ULTRASLOW AND STOPPED LIGHT IN METAMATERIALS

Kosmas L. Tsakmakidis

Submitted for the degree of
Doctor of Philosophy

UNIVERSITY OF
SURREY

Advanced Technology Institute
School of Electronics and Physical Sciences
University of Surrey
United Kingdom

December 2008

© Kosmas L. Tsakmakidis
Declaration

This thesis is submitted for the degree of Doctor of Philosophy at the University of Surrey. The presented work has been carried out under the supervision of Professor Ortwin Hess, Professor Michael Kearney, at the Advanced Technology Institute and Department of Physics, School of Electronics and Physical Sciences, University of Surrey. This thesis is a presentation of my original research work. No part of this thesis has been submitted in any form for any other degree, diploma or any other qualification. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.
Do not all charms fly
At the mere touch of cold philosophy?
There was an awful rainbow once in heaven:
We know her woof, her texture; she is given
In the dull catalogue of common things.
Philosophy will clip an Angel's wings,
Conquer all mysteries by rule and line,
Empty the haunted air, and gnomed mine -
Unweave a rainbow, as it erewhile made
The tender-person'd Lamia melt into a shade.

John Keats (1795 – 1821), *Lamia*

(The famous, romantic English poet accusing
Sir Isaac Newton of destroying all the poetry
of the rainbow by reducing it to its prismatic colors!)
Acknowledgements

I would herein like to take the opportunity and express my sincere gratitude to several people who have, in one way or another, helped me during the long course of this work. First of all, I am indeed indebted to Professor Ortwin Hess for the support and trust with which he has surrounded this work, the very many and useful discussions, his guidance and mentorship. I have had (and hope to continue having for a long more time) a great time working with him and his group, without whom I would have never even entered the realms of ‘metamaterials’ and ‘slow light’. I am, also, truly grateful to several previous members of the Theory and Advanced Computation (TAC) group at the Advanced Technology Institute (ATI) of the University of Surrey (UniS), including Dr Christian Herman, Dr Andreas Klaedtke, Dr Cécile Jamois and Dr Klaus Boehringer for the daily supervision and assistance during the early stages of this work. I am certainly indebted to my dear friend, Dr Durga Aryal, for the many discussions and opinions we have exchanged on numerous scientific issues, and for his friendship (and for lending me his credit card in several scientific conferences!). I would also like to extend my thanks to Dr Joachim Hamm for the computational support during the course of my thesis, and to our two new PhD students, Mr James Cook and Mr Edmund Kirby, for elaborating and extending many issues discussed in this work. Finally, it is my pleasure to express my thankfulness to (collectively) all the previous MSc and PhD students of the TAC group, all of whom have contributed in my pleasant and (hopefully) productive stay in England.

Last but not least, I would like to thank my parents, Lazaros and Maria, and brother Lefteris for all the moral support and encouragement they have abundantly provided me with during my doctoral study. This thesis is, indeed, dedicated to them all.

Guildford, December 2008
Kosmas L. Tsakmakidis
Contents

Acknowledgements ................................................................................................ iv

1 Abstract .................................................................................................................. 1

2 Introduction ........................................................................................................... 2

3 Vectorial Field Theory and Modelling of 3-D Dielectric Waveguides 9
   3.1 Maxwell’s Equations of Electromagnetism ......................................................... 9
   3.2 The Wave Equations .......................................................................................... 11
      3.2.1 Vectorial Wave Equation for the Electric Field E ........................................ 11
      3.2.2 Vectorial Wave Equation for the Magnetic Field H ..................................... 13
   3.3 Boundary Conditions for the Electromagnetic Fields ....................................... 15
   3.4 Exact Analysis of a Three-Layer Slab Waveguide ............................................ 17
   3.5 Effective Index Method .................................................................................... 19
   3.6 Marcatili’s Method ........................................................................................... 23
   3.7 Finite-Difference Frequency-Domain Mode-Solvers ......................................... 27
      3.7.1 Vectorial Wave Equations for the E-Field Components ............................... 27
      3.7.2 Vectorial Wave Equations for the H-Field Components ............................... 29
      3.7.3 Semi-Vectorial Wave Equations .................................................................. 31
      3.7.4 Finite-Difference Discretization .................................................................. 33
      3.7.5 Numerical Examples of the FV-FDFD Mode-Solver .................................... 36

4 Conventional and Nonstandard Finite-Difference Time-Domain Method 41
   4.1 The Yee Algorithm ........................................................................................... 41
   4.2 Stability of the Yee Algorithm .......................................................................... 44
   4.3 ‘Exact’ and ‘Nonstandard’ Finite-Differences .................................................... 47
   4.4 Generalization to Three Dimensions .................................................................. 49
4.4.1 Standard Wave-Equation Finite-Difference Algorithm .................................. 49
4.4.2 Non-Standard Finite-Difference Time-Domain Algorithm ............................. 50
4.5 Application of the NS-FDTD Method in the Simulation of a Three-Dimensional Dielectric Waveguide.............................................................................................. 53

5 Light Propagation in Negative-Refractive-Index Metamaterials and Waveguides 60
5.1 A Brief History of Metamaterials .............................................................................. 61
5.2 Sign of the Refractive Index and Energy Density Expression in Passive ‘Double Negative’ Metamaterials ................................................................. 63
5.3 Refraction and E-H-k Vector Triad inside a ‘Double Negative’ Metamaterial .... 65
5.4 Fresnel’s Formulas for Plane Wave Incidence at a Planar RH/LH Media Interface 69
5.5 Metamaterial-Enabled ‘Perfect’ Lens ....................................................................... 71
5.6 Surface Plasmon Polaritons in Asymmetric DNGM Slab Heterostructures .... 74
5.6.1 Surface Plasmon Polaritons at a Planar LH/RH Interface ................................. 75
5.6.2 Surface Plasmon Polaritons in asymmetric LH Slab Waveguides ................... 78
5.6.3 Summary of the Identified SPP Solutions ..................................................... 98
5.7 Oscillatory Guided Modes Supported by Asymmetric DNGM Slab Heterostructures .... 99

6 Design Methods for Constructing Metamaterials 107
6.1 Metamaterials with Negative Effective Permittivity in the Microwave Regime ........ 108
6.2 Metamaterials with Negative Effective Permeability in the Microwave Regime .... 112
6.3 Intrinsically Lossless Magnetic Metamaterials ......................................................... 114
6.3.1 Case of ‘Isolated’ Unit Cells ........................................................................... 115
6.3.2 Calculations of Active Powers in the Equivalent Electrical Circuits .............. 129
6.3.3 Case of ‘Tightly Coupled’ Unit Cells ........................................................... 135
6.3.4 Remarks on the Working Principle of the Magnetically Lossless 2-DEG Configuration ........................................................................................................... 143
6.3.5 Issues with the Homogenisation Method for the Case of ‘Tightly Coupled’ Unit Cells ................................................................. 148
6.3.6 Systems with M Degrees of Freedom ........................................................... 152
7 'Trapped Rainbow' Stopping of Light in Metamaterials

7.1 Broadband Stopping of Light in Metamaterial Waveguides 

7.2 Derivation of Spatiotemporal Field-Component Equations Used in the
Adiabatic Variation

7.3 Derivation of the Expressions for Light-Ray
Goos-Hänchen Spatial Displacements

7.4 Derivation of the Relation between the Total Time-Averaged Power Flow and
the Effective Thickness of the Waveguide

7.5 Characteristic Impedance of Left- and Right-Handed Waveguides

8 Conclusions and Further Work

References

Publications
1 Abstract

The scope of the present doctoral thesis has been the conception of a novel and efficient method for decelerating, over a range of frequencies, and completely ‘stopping’ light (zero group velocity, $v_g = 0$) inside solid-state materials, at room temperature. To this end, we analytically show that an adiabatically tapered waveguide having a core of a lossless negative refractive index (NRI) metamaterial (MM) and claddings made of normal dielectrics can ‘trap’ a light pulse in such a way that each individual frequency component of the pulse is stopped at a different point along the waveguide, forming what we have called a ‘trapped rainbow’. Crucially, it is shown that light can efficiently be in-coupled inside such a waveguide heterostructure from a normal dielectric waveguide, since with a suitable design one can achieve simultaneous thickness-, mode- and characteristic-impedance-matching between the two waveguides. A pertinent analysis reveals that the optical path length of a ‘trapped’ light ray (associated with a particular frequency component of the pulse), as well as the corresponding effective thickness of the NRI waveguide itself, become exactly zero. The ray circulates at the point where it is trapped in such a way that its trajectory forms what we have called (in view of its characteristic hourglass form) an ‘optical clepsydra’.

Furthermore, we introduce a novel methodology that allows for obtaining ultra-low- or zero-loss magnetic metamaterials over a continuous range of frequencies. We analytically prove that a higher-degrees-of-freedom MM design methodology based on equivalent electrical circuits with more than one mesh leads to metamaterial magnetism with either ultra-high figures-of-merit or with perfectly lossless performance over a broad range of frequencies. The so-obtained lossless metamaterial magnetism has a truly intrinsic character, and as such is scalable and can be implemented at any frequency regime, from the radio up to the optical domain.
2 Introduction

An important scientific breakthrough in modern electromagnetics and optics has been the conception [1]-[4] and practical implementation [5]-[15] of materials exhibiting simultaneously negative electric permittivity, magnetic permeability and refractive index, known also as “double-negative” (DNG) or “left-handed” (LH) or, simply, “negative-refractive-index” (NRI) “metamaterials” (MMs). Their conceivable strong economic and social impact, owing to their potential applicability in diverse realms of science, such as telecommunications, radars and defence, nanolithography with light, microelectronics, medical imaging, and so forth, has lately prompted an overwhelming excitement within the wider scientific community [16]-[19]. Indeed, according to recent bibliometric data [20], metamaterials – together with the closely related field of plasmonics – are nowadays one of the largest (in terms of published research output) realms of contemporary physics, even larger than well-known, high-profile areas of (theoretical or experimental) physics, such as string theory or high-temperature superconductivity.

