^{j n_e f f L}}$.

Thus we can obtain the transfer matrix for the ring resonator as a 2×2 component

$$\begin{bmatrix}
  s_1 + \frac{k_1^2 s_2 H^2}{1 - s_1 s_2 H^2} & \frac{k_1 k_2 H}{1 - s_1 s_2 H^2} \\
  \frac{k_1 k_2 H}{1 - s_1 s_2 H^2} & s_2 + \frac{k_2^2 s_1 H^2}{1 - s_1 s_2 H^2}
\end{bmatrix}
\begin{bmatrix}
  I_1 \\
  I_4
\end{bmatrix} =
\begin{bmatrix}
  O_3 \\
  O_2
\end{bmatrix} \quad (1.309)$$

where $H = e^{-\gamma L e^{j n_e f f L}/2}$. The CPA requires $O_3 = 0$ when $I_1 = I_4$. Thus at a symmetric case that $s_1 = s_2 = s$, $k_1 = k_2 = k$, it reduces to

$$s + \frac{k^2 s H^2}{1 - s^2 H^2} = -\frac{k^2 H}{1 - s^2 H^2} \quad (1.310)$$
since \( s = \sqrt{1 - \alpha} \), \( k = j \sqrt{\alpha} \), where \( \alpha \) is the power split ratio. By substituting \( k^2 = s^2 - 1 \), we can obtain

\[
H s^2 - (1 - H^2) s - H = 0
\]  

(1.311)

thus

\[
(s - H)(s + \frac{1}{H}) = 0
\]

(1.312)

Then we can obtain the requirement for CPA is

\[
s = H
\]

(1.313)

So far, we have demonstrated the CPA condition, that is when \( I_1 = I_4 \), then \( O_3 = 0 \); when \( I_1 = -I_4 \), then \( O_3 \neq 0 \).

### 1.3 Third Harmonic Generation Experiment

Third harmonic generation (THG) is the third-order nonlinear optical process that three photons at frequency \( \omega \) are annihilated and one photon of frequency \( 3\omega \) is created. Fig. 1.12 shows the energy-level description of THG process. Our nonlinear optical microscopy setup uses a 76MHz train of 792 nm pulses from a Ti:sapphire laser (Coherent Mira). The intensity of laser pulses is adjusted by a variable neutral density filter and the pulse width is compressed to 50 fs by a Brewster-angle prism pair. The laser is transmitted through a dichroic mirror and focused by a lens
(f=50mm) onto the sample with a typical average power of 100 mW at a 0° incidence angle (normal incidence). The sample is mounted on a mated orthogonal-translation (Klinger) and rotation stage (Thorlab, CR1-Z7), which permitted 2D scanning as well as measurements of the rotational anisotropy of the signals from the sample. The reflected signal is passing through the focused lens again and reflected by the dichroic mirror, and then a Pellin-Broca prism (Thorlab, material: UV-grade silica) to spatially filter out the fundamental beam and a monochromator before being detected by a photomultiplier tube (Hamamatsu). The Pellin-Broca prism is mounted on a rotatable stage so that one can select a wavelength to be detected. The monochromator, whose throughput is approximately 10% due to its small bandwidth window, reduces background signal significantly to the level of the dark count.

**Software for Micron-Scale Scanning System**

Motorized 2D-scanning-translational and rotational stages are key equipments to image a nonlinear optical response from a micron-sized sample. Full automation of both motorized stages and a photon counter speeds up the rate of data collection and im-
Figure 1.13: Schematic diagram of the THG measurement setup

Figure 1.14: Photograph of a custom nonlinear optical scanning microscope for SHG, THG, and photoluminescence imaging experiments.
proves the accuracy and reproducibility. In order to control the stages and provide user-friendly interface, a developed Labview software programs was used, which is shown in Fig. 1.15.

In the “area scan” mode, one can scan an area of the sample and change the orientation. “X-Start”, “X-End”, “Y-Start”, and “Y-End” determine the area for scan. “X-interval” and “Y-interval” determine the scanning resolution. All the other parameters, “Disc level”, “dwell time”, and “integration time” are related to the photon counter control. There are also parameters to record external information such as “beam power”, “pol”, “Sample”, etc. This information is recorded in the first several lines in the data file.
1.4 Fabrication Process

Here, we briefly describe the fabrication process used in silicon component fabrication that was performed at Brookhaven National Lab. The wafer we used are SOI wafers consisting of a 220 nm Si top layer on a 3 μm SiO₂ layer and mounted on a 625 μm Si carrier wafer. This wafer is usually cut into pieces with a typical size of 1cm × 1cm.

The silicon wafer is then immersed and sonicated in acetone, followed by methanol, and then isopropanol alcohol for 5 minutes each in order to clean up the surface. Compressed N₂ is used for drying the wafer. Then the chip is baked at 450° for 30 minutes to remove the surface moisture. The chip is going through a O₂ plasma cleaned at 100 W for 2 minutes. A hexamethyldisilazane (HMDS) monolayer is then applied on top of the wafer chips utilizing a Yield Engineering Systems HMDS oven.

The chips are then spin coated with Dow Corning XR1541-006 at 4500 rpm for 60 seconds so that the photoresist can be uniformly placed on top of the chip. Three drops are typically used for a 1cm × 1cm sample. A XR1541-006 flowable oxide composed of a 6% concentration of HSQ is used as a negative photoresist. After coating the resist, the chip is then baked at 80° for 4 minutes to ensure that all solvent is removed from the resist.

The chip is then exposed to a electron-beam lithography (JEOL 6300FS) process. A 100 kV acceleration voltage, a 2 nm spot size and a 4 nA beam current are then used for the setup.

After electron-beam exposure, we then use a developer which is a mixture of 1.5 grams of NaOH and 6 grams of NaCl and 150 mL of de-ionized (DI) water. The developer
is then split into two equal parts. The chip is then put into the first developer and sonicate for 5 seconds, and then place into the second developer and sonicate for 5 seconds, and then stayed in the second developer for 10 minutes. After the developing process, the chip is placed in a DI water, sonicate for 5 second, and stay for 10 minutes. SF₆/O₂ is used with a pre-cooled process to be −100° for etching process under Oxford Plasmalab etcher. The final step is to deposit a layer of SiO₂ using Trion Orion III PECVD tool. A diamond pen is usually used for getting a clean edge on the purpose of edge coupling.

Fig. 1.16 shows an example of resulting fabricated components. These elementary components are used to validate our design models.

Figure 1.16: SEM images of (a) a delay line (b) a ring resonator.

1.5 Other Tools Used

In fact, there are still many other tools, which are very important as well and have been utilized in my research, but are not mentioned above. For example (1) effective index method for calculation propagation constant and dispersion in Si waveguides (2) Coupling theory for calculating the coupling loss between the waveguide and ring
resonator (3) Optsim for integrated photonic circuit simulation and lightwave system simulation, and so on. For simplicity these are not mentioned here but they are also very important for the projects.
Chapter 2

*Numerical Surface-Corrected Nonlocal Electrodynamic Model for Nanophotonic Structure*

Recent work has shown the significance of a nonlocal dielectric response and surface correction for nanometer-length-scale plasmonic structures. In this chapter, we propose a new surface hydrodynamic model, which incorporates such nonlocality and surface correction. Our approach, which is based on the hydrodynamic model (HDM), uses a numerical approach based on full-wave numerical computation and thus enables calculation of surface plasmon polaritons (SPP) fields. The model also introduces a discontinuity in the normal component of the electric displacement at the interface by taking into account the change in the electron distribution at the interface. The method makes use of the Feibelman $d$-parameter approach for the surface correction by equating the corrected Fresnel reflection of a p-polarized incident field across a planar interface. To demonstrate our method, we first examine numerical calculation of SPP propagation at a single interface structure; this work is then followed by a demonstration for a set of more complex thin-film structures. The latter demonstrates that our method is suitable for use in numerical modeling of complex nanophotonic structures.
2.1 Introduction

Recent studies [67, 73–75] have shown that in plasmonic guided-wave structures accurate simulation of polaritonic physics is limited in many cases by the importance of spatial dispersion or nonlocal polarization, particularly for frequencies near the surface plasmon resonance frequency. These structures extend from simple metal/dielectric interfaces [67, 74] to more complex guided-wave layered structures [73, 75]. Increasingly the interests for plasmonic design in such complex spatially dispersive structures require rigorous computational methods. In fact, the emergence of new nanofabrication tools has enabled the physical realization of nanoscale plasmonic systems for a variety of potential applications [76–85]. Since surface plasmon polaritons (SPPs) remain tightly confined beyond the free-space diffraction limit, they offer unique prospects for applications requiring ultrasmall photonic systems, e.g., nanoscale resolution optical imaging systems [81], deeply subwavelength-size waveguides [79, 80]. Spatial dispersion effects become significant for near-atomic length scale structures [86–90] and also even in planar interfacial structures [67, 73–75], thus limiting the use of the classical electromagnetic approach for simulation for device design [73, 87–90].

While time-dependent density-functional theory (TDDFT) [91–94], in principle, can bridge the gap between classical methods and those for nanoplasmonics, its explicit application is, in practice, limited to systems containing only a few clusters due to computational constraints. On the other hand, the semiclassical approach to this problem, that is embodied by the hydrodynamic model (HDM) [95, 96], has attracted
much interest, since it retains the physical insight and predictive value of classical electrodynamics, as well as its computational efficiency. Major advances in the hydrodynamic model, using, for example, concepts such as quantum pressure, have been made in order to deal with issues such as developing a more accurate prediction of the reflection at an interface [67, 97] or a more accurate prediction of dispersive behavior of SPPs near their plasmon resonance frequency [73, 98]. But despite these major advances, additional challenges remain. These challenges, which include the (1) effects of the spill-out of conduction electrons beyond the material boundaries [99, 100], and (2) incomplete accounting of surface-enabled plasmon damping by electron-hole pair creation [74, 101, 102], can be simultaneously overcome by extending the applicability of Feibelman’s $d$-parameters [68]. This latter development has been inspired, in part by recent computational [103] and analytical methods [86]. As the scale of metallic photonic structures begin to decrease to subwavelength dimensions, the effect of the surface contribution to nonlocal polarization has become more important. Hence in solving hydrodynamic problems, it is important to take into account the nature of the surface electrons more accurately.