The fact that the refractive index, $n$, of a passive medium becomes negative when its electric permittivity, $e$, and magnetic permeability, $\mu$, are simultaneously (at the same frequency region) negative was established by a Russian physicist, Victor Veselago, some four decades ago [1], [2]. Veselago also noticed that, amongst others, a negative $n$ implies a left-handed E-H-k vector triad (k being the wavevector) and antiparallel phase and group velocities for a propagating plane electromagnetic wave, as well as a reversal of a number of well-known electromagnetic effects, including the direction of refraction of a plane wave inside a DNG medium. The latter effect leads directly to the ability of a planar slab made of a DNG medium to act as lens, bringing to a double focus an electromagnetic wave radiated by an object source, once inside the DNG slab and a second time outside the opposite (to the side nearest to the source) side of the slab.

Despite these remarkable findings, Veselago’s work on DNG media went largely – but not entirely [21] – unnoticed and, with time, was forgotten by the scientific community. The reasons should probably be traced to the different priorities that were laid down for optics and electromagnetics research at the time, and also because Veselago’s work might have been regarded as “too theoretical”, since he did not
accompany his observations by clear recipes as to how such DNG materials could be fabricated and tested. It is interesting to note here that, amongst others, Kock also considered — two decades before Veselago — the possibility of constructing “artificial dielectrics” [22]. In particular, he realised that by suitably arranging metallic inclusions of various shapes and arrangements in a dielectric host, one could manipulate the effective refractive index of the resulting effective medium. Kock deployed this concept in the construction of lightwave microwave lenses used in antennas applications. While a number of scientists and engineers did extend and built up on Kock’s work on artificial dielectrics [23], and while it appears that some scientists have indeed considered the possibility of antiparallel phase and group velocities, as well as ‘negative refraction’ as early as 1904 [24], it is probably fair to state that prior to Veselago there has never been a systematic study of “artificial dielectrics” exhibiting concurrently negative (effective) $\varepsilon$ and $\mu$, nor any identification of the important consequences that such a concurrence implies — negative refractive index, backward phase, left-handed triad, flat lens, reversed Vavilov-Čerenkov effect, reversed Doppler effect, and so forth.

A revived interest in these materials occurred nearly a decade ago, when Sir John Pendry theoretically proposed practical methods with which one could ‘build’ negative-$\varepsilon$ [3] and negative-$\mu$ [4] materials — the building blocks for a NRI metamaterial. Soon afterwards his proposals were experimentally tested and verified, with the first NRI medium been demonstrated in 2000 [5]. At the same year Pendry made the remarkable prediction that a slab made of NRI medium with $n = -1$, surrounded by air, can act as a ‘perfect’ lens, not limited by diffraction, i.e. a lens that can enable ‘perfect’ resolution and restoration (at the image plane) of even the tiniest details of an object [25]. Although that suggestion — along with the possibility of even obtaining NRI materials and observing negative refraction — was initially met with scepticism by some scientists [26]-[29], it was soon proved conclusively by means of theory [30], [31], simulations [32], [33], and further experiments [8], [9] that NRI materials, negative refraction and even a ‘superlens’ do not violate any physical law and can, thus, be constructed, implemented and harnessed.

It is by now well-established that not only do such materials exist, but that they can even be constructed to exhibit broadband behaviour [11], [16]-[18]— in fact, they may even possess infinite bandwidth —, scaled down to optical frequencies and be built in three dimensions [34], as well as allow for efficient and fast tuning and switching
It has, also, been established that, by using suitable electronic or optical gain media to compensate for the occurrence of losses, DNG media can be made lossless over a broad but finite – frequency region. Therefore, metamaterials exhibiting unusual (not necessarily only negative) values for their effective magnetic permeability and/or electric permittivity do open a completely new perspective to the optical world, enabling functionalities unattainable using conventional (mostly dielectric) media, such as optical nanocircuits, or functionalities that were previously thought to be impossible, such as focusing of electron de Broglie waves by sharp p-n junctions in graphene or even ‘invisibility cloaks’.

During the same period and in parallel with the advances that were occurring in the realm of metamaterials, another exciting and wide-ranging area of contemporary research has also been developing. The goal in that field was to produce ultra-slow or even completely ‘stopped’ light, i.e. light that propagates in a medium with a velocity much smaller compared to the speed of light in vacuum. Apart from being an exciting objective in itself, the reason for pursuing such a research was partly motivated by the realisation that ‘slow’ light could enable much more efficient manipulation, routing and switching of optical packets of information in optical-fibre networks, and could even allow for all-optical computing and memories that could one day replace their electronic counterparts. However, ‘stopping’ photons (the fastest ‘entities’ in the universe) is extremely challenging and, indeed, for decades it was thought that “optical data cannot be stored statically and must be processed and switched on the fly”.

Undeniably, the absence of any form of interaction between photons and other elementary particles, as well as their enormous speed, makes confining them to a finite volume by reducing their velocity down to zero excessively difficult. However, in recent years, scientists were able to annul such assertions and proved conclusively that it is, indeed, possible to bring light to a complete standstill. Amongst others, electromagnetically induced transparency (EIT), quantum-dot semiconductor optical amplifiers (QD-SOAs), photonic crystals (PhCs), coherent population oscillations (CPOs), stimulated Brillouin scattering (SBS) and surface plasmon polaritons (SPPs) in metallodielectric waveguides have been proposed as means of producing ‘slow’ or even ‘stopped’, stored and regenerated light.
Unfortunately, most of these methods bear inherent limitations that may hinder their practical deployment. For instance, EIT uses ultracold atomic gases and not solid state materials, QD-SOAs usually allow for only modest delays but for potentially ultra-broadband light pulses, CPOs and SBS are very narrowband owing to the narrow transparency window of the former and the narrow Brillouin gain bandwidth (around 30 MHz in standard single-mode optical fibres) of the latter, SPPs are very sensitive to surface roughness and are relatively difficult to excite, while PhCs are normally highly multimodal; this, combined with the strong impedance mismatch in the ‘slow-light regime’ causes launching the incoming light energy to a single, slow mode alone overly difficult.

In an effort to overcome these hindrances a fundamentally new and auspicious approached has been recently reported [55], [56], which indeed constitutes the prime scope of the work presented herein. This approach, which will be presented in detail in chapter 7, invokes solid-state materials and, as such, is not subject to low-temperature or atomic coherence limitations. Moreover, it inherently allows for high in-coupling efficiencies and broadband function, since the deceleration of light does not rely on refractive index resonances. In particular, we shall show that an axially varying heterostructure with a core of NRI metamaterial can be used to efficiently and coherently bring light to a complete standstill. In sharp contrast to the previously mentioned schemes for decelerating and storing light, the present one simultaneously allows for high in-coupling efficiencies and broadband, room-temperature, operation. Our analysis will reveal that at a critical point the effective thickness of the metamaterial waveguide reduces to zero, preventing the lightwave to propagate further. At this point, a light ray is found to be permanently trapped, its trajectory forming a double light-cone that we call an ‘optical clepsydra’. Each frequency component of a wave packet is stopped at a different guide thickness, leading to the spatial separation of the packet’s spectrum and the formation of a ‘trapped rainbow’. Our results, thus, bridge the gap between the realms of metamaterials and slow light, and may open a new host of combined investigations. Such macroscopic control of photons could, also, conceivably find applications in optical data processing and storage or in the realisation of quantum optical memories.

This work begins, in chapter 3, with an introduction into dielectric waveguide theory [57], which we shall find useful when we later examine, in chapter 7, the slowing and stopping of light in NRI metamaterial waveguides. Chapter 3 introduces
the relations that describe the vectorial field propagation in waveguides, and summarises several semi-analytic methods that are sometimes used in the investigation of such structures. Particular emphasis is paid upon the derivation and finite-difference discretisation of the vectorial wave equations, as well as on the development of fully-vectorial (FV) mode-solvers [58], which nowadays constitute an indispensable tool in the analysis of three-dimensional waveguides. We compare a number of results obtained by our so-developed FV mode-solvers with published data (obtained with FV mode-solvers or with other numerical techniques) and we report good agreement between the two.

Chapter 4 introduces the finite-difference time-domain (FDTD) method for the numerical solution of Maxwell’s equations of electromagnetism [59]-[61]. This method is widely-used by the scientific community in assessing and optimising the performance of optical devices, and is also extensively deployed in the analysis of NRI metamaterials. We deploy this method later in chapter 5 to show that a NRI slab can bring light to a double focus, but in chapter 4 we are mainly concerned with the theoretical aspects behind the operation of this method – in particular, its stability. We, further, study the recently introduced, exciting concepts of ‘exact’ and ‘nonstandard’ finite-differences [62], which can lead to enhanced computational accuracy and overall algorithmic efficiency. Indeed, we compare the standard and the higher-order (so-called, ‘nonstandard’ [63]) FDTD algorithm in the simulation of optical propagation inside a three-dimensional dielectric waveguide, and report that improved accuracy or substantial computational savings (in terms of time and/or memory) are obtained when the nonstandard algorithm is deployed.