In this chapter, we use full-wave numerical methods and augment the standard hydrodynamic method with an additional surface continuity equation so as to obtain a new surface hydrodynamic model (SHDM). We investigate this method for use with SPPs at single and multiple interfaces, i.e., structures, which have been examined by others for studies of spatial dispersion or localization in plasmonic problems [67, 73–75]. We begin by considering a a notional diagram described by $2.1(a)$ of the equilibrium $n_0(r)$, perturbed $n(r)$, and induced $\delta n(r)$ electron densities, which
leads to an induced surface density pressure (2.1(b)); 2.1(c) indicates the relevant parameters, which will be described below. To examine our modified hydrodynamic model, we first determined an effective perturbation of the dielectric constant for the metal at boundary condition, which is denoted by the discontinuity of the surface dielectric constant $\Delta \epsilon$ (see sketch in 2.1). This term is obtained via the reflectance at a planar interface as calculated using the well-known Feibelman $d$-parameter approach and its dependence on excess surface electron density, $\delta n$ [67]. This discontinuity in the dielectric constant describes the spill-out information, i.e., the excess surface electron density, which is encoded in the $d$-parameters. The deduced parameter $\Delta \epsilon$ is then used to impose a modified boundary condition for the displacement $\Delta D_n$ at the interface. We have implemented this numerical approach using COMSOL by employing an effective surface current flowing along the boundary surface.

To show the accuracy of our nonlocal surface term $\Delta \epsilon$, the results of the SHDM model with this nonlocal term are compared to an analytical model proposed and tested by Khurgin [74] for a model interface. Finally, it is shown that our approach and its numerical implementation can be applied more readily to a structurally more complex class of problems than the $d$-parameter approach [68]. Specifically, we have studied surface plasmon-polariton (SPP) dispersion curves in double-interface structures and compared the results obtained using the local-response approach and standard nonlocal HDM. We note, however, that the Mortensen Group [86, 104] has recently extended the $d$-parameter paradigm, which also rigorously remedies this deficiency through the incorporation of the Feibelman $d$-parameters. Our approach can also achieve more accurate results near plasmon resonance frequency than the tra-
ditional hydrodynamic method; this region of frequency space has typically a more complex behavior for the SPP dispersion relation [73, 98, 102].

Figure 2.1: Notional diagram of (a) the equilibrium and non-equilibrium near-surface electron density, including $d_\perp$ parameter; (b) HDM pressure-driven electrons. Panel (c) provides a schematic sketch indicating the polarization quantities used in this paper. The spill-over electron density $\delta n(r)$ via the $d$-parameter gives rise to a non-equilibrium discontinuity in the dielectric constant, which is used to describe the nonlocal response of the metal-dielectric system.
2.2 Basic Electromagnetic Framework

Several [74, 95, 98, 105–109] approaches have been derived to examine the response of a material in the presence of nonlocal polarization. In particular for the case of a classical medium, the material electromagnetic polarization depends not only on the local electric field but also that, which is typically removed by atomic distances from a location of interest. In this case, the electric field $E(r)$ in the medium is then formally governed by

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

(2.1)

where $\omega$ is the angular frequency of the incoming radiation and $\epsilon(r, r')$ is the nonlocal dielectric response function.

To investigate this problem in detail, we start with the well-known hydrodynamic framework [98, 110, 111]. In this formulation, it is found that a wavevector dependence of the degree of localization arises from the spatial dispersion produced by the introduction of a random carrier velocity and, hence, a finite particle pressure. Thus, sensibly the hydrodynamic model will reduce to the usual local dielectric case in the cold plasma approximation, where $v_m = 0$ and where $v_m$ is the mean velocity of the carriers. The usual hydrodynamic pressure gradient, which is obtained from the functional derivative of the universal energy functional includes the exchange, correlation, and kinetic energy of the system. Note that the hydrodynamic model can then introduce a nonlocal polarization by means of a pressure gradient.
For certain metals, e.g., silver, the contribution from the screened $d$ electrons is important [67, 112]. The total polarization $P$ can then be treated as having contributions from both bound ($P_b$) and free electrons ($P_f$), i.e., $P = P_b + P_f$. For many cases [94], the hydrodynamic contribution to the nonlocality of the permittivity stems predominantly from free electrons. Hence, the nonlocality induced by the free electrons can be described by a Navier-Stokes-like equation [67, 95, 96]

$$\beta^2 \nabla (\nabla \cdot P_f) + \nabla P_f + \gamma \nabla E = \epsilon_0 \omega_p^2 E$$  \hspace{1cm} (2.2)

where $\gamma$ is the damping factor, due to collisions within the electron gas, and $\omega_p$ is the plasmon frequency. The divergence term in this expression is a result of the internal kinetic energy of the electron gas and $\beta$ is a phenomenological nonlocal parameter proportional to the Fermi velocity $v_F$, usually appearing as $\beta = \sqrt{3/5}v_F$ [67, 74]. This term describes the nonlocal component of the polarization. Recall that the polarization of a bound electron can be simply described by $P_b = \chi_b \epsilon_0 E$. The hydrodynamic model can then be fully implemented by combining the nonlocal polarization relation with the standard Maxwell’s equation [111]

$$\nabla \times \nabla \times E = \frac{\omega^2}{c^2} E + \omega^2 \mu_0 P$$  \hspace{1cm} (2.3)
2.3 Feibelman \( d \) parameters

In many device applications, it is necessary to examine the classical local response (LR) of light scattering at a basic optical interface. For plasmonic problems, the planar interface at \( z = 0 \) is separated into a metallic \((z < 0)\) and a dielectric \((z > 0)\) regions with LR bulk dielectric functions \( \epsilon_m(\omega) \) and \( \epsilon_d(\omega) \), respectively (see sketch in 2.1(a)). This separation implies that the induced charge density \( \delta n(r) \), is confined strictly at the interface and is expressed as a product of the Dirac delta function and the surface charge density, such that \( \delta n(r) = \delta(z) \sigma(x,y) \), where \( r \) is the position vector \( r(x,y,z) \). In fact, such interfacial confined charge distributions are known to be present in a wide variety of material systems [113]; see for example the sketch in 2.1(a).

A powerful method to address such interface problems is to make use of the Feibelman \( d \)-parameter approach. This approach was demonstrated in Feibelman’s seminal work on planar semi-infinite systems [68, 112, 114]. In this approach, the first-order extension of the flat charge distribution introduces two auxiliary quantities, \( d_\perp \) and \( d_\parallel \), which parametrize the first moments of the induced charge \( \delta n \) and the parallel component of the induced current density \( j_\parallel \)

\[
\begin{align*}
  d_\perp(\omega) & = \frac{\int dz \delta n(z)}{\int dz \delta n(z)} \\
  d_\parallel(\omega) & = \frac{\int dz z \partial j_\parallel/\partial z}{\int dz z \partial j_\parallel/\partial z}
\end{align*}
\]

Both quantities define characteristic length scales for the dynamical problem associ-
ated with the centroid of induced charge, $d_\perp$, and the normal derivative of induced charge, $d_\parallel$. Notably, $d_\parallel$ vanishes for a charge-neutral strictly planar interface [112], leaving $d_\perp$ as the main quantity of interest for intrinsic quantum corrections. Note that the $d$-parameters are implicit functions of both wavevector $k$ and frequency $\omega$. The $k$ dependence of the $d$-parameter is found to be weak [86]. Thus in this chapter, we treat $d = d(k = 0)$. The non-uniform distribution of electrons, represented by Feibelman’s $d$-parameters, has been shown [68, 112, 114] to give corrections on optical properties beyond those obtained in the classical description, e.g., the Fresnel equations and the plasmon dispersion.

The $d$-parameter of a silver surface is shown as a function of wavelength in 2.2(a), a result, which is obtained from time-domain local density approximation (TDLDA) calculations assuming a Wigner-Seitz radius $r_s = 3$ [115, 116]; bulk damping is neglected in this calculation since it does not significantly influence the $d$-parameter [117].

## 2.4 Surface Hydrodynamic Model

Our surface hydrodynamic model can be described briefly as follows: We solve the Maxwell and hydrodynamic equations, with the addition of a surface contribution due to spill-out effects, using the $d$-parameters calculated from TD-LDA results. We begin by first considering the standard HDM. In the standard HDM, which is explored in many investigations of nonlocal plasmonic response [67, 95, 96], a hard-wall boundary condition, $P_n = 0$, is applied at $z = 0$, where $P_n$ is the normal
component of electric polarization \( \mathbf{P} \). The standard HDM has two main limitations: (1) It does not describe the electron spill-over beyond the metal/dielectric interface, and (2) it does not accurately deal with the plasmonic damping induced from electron-hole excitation. In particular, the \( d \)-parameters produced by a standard HDM can be represented as 
\[
d_{||}^{HDM} = 0, \quad d_{\perp}^{HDM} = -v_F \sqrt{\frac{3}{5(\omega_p^2 - \omega^2)}} \quad [86].
\]
The consideration of the interfacial electron distribution suggests, as a first-order correction, to include a discontinuity in the electric displacement at the interface; hence the introduction of the surface HD model. In order to obtain this correction, we set the reflection coefficient obtained from the surface-HDM equal to the reflection coefficient obtained from the \( d \)-parameter approach. The process of equating the reflection coefficients in the two models is also suggested in a recent work [103].

In order to obtain the discontinuity in the normal component of the electric displacement for the boundary condition given by 
\[
D_n^d = D_{n}^m + \Delta \epsilon \epsilon_0 E_{n}^m,
\]
we use the calculations given in the Appendix A. This approach allows us to obtain the corrected Fresnel reflection coefficients for the surface HDM

\[
r_{\text{SHDM}} = \frac{ek_x - \epsilon_d k_t + i\epsilon \epsilon_d \Omega^*}{ek_x + \epsilon_d k_t - i\epsilon \epsilon_d \Omega^*},
\]

(2.6)

and the \( d \)-parameter approach [114]

\[
r_{D} = \frac{ek_x - \epsilon_d k_t - ik^2_x (\epsilon - \epsilon_d)d_{\perp}}{ek_x + \epsilon_d k_t + ik^2_x (\epsilon - \epsilon_d)d_{\perp}},
\]

(2.7)
where for a \( p \)-polarized incidence beam,

\[
\Omega^* = \frac{k_x^2}{k_t} \left( \frac{1}{\epsilon} - \frac{1}{\epsilon_b(1 + \Delta\epsilon)} \right),
\]

(2.8)

\[
k_t^2 = \frac{1}{\beta^2} \left( \frac{\omega_p^2}{1 + \chi_b} - \omega^2 - i\gamma\omega \right) + k_x^2,
\]

(2.9)

\( \epsilon \) and \( \epsilon_b \) are the local dielectric constants of the metal and dielectric constant of the bound electrons in the metal, respectively, \( k_t = \sqrt{k_x^2 - \epsilon k_0^2} \), \( k_x \) is the \( x \)-component of the wave vector \( \mathbf{k} \), and \( k_0 = \omega/c \). Note that for optical frequencies close to the plasmon frequency, \( k_x \) is much smaller than \( k_t \), thus to a good approximation \( k_t = k_t(k_x = 0) \).

Now by equating the reflection coefficients obtained via the surface HDM and \( d \)-parameter approach, i.e., \( r^{\text{SHDM}} = r^d \), we obtain the induced discontinuity of the dielectric function:

\[
\Delta\epsilon = \frac{1}{\epsilon_b} \frac{\epsilon}{\left[ 1 + \left( \frac{\epsilon - \epsilon_d}{\epsilon_d} \right) d \cdot k_t \right]} - 1
\]

(2.10)

This equation can then be used, in the case of a medium having spatial dispersion, to calculate its behavior in the presence of plasmonic waves.

### 2.5 Application to Planar Interfaces

As indicated above, our optical structures in this paper focus on those with basic planar interfaces. Such structures have been not only used recently for guided plasmon-
polariton devices, but, in addition, have been the subject of several basic studies, in which the nonlocal polarization has been shown to be significant in a plain interface structure [74, 117]. Our first test structure uses a basic planar interface of metal on a planar dielectric crystal. In particular, our structure employed a Ag/GaN interface due to its relevance for certain plasmonics problems and as well as its established optical properties. In our calculations, the relevant parameters for Ag follow those used by the Khurgin [81], such as the susceptibility of the bound electron $\chi_b = 3.1$, the damping constant $\gamma = 3.2 \times 10^{13} \text{s}^{-1}$, the plasma frequency $\omega_p = 1.14129 \times 10^{14} \text{Hz}$, and the Fermi velocity $v_F = 1.4 \times 10^{8} \text{cm s}^{-1}$. Note also that the relative permittivity of GaN is $\epsilon_d = 5.29$ [74].

To implement the surface-hydrodynamic model, we used COMSOL to solve the nonlocal wave equation using a finite-element-method calculation. The geometry of a single interface of bulk materials was chosen to be a 0.2$\mu$m thickness of bulk silver.
on GaN, with a 1\(\mu\)m overall propagation length. These dimensions are thus suitable for testing the behavior of the SPP modal wave. In particular, the path length is long enough to allow propagation of a stable mode and yet damp out any other transient and radiating modes. The width of the structure is also sufficiently long to prevent coupling between the SPP propagating along the left vertical interface and that along the right vertical interface (not shown). In addition, the length of the structure is sufficiently long to prevent interference from the SPP back-reflections emanating from the top interface. A continuous-wave optical beam was then chosen to irradiate this interface structure, i.e., end-fire coupling [118], so as to excite a single SPP mode, propagating along the interface. This configuration was also found to efficiently excite the interfacial SPP mode, as is shown in our simulation results presented in 2.3(b). Note that the simulated structure is chosen so as to allow symmetry considerations to reduce the simulation time. Specifically, the dashed line running along the right edge of 2.3(b) serves as a mirror symmetry.

With the geometry fixed, the desired modified nonlocal equation, i.e., Maxwell’s equation with a nonlocal polarization term Eq. (2.3), was implemented. Additionally a second coupled equation, Eq. (2.2), was used in order to get the nonlocal dependence of the polarization \(\mathbf{P}\) on the electric field \(\mathbf{E}\). These equations are then solved through an iterative solver of the field. Note also that the wave equation was modified through use of the \(d\)-parameter to take in account the surface modification of the dielectric constant, \(\Delta\epsilon\); this modification is implemented as a surface current that flows along the boundary surface. Symmetry was also used to simplify our calculations as mentioned earlier in the paper. Finally it is important to note that our
approach was validated by a comparison of our numerical calculations of the HDM model and those obtained with the same model using the analytical approach of the Mortensen Group [73].

Using the interfacial structure and procedure described above, the behavior of the SPP mode was investigated for each of three common polarization models, i.e., two cases having significant nonlocal contributions (the hydrodynamic and the surface-hydrodynamic models) and one case having only a local contribution (the Drude model), were investigated; in some cases a comparison with the recent analytical approach by Khurgin was also made. The wave vector of the SPP mode was readily obtained through the fitting of the field at the interface using a nonlinear regression method with MATLAB. Note that our simulation uses a 2D fullwave calculation [119] in the frequency domain. Also in our calculation, convergence was obtained by reducing the grid size until the field distribution no longer changed significantly between successive “runs”. However in our calculations for the results presented here, a nonuniform mesh strategy was also used. The grid size needed to achieve convergence was found to be $< 1$ nm. However, since nonlocal polarization was found to always dominate near the domain boundary, that is, below the skin depth, a grid size $< 0.1$ nm was used near the boundary. This nonuniform method enabled a larger grid size to be used inside most of the domain, thus increasing the computational efficiency of the calculation. The simulation results based on this mesh strategy were sufficient to give convergent results. With these computational tools in place, our calculation proceeded by determining the dispersion relation and fields for the local, the hydrodynamic, and the surface-hydrodynamic models.
Our calculation of $\Delta \epsilon$, based on the TDLDA results of $d$-parameter, shows that both the magnitude of the real part and the imaginary part of $\Delta \epsilon$ decreases when the wavelength increases, and converges to 0 at the long wavelength limit. In particular, for a wavelength range from 400 nm to 600 nm, $\Delta \epsilon$ varies from -0.3 to -0.05, as shown in 2.2(b).

An examination of 2.3(a) shows the salient details of the three models. First, all modes exhibit the same trend in their variation with the SPP wave vector $k'$. In addition, each model yields nearly the same long wavelength behavior. However, near the plasmon resonance frequency, each of the models has important differences from each other. We see from 2.3(a) that for the case of $\omega < \omega_{\text{SPP}}$ near the surface plasmon resonance frequency, the relations $\lambda_{\text{loc}} < \lambda_{\text{SHD}} < \lambda_{\text{Kh}} < \lambda_{\text{HD}}$, where $\lambda_{\text{loc}}$, $\lambda_{\text{SHD}}$, $\lambda_{\text{Kh}}$, and $\lambda_{\text{HD}}$ compare the SPP wavelengths derived using local, the SHDM, the Khurgin analytical, and the HD models, respectively. For the case of the classical local and the Khurgin models the width of the SPP energy resonance as well as the maximum SPP wave vector depend on the degree of damping.

Additional observations can be extracted from the dispersion curves in 2.3(a). For example, it is seen that the hydrodynamic model has consistently a larger SPP wavelength, and this effect becomes significantly larger closer to the plasmon frequency. In addition, the dispersion behavior of the SPP continues to persist, i.e., increasing SPP wave vector $k'$ with increasing SPP energy, even when the frequency goes above the plasmon frequency, a behavior different from the local case, in which the dispersion becomes “anomalous”. This behavior agrees with other observations of the nonlocal problem [73], and is treated as a competition between the nonlocal
Figure 2.3: (a) Comparison of the theoretical dispersion relations of SPPs at an Ag/GaN interface as predicted by a local dielectric model (gray line), an analytical result from [74] (green line), the hydrodynamic dielectric model (blue triangles), and the surface-hydrodynamic model (present work - red circle). For a SPP energy in the vicinity of the surface plasmon resonance frequency, e.g., 3.02eV, the surface-hydrodynamic model predicts that the effective wavelength of the SPP mode lies between that obtained with the hydrodynamic model and that from the standard local dielectric model. (b) Plot of the simulated electric-field distribution along the interface in the presence of nonlocal polarization terms at a wavelength of 440nm using the surface-HDM model for a single Ag/GaN interface.
polarization associated with the movement of free-electrons and the loss that impairs the movement of free electrons, i.e., high collision frequencies.

The surface-hydrodynamic model also predicts a larger effective wavelength than does the classical method near the plasmon resonance frequency; note that this difference is not significant at long wavelengths. In addition, the dispersion predicted by the surface-hydrodynamic model lies between that predicted from the hydrodynamic model and that of the local dielectric model. It is also of interest to compare our simulation results with the analytical results of Khurgin [74], as shown in 2.3(a). The Khurgin model also gives a larger SPP effective wavelength when the optical frequency is near the surface plasmon resonance frequency. Note that the results of the surface-hydrodynamic model agree well with the analytical results, especially in the range from 2.9eV to 3eV, in which the standard HDM diverges from the analytical results.

2.6 Solution for a Thin Film Stack

Thus far we have demonstrated the application of our surface HDM, using fullwave computation, to an important, albeit relatively simple, nanostructure, i.e., a single interface. However, it is also important to demonstrate the application of our surface HDM to more complex and potentially more practical nanostructures. Our choice for one such structure is an insulator/metal/insulator (IMI) stack. Such an IMI stack is used frequently in plasmonics because of its strong field confinement, which is of interest for applications in waveguiding [120–122].
Figure 2.4: Comparison of the relation of SPPs at a Ag/GaN thin film stack with thickness (a) $w = 50 \text{ nm}$ and (b) $w = 20 \text{ nm}$ using the local Maxwell-equation formalism (gray), HDM (blue) and surface-HDM (red). The $x$-axis is the propagation constant $k'$ for the main figures and decay constant $k''$ for the inset figures. Note that for the bottom sketches the SPP propagation is from left to right.
To illustrate, we describe our method for the case of an IMI structure using the same materials chosen above, Ag and GaN. In particular, a silver thin film is sandwiched between two GaN thin films. This structure allows us to investigate confinement effects and the application of our model as a function of metal thickness, $w$. This approach to the calculation follows the method used above for the single interface problem with an excitation source to the left of the structures in 2.4 to excite mirror-symmetric (i.e., mirror symmetry along the length of the metal waveguide) SPP modes at the interface, and using a similar mesh structure for the planar interface case. The nonlocal response of this layered Ag waveguide was examined for varying widths ranging from 20 nm to 100 nm, and 1μm in length. We consider here two specific illustrative cases, i.e., those corresponding to the 20 nm and 50 nm waveguides widths. A typical ratio of the $d$-parameter to the thickness is $\sim 0.005$, thus the two interfaces are sufficiently apart and thus can be considered as isolated systems with respect to their electronic response, thereby enabling the applicability of our surface hydrodynamic model. Hence, it is important to keep in mind that the use of our modified approach necessitates that this ratio be $\ll 1$, a limitation that is typical in most electromagnetic calculations since the $d$-parameter encapsulates the interfacial electronic response. As can be seen in 2.4, the SPP mode propagating along a 50 nm width thin film, has an increasing propagation constant with increasing frequency. The dispersion relation for this 50 nm-thickness sample is similar to the dispersion relation for a single interface SPP, since the SPP coupling between the two interfaces is weak for a large-thickness metal film.

For a small-thickness (20 nm) IMI stack, the dispersion relations of both the
standard HDM and surface HDM have greater differences in the propagation constant compared with that of the local case, since nonlocality, whose effects becomes more prevalent with increasing confinement, becomes more significant. The standard HDM shows a blue shift with respect to the standard HDM. On the other hand, the decay constant obtained using standard HDM was found to be smaller than for the classical case because of nonlocality. In addition, the decay constant of SHDM was also shown to be significantly larger than that of the standard HDM model since the SHDM introduces surface damping through the imaginary part of the discontinuity $\Delta \epsilon$. As a result of tradeoff between nonlocality and surface correction, the propagation constant of the surface HDM is larger than the local case, and the decay constant of surface HDM is larger than the local case.