In chapter 5 we provide an introduction into the theory of light propagation inside passive NRI bulk metamaterials and waveguides. We establish a number of crucial issues pertaining to the properties of such media, such as why does the sign of the refractive index become negative when both $\varepsilon$ and $\mu$ are negative, to which direction does a plane electromagnetic wave refract inside an isotropic NRI medium, and so forth. Upon deriving the Fresnel equations for the reflection and transmission coefficients of propagating and evanescent waves at the interface between a normal dielectric and a NRI medium, we study in some detail one of the most remarkable properties of NRI metamaterials, namely the fact that a planar slab made of such a material can act as a ‘perfect’ lens, able to bring into focus and restore both the propagating and the evanescent components of an object source. Since surface waves
that decay evanescently at the two interfaces of a NRI slab play a crucial role to the manifestation of the ‘perfect’ lensing effect, we then proceed by studying in detail and classifying all surface plasmon polariton (SPP) modes supported by such slab structures [64]. We, also, study the oscillatory modes propagating in these waveguides and identify that some of them can attain zero group velocity, exist alone, and be efficiently excited – results that will be proven useful in the discussions that will be presented later, in chapter 7.

Chapter 6 is concerned with one of the most important contributions of this work, that is how to create an ultra-low- or zero-loss metamaterial over a continuous range of frequencies [65]. We start by reviewing some central previous works on how to construct metamaterials exhibiting negative electric permittivity or magnetic permeability in the microwave regime. We then introduce a novel methodology, utilizing equivalent electric circuit configurations with multiple degrees of freedom, which is shown to result in ultra-low-loss metamaterials exhibiting negative magnetic permeability. We show that with the novel metamaterial configurations one is able to obtain metamaterial magnetism with large figures-of-merit (defined as the ratio of the real part of $\mu$ to imaginary part of $\mu$). Furthermore, when the new magnetic metamaterials are made anisotropic, it is shown that we can even obtain artificial magnetic “gain” over a range of frequencies at a specified direction, at the cost of having increased (magnetic) losses at another direction. Thus, this methodology opens a ‘window’ for harnessing intrinsically zero-loss metamaterial magnetism – in fact, even with artificial “gain” – over a frequency range of frequencies, without having to use gain media to compensate for the losses.

Finally, this work concludes by introducing and presenting a novel and important potential application that these low- or zero-loss metamaterials can find, namely in decelerating and storing broadband light inside NRI waveguide heterostructures. In particular, we show that an electromagnetic pulse propagating along an adiabatically, axially varying waveguide with a NRI core (or cladding) can be entirely stopped and stored inside it, with each frequency component stopping at a different point inside the tapered waveguide. This configuration, thus, leads to the spatial separation and ‘trapping’ of the frequency components of a pulse, forming what we have called a ‘trapped rainbow’. Conceivably, such a functionality may enable ultra-compact dispersion compensators, dramatically enhanced (low-intensity/few-photon) nonlinear
effects, as well as ultimately lead to applications in all-optical routing, switching and data storage.
3 Vectorial Field Theory and Modelling of 3-D Dielectric Waveguides

The basic principles and underlying field theory of two- (2-D) and three-dimensional (3-D) dielectric waveguides are introduced in this chapter. First, we summarize Maxwell’s equations, the vectorial wave equations and the associated boundary conditions for the electromagnetic fields. We, then, concisely describe several analytic methods, including the effective index method and Marcatili’s method, commonly deployed for prompt investigations of 3-D waveguiding structures. Finally, we describe in some detail the rigorous semi- (SV) and fully-vectorial (FV) finite-difference frequency-domain (FDFD) mode-solving techniques for the acquisition of the electromagnetic eigenmodes in a dielectric waveguide with an arbitrary 2-D cross-section.

3.1 Maxwell’s Equations of Electromagnetism

For a homogeneous, isotropic, non-dispersive, linear and achiral medium the electric field $\mathbf{E}$ (V/m), magnetic field $\mathbf{H}$ (A/m), electric flux density $\mathbf{D}$ (C/m$^2$) and magnetic flux $\mathbf{B}$ (A/m$^2$) are related, in the time-domain, through the following macroscopic constitutive relationships [66], [67]:

\[
\begin{align*}
\mathbf{D} &= \varepsilon \mathbf{E}, \\
\mathbf{B} &= \mu \mathbf{H},
\end{align*}
\]  

(3.1) \hspace{1cm} (3.2)

where the electric permittivity $\varepsilon$ and the magnetic permeability $\mu$ are defined as:

\[
\begin{align*}
\varepsilon &= \varepsilon_0 \varepsilon_r, \\
\mu &= \mu_0 \mu_r.
\end{align*}
\]  

(3.3) \hspace{1cm} (3.4)

In these expressions, $\varepsilon_0$ and $\mu_0$ are, respectively, the electric permittivity and magnetic permeability of vacuum. Moreover, with $\varepsilon_r$ and $\mu_r$ we denote the relative
(to the vacuum) permittivity and permeability of the material. As we shall see later, in chapter 5, the relative permeability \( \mu_r \) of several artificial, non-magnetic, dielectrics may exceed 1, or even become negative (e.g. \( \mu_r = -1 \)). Nonetheless, throughout the present chapter we shall assume that for all non-magnetic dielectrics investigated, \( \mu_r \) is exactly equal to 1. Denoting the velocity of light in vacuum with \( c \), one may then obtain:

\[
\varepsilon_0 = \frac{1}{c^2 \mu_0} \equiv 8.85 \times 10^{-12} \text{ F/m},
\]

(3.5)

\[
\mu_0 = 4\pi \times 10^{-7} \text{ H/m}.
\]

(3.6)

It should also be noted here that in a conductive dielectric material the current density \( \mathbf{J} \) (A/m\(^2\)) relates to the electric field \( \mathbf{E} \) through the following relationship:

\[
\mathbf{J} = \sigma \mathbf{E},
\]

(3.7)

where \( \sigma \) is the conductivity of the material.

Armed with the above relations, we may now proceed to analytically describe the interactions between the electric and magnetic fields in such media. In general, the electromagnetic fields satisfy the following set of coupled, first-order, differential equations, introduced by Maxwell more than a century ago [68]:

\[
\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t},
\]

(3.8)

\[
\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J}.
\]

(3.9)

The above two equations are supplemented by the following two independent equations, which imply the absence of magnetic monopoles (Eq. (3.10)) and Gauss’s law (Eq. (3.11)):

\[
\nabla \cdot \mathbf{B} = 0,
\]

(3.10)

\[
\nabla \cdot \mathbf{D} = \rho.
\]

(3.11)
In deriving Eq. (3.11), it was assumed that the current density \( J \) is related to the charge density \( \rho \) (C/m\(^2\)) as follows:

\[
\nabla \cdot J = \frac{\partial \rho}{\partial t}.
\]

(3.12)

### 3.2 The Wave Equations

Let us now assume that an electromagnetic field oscillates at a single angular frequency \( \omega \) (rad/s). Then, the electric and magnetic field, as well as the electric and magnetic flux density, each designated with a vector \( \mathbf{A} \), may be expressed as:

\[
\mathbf{A}(r, t) = \text{Re}\{\mathbf{A}(r)\exp(j\omega t)\}.
\]

(3.13)

In the following, and for the sake of simplicity, we will denote the phasors \( \mathbf{E}, \mathbf{H}, \mathbf{D} \) and \( \mathbf{B} \) as \( \mathbf{E}, \mathbf{H}, \mathbf{D} \) and \( \mathbf{B} \), respectively. Based on Eq. (3.13), we can rewrite Eqs. (3.8) to (3.11) in the frequency domain as:

\[
\nabla \times \mathbf{E} = -j\omega \mathbf{B} = -j\omega \mu_0 \mathbf{H},
\]

(3.14)

\[
\nabla \times \mathbf{H} = j\omega \mathbf{D} = j\omega \varepsilon \mathbf{E},
\]

(3.15)

\[
\nabla \cdot \mathbf{H} = 0,
\]

(3.16)

\[
\nabla \cdot (\varepsilon, \mathbf{E}) = 0,
\]

(3.17)

where it was assumed that \( \mu_r = 1 \) and \( \rho = 0 \).