In addition, it was found that the difference between the standard HDM and the surface-HDM method decreases as the wavelength increases. This decrease occurs because the surface correction becomes weaker at longer wavelengths and, in fact, vanishes at the long wavelength limit as discussed above. Finally, both standard HDM and surface HDM converge to the classical limit at long wavelengths.

To summarize, for the case of nanolayered films such as discussed in this section, we observe (1) a redshift in the dispersion for the SHDM compared to standard HDM, (2) the effective wavelength of SHDM is smaller compared with that of standard HDM, and (3) the decay constant observed in the SHDM is larger in comparison with that of the standard HDM.
2.7 Conclusion

We have developed a new surface-hydrodynamic model to enable numerical simulation of plasmonic devices in the presence of significant nonlocal polarization in the material medium. Such numerical methods are of pressing importance for calculations of SPPs on media with complex geometries and materials. Our model is based on the standard hydrodynamic model and enables an accurate response near plasmon resonance frequency by taking into account the change in electron density at the interface, i.e., electron spill-out, through $\Delta \epsilon$, which is obtained from the accumulated $d$-parameter under consideration. In developing this model, we thus demonstrate the extent to which material nonlocality plays an important role in the SPP behavior near the surface plasmon resonance frequency. Our surface-hydrodynamic model agrees well with the analytical method, which is distinct from the standard HDM. Specifically the discontinuity in the electric displacement field increases from zero at the long-wavelength limit to a finite value when the wavelength decreases to its value in the vicinity of the surface plasmon resonance frequency. In short, this model allows accurate prediction of the behavior of SPPs, which also agrees with a previous analytical result. Our approach also provides a new method for accurate simulation of plasmonic devices, which may be too complex for direct use of an analytic model. In conclusion, we have shown that our numerical approach is useful for plasmon calculations at wavelengths near the surface plasmon resonance where nonlocal effects must be taken into consideration, and for the case of waveguided-nanostructure-supported SPPs, e.g., IMI or MIM, metal tips, and hybrid plasmonic structures.
2.8 APPENDIX A. Modified Reflection Coefficient for the Surface Hydrodynamic Model

In order to obtain the corrected Fresnel reflection coefficient for the surface HDM, $r_{SHDM}$, we follow the derivation in [67]. In a metal, Maxwell’s equations can be written

\begin{align*}
\nabla \times E &= i\omega \mu_0 H \quad (2.11) \\
\nabla \times H &= -\omega \epsilon_0 \epsilon (E - \alpha \nabla (\nabla \cdot E)) \quad (2.12)
\end{align*}

where $\epsilon$ is the local relative permittivity of the metal, and

\[
\alpha = \frac{\beta^2}{\frac{\omega_p^2}{1 + \chi_0} - \omega^2 - i \gamma \omega} \quad (2.13)
\]

Thus the solution in the metal can be divided into transverse and longitudinal modes.

Let us now consider a $p$-polarized plane wave incident from the dielectric medium ($z > 0$) with permittivity, $\epsilon_d$, and reflected by a metallic interface located at $z = 0$. 

104
The electric field $\mathbf{E}$ and magnetic field $\mathbf{H}$ in the dielectric region can be written as

$$H_y = (e^{-ik_z z} + re^{ik_z z})e^{ik_x x - i\omega t} \quad (2.14)$$

$$E_x = \frac{ik_z}{i \omega \epsilon_0 \epsilon_d} (re^{ik_z z} - e^{-ik_z z})e^{ik_x x - i\omega t} \quad (2.15)$$

$$E_z = -\frac{k_z}{\omega \epsilon_0 \epsilon_d} (1 + r) \quad (2.16)$$

where $k_z = \sqrt{\epsilon_d k_0^2 - k_x^2}$. The fields in metal can be written as

$$H_y = Ae^{k_l z}e^{ik_x x - i\omega t} \quad (2.17)$$

$$E_x = \left( \frac{k_l}{i \omega \epsilon_0 \epsilon} Ae^{k_l z} + Be^{k_l z} \right) e^{ik_x x - i\omega t} \quad (2.18)$$

$$E_z = \left( -\frac{ik_x}{i \omega \epsilon_0 \epsilon} Ae^{k_l z} + \frac{k_l}{ik_x} Be^{k_l z} \right) e^{ik_x x - i\omega t} \quad (2.19)$$

Thus the continuity conditions for $H_y$ and $E_x$ are given by

$$1 + r = A \quad (2.20)$$

$$(r - 1) \frac{ik_z}{\epsilon_d} = \frac{k_l}{\epsilon} A + i \omega \epsilon_0 B \quad (2.21)$$

Considering the discontinuity of the dielectric constant as $\Delta \epsilon$, we can write the normal component of the polarization as $P_{fz} = \epsilon_0 \Delta \epsilon E_{mz}$. Since $P_{fz} = -\frac{1}{i \omega} \partial_x H_y - \epsilon_0 (1 + \chi_b) E_z$

we obtain

$$A \left\{ k_x \left( -1 + \frac{1 + \chi_b}{\epsilon} (1 + \Delta \epsilon) \right) \right\} = B \epsilon_0 \omega \left\{ (1 + \chi_b) \frac{k_l}{ik_x} \right\} (1 + \Delta \epsilon) \quad (2.22)$$
Combining (2.16), (2.17), and (2.18), we obtain the Fresnel reflection coefficient for the surface HDM model as $r_{\text{SHDM}}$,

$$r_{\text{SHDM}} = \frac{ik_z}{\epsilon_d} + \frac{k_t}{\epsilon} - \Omega^* \frac{k_t}{\epsilon_d} - \frac{k_t}{\epsilon} + \Omega^*$$

(2.23)

where $\Omega^* = k_t^2 \left( \frac{1}{\epsilon} - \frac{1}{\epsilon_b(1+\Delta\epsilon)} \right)$, $k_t = \sqrt{k_x^2 - \epsilon k_0^2}$, and $k_t^2 = \frac{1}{\beta^2} \left( \frac{\omega^2}{1+\chi_b} - \omega^2 - i\gamma\omega \right) + k_x^2$.

Note that the reflection coefficient for a two-interface case, i.e., a dielectric-metal-dielectric slab system, can be derived in a straightforward manner although it results in a more complex form than that of a single-interface case. Since the skin depth is very small compared to the metal thickness considered in this chapter, the reflection coefficient for two-interface case will be reduced to that of a single-interface.

2.9 APPENDIX B. Numerical Implementation of the Surface Current

The surface current is defined via the boundary condition $\mathbf{n} \times (\mathbf{H}_2 - \mathbf{H}_1) = J_{s0}$, where 1 and 2 denote the metallic and dielectric regions. Based on the Maxwell equation relation $\nabla \times \mathbf{H} = \mathbf{J} + \partial \mathbf{D}/\partial t$, the discontinuity in the displacement, which may be expressed as $\mathbf{D}_2 = (1 + \frac{\Delta\epsilon}{\epsilon_b})\mathbf{D}_1$ can be equivalently written as $\nabla \times (\mathbf{H}_2 - \mathbf{H}_1) = \frac{\Delta\epsilon}{\epsilon_b} \nabla \times \mathbf{H}_1$. Thus, at the interface, we find that the surface current is given by $J_{s0} = \frac{\Delta\epsilon}{\epsilon_b} \nabla \times \mathbf{H}_1$. To implement this relation into COMSOL, the surface current modules $J_{s,x}$ and $J_{s,y}$ should be reset to the following relations

$$-\text{up}(\text{emw} \cdot \text{Hz}) * \text{epsilon}_\text{dif}/(\text{eps}_{-b}) * (-ny)$$
\[-up(emw.Hz) * epsilon\_dif/(eps\_b) * (nx);\]
Chapter 3

Layer Dependence of Third-Harmonic Generation in Thick Multilayer Graphene

We report layer-dependent third-harmonic (TH) generation in thick multilayer graphene (MLG) on quartz substrate, wherein the graphene layer number $N$ varies from $\sim 3$ to 50. Theoretical predictions of the layer-dependent TH signal are compared with experiment, including those measured using atomic-force and optical microscopy. We find the optimal TH signal for $N \approx 24$, in good agreement with two theoretical models: one in which the graphene layers are assumed isolated and one in which the multilayer graphene is assumed as one continuous thin-film material. In the presence of absorption, the coherence of the nonlinear optical sources arising from discrete graphene layers yields a distinctive layer-dependent TH signal that establishes the validity of the stacking property of MLG, i.e., $\sigma_{N,g} = N \sigma_0 f$, where $\sigma_0 = e^2/4\hbar$ and $f$ is a wavelength-dependent parameter. Our result will lead to the development of ultrathin composite graphene films having large third-order optical nonlinearities thereby opening a vast array of potential applications in optoelectronics and optical devices.
3.1 Introduction.

Single-layer graphene has become a subject of intense interest and study because of its remarkable 2D electronic, optical, mechanical, and thermal properties. Graphene offers a strong doping-dependent absorption edge and pronounced excitonic effects [47, 48]. In addition, measurements of optical carrier generation in graphene have also led to the observation of strong hot-electron photoluminescence [123, 124]. However, in addition to this interest in monolayer samples, multilayer graphene (MLG) also offers an array of properties that are of interest for optical physics and devices [125–128]. These properties include relatively strong optical absorption for visible and IR wavelengths, which underscores the importance of the linear optical properties in these thick samples.

Similarly, interest is developing in nonlinear optical effects in graphene of all degrees of thickness [65, 129], and in the optimization or enhancement of its nonlinear optical response. Thus as an example, recently theoretical investigation of the nonlinear optical effects arising from interband electronic transitions in monolayer samples have revealed that, despite its single-atomic-layer thickness, its nonlinear optical response is particularly strong [130]. Researchers have shown that its second-order-nonlinear optical response is intrinsically weak [131–133], because of the centrosymmetric nature of an ideal free-standing monolayer of graphene. On the other hand, third-order nonlinear optical effects in graphene are symmetry-allowed, thus leading to studies of several third-order processes including saturable absorption [134, 135], four-wave mixing [136–140], and TH generation in few-layers graphene.
electric field induced second-harmonic generation. Important developments in the quantum theory of third-order processes in graphene have recently been presented [146–148] as well as a quasiclassical non-perturbative treatment of the nonlinear response of graphene [149]. It was recently predicted that the third-order nonlinear response of MLG is strongly dependent on the layer number and that this effect arises from the interplay of coherence and loss [65, 129].

This layer dependence allows the possibility of obtaining the optimal thickness of MLG such that it would yield the highest overall third-order nonlinear susceptibility. In fact, Mikhailov has shown theoretically that the optimal number of layers for four-wave mixing is \( N \sim 29 \) layers [65]. In addition to layer dependence as a means of enhancing the nonlinearity, other techniques may be employed to optimize the nonlinear signal such as metasurfaces and cavity enhancement [150, 151]. It should be pointed out, however, that layer-dependent second-order nonlinear measurements have been carried out in noncentrosymmetric 2D layered crystals [152–154]. While the incoherent Raman scattering process has hinted at the existence of an optimal layer number for MLG [155, 156], no systematic layer-dependent measurements of coherent third-order processes have been carried out in thick MLG.

In this paper, we report on measurements of the layer-number-dependent TH generation in thick MLG, with layer number ranging from \( N \sim 3 \) to 50. The fundamental wavelength of our optical pump corresponds to a TH photon energy of \( \sim 4.7 \) eV that allows access to the M point and its vicinity for the case of single-layer graphene. In this region, the usual linear dispersion flattens, and thus the density of states is high because of the van Hove singularity at the saddle point, thereby enhancing the
Figure 3.1: (a) Linear transmission, reflection, and absorption coefficients for MLG between air and quartz substrate vs. layer number $N$ for the interface (blue curves) and slab (red curves) models. (Left inset) Interface model: MLG is treated as a single source of interface current via the surface conductivity $\sigma_{N,g}$. The TH reflected and transmitted waves, $E_{3\omega,R}$ and $E_{3\omega,T}$, are generated via the third-order conductivity $\sigma_{N,g}^{(3)}$. (Right inset) Slab model: the MLG is treated as a single slab with third-order nonlinearity $\chi^{(3)}$. Four TH waves are generated: $E_{3\omega,R}$, $E_{3\omega,T}$, $E_{3\omega,M}$ and $E_{3\omega,M'}$, which denote the reflected, transmitted, internal down- and up-going waves, respectively. (b) Reflected THG signal predicted by the interface and slab models for MLG. Inset: prediction by slab model for a large $N$. 
graphene nonlinear optical susceptibility via a three-photon resonance [129]. Our measurements show that the TH signal peaks at a MLG thickness of $N = 24$ due to the balance between optical absorption and number of emitting layers. We show excellent agreement between measurements and two seemingly divergent but equivalent theoretical approaches: one in which the graphene layers are assumed to be discrete and isolated coherent TH sources and one in which the MLG is treated as one continuous nonlinear thin film. Our measurements are carried out by comparing high-quality thickness-dependent TH microscopy images to atomic force and optical microscopy scans.

### 3.2 Theoretical Considerations.

To describe the nonlinear optical response of MLG, we consider two methods indicated in Fig.3.1. A schematic of the first model, developed by Mikhailov [65], which we denote as the interface model, is shown in Fig.3.1(a) left inset; this model approximates the MLG as an infinitely thin stack of isolated graphene layers between two dielectric media (e.g., air and substrate). At the interface of these two media, the incoming waves induce a surface current at the MLG stack through its optical conductivity. This approach treats the composite structure an effective 2D system because of its essentially pure in-plane electronic response. Albeit an approximation, the interface model offers an elegant approach that provides a remarkably useful prediction of the layer dependence of the nonlinear optical response of MLG. In the second approach (Fig.3.1(a) right inset), the nonlinear optical process occurs within
a slab, as previously considered in the classic paper by Bloembergen and Pershan [157]. In contrast to the discrete atomic layers employed in the interface model, the MLG is regarded simply as a continuous thin film with a thickness defined by the product $Nd$ where $d$ is thickness of monolayer graphene. This approach takes into account the individual phases of the pump and emitted waves within the film and provides a general solution for the third-order nonlinear response of the bulk material. These two approaches both allow us to treat the nonlinear response and to compare prediction with experiment so as to understand the 2D to 3D, i.e., ultrathin film to bulk, transition clearly.

The optical property of single-layer graphene is dictated primarily by its optical conductivity $\sigma_g$ that arises from intra-band and inter-band transitions [64]; it may be written as $\sigma_g = \sigma_0 f$, where $\sigma_0 = e^2/\hbar = c\alpha/4$ is the universal optical conductivity of graphene [63], $\alpha = e^2/\hbar c$ is the fine structure constant, and $f$ is a dimensionless frequency-dependent parameter. Over a wide photon energy range of 0.5 to 1.5 eV, $f \approx 1$ [64], whose value increases to $\sim 4$ near 4.62 eV [47]. Within the interface model [65], the MLG is treated as an atomically thin medium such that $Nd \ll \lambda$, where $\lambda$ is the pump wavelength so that all layers emit coherently. Thus, its optical conductivity can be simply described as $\sigma_{Ng} = N\sigma_g$ when $\hbar \omega > 1$eV [158]. Following Mikhailov’s method, we obtain the linear and nonlinear optical fields at the MLG and bounding media by solving two separate, i.e., linear and nonlinear, sets of Maxwell equations and the relevant boundary conditions. We assume that the bounding media of the MLG are air and quartz substrate having refractive index $n$. For a normally incident field $E_{\omega,0}$, the amplitude of the electric field acting on the multilayer graphene
sheet becomes $2E_{\omega,0}D_{\omega}^{-1}$, where $D_{\omega} = 1 + n_{\omega} + N\pi\alpha f_{\omega}$, and $\Omega = \{\omega, 3\omega\}$ denotes fundamental and TH waves. The maximum thickness is of our MLG is of the order of $\sim 10$ nm, which is much smaller than the wavelengths considered here. Consequently, each graphene layer experiences the same effective driving field, denoted by $E_{\omega,\text{eff}}$. The corresponding transmittance, reflectance, and absorbance factors are found to be $T = 4n_{\omega}D_{\omega}^{-2}$, $R = (n_{\omega} - 1 + N\pi\alpha f_{\omega})^2D_{\omega}^{-2}$, and $A = 2N\pi\alpha f_{\omega}D_{\omega}^{-2}$, respectively.

Fig. 3.1(a) illustrates all three quantities above as a function of the layer number $N$ under the assumption $f_{\omega} = 1$.

The third-order current induced in each layer is given by $j^{(3)} = \sigma^{(3)}\omega^{3}E_{\omega,\text{eff}}$. The induced oscillating currents from the $N$ layers add coherently and a secondary wave at frequency $3\omega$ is emitted, i.e., reflected or transmitted, having an amplitude $E_{3\omega} = -4N\pi j^{(3)}c^{-1}D_{3\omega}^{-1}$, which may be written as

$$E_{3\omega,(R,T)} = -32N\pi E_{\omega,0}^{3}c^{-1}D_{3\omega}^{-1}\omega^{-3}$$

Eq. (3.1) predicts that for small $N$, the reflected signal $I_{3\omega} \propto |E_{3\omega,R}|^2$ increases quadratically; as $N$ increases further, the signal increases to a maximum value then decreases monotonically.

In the slab model, we treat the MLG as a nonlinear slab media with thickness $Nd$, where $d = 0.335$ nm is the thickness of single-layer graphene [48, 159]. Normal incidence is assumed but explicit solutions for the case of non-normal incidence are shown in [129]. Briefly, the linear and harmonic optical fields given by $E_{\omega,i}$, and $E_{3\omega,i}$, respectively, are derived using Maxwell’s equations, with $i = R, M, M'$, and $T$ indicat-
ing the reflected, internal downward-going, internal upward-going, and transmitted fields, respectively. The nonlinear polarization is given by $P^{NL}(3\omega) = \chi(E_{\omega,M})^3$ where $\chi_1 \equiv \chi^{(3)}_{xxxx} = \chi^{(3)}_{yyyy}$ that describes the pure in-plane nonlinear response.

Solving the nonlinear Maxwell’s equations yields the nonlinear electric fields $E_{3\omega,i}$. The reflected TH wave has the complex amplitude

$$E_{3\omega,R} = \frac{4\pi P^{NL}}{n_M^2 - n_S^2} \times \frac{A(n_T - n_S)n_M - B(n_Sn_T - n_M^2)}{e^{2i\phi_M}(1 - n_M)(n_T - n_M) - (1 + n_M)(n_T + n_M)}$$

where $A = 1 + e^{2i\phi_M} - 2e^{i(\phi_M + \phi_S)}$, $B = 1 - e^{2i\phi_M}$; $\phi_S = 3n_S\omega d/c$ and $\phi_M = 3n_M\omega d/c$ are the phase shifts of the inhomogeneous and homogeneous waves, respectively; $n_S = n_{g,\omega}$ and $n_M = n_{g,3\omega}$ are the refractive indices of graphene at the fundamental and TH frequencies, respectively; and $n_T = n_{3\omega}$ is the refractive index of the substrate at the TH frequency.

Prior to examining the nonlinear properties of MLG, we calculated the predicted reflection, transmission, and absorption coefficients of fundamental light for the slab model of MLG on z-cut quartz substrate, i.e., the quantities $R = |E_{\omega,R}/E_{\omega,0}|^2$, $T = n_\omega|E_{\omega,T}/E_{\omega,0}|^2$, and $A = 1 - T - R$, respectively. Fig. 3.1(a) shows a comparison of the interface and slab models, which reveals that the two models are fairly consistent even up to $N = 200$ although small differences in the optical parameters become apparent starting near $N = 20$. The interface model makes use of the stacking property of the universal conductivity of graphene, i.e., $\sigma_{N,g} = N\sigma_g$ [158], as its basic optical parameter with $f_\omega = 1$, whereas the slab model makes use of the refractive index of graphene, $n_{g,\omega} \approx 3 + 1.5i$ [160]. Furthermore, as is the case for the interface model,
Figure 3.2: Experimental setup. The fundamental beam from a Ti:sapphire laser (TS) passes through dichroic mirror (DC) and is focused on the MLG (G) sample via a microscope objective (MO), which generates a TH signal that is reflected by the DC into a Pellin-Broca prism (BP), reflected by mirror (M), and focused by a lens (L) into the monochromator (MC). The detection system consists of a photomultiplier tube (PMT) and a photon counter (PC). Left inset: cubic-power dependence of TH signal; right inset: TH spectrum.

the slab model results agree well with measured layer-dependent optical properties of free-standing few-layer graphene [66]. Turning now to the nonlinear properties, the THG signal predicted by the slab model for MLG with N ranging from 0 to 200 is displayed in Fig.3.1(b). For the TH harmonic calculations we used, in addition, \( f_{3\omega} = 4 \) and \( n_{g,\omega} \approx 2 + 3i \) [160]. The layer dependence of the TH signal exhibits the general behavior as discussed earlier for the interface model, namely the rise with layer number until absorption at both fundamental and TH frequencies overwhelms this increase and yields a peak in the curve. Both models predict a peak near \( N \approx 24 \) although differences between the two models begin to emerge near \( N \approx 10 \) because at this point the overall phases accumulated by both fundamental and TH waves become significant. At higher layer numbers shown in Fig.3.1(b) inset, the TH
signal predicted by the slab model shows a secondary maxima near \( N \approx 300 \), which is simply an interference effect. At even higher layer numbers the TH signal becomes constant as expected because of the finite escape depth of the TH wave, which is of the order \( \sim 10 \) nm.