#### 3.2.1 Vectorial Wave Equation for the Electric Field \( \mathbf{E} \)

We, now, derive in some detail the fully-vectorial (FV) wave equation for the electric field. This equation will be of particular relevance when, later in this chapter, we shall be concerned with the various mode-solving techniques used in the analysis of 3-D dielectric waveguides. Applying the vectorial rotation operator \( \nabla \times \) to Eq. (3.14), we obtain:
Using the vectorial formula:

\[ \nabla \times (\nabla \times \mathbf{A}) = \nabla (\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}, \]  

with the symbol \( \nabla^2 \) denoting the 3-D Laplacian operator, we can rewrite the left-hand side of Eq. (3.18) as:

\[ \nabla (\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E}. \]  

We note at this point that Eq. (3.17) can also be written as:

\[ \nabla \cdot (\varepsilon_r \mathbf{E}) = \nabla \varepsilon_r \cdot \mathbf{E} + \varepsilon_r \nabla \cdot \mathbf{E} = 0, \]  

from where one may obtain:

\[ \nabla \cdot \mathbf{E} = -\frac{\nabla \varepsilon_r \cdot \mathbf{E}}{\varepsilon_r}. \]

Therefore, the left-hand side of Eq. (3.18) becomes:

\[ -\nabla \left( \frac{\nabla \varepsilon_r \cdot \mathbf{E}}{\varepsilon_r} \right) - \nabla^2 \mathbf{E}. \]  

On the other hand, by means of Eq. (3.15), the right-hand side of Eq. (3.18) becomes:

\[ k_0^2 \varepsilon_r \mathbf{E}, \]

where \( k_0 \) is the wavenumber in vacuum and is expressed as:
Thus, for a dielectric medium with relative permittivity \( \varepsilon_r \), the vectorial wave equation for the electric field \( \mathbf{E} \) is given by:

\[
\nabla^2 \mathbf{E} + \nabla \left( \frac{\nabla \varepsilon_r}{\varepsilon_r} \cdot \mathbf{E} \right) + k_0^2 \varepsilon_r \mathbf{E} = 0. \tag{3.26}
\]

Moreover, if we assume that the wave number \( k \) in the aforesaid medium is:

\[
k = k_0 n = k_0 \sqrt{\varepsilon_r} = \omega \sqrt{\varepsilon_0 \mu_0} = \omega \sqrt{\mu_0},
\]

we may rewrite Eq. (3.26) as:

\[
\nabla^2 \mathbf{E} + \nabla \left( \frac{\nabla \varepsilon_r}{\varepsilon_r} \cdot \mathbf{E} \right) + k^2 \mathbf{E} = 0. \tag{3.28}
\]

For a homogeneous medium, i.e. one in which the relative permittivity \( \varepsilon_r \) is constant, the vectorial wave equation can be reduced to the well-known, scalar, Helmholtz equation for the \( \mathbf{E} \)-field:

\[
\nabla^2 \mathbf{E} + k^2 \mathbf{E} = 0. \tag{3.29}
\]

### 3.2.2 Vectorial Wave Equation for the Magnetic Field \( \mathbf{H} \)

In the analysis of dielectric waveguides, the fully-vectorial wave equation for the magnetic field is sometimes preferred over the corresponding equation for the electric field, owing to the absence of field discontinuities at the dielectric interfaces that could hinder high-accuracy calculations. Here, we show how the FV wave equation for the \( \mathbf{H} \)-field can be obtained from Eqs. (3.14) to (3.17). To this endeavour, we start by applying the vectorial rotation operator \( \nabla \times \) to Eq. (3.15):

\[
k_0 = \omega \sqrt{\varepsilon_0 \mu_0} = \frac{\omega}{c}. \tag{3.25}
\]
\( \nabla \times (\nabla \times \mathbf{H}) = j\omega \varepsilon_0 \nabla \times (\varepsilon, \mathbf{E}). \)  

(3.30)

We, now, successively have:

\[
\nabla (\nabla \cdot \mathbf{H}) - \nabla^2 \mathbf{H} = j\omega \varepsilon_0 (\nabla \varepsilon_r \times \mathbf{E} + \varepsilon_r \nabla \times \mathbf{E})
= j\omega \varepsilon_0 (\nabla \varepsilon_r \times \mathbf{E}) + j\omega \varepsilon_0 \varepsilon_r (-j\omega \mu_0 \mathbf{H})
= j\omega \varepsilon_0 (\nabla \varepsilon_r \times \mathbf{E}) + k_0^2 \varepsilon_r \mathbf{H}.
\]

(3.31)

Based on Eqs. (3.15) and (3.16), one may further obtain:

\[
\mathbf{E} = \frac{1}{j\omega \varepsilon_0 \varepsilon_r} \nabla \times \mathbf{H},
\]

(3.32)

which, inserted into Eq. (3.31), leads to:

\[
\nabla^2 \mathbf{H} + \frac{\nabla \varepsilon_r \times (\nabla \times \mathbf{H})}{\varepsilon_r} + k_0^2 \varepsilon_r \mathbf{H} = 0.
\]

(3.33)

Using Eq. (3.27), we can rewrite Eq. (3.33) as:

\[
\nabla^2 \mathbf{H} + \frac{\nabla \varepsilon_r \times (\nabla \times \mathbf{H})}{\varepsilon_r} + k^2 \mathbf{H} = 0.
\]

(3.34)

For a homogeneous medium, i.e. one in which the relative permittivity \( \varepsilon_r \) is constant, the vectorial wave equation can be reduced to the well-known, scalar, Helmholtz equation for the \( \mathbf{H} \)-field:

\[
\nabla^2 \mathbf{H} + k^2 \mathbf{H} = 0.
\]

(3.35)

For a dielectric waveguide whose structure is uniform in the \( z \) direction (i.e. longitudinally invariant in \( z \)) the derivative of an electromagnetic field component with respect to \( z \) may be written as:
where \( \beta \) is the longitudinal propagation constant of the field. The ratio of the propagation constant \( \beta \) to the wave number in vacuum, \( k_0 \), is usually referred to as the effective index:

\[
n_{\text{eff}} = \frac{\beta}{k_0}.
\]  

(3.37)

The physical meaning of the propagation constant \( \beta \) is the phase accumulated (or rotated) per unit propagation distance. Therefore, the effective index, \( n_{\text{eff}} \), can be interpreted as the ratio of a phase rotation in a medium to the phase rotation in vacuum.

We may summarize the Helmholtz equations for the electric and magnetic field as:

\[
\nabla_\perp^2 U + (k^2 - \beta^2) U = 0,
\]  

(3.38)

or:

\[
\nabla_\perp^2 U + k_0^2 (\varepsilon - n_{\text{eff}}^2) U = 0,
\]  

(3.39)

where \( U = E \) or \( H \), and \( \nabla_\perp^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \).

### 3.3 Boundary Conditions for the Electromagnetic Fields

The boundary conditions required for the electromagnetic fields, in the case of two adjoining dielectric media 1 and 2, are summarized as follows:

(a) Tangential components of the electric fields are continuous, such that:

\[
E_{1r} = E_{2r}.
\]  

(3.40)
(b) When no current flows at the interface of two media the tangential components of the magnetic fields are continuous, such that:

\[ H_{1t} = H_{2t}. \]  

(3.41)

By contrast, when a current flows at the interface, the magnetic field is discontinuous. The magnetic components at each side of the interface are related to the current density as follows:

\[ H_{1t} - H_{2t} = J_S. \]  

(3.42)

Since the magnetic field and the current are perpendicular to each other, the vectorial representation is:

\[ \mathbf{n} \times (\mathbf{H}_1 - \mathbf{H}_2) = \mathbf{J}_S, \]  

(3.43)

where the vector \( \mathbf{n} \) is the unit normal vector at the interface of the two dielectrics.

(c) When there is no charge on the interface, the normal components of the electric flux densities are continuous, such that:

\[ D_{1n} = D_{2n}. \]  

(3.44)

By contrast, when there are charges on the surface, the electric flux densities are discontinuous and are related to the surface charge density, \( \sigma_S \), as follows:

\[ D_{1n} - D_{2n} = \sigma_S. \]  

(3.45)

(d) Normal components of the magnetic flux densities are continuous, such that:

\[ B_{1n} = B_{2n}. \]  

(3.46)
Before discussing the numerical, mode-solving, techniques for the study of dielectric waveguides it is instructive at this point to describe the corresponding analytic methodologies. First, we highlight an exact method for the analysis of a three-layer, asymmetric, slab waveguide. We shall return to this method later, in chapter 6, when we shall discuss the mode theory of generalized waveguides with negative refractive index. Here, we shall instead examine two further common methods, namely the effective index method and Marcatili’s method, which are commonly deployed for the mode analysis of rectangular dielectric waveguides with 2D cross-section.

3.4 Exact Analysis of a Three-Layer Slab Waveguide

Let us consider a three-layer, planar, dielectric slab waveguide [57] with refractive indices $n_1$, $n_2$, and $n_3$, as illustrated in Fig. 3.1. The direction of propagation is along the $+z$ direction. The structure is uniform in the $y$ and $z$ directions (i.e.: $\partial/\partial y = 0$ and $\partial/\partial z = -j\beta$, Eq. (3.36)). Region 2 is the core layer that has refractive index higher than those of the cladding layers. Since media 1 and 3 are semi-infinite (and uniform) and the tangential electromagnetic field components are, indeed, connected at the

![Figure 3.1: Schematic illustration of a three-layer, asymmetric, dielectric waveguide.](image-url)
interfaces between adjacent media, as mentioned in the previous paragraph, it suffices to start our analysis with the scalar Helmholtz equations (3.38) or (3.39) applied to each dielectric region.

Next, we discuss the two types of modes that propagate in this structure. To this end, we note from Eqs. (3.26) and (3.33) that the transverse electric and magnetic field components decouple owing to the absence of variation in the materials' refractive indices along the y direction. It, then, swiftly turns out directly from Maxwell's equations (3.14) and (3.15) that we may distinguish between two different electromagnetic modes: the transverse electric mode (TE mode), in which the electric field is not in the longitudinal direction ($E_z = 0$) but in the transverse direction ($E_y \neq 0$), and the transverse magnetic mode (TM mode), in which the magnetic field is not in the longitudinal direction ($H_z = 0$) but in the transverse direction ($H_y \neq 0$).