### 3.3 Experiment.

The MLG flakes are produced by mechanical exfoliation, which are then transferred to 1-mm-thick quartz substrates. Flake thicknesses are determined by their optical contrasts and atomic force microscopy. Each sample is mounted on a mated orthogonal-translation stage, which permits 2D scanning of the sample. Fig.3.2 illustrates our experimental setup: a Ti:sapphire laser produces 50-fs, 792-nm pulses at 76MHz repetition rate, which are transmitted through a dichroic mirror and focused on the sample with a typical average power of 100 mW at normal incidence angle. The TH signals are spectrally filtered via a Pellin-Broca prism and a monochromator, whose output is directed to a photon counting system. The measured spot size of the TH signal is \( \approx 1.8 \mu\text{m} \) thereby establishing the spatial resolution of our measurements. No optical or thermal-induced degradation of graphene during irradiation within the measurement period is observed. The TH signal exhibits a cubic power dependence on the incident power thus fully confirming the nature of the nonlinearity of the TH signal (inset), and is spectrally centered at 264 nm (sub-inset).
3.4 Results and Discussion.

In order to obtain maximum information on the MLG sample, only selected regions containing a large range of layer thicknesses are investigated. The results for a region of dimensions of $145 \times 95\mu m$ on a typical MLG sample are shown in Fig.3.3(a)-(b) using optical and TH microscopy. The results show clearly how changes in layer numbers in the sample yield drastic changes in the TH signal. We first qualitatively discuss key observations. The darkest region (region A) in the optical image coincides with the bare quartz region, which shows very small TH signals. The brightest region (region B) in the optical image correspond to a very large layer graphene region consisting of several 10s of layers, which yields a $10 \times$ improvement of the TH signal.
compared to that of the pure substrate. A region of intermediate brightness (region C) in the center of the optical image, which has a thickness of ~20 monolayers, yields a significant signal improvement, which is about $10^4 \times$ stronger than that of the bare substrate. These noteworthy results indicate a complex signal dependence on graphene layer number: both bare and high-layer areas yield small TH signals, while the intermediate layer number regions exhibit significant TH signals.

![Normalized layer-dependent TH generation in MLG](image)

**Figure 3.4:** Normalized layer-dependent TH generation in MLG. Circles: Experimental data. Blue and red curves: theoretical predictions using the interface and slab models, respectively.

Finally, we carried out a systematic study of the TH intensity *versus* local sample thickness, and cross-calibrated our results with atomic force microscopy (AFM). 

**Fig.3.3(c)-(d)** shows a direct comparison between THG, AFM, and optical mi-
croscopy, including line scans in Fig.3.3(d). Before summarizing our results, we calibrated the third-order susceptibility of graphene by comparing the THG signal from a 5-layer graphene sample to that from an silicon sample having a 285-nm-thick oxide layer [129, 136]. Using the slab model described above and a separate similar treatment for the SiO$_2$/Si surface, we obtain $|\chi^{(3)}_{5,9}/\chi^{(3)}_{Si}| = 0.12$. The layer-dependent TH measurements is presented in Fig.3.4, including theoretical results for comparison. We see that the TH signal exhibits a peak at $N \approx 24$. Note that the experimental results not only agree with theory with respect to the peak location but they also agree with respect to the relative intensities for different layer numbers.

![Figure 3.5: (a) MLG sample on a quartz substrate (b) Silicon/SiO$_2$ sample used for calibration.](image)

Before concluding, we offer a few remarks regarding the nature of the nonlinear process in MLG. Although the interface model neglects the optical phases at the fundamental and third-harmonic waves, it provides a more direct assessment on the validity of the stacking property of conductivity of MLG. Such assessment is not possible via the slab model although it can more accurately account for the optical phases of the two waves and, ultimately, the layer dependence of the TH signal. For optical frequencies used in our experiment, our TH signal at 4.7 eV, corresponds to the M point in the graphene band structure. Thus, only an interband transition plays an important role since this transition is the only resonant spectral feature near
the M point, where a high density of states is present. The photon energies of both fundamental and TH waves can access states that are much higher in energy than those near the Dirac point in graphene. Thus it can be expected that the interaction between layers should be weak, a conclusion reached from other works [161, 162] and consequently, the treatment of optical conductivity for MLG as a product of layer number and the optical conductivity of monolayer graphene, a core assumption in the Mikhailov model, i.e., weak interlayer interaction. The result is surprisingly valid even up to large number of layers considered in our work. Finally, we note that the so-called “optimal” layer number that yields the most significant nonlinear optical response depends on several factors, including the geometry (e.g., incidence angle), the bounding media, and, importantly, the type of the nonlinear optical process. The existence of this optimal graphene nonlinearity opens an avenue for engineering composite, ultrathin MLG films for nonlinear optical and optoelectronic devices.

In conclusion, we have demonstrated optical third-harmonic generation to be a deterministic probe of the nonlinear optical property of multilayer graphene. The TH signal of MLG exhibits a complex dependence on its layer number. We found that the optimal TH signal occurs for \( N \approx 24 \), in good agreement with both interface and slab models. Based on a direct comparison with the interface model, the coherence of the TH signal arising from the different graphene layers is consistent with the stacking property of the conductivity of MLG, in which \( \sigma_{N,g} = N\sigma_g \) even for the case of large layer numbers.
Chapter 4

Differential Phase-Shift Keying Demodulation by Coherent Perfect Absorption in Silicon Photonics

We demonstrate a novel differential phase-shift keying (DPSK) demodulator based on coherent perfect absorption (CPA). Our DPSK demodulator chip device, which incorporates a silicon ring resonator, two bus waveguide inputs, and monolithically integrated detectors, operates passively at bit rate of 10 Gbps at telecommunication wavelengths, and fits within a mm-scale footprint. Critical coupling is used to achieve efficient CPA by tuning the gap between the ring and bus waveguides. The device has a vertical eye opening of 12.47 mV and a quality factor exceeding $3 \times 10^4$. The fundamental principle behind this photonic circuit can be extended to other formats of integrated demodulators.

4.1 Introduction

Differential phase shift keying (DPSK) has become a promising modulation format for on-chip optical communications. It offers several key advantages for data communication systems including improved resistance to nonlinear effects and a $\sim 3$-dB enhanced sensitivity over that of intensity modulated direct detection (IMDD) sys-
tems [163]. It also offers a relatively high extinction ratio and dispersion tolerance [164]. These advantages have led to DPSK-based transceivers being actively considered for short-haul optical communication systems. Many approaches to DPSK demodulation have been implemented, which include controlling the birefringence of a fiber loop mirror filter [165], employing delayed path-length based Mach-Zehnder interferometer [166], and using a compact preamplified receiver based on Gaussian filtering [167].

Because of the aforementioned advantages, there has been a growing interest in moving toward a fully integrated chip-based systems for DPSK modulation and demodulation. Of the possible materials platforms for such integration, silicon photonics is particularly attractive due to its great potential for high quality large scale systems [168–171]. In fact, a Si photonic DPSK demodulation scheme based on ring resonators has been proposed recently. This scheme uses coherent perfect absorption (CPA) [172], where absorption in such a system can be coherently controlled by manipulating internal field interference by tuning the absorption and scattering loss of the ring resonator, which in turn modifies the critical coupling of the ring with bus and receiver waveguide [72, 173]. Such CPA-based devices have the advantages of independent switching of adjacent resonances, direct heterodyne detection and smaller footprints of the fabricated devices [72].

In this chapter, we report the fabrication and testing of a fully integrated demodulator optical circuit. The demodulator builds on a new approach for a CPA-based integrated optical modulator using a ring resonator [72] and experimentally demonstrated, similarly shaped racetrack resonator in silicon-on-insulator (SOI) platform at
telecom wavelengths [173]. The chip is designed for a bit rate of 10 Gbps at C-band communication wavelengths using a 1-bit integrated delay line between the two bus waveguide inputs of the ring resonator. The system output is measured at 10 Gbps, including its performance due to variation of bit rate and duty ratio (DR).

4.2 Device Design

A schematic layout of the experimental setup is shown in Fig.4.1(a). Here the fundamental operating parameters are briefly laid out; a more extensive discussion of the design of this DPSK demodulator has been presented previously [72, 173]. In particular, the input to the ring resonator is evenly split into two signals, with a phase shift ($\Delta \phi$) introduced into one of them. The two signals are then coupled to the ring through the two bus waveguides. The mode profile in each waveguide is shown in Fig.4.1(b). To achieve efficient power transfer between the ring and the waveguide, the device must operate such that the steady-state power to the resonator coupling from the two bus waveguides ($2\gamma_c$) equals the steady-state power loss rate inside the ring due to material absorption and scattering loss ($\gamma_i$), which is a condition known as critical coupling [174, 175]. Under this condition, the transmitted signal ideally drops to zero. In our device, we achieve the critical coupling condition by changing the gap between the waveguide and the ring, which essentially controls the coupling parameter ($\gamma_c$). Using temporal coupled mode theory [174], we find that the steady
Figure 4.1: (a) Experimental setup used to characterize the demodulator circuit. The input light from a tunable laser source (TLS) is edge coupled to the chip using a 2.5 μm spot lens-tapered fiber (LTF). A polarization rotator (PR) is used to maximize the current output. An inverted taper of 180 nm width is used to couple the light from the fiber to the SOI chip. The schematic of the photonic circuit shows the Y-branch where one signal branch is routed directly to the ring, while the other signal goes through a bit-rate-dependent delay line (DL). The two signals are coupled into the ring resonator. Four photodiodes (PDs) collect the photocurrent from the two signal buses and the two ring resonator ports. The photocurrents are measured with the help of a digital multimeter (DM) through a 150 μm pitch electrical probe. (b) Mode profile of the fundamental TE mode of the 500 nm x 220 nm Si waveguide at 1550 nm. (c) Optical microscope image of one 10 Gbps device. (d) Power transmission through the signal bus and the ring resonator as a function of phase shift ($\Delta \phi$). (e) Theoretical DPSK bit sequence and simulated resultant power absorption. The blue curve represents power through the signal bus while the red curve represents power inside the ring. When $\Delta \phi = 0$, the two inputs are identical and the ring is in resonance resulting in full power loading of the ring and extinguishing power in the signal bus. This situation is reversed for $\Delta \phi = \pi$. 
state power transmission through any of the signal waveguides varies as [72]

\[ P_{\text{signal}} = \frac{P_0}{2} \left( \frac{\gamma_i^2 + \frac{\Delta^2}{4} - \gamma_c \gamma_i \cos(\Delta \phi) + \Delta \omega^2 \pm 2 \gamma_c \Delta \omega \sin(\Delta \phi)}{\Delta \omega^2 + \frac{1}{4} \gamma_{\text{tot}}^2} \right) \]  

(4.1)

where \( \Delta \omega = \omega - \omega_0 \), \( \omega \) is the angular frequency, \( \omega_0 \) is the resonant angular frequency, and \( \gamma_{\text{tot}} = 2 \gamma_c + \gamma_i \) is the total loss rate of the cavity. Under the critical coupling condition (\( \gamma_i = 2 \gamma_c \)) and on-resonance (\( \Delta \omega = 0 \)), the above equation simplifies to [72, 172, 176]

\[ P_{\text{signal}} = \left( \frac{P_0}{2} \right) \sin^2(\Delta \phi/2), \]  

while the power inside the ring varies as

\[ P_{\text{RR1}} = P_{\text{RR2}} = P_0 \cos^2(\Delta \phi/2) \]  

as shown in Fig.4.1(d). Therefore, we made use of differential detection by measuring the relative phase change (\( \Delta \phi \)) between successive modulated phases of the received signal. A delay line is integrated into the chip that corresponds to one bit interval. The bit rate (\( 1/\Delta T \)) is related to the phase difference through the length of the delay line (\( \Delta L \)) and can be expressed as

\[ \Delta \phi = \frac{2 \pi c \Delta T}{\lambda n_g}, \]

where \( c \) is the speed of light, \( \lambda \) is the resonant wavelength of the ring and \( n_g \) is the waveguide group index. The two signals are then coupled inside the ring resonator through CPA effect such that based on the value of the phase difference (\( \Delta \phi \)) of two successive bits, the demodulator will have power at either the bus waveguide or the ring itself (no power through the bus waveguide). The patterned circuit used here was fabricated in a commercial Si photonics foundry at the Singapore Institute of Microelectronics (IME). The structure is comprised of a 220 nm silicon device layer with a 2 \( \mu m \) buried SiO\(_2\) layer based on standard SOI materials. The entire chip is clad with SiO\(_2\). Fig.4.1(c) shows an optical microscope image of a representative example of a 10 Gbps device along with a schematic of the layout of the device. The
loss coefficient $\alpha$ of the fundamental waveguide TE mode due to material absorption as calculated using RSoft FemSIM is 0.049 cm$^{-1}$. To first-order approximation, the loss rate is related to $\alpha$ according to $\gamma_i = \frac{2\pi R\alpha v_g}{L_{tot}}$ and the coupling rate to the length of the coupling region ($L_c$ according to $\gamma_c = \sin^2(\kappa L_c) v_g / L_{tot}$, where $R$ is the radius of the ring, $v_g$ is the group velocity, $L_{tot}$ is the total linear length of the waveguide, and $\kappa$ is the bus-to-resonator coupling coefficient [177]. Using a coupling gap of 434 nm, we find that the device is nearly critically coupled with $\gamma_i \approx 2.3\gamma_c$. The calculated delay based on the waveguide geometry and material and waveguide dispersion was $\sim 0.1$ ns, corresponding to a bit rate of 10 Gbps.

As shown in Fig.4.1(a) a Y-branch and delay line (DL) are used to split the input signal into the two paths, such that a relative delay of one bit slot is imparted. The waveguide and the ring are comprised of oxide-clad Si channel waveguides with cross-sectional dimensions of 500 nm $\times$ 220 nm. In Fig.4.1(b), we plot the mode profile of the fundamental TE mode of the Si waveguide, calculated using RSoft FemSIM, and is used to compute the waveguide group index ($n_g = 4.14$ at $\lambda = 1.55$ $\mu$m). Thus, the corresponding delay length ($\Delta L = c\Delta T / n_g$) for $1 / \Delta T = 10$ Gbps is $\Delta T \approx 7.2$ mm. For detection of the DPSK signals, four identical Ge p-i-n photodiodes are incorporated onto the chip because of their high speed, broad detection spectrum and the compatibility of their processing (primarily thermal management in growth) compatibility with Si CMOS technology. The Ge-photodiode length has been reduced to increase the bandwidth of these photodetectors. Eight Al contacts were fabricated to measure the electrical photo current out of the chip. There are four 10 Gbps devices with varying gaps from 404-494 nm between the bus waveguides and the ring. This
results into varying coupling coefficient \((\gamma_c)\) values ranging from 2 cm\(^{-1}\) to 6 cm\(^{-1}\).

4.3 Device Performance

To investigate a priori the DPSK demodulation scheme response of our proposed structure 2D-FDTD simulations (RSoft FullWave) are performed using the exact on-chip dimensions of the ring. In this simulation, an arbitrary eight-bit sequence was launched into the structure, resulting in the powers shown (Fig.4.1(e)) for the photodiodes labeled signal and RR1 in Fig.4.1(a).

When \(\Delta\phi = 0\), the two inputs are identical and the ring is in CPA resonance condition, resulting in power getting absorbed inside the ring, and nothing is transmitted out of the signal bus. When \(\Delta\phi = \pi\), the ring is not in CPA resonance condition and thus the power of the ring goes down and the transmitted power in the signal bus goes up. Thus, for DPSK signals, where the bit information is stored inside the relative phase of the successive bits, this device can extract that information using the delay line and its RR-CPA response. The step response in the power signal corresponds to the stabilization of the round-trip transit time of the electric field inside the ring resonator.

Prior to high-frequency testing of DPSK demodulation, the transmission spectrum of the device was measured via the experimental setup outlined in Fig.4.1(a). The measurement was performed with a continuous wave tunable laser source (TLS) (APEX AP3350A) set at 0 dBm and the polarization rotator was set to ensure maximum signal strength at the photodiodes. The input light from the TLS was edge
coupled onto the chip using lensed-tapered fiber (LTF) that yields a 2.5 μm spot size. An inverted taper of 180 nm width was used to couple the light from the fiber to the 500 nm wide waveguide. The generated photocurrent was measured using a MATLAB program to control an APEX AP1000-8 Mainframe controller with an Agilent 34410A multimeter.

The resultant photocurrent transmission spectra are shown in Fig.4.2 for the delayed signal and the RR2 ports with 0.02 nm resolution. With a group index of 4.14 and a circumference of 214.25 μm, the free spectral range (FSR) of the ring resonator is calculated as 2.7 nm at 1.55 μm. The calculated Q factor \( \left( \frac{\tau \omega_0}{2} \right) \) for the ring resonator is obtained from a FDTD simulation to be \( 2.8 \times 10^4 \). From Fig.4.2, the measured FSR is found to be 2.68 nm for a device with a 434 nm gap between bus and ring. The top-right inset of Fig.4.2 shows a dip in the delay signal in the green shaded region of the figure with a 0.001 nm resolution. The measured quality \( (Q) \) factor of the ring \( \left( \frac{\lambda_{res}}{\Delta \lambda} \right) \) as seen in the upper right inset of Fig.4.2 is \( 3.2 \times 10^4 \), in good agreement with the calculated value.

In addition, the wavelength dependence of the phase difference between two arms is calculated to be \(-24.6\pi\) rad/nm based upon the usual expression

\[
\frac{\partial}{\partial \lambda} \Delta \phi = \frac{-2\pi \Delta d}{\lambda^2} n_g.
\]

Thus, the phase difference changes by \(2\pi\) every 0.08 nm of wavelength, which agrees closely with the measured phase change of 0.07 nm from the photocurrent values as shown in the bottom-right inset of Fig. 4.2.
Figure 4.2: Measured photocurrent vs. wavelength for the signal and the RR1 port for a 10 Gbps device with a 434 nm bus and ring gap. The top right inset shows one resonant dip of the delayed signal port used to calculate the $Q$-factor of the device. The bottom right inset shows the photocurrent inside the ring used to calculate the wavelength change for a $2\pi$ phase shift.

Figure 4.3: Experimental setup for eye diagram. PPG: Pulse Pattern Generator, RFA: RF Amplifier, TLS: Tunable Laser Source, PR: Polarization Rotator, PM: Phase Modulator, EDFA: Er-Doped Fiber Amplifier, OGF: Optical Grating Filter, CSG: Clock Signal Generator, DSA: Digital Serial Analyzer.
4.4 DPSK Demodulation

In order to assess the quality of our demodulator circuit, the eye diagrams of the received signal were measured for different wavelengths, data rates and duty ratios. The experimental setup for eye extraction is shown in Fig.4.3. In this set up, a pulse pattern generator is used to generate a random bit sequence with a bit pattern length of $2^{15} - 1$ and a peak-to-peak voltage of 0.95 V at a particular clock frequency set by the clock signal generator. The bit sequence is fed to the phase modulator through a RF amplifier. RF amplification is required since the voltage required for the phase modulator for a $\pi$ phase shift exceeds that available from the PPG; i.e., 7 V for the former vs 2 V for the latter. The phase modulator modulates the input optical phase based on the bit pattern. The laser light is fixed at a resonant wavelength of 1551.88 nm. For a ring of a given radius, the resonant wavelength leads to the best performance for demodulation. The light then passes through the chip so as to demodulate the encoded signal. The photodiodes are reverse biased at 1.34 V through a bias-tee for high speed operation. Also, the laser light is amplified using an EDFA and optical grating filter before being its input to the chip.

A typical eye diagram after demodulation is shown in Fig.4.4(a). Open eyes were obtained at a data rate of 10 Gbps. The eye amplitude is measured to be 12.47 mV. Note the rise and the fall times of this eye are $\sim$0.09ns. These settings thus limit the device data rate to $\sim$10 Gbps. This speed limitation comes from stabilization time of the ring and the LC or RC constant of the external electrical measurement system. Fig.4.4(a) shows eye diagrams with different data rates of 9, 9.5, 10, 10.5,
Figure 4.4: (a) Eye diagrams for data rates of 9Gbps, 9.5Gbps, 10 Gbps, 10.5 Gbps, and 11 Gbps, which result in a measured eye amplitude of 9.9, 9.915, 12.47, 9.915, and 9.915 mV, respectively. As the data rate deviates from the optimal rate, the eyes degrade and appear more closed and distorted. The minimum BER and maximum extinction ratio occur at the designed data rates, i.e., 10 Gbps, which demonstrates the applicability of our device for high speed data rates. Note that shorter delay lines can be expected to yield data rates of 20 Gbps.

In addition to frequency response of the eye diagrams, we also investigated their response to the relative abundance of the symbols. The top eye diagram in Fig.4.4(b)
is obtained with a duty ratio, i.e., ratio of ‘0’ bit to ‘1’ bit in the randomly generated bit pattern of 50%. Operation under different duty ratios was then examined, namely, 75% and 87.5% by increasing the number of ‘1’ bits in the input bit stream. Since DPSK demodulation operates by decoding the message from phase difference of successive bits, it results in a higher number of ‘1’ bits than that of ‘0’ bits as shown in the bottom two eye diagrams of Fig.4.4 for duty ratios of 75% and 87.5%, respectively. The resulting dependence of measured extinction ratio on the occupation ratio of bit 1, as shown in Fig.4.4, is weak. Finally, we attribute the rippled top (bit 1) and button (bit 0) of the eye diagrams from the electrical stabilization time of the ring. Note that this effect has also been described by Dong et al. [178]. This effect can be stabilized with an electrical filter.