In this paragraph, we shall focus on the somewhat more involved TM case. From Maxwell's equations it turns out that, in this case, it is $H_x = H_z = E_y = 0$. The two electric field components are given by:

$$E_x = \frac{\beta}{\omega \varepsilon_0 \varepsilon_r} H_y,$$

$$E_z = -\frac{j}{\omega \varepsilon_0 \varepsilon_r} \frac{\partial H_y}{\partial x},$$

and the sole magnetic field component, $H_y$, fulfills the scalar wave equation:

$$\frac{d^2 H_y}{dx^2} + k_0^2 (\varepsilon_r - n_{eff}^2) H_y = 0.$$

The principal magnetic field component in the three dielectric regions can, therefore, be expressed as:

$$H_y(x) = \begin{cases} C_1 \exp(y_1 x), & x \leq 0 \\ C_2 \cos(y_2 x + \alpha), & 0 \leq x \leq W, \\ C_3 \exp[-y_3(x-W)], & x \geq W \end{cases}$$
where $\gamma_1 = k_0(n_{\text{eff}}^2 - n_1^2)$, $\gamma_2 = k_0(n_2^2 - n_{\text{eff}}^2)$ and $\gamma_3 = k_0(n_{\text{eff}}^2 - n_3^2)$.

Imposing the boundary conditions on the tangential field components at $x = 0$ and $x = W$, results in the following characteristic equation:

$$\gamma_2 W = \arctan\left(\frac{\varepsilon_{r2}}{\varepsilon_{r1}} \frac{\gamma_1}{\gamma_2}\right) + \arctan\left(\frac{\varepsilon_{r2}}{\varepsilon_{r3}} \frac{\gamma_3}{\gamma_2}\right) + m\pi \quad (m = 0, 1, 2 \ldots). \quad (3.51)$$

Using the identity:

$$\arctan\left(\frac{y}{x}\right) = \frac{\pi}{2} - \arctan\left(\frac{x}{y}\right),$$

one may also obtain:

$$\gamma_2 W = -\arctan\left(\frac{\varepsilon_{r1}}{\varepsilon_{r2}} \frac{\gamma_1}{\gamma_2}\right) - \arctan\left(\frac{\varepsilon_{r3}}{\varepsilon_{r2}} \frac{\gamma_3}{\gamma_2}\right) + (m + 1)\pi, \quad (m = 0, 1, 2 \ldots). \quad (3.53)$$

Following a similar (dual) analysis for the TE modes, for which $E_x = E_y = H_y = 0$, one discovers that the so obtained characteristic equation is akin to Eqs. (3.51) and (3.53), with the sole difference being that the ratio of the relative permittivities appearing in the argument of the inverse tangents is, now, removed.

### 3.5 Effective Index Method

For the analysis of dielectric waveguides that are not uniform in the transverse, $y$, direction (i.e. they have a 2-D cross-section), the analytic method reported before cannot be efficiently applied owing to the difficulty of matching the tangential field components at the dielectric corners. A quasi-analytic method, known as Marcatili’s method [69], is usually deployed in these situations and will be described in some detail in the next paragraph. Here, we shall focus on the effective index method, which allows one to analyse 2-D cross-sectional dielectric waveguides by simply repeating the 1-D slab waveguide analysis highlighted above.
Figure 3.2: Schematic illustration of the "effective index method" concept.

Figure 3.2 shows an example of a 3-D optical waveguide (uniform in z direction) and concisely illustrates the concept of the effective index method. We start by considering the scalar wave equation:

\[ \frac{\partial^2 \varphi(x, y)}{\partial x^2} + \frac{\partial^2 \varphi(x, y)}{\partial y^2} + k_0^2 (\varepsilon, (x, y) - n_{\text{eff}}^2) \varphi(x, y) = 0, \]  

(3.54)
where \( n_{\text{eff}} \) is the eigenmode's effective index to be obtained. We, then, separate the wave function, \( \varphi(x, y) \), into two functions:

\[
\varphi(x, y) = f(x) \cdot g(y). \tag{3.55}
\]

Equation (3.55) implies that there is no coupling between the transverse \((x \text{ and } y)\) electromagnetic field components. This assumption, strictly speaking, is not true. However, it may lead to acceptable results on the condition that the dominant electric and magnetic field components are well confined inside the central/guiding layer (core), having negligible magnitude near the dielectric corners. As a result, the present method leads to reasonable results, typically, for the first couple of eigenmodes, or for weakly-guiding structures, for which the field-coupling at the corners is small.

Substituting Eq. (3.55) into Eq. (3.54) and dividing the resultant equation by the wave function \( \varphi(x, y) \), we obtain:

\[
\frac{1}{f(x)} \frac{d^2f(x)}{dx^2} + \frac{1}{g(y)} \frac{d^2g(y)}{dy^2} + k_0^2 (\varepsilon_r(x, y) - n_{\text{eff}}^2) = 0. \tag{3.56}
\]

Setting the sum of the second and third terms of Eq. (3.56) equal to \( k_0^2 N^2(x) \), we arrive at:

\[
\frac{1}{g(y)} \frac{d^2g(y)}{dy^2} + k_0^2 \varepsilon_r(x, y) = k_0^2 N^2(x). \tag{3.57}
\]

Inspection of Eqs. (3.56) and (3.57) reveals that the sum of the first and fourth terms is equal to \(-k_0^2 N^2(x)\):

\[
\frac{1}{f(x)} \frac{d^2f(x)}{dx^2} - k_0^2 n_{\text{eff}}^2 = -k_0^2 N^2(x). \tag{3.58}
\]

Therefore, via the aforementioned procedures, we obtain the following two independent equations:
\[
\frac{d^2 g(y)}{dy^2} + k_o^2 [e_r(x,y) - N^2(x)] g(y) = 0, \tag{3.59}
\]

and:

\[
\frac{d^2 f(x)}{dx^2} + k_o^2 [N^2(x) - n_{\text{eff}}^2] f(x) = 0. \tag{3.60}
\]

Based on Eqs. (3.59) and (3.60), and on Fig. 3.2, the eigenmode’s effective index calculation procedure can be summarized as follows:

(a) First, the 2-D cross-sectional dielectric waveguide is replaced with a combination of 1-D slab waveguides, as illustrated in Fig. 3.2.

(b) For each slab waveguide, we calculate the effective index along the \( y \) axis, based on the methodology described in section 3.4.

(c) We derive an “equivalent” slab waveguide by placing the effective indices calculated in (b) along the \( x \)-axis.

(d) Finally, we calculate the investigated eigenmode’s effective index by solving (using again the methodology described in section 3.4) the model obtained in (c).

It is important to note here that, for the quasi TE-like eigenmode of the 3-D waveguide shown in Fig. 3.2 (\( E \)-field parallel to the \( x \)-axis), one should first pursue the TE-mode analysis, followed by a corresponding \( TM \)-mode analysis, since in the final “equivalent” slab structure, the \( E_x \)-field will be perpendicular to the dielectric interfaces; hence the mode will, now, appear as a TM one. In a similar vein, for the quasi TM-like eigenmode (\( E \)-field parallel to the \( y \)-axis), one should first pursue the TM-mode analysis, followed by a \( TE \)-mode analysis, since in the “equivalent” slab waveguide, shown in the last part Fig. 3.2, the \( E_y \)-field will be parallel to the dielectric interfaces; hence the mode will appear as a TE one.
3.6 Marcatili’s Method

Proposed by Marcatili, who at that time, exactly 40 years ago [69], was working at Bell Labs, the method that we shall discuss in this section, is still in wide use for prompt investigations of various types of rectangular dielectric waveguides.

Figure 3.3 illustrates a cross-sectional view of a, so-called, buried optical waveguide. The core has refractive index \( n_1 \), width \( 2a \), and height \( 2b \). It is surrounded by cladding that has a refractive index \( n_2 \). Likewise the effective-index method, it is again assumed that the dominant components of the electric and magnetic fields are well-confined to the core and do not exist at all in the four shaded regions, shown in Fig. 3.3. Therefore, the continuity conditions for the tangential components need only be imposed at the interfaces of regions 1 and 2, 1 and 3, 1 and 4, and 1 and 5. The inherent limitations of the aforementioned assumption are essentially the same with the ones highlighted in section 3.5.

Let us focus on the \( E_{pq}^x \) eigenmode, which has \( E_x \) and \( H_y \) as principal field components. Here, \( p \) and \( q \) are integers that, respectively, correspond to the number of peaks of optical power in the \( x \) and \( y \) directions. Thus, unlike ordinary mode-orders, which begin from 0 (see e.g. Eqs. (3.51) and (3.53)), the present ones begin from 1.

![Figure 3.3: Geometry considered for the analysis of Marcatili’s method.](image-url)
The electric field of the $E_{pq}^{x}$ eigenmode is assumed to be polarized in the $x$ direction, which results in $E_y = 0$. Since the structure of the investigated waveguide is assumed to be invariant in the $z$ direction, the derivative with respect to $z$ is replaced by $-j\beta$ (Eq. 3.36)). Then, after some algebraic manipulations starting from Maxwell's equations (3.8)-(3.10), one may obtain:

$$H_x = \frac{1}{\omega \mu_0 \beta} \frac{\partial^2 E_x}{\partial x \partial y}, \quad (3.61)$$
$$H_y = \frac{1}{\omega \mu_0 \beta} \left( \beta^2 E_x - \frac{\partial^2 E_x}{\partial x^2} \right), \quad (3.62)$$
$$H_z = \frac{1}{j \omega \mu_0} \frac{\partial E_x}{\partial y}, \quad (3.63)$$
$$E_z = \frac{1}{j \beta} \frac{\partial E_x}{\partial x}. \quad (3.64)$$

It, further, swiftly turns out that the principal field component, $E_x$, satisfies the 2-D scalar wave equation:

$$\frac{\partial^2 E_x}{\partial x^2} + \frac{\partial^2 E_y}{\partial y^2} + k_0^2 (\varepsilon_x - n_{eff}^2) E_x = 0. \quad (3.65)$$

Since the dominant electric field component is a solution of Eq. (3.65), it should be described by the following expressions throughout all dielectric regions:

$$H_y(x) = \begin{cases} 
C_1 \cos(k_x x + a_1) \cos(k_y y + a_2), & -a \leq x \leq a, \quad -b \leq y \leq b \\
C_2 \cos(k_x x + a_1) \exp[-\gamma_y (y - b)], & -a \leq x \leq a, \quad b \leq y \\
C_3 \cos(k_x x + a_2) \cos(k_y y + a_1), & -b \leq x \leq a, \quad -b \leq y \leq b \\
C_4 \exp[\gamma_x (x - a)] \cos(k_y y + a_1), & a \leq x, \quad -b \leq y \leq b, \quad (3.66) \\
C_5 \cos(k_x x + a_1) \exp[\gamma_y (y + b)], & -a \leq x \leq a, \quad y \leq -b \\
C_6 \exp[\gamma_x (x + a)] \cos(k_y y + a_2), & x \leq -a, \quad -b \leq y \leq b
\end{cases}$$

Upon substituting the above functions into the wave equation (3.65), we obtain the following relations for the various wave numbers:
\[ k^2_x + k^2_y + \beta^2 = \kappa_0^2 n_1^2, \quad (3.67) \]
\[ k^2_x - \gamma^2_y + \beta^2 = \kappa_0^2 n_2^2, \quad (3.68) \]
\[ \gamma^2_x + k^2_y + \beta^2 = \kappa_0^2 n_2^2. \quad (3.69) \]

Subtracting Eq. (3.67) from Eq. (3.69), and Eq. (3.67) from Eq. (3.68), we arrive at the following expressions for the two decay constants:

\[ \gamma_x^2 = k_0^2 (n_2^2 - n_3^2) - k_x^2, \quad (3.70) \]
\[ \gamma_y^2 = k_0^2 (n_1^2 - n_2^2) - k_y^2. \quad (3.71) \]

The next step in the presented methodology is to impose the boundary conditions, specified by Eqs. (3.40) and (3.41), on the tangential \( E_x \) and \( H_z \) field components. By doing so at the \( y = b \) boundary, we obtain the following characteristic equation:

\[ k_y b + \alpha_z = \arctan \left( \left| \frac{\gamma_y}{k_y} \right| \right) + m_1 \pi, \quad (m_1 = 0, 1 \ldots). \quad (3.72) \]

In a similar vein, by imposing Eqs. (3.40) and (3.41) on \( E_x \) and \( H_z \) at the \( y = -b \) dielectric interface, we get:

\[ k_y b - \alpha_z = \arctan \left( \left| \frac{\gamma_y}{k_y} \right| \right) + m_2 \pi, \quad (m_2 = 0, 1 \ldots). \quad (3.73) \]

By adding the two last equations, one arrives at:

\[ k_y b = \arctan \left( \left| \frac{\gamma_y}{k_y} \right| \right) + \frac{1}{2} (m - 1) \pi, \quad (m = 1, 2 \ldots). \quad (3.74) \]

It, now, remains to implement the same conditions at the remaining two boundaries, namely \( x = a \) and \( x = -a \). By doing so at the first one, we obtain after some mathematical manipulations:
\[ k_x a + \alpha_1 = \arctan \left( \frac{(k_0^2 n_1^2 - k_y^2) \gamma_x}{(k_0^2 n_2^2 - k_y^2) k_x} \right) + n_1 \pi, \quad (n_1 = 0, 1 \ldots). \]  

(3.75)

On the other hand, from the boundary \( x = -a \), we have:

\[ k_x a - \alpha_1 = \arctan \left( \frac{(k_0^2 n_1^2 - k_y^2) \gamma_x}{(k_0^2 n_2^2 - k_y^2) k_x} \right) + n_2 \pi, \quad (n_2 = 0, 1 \ldots). \]  

(3.76)

Thus, upon adding Eqs. (3.75) and (3.76), we finally get:

\[ k_x a = \arctan \left( \frac{(k_0^2 n_1^2 - k_y^2) \gamma_x}{(k_0^2 n_2^2 - k_y^2) k_x} \right) + \frac{1}{2} (n-1) \pi, \quad (n=1, 2 \ldots). \]  

(3.77)

Since for most practical cases it is: \( k_0 n_{1,3} \gg k_y \), Eq. (3.77) further simplifies to:

\[ k_x a = \arctan \left( \frac{e_{rl} \gamma_x}{e_{r2} k_x} \right) + \frac{1}{2} (n-1) \pi, \quad (n=1, 2 \ldots). \]  

(3.78)

We may, now, summarize Marcatili's methodology for calculating the longitudinal propagation constant \( \beta \) of an eigenmode supported by a rectangular waveguide, as follows: First, we determine the value of \( k_x \) by numerically solving Eq. (3.78). Similarly, we obtain \( k_y \) by making use of Eqs. (3.71) and (3.74). Finally, we immediately obtain the value of \( \beta \) from Eq. (3.67).

It is noteworthy that while Eq. (3.74) corresponds to the characteristic equation of a TE mode in a three-layer slab waveguide parallel to the \( x \) axis, Eq. (3.78) corresponds to the characteristic equation of a TM mode in a three-layer slab waveguide parallel to the \( y \) axis. The situation is reminiscent of that described in the last paragraph of section 3.5.

Following closely a dual analysis, one may yet obtain analogous results for the other mode, \( E_{pq}^\nu \), in which the magnetic field is assumed to be polarized in the \( x \) direction, resulting in \( H_y = 0 \).
3.7 Finite-Difference Frequency-Domain Mode-Solvers

The introduction of semi-vectorial finite-difference (SV-FD) mode-solving techniques by Stern [70]-[72] led to the development of numerically efficient waveguide analysis methods, which take polarization into consideration and providing fairly accurate results. Therefore, they are widely used in the computer-aided design (CAD) analyses of dielectric waveguiding structures with (rectangular) 2-D cross-section. Nonetheless, for increased accuracy, particularly when high-index contrast or circular cross-sectional structures (e.g. fibres) are considered, one inevitably has to resort to fully-vectorial computations, which take field-coupling into account as well.

In this section we will start by deriving the vectorial wave equations for each transverse electromagnetic field component. As a next step, the so-called semi-vectorial wave equations are obtained by ignoring the coupling terms. We shall proceed by formulating the finite-difference approximations for both types of wave equations. At this point we describe the structure of the resultant sparse matrix and highlight the procedure for solving the accompanying eigenvalue matrix equation [58]. In the last part of the section, exemplary results will be presented and compared with published work.

3.7.1 Vectorial Wave Equations for the E-Field Components

As we have already shown, the vectorial equation for the electric field \( \mathbf{E} \) is (Eq. (3.26)):

\[
\nabla^2 \mathbf{E} + \nabla \left( \frac{\nabla \varepsilon_r}{\varepsilon_r} \cdot \mathbf{E} \right) + k_0^2 \varepsilon_r \mathbf{E} = 0. \tag{3.79}
\]

Let us consider a structure uniform in the \( z \) direction. In this case, the derivative of the relative permittivity with respect to \( z \) will be zero:

\[
\frac{\partial \varepsilon_r}{\partial z} = 0. \tag{3.80}
\]
Thus, the second term in Eq. (3.79) can be written as:

$$\nabla \left( \frac{\nabla \varepsilon_r(\varepsilon, E)}{\varepsilon_r} \cdot \mathbf{E} \right) = \nabla \left( \frac{1}{\varepsilon_r} \frac{\partial \varepsilon_r}{\partial x} E_x + \frac{1}{\varepsilon_r} \frac{\partial \varepsilon_r}{\partial y} E_y \right). \quad (3.81)$$

Upon substituting Eq. (3.81) into Eq. (3.79), we separate Eq. (3.79) into the $x$ and $y$ components. Thus, we obtain the following vectorial wave equations for the $E_x$ and $E_y$ field components:

$$\frac{\partial^2 E_x}{\partial x^2} + \frac{\partial}{\partial x} \left( \frac{1}{\varepsilon_r} \frac{\partial \varepsilon_r}{\partial x} E_x \right) + \frac{\partial^2 E_x}{\partial y^2} + \frac{\partial^2 E_x}{\partial z^2} + k_0^2 \varepsilon_r E_x + \frac{\partial}{\partial x} \left( \frac{1}{\varepsilon_r} \frac{\partial \varepsilon_r}{\partial y} E_y \right) = 0, \quad (3.82)$$

$$\frac{\partial^2 E_y}{\partial x^2} + \frac{\partial^2 E_y}{\partial y^2} + \frac{\partial}{\partial y} \left( \frac{1}{\varepsilon_r} \frac{\partial \varepsilon_r}{\partial y} E_y \right) + \frac{\partial^2 E_y}{\partial z^2} + k_0^2 \varepsilon_r E_y + \frac{\partial}{\partial y} \left( \frac{1}{\varepsilon_r} \frac{\partial \varepsilon_r}{\partial x} E_x \right) = 0. \quad (3.83)$$