We have also characterized the eye diagram quality in the absence of the EDFA, which means that a decreasing in power input from 12.5 dBm to 9 dBm. The eye diagram gets significant improved with the help of EDFA, which is also an estimation of better BER. However, the improvement from using an EDFA is limited because of local chip heating.

The above results demonstrate the robustness of our DPSK demodulator at different input levels. Note that this characterization of this monolithic integration of the PD with passive a CPA DPSK demodulator is done without the use of post chip electrical amplification. In addition, it is anticipated that the performance of the DPSK device can be further improved with the help of post processing. More generally, our results show clearly that coherent absorption can be used to realize a functional and robust demodulator method for high-speed data transport. Further our results show
that this device is readily fabricated in a commercial foundry.

4.5 CPA-DPSK with Grating Coupler

Besides, we also designed and fabricated a CPA based DPSK demodulator with grating coupler Fig.4.5. This device is designed to have the DPSK demodulation function, which is based on the coherent perfect absorption. This chip is designed within a 605 × 410 μm area, and use the same structure of the ring resonator as described previously in this chapter, with a circumference of 214.25μm. The delay line is also 7.2 mm in length for operating at 10 Gbps. The gap distance between the ring resonator and the bus waveguide is 434 nm, which is chosen based on the previous experiment results. The designed structure is shown in Fig.4.5(a).

The photonic devices were fabricated using the NanoSOI MPW fabrication process by Applied Nanotools Inc. (Edmonton, Canada) which is based on direct-write 100 keV electron beam lithography technology. Silicon-on-insulator wafers of 200 mm diameter, 220 nm device thickness and 2 μm buffer oxide thickness are used as the base material for the fabrication.

A CW laser source is input in the demodulator to measure the transmission spectrum using a similar setup as shown in Fig.4.2. The resultant photocurrent transmission spectra are shown in Fig.4.5(c). The measured FSR is found to be the same as the previous design with edge coupling, which also agrees with analytical results.
Figure 4.5: (a) Structure of CPA-DPSK with grating coupler (b) TE grating coupler (c) Transmission spectra of CPA-DPSK with grating coupler.
4.6 Optsim Demonstration of CPA

The coherent perfect effect is also demonstrated using Optsim for a system level simulation. Fig.4.6 shows the diagram we used in Optsim for testing the performance of a CPA based DPSK system. A PRBS8 signal is generated from a pseudorandom binary sequence generator (PRBS1) and pass through a T-flip-flop (TFF2) for precoding for DPSK modulation. This logical signal then transfer to a voltage signal by a voltage driver (RC_Driver1). The voltage output is then applied on a extensional modulator (ExMod1) to induce the phase modulation on a CW source (CWLaser1). This DPSK modulated optical signal is split into two arms with a 0.01 ns (corresponding to 100G bits/s) from a fiber delay, and then go into the two input ports of a ring resonator. The ring resonator is represented by a combination of two $2 \times 2$ directional couplers and two straight waveguides. The length of the two waveguides represents the circumstance of the ring resonator. The coupling loss of ring is controlled by the coupling coefficient of the directional coupler. And the intrinsic loss of the ring is controlled by an additional loss modules. The light coming from the through port is then pass to a photodiodes following by a transimpedance amplifier (TIA). A set of monitors with different function are used for displaying signals.

Fig.4.7(b) shows that the output signal is perfectly matched with the original logical signal Fig.4.7(a), which demonstrates the CPA-DPSK system. The eye-diagram of the CPA-DPSK system is shown in Fig.4.7(c). The bit error rate (BER) is measured to be $10^{-9}$, and the Q-square factor is measured to be $Q^2 = 15.6$. 

136
Figure 4.6: Optsim setup for CPA-DPSK demonstration
Figure 4.7: (a) Original signal coming from a PRBS generator. (b) Demodulated DPSK voltage signal. (c) Eye diagram of the output DPSK demodulated voltage signal.
5.1 Conclusion

We are living in an age of “nano”. Thus nanophotonics is an new exciting frontier that captures the imaginations of people worldwide, creates exciting opportunities and enables new technologies. The key fact is that nanophotonics deals with interactions between light and matter at a scale shorter than the wavelength of a light itself. In this thesis, we investigate three topics of nanophotonics, i.e., plasmonics, graphene, and silicon photonics, to explore the applications of nanophotonics.

In Chapter 1, we summarize the fundamental theories and concepts of a nonlinear slab model, a MLG interface model, a surface hydrodynamic model, and coherent perfect absorption effect that are essential to this thesis work. We also provided a description of various simulation tools, i.e., finite element method, and finite difference time domain method, used to perform calculations regarding propagation of electromagnetic waves in the following chapters. We then discuss a more detailed description of THG experiment setup used in Chapter 3. Finally, we also provide the processing steps for fabrication of silicon photonics circuits, as performed at Brookhaven National Laboratories.
Plasmonics is a very important concept in the field of nanophotonics, where we are concerned primarily with the manipulation of light at the nanoscale. In Chapter 2, we have developed a new surface-hydrodynamic model to enable numerical simulation of plasmonic devices in the presence of significant nonlocal polarization in the material medium. Such numerical methods are of pressing importance for calculations of SPPs on media with complex geometries and materials. Our model is based on the standard hydrodynamic model and enables an accurate response near plasmon resonance frequency by taking into account the change in electron density at the interface, i.e., electron spill-out, which is obtained from the accumulated d-parameter under consideration. Our surface-hydrodynamic model agrees well with the analytical method, which is distinct from the standard HDM. Our approach also provides a new method for accurate simulation of plasmonic devices, which may be too complex for direct use of an analytic model.

Graphene is an important material in nanophotonics and attracts a lot of interest in both its linear and nonlinear optical properties. In Chapter 3, we have demonstrated optical third-harmonic generation to be a deterministic probe of the nonlinear optical property of multilayer graphene. The TH signal of MLG exhibits a complex dependence on its layer number. We found that the optimal TH signal occurs for \( N \approx 24 \), in good agreement with both interface and slab models. The coherence of the TH signal arising from the different graphene layers is consistent with the stacking property of the conductivity of MLG, in which \( \sigma_{N,g} = N\sigma_g \) even for the case of large layer numbers. Our result will lead to the development of ultrathin composite graphene films having large third-order optical nonlinearities thereby opening a vast
array of potential applications in optoelectronic and optical devices.

Silicon photonics is an important aspect of nanophotonics. DPSK is an increasingly important optical modulation format for long-haul fiber communications and short-haul on-chip communications. In Chapter 4, we make design and demonstration of a differential phase shift keying modulator on the base of coherent perfect absorption effect. We have thoroughly designed such circuit, simulated and demonstrated the idea with extensive simulations, including FDTD simulations and Optsim simulations. The experiment results show that we can indeed demodulate a DPSK encoded signal at a data rate of 10 Gbps.

5.2 Future Works

We have demonstrated CPA-based DPSK demodulation for a data rate of 10Gbps. In the near future, we will explore a higher bit rate device, with a shorter path length difference. In addition, we have demonstrated the working principle of CPA-based DPSK using Optsim, a system level simulation software. In the future, we wish to make a more accurate simulation combining both system level simulation Optsim and device level simulation RSOFT.

We have investigated the THG in multilayer graphene, and found an optimal layer is giving the highest THG signal from MLG. In the future, we will explore enhancing the THG by stacking the MLG/dielectric system. From the phase match point of view, this system is working like a quasi-phase match method, which can enhance the THG signal. A theoretical approach is also needed to develop to consider the
coherent combination from each MLG/dielectric subsystem.

We have proposed a numerical approach – surface hydrodynamic model, which incorporates nonlocality and surface correction. We have also demonstrated its ability for complex structures, i.e., a dielectric/metal/dielectric system. In the future, we will explore the applications of the surface hydrodynamic model in more complex plasmonic structures, e.g., V-groove waveguides and hybrid waveguides.
Other Relevant Works

Two-color Field Enhancement at an STM Junction for Spatiotemporally Resolved Photoemission

In this work, we report measurements and numerical simulations of ultrafast laser-excited carrier flow across a scanning tunneling microscope (STM) junction. The current from a nanoscopic tungsten tip across a ~1 nm vacuum gap to a silver surface is driven by a two-color excitation scheme that uses an optical delay-modulation technique to extract the two-color signal from background contributions. The role of optical field enhancements in driving the current is investigated using density functional theory and full three-dimensional finite-difference time-domain computations. We find that simulated field-enhanced two-photon photoemission (2PPE) currents are in excellent agreement with the observed exponential decay of the two-color photoexcited current with increasing tip-surface separation, as well as its optical-delay dependence. The results suggest an approach to 2PPE with simultaneous subpicosecond temporal and nanometer spatial resolution.
Figure 5.1: (a) Normalized, time-averaged electric field intensity spatial distribution in response to two-color excitation for a 55 nm radius tungsten tip, and (b) the plasmonic mode intensity at the midpoint between the tip apex and the substrate as a function of the tip radius. The tip apex-substrate spacing is 2 nm in both (a) and (b).

Figure 5.2: (a) Time-resolved 2C-SPPX photoemission signal (open circles). Solid line: delay-time derivative of the 3D-FDTD-simulated 2C-SPPX signal \( \frac{d\gamma}{dt_d} \), \( \gamma \) is defined in the main text). (b) Measured STM current (circles), 2C-SPPX signal (squares) and \( \gamma \) (proportional to \( \frac{d\gamma}{dt_d} \) at fixed delay, triangles) versus tip-sample distance. Solid lines are single-exponential-decay fits.
Bibliography


132. Dean, J. J. & van Driel, H. M. Second harmonic generation from graphene and


134. Bao, Q. *et al.* Atomic-layer graphene as a saturable absorber for ultrafast pulsed


137. Gu, T. *et al.* Regenerative oscillation and four-wave mixing in graphene opto-


Thourhout, D. Electrically tunable optical nonlinearities in graphene-covered
SiN waveguides characterized by four-wave mixing. *ACS Photonics* **4**, 3039–
3044 (2017).

140. Harutyunyan, H., Beams, R. & Novotny, L. Controllable optical negative re-

141. Kumar, N. *et al.* Third harmonic generation in graphene and few-layer graphite

142. Bykov, A. Y., Murzina, T. V., Rybin, M. G. & Obraztsova, E. D. Second
harmonic generation in multilayer graphene induced by direct electric current.

second-harmonic generation induced by electric current in graphene on Si and

144. Margulis, V. A., Muryumin, E. & Gaiduk, E. Electric-field-induced optical


Appendix. Publication

Journals


Conferences