Furthermore, we note that:

$$\frac{\partial}{\partial x} \left( \frac{1}{\varepsilon_r} \frac{\partial \varepsilon_r}{\partial x} (\varepsilon, E_x) \right) = \frac{\partial^2 E_x}{\partial x^2} + \frac{\partial}{\partial x} \left( \frac{1}{\varepsilon_r} \frac{\partial \varepsilon_r}{\partial x} E_x \right), \quad (3.84)$$

and:

$$\frac{\partial}{\partial y} \left( \frac{1}{\varepsilon_r} \frac{\partial \varepsilon_r}{\partial y} (\varepsilon, E_y) \right) = \frac{\partial^2 E_y}{\partial y^2} + \frac{\partial}{\partial y} \left( \frac{1}{\varepsilon_r} \frac{\partial \varepsilon_r}{\partial y} E_y \right). \quad (3.85)$$

Considering also Eq. (3.36), we may hence rewrite Eqs. (3.82) and (3.83) in the following, compact, form:

$$\frac{\partial}{\partial x} \left( \frac{1}{\varepsilon_r} \frac{\partial \varepsilon_r}{\partial x} (\varepsilon, E_x) \right) + \frac{\partial^3 E_x}{\partial x^3} + (k_0^2 \varepsilon_r - \beta^2) E_x + \frac{\partial}{\partial x} \left( \frac{1}{\varepsilon_r} \frac{\partial \varepsilon_r}{\partial y} E_y \right) = 0, \quad (3.86)$$

$$\frac{\partial^2 E_x}{\partial x^2} + \frac{\partial}{\partial y} \left( \frac{1}{\varepsilon_r} \frac{\partial \varepsilon_r}{\partial y} (\varepsilon, E_y) \right) + (k_0^2 \varepsilon_r - \beta^2) E_y + \frac{\partial}{\partial y} \left( \frac{1}{\varepsilon_r} \frac{\partial \varepsilon_r}{\partial x} E_x \right) = 0. \quad (3.87)$$
The last term in each of Eqs. (3.86) and (3.87) corresponds to the coupling between the \( x \)-directed electric field component, \( E_x \), and the \( y \)-directed electric field component, \( E_y \).

### 3.7.2 Vectorial Wave Equations for the H-Field Components

Let us next derive the corresponding components of the vectorial wave equation for the magnetic field \( \mathbf{H} \). As it has already been shown, the vectorial equation for the magnetic field, is (Eq. (3.33)):

\[
\nabla^2 \mathbf{H} + \frac{\nabla \varepsilon_r}{\varepsilon_r} \times (\nabla \times \mathbf{H}) + k_0^2 \varepsilon_r \mathbf{H} = \mathbf{0}. \tag{3.88}
\]

Here, the second term of Eq. (3.88) is investigated in detail. To this end, we recall that the considered structures are uniform in the \( z \) direction and that Eq. (3.80) holds. Hence, if \( \mathbf{i}, \mathbf{j}, \mathbf{k} \) are, respectively, the unit vectors in the \( x, y \) and \( z \) directions, we may obtain:

\[
\nabla \varepsilon_r \times (\nabla \times \mathbf{H}) = \begin{vmatrix}
\mathbf{i} & \mathbf{j} & \mathbf{k} \\
\frac{\partial \varepsilon_r}{\partial x} & \frac{\partial \varepsilon_r}{\partial y} & 0 \\
(\nabla \times \mathbf{H})_x & (\nabla \times \mathbf{H})_y & (\nabla \times \mathbf{H})_z
\end{vmatrix} = \frac{\partial \varepsilon_r}{\partial y} (\nabla \times \mathbf{H})_x \mathbf{i} - \frac{\partial \varepsilon_r}{\partial x} (\nabla \times \mathbf{H})_y \mathbf{j} + \left\{ \frac{\partial \varepsilon_r}{\partial x} (\nabla \times \mathbf{H})_y - \frac{\partial \varepsilon_r}{\partial y} (\nabla \times \mathbf{H})_x \right\} \mathbf{k}, \tag{3.89}
\]

where we have used the following expressions:

\[
(\nabla \times \mathbf{H})_x = \frac{\partial H_y}{\partial y} - \frac{\partial H_z}{\partial z}, \tag{3.90}
\]

\[
(\nabla \times \mathbf{H})_y = \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x}. \tag{3.91}
\]
\((\nabla \times \mathbf{H})_x = \frac{\partial H_y}{\partial y} - \frac{\partial H_x}{\partial x}\). \hspace{1cm} (3.92)

Substitution of Eqs. (3.90) - (3.92) into Eq. (3.89), results in:

\[
\nabla \varepsilon \times (\nabla \times \mathbf{H}) = \frac{\partial \varepsilon_r}{\partial y} \left( \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} \right) i - \frac{\partial \varepsilon_r}{\partial x} \left( \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} \right) j
\]

\[
+ \left\{ \frac{\partial \varepsilon_r}{\partial z} \left( \frac{\partial H_z}{\partial x} - \frac{\partial H_x}{\partial z} \right) - \frac{\partial \varepsilon_r}{\partial y} \left( \frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} \right) \right\} k.
\hspace{1cm} (3.93)

Substituting Eq. (3.93) into Eq. (3.88) and separating the result into the \(x\) and \(y\) components, one obtains the following vectorial wave equations for each magnetic field component:

\[
\frac{\partial^2 H_x}{\partial x^2} + \frac{\partial^2 H_x}{\partial y^2} - \frac{1}{\varepsilon_r} \frac{\partial}{\partial y} \left( \frac{\partial H_x}{\partial y} - \frac{\partial H_y}{\partial x} \right) + \frac{k_0^2 \varepsilon_r}{\varepsilon_r} H_x + \frac{1}{\varepsilon_r} \frac{\partial H_y}{\partial x} = 0,
\hspace{1cm} (3.94)
\]

\[
\frac{\partial^2 H_y}{\partial x^2} - \frac{1}{\varepsilon_r} \frac{\partial}{\partial x} \left( \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} \right) + \frac{k_0^2 \varepsilon_r}{\varepsilon_r} H_x + \frac{1}{\varepsilon_r} \frac{\partial H_x}{\partial y} = 0.
\hspace{1cm} (3.95)
\]

Furthermore, because it is:

\[
\varepsilon_r \frac{\partial}{\partial y} \left( \frac{1}{\varepsilon_r} \frac{\partial H_x}{\partial y} \right) = \frac{\partial^2 H_x}{\partial y^2} - \frac{1}{\varepsilon_r} \frac{\partial}{\partial y} \left( \frac{\partial H_x}{\partial y} - \frac{\partial H_y}{\partial x} \right),
\hspace{1cm} (3.96)
\]

and:

\[
\varepsilon_r \frac{\partial}{\partial x} \left( \frac{1}{\varepsilon_r} \frac{\partial H_y}{\partial x} \right) = \frac{\partial^2 H_y}{\partial x^2} - \frac{1}{\varepsilon_r} \frac{\partial}{\partial x} \left( \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} \right),
\hspace{1cm} (3.97)
\]

we may rewrite Eqs. (3.94) and (3.95) in a more compact form, as follows:

\[
\frac{\partial^2 H_x}{\partial x^2} + \varepsilon_r \frac{\partial}{\partial y} \left( \frac{1}{\varepsilon_r} \frac{\partial H_x}{\partial y} \right) + (k_0^2 \varepsilon_r - \beta^2) H_x + \frac{1}{\varepsilon_r} \frac{\partial H_y}{\partial x} = 0,
\hspace{1cm} (3.98)
\]

...
where we have made use of Eq. (3.36).

Likewise what we mentioned in the corresponding equations for the electric field $E$, the last term in each of Eqs. (3.98) and (3.99) corresponds to the coupling between the $x$-directed magnetic field component, $H_x$, and the $y$-directed magnetic field component, $H_y$.

### 3.7.3 Semi-Vectorial Wave Equations

In the propagation equations of electromagnetic fields inside dielectric waveguides, the terms corresponding to the interaction between the $x$-directed electric field component, $E_x$, and the $y$-directed electric field component, $E_y$, i.e.:

\[
\frac{\partial}{\partial x} \left( \frac{1}{\varepsilon_r} \frac{\partial H_y}{\partial y} \right) \quad \text{and} \quad \frac{\partial}{\partial y} \left( \frac{1}{\varepsilon_r} \frac{\partial E_x}{\partial x} \right)
\]

as well as the terms corresponding to the interaction between the $x$-directed magnetic field component, $H_x$, and the $y$-directed magnetic field component, $H_y$, i.e.:

\[
\frac{1}{\varepsilon_r} \frac{\partial E_y}{\partial y} \frac{\partial H_y}{\partial x} \quad \text{and} \quad \frac{1}{\varepsilon_r} \frac{\partial E_x}{\partial x} \frac{\partial H_x}{\partial y}
\]

are quite often small. Ignoring these terms that account for the interaction, we can decouple the vectorial wave equations for the $x$- and $y$-directed field components and reduce them to semi-vectorial wave equations, which are considerably easier to solve and preserve the polarization nature of the electromagnetic field, i.e. they take into account the discontinuities of an electric field component normal to a dielectric interface. As shown in Fig. 3.4, the semi-vectorial analyses may be divided into the quasi-TE mode analysis, in which the principal field component is $E_x$ or $H_y$, and the quasi-TM mode analysis, in which the principal field component is $E_y$ or $H_x$. 

31
According to what was mentioned above, the semi-vectorial wave equations in the electric and magnetic field representation, respectively, for the quasi-TE mode are the following:

\[
\frac{\partial}{\partial x} \left\{ \frac{1}{\epsilon_r} \frac{\partial}{\partial x} (\epsilon_r E_x) \right\} + \frac{\partial^2 E_x}{\partial y^2} + (k_0^2 \epsilon_r - \beta^2) E_x = 0, \quad (3.100)
\]

\[
\epsilon_r \frac{\partial}{\partial x} \left( \frac{1}{\epsilon_r} \frac{\partial H_x}{\partial x} \right) + \frac{\partial^2 H_x}{\partial y^2} + (k_0^2 \epsilon_r - \beta^2) H_x = 0. \quad (3.101)
\]

Likewise, the corresponding semi-vectorial wave equations in the electric and magnetic field representation, respectively, for the quasi-TM mode are:

\[
\frac{\partial^2 E_y}{\partial x^2} + \frac{\partial}{\partial y} \left\{ \frac{1}{\epsilon_r} \frac{\partial}{\partial y} (\epsilon_r E_y) \right\} + (k_0^2 \epsilon_r - \beta^2) E_y = 0, \quad (3.102)
\]

\[
\frac{\partial^2 H_y}{\partial x^2} + \epsilon_r \frac{\partial}{\partial y} \left( \frac{1}{\epsilon_r} \frac{\partial H_y}{\partial y} \right) + (k_0^2 \epsilon_r - \beta^2) H_y = 0. \quad (3.103)
\]
One may notice at this point that in Eqs. (3.100) – (3.103), the derivatives of the relative permittivity, \( e_r \), with respect to the \( x \) and \( y \) coordinates are taken into consideration. If we, further, assume that these derivatives are zero, i.e.:

\[
\frac{\partial e_r}{\partial x} = 0 \quad \text{and} \quad \frac{\partial e_r}{\partial y} = 0,
\]

the semi-vectorial wave equations can be further reduced to the scalar wave equation:

\[
\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} + (k_0^2 e_r - \beta^2) \phi = 0,
\]

where \( \phi = E \) or \( H \), is a wave function that designates a scalar field component.

### 3.7.4 Finite-Difference Discretization

Let us rewrite Eqs. (3.86) – (3.87) in the following, compact and more convenient, matrix form:

\[
\begin{pmatrix}
A_{xx} & A_{xy} \\
A_{yx} & A_{yy}
\end{pmatrix}
\begin{pmatrix}
E_x \\
E_y
\end{pmatrix} = \beta^2
\begin{pmatrix}
E_x \\
E_y
\end{pmatrix}
\]

where the differential operators \( A_j (i = x, y \text{ and } j = x, y) \) are defined as follows [58]:

\[
A_{xx} E_x = \frac{\partial}{\partial x} \left[ \frac{1}{n^2} \frac{\partial(n^2 E_x)}{\partial x} \right] + \frac{\partial^2 E_x}{\partial y^2} + n^2 k_0^2 E_x,
\]

\[
A_{yy} E_y = \frac{\partial^2 E_y}{\partial x^2} + \frac{\partial}{\partial y} \left[ \frac{1}{n^2} \frac{\partial(n^2 E_y)}{\partial y} \right] + n^2 k_0^2 E_y,
\]

\[
A_{xy} E_y = \frac{\partial}{\partial x} \left[ \frac{1}{n^2} \frac{\partial(n^2 E_y)}{\partial y} \right] - \frac{\partial^2 E_y}{\partial x \partial y},
\]

\[
A_{yx} E_x = \frac{\partial}{\partial y} \left[ \frac{1}{n^2} \frac{\partial(n^2 E_x)}{\partial x} \right] - \frac{\partial^2 E_x}{\partial y \partial x}.
\]
Similar formulations may be applied for the magnetic field, following Eqs. (3.98) – (3.99).

We may efficiently solve Eq. (3.106) by means of a finite-difference mode-solving technique [58]. In the finite-difference methodology [73], the continuous space is replaced by a discrete lattice structure defined in the computational domain. The fields at the lattice point \( x = m\Delta x \) and \( y = n\Delta y \) are represented by their discrete counterparts. The differential operators appearing in Eqs. (3.107) – (3.110) are also approximated by finite-difference approximations. The discrete form of these operators can, in fact, be found in a straightforward way by noting that at the dielectric interfaces, although the normal electric field components are discontinuous, the displacement vectors \( n^2\hat{E}_i \ (i = x, y) \) appearing in Eqs. (3.107) – (3.110) are continuous. As a result, a central finite-difference scheme can be applied directly, without invoking any special treatments.

After some algebraic manipulations, the expressions for the differential operators \( A_{xy} \ (i = x, y \text{ and } j = x, y) \) are found to be the following [58], [74]:

\[
A_{xx}\hat{E}_x = \frac{T_{m+1,n}\hat{E}_x(m+1,n) - (2 - R_{m+1,n} - R_{m-1,n})\hat{E}_x(m,n) + T_{m-1,n}\hat{E}_x(m-1,n)}{\Delta x^2} + \frac{\hat{E}_x(m,n+1) - 2\hat{E}_x(m,n) + \hat{E}_x(m,n-1)}{\Delta y^2} + n^2(m,n,l)k_0^2\hat{E}_x(m,n)
\]

\[
A_{yy}\hat{E}_y = \frac{T_{m,n+1}\hat{E}_y(m,n+1) - (2 - R_{m,n+1} - R_{m,n-1})\hat{E}_y(m,n) + T_{m,n-1}\hat{E}_y(m,n-1)}{\Delta y^2} + \frac{\hat{E}_y(m+1,n) - 2\hat{E}_y(m,n) + \hat{E}_y(m-1,n)}{\Delta x^2} + n^2(m,n,l)k_0^2\hat{E}_y(m,n)
\]

(3.111)

where:

\[
T_{m\pm1,n} = \frac{2n^2(m\pm1,n)}{n^2(m\pm1,n) + n^2(m,n)}, \quad R_{m\pm1,n} = T_{m\pm1,n} - 1
\]

(3.112), (3.113)

are, respectively, the transmission and reflection coefficients across the index interfaces between \( m\Delta x \) and \( (m+1)\Delta x \). For the \( A_{xy} \) operator, we have:

\[
A_{xy}\hat{E}_y = \frac{T_{m,n+1}\hat{E}_y(m,n+1) - (2 - R_{m,n+1} - R_{m,n-1})\hat{E}_y(m,n) + T_{m,n-1}\hat{E}_y(m,n-1)}{\Delta y^2} + \frac{\hat{E}_y(m+1,n) - 2\hat{E}_y(m,n) + \hat{E}_y(m-1,n)}{\Delta x^2} + n^2(m,n,l)k_0^2\hat{E}_y(m,n)
\]

(3.114)
where:

\[ T_{m,n±1} = \frac{2n^2(m,n±1)}{n^2(m,n±1) + n^2(m,n)} \], \quad (3.115) 

\[ R_{m,n±1} = T_{m,n±1} - 1, \quad (3.116) \]

are, respectively, the transmission and reflection coefficients across the index interfaces between \( n\Delta y \) and \((n+1)\Delta y \). In a similar vein, the remaining two operators, \( A_{xy} \) and \( A_{yx} \), are expressed by:

\[
\begin{align*}
A_{xy}\hat{E}_y &= \frac{1}{4\Delta x\Delta y} \left[ \left[ \frac{n^2(m+1,n+1)}{n^2(m+1,n)} - 1 \right] \cdot \hat{E}_y(m+1,n+1) \\
&- \left[ \frac{n^2(m+1,n-1)}{n^2(m+1,n)} - 1 \right] \cdot \hat{E}_y(m+1,n-1) \\
&- \left[ \frac{n^2(m-1,n+1)}{n^2(m-1,n)} - 1 \right] \cdot \hat{E}_y(m-1,n+1) \\
&+ \left[ \frac{n^2(m-1,n-1)}{n^2(m-1,n)} - 1 \right] \cdot \hat{E}_y(m-1,n-1) \right], \\
A_{yx}\hat{E}_x &= \frac{1}{4\Delta x\Delta y} \left[ \left[ \frac{n^2(m+1,n+1)}{n^2(m,n+1)} - 1 \right] \cdot \hat{E}_x(m+1,n+1) \\
&- \left[ \frac{n^2(m-1,n+1)}{n^2(m,n-1)} - 1 \right] \cdot \hat{E}_x(m-1,n+1) \\
&- \left[ \frac{n^2(m+1,n-1)}{n^2(m,n+1)} - 1 \right] \cdot \hat{E}_x(m+1,n-1) \\
&+ \left[ \frac{n^2(m-1,n-1)}{n^2(m,n-1)} - 1 \right] \cdot \hat{E}_x(m-1,n-1) \right].
\end{align*}
\]

Calculating Eqs. (3.111) - (3.118) for each node in the discrete computational space, results in the following eigenvalue matrix equation:
\[ [A] \{ \phi \} = \beta^2 \{ \phi \}, \]  

(3.119)

where \( \beta^2 \) is an eigenvalue and \( \{ \phi \} \) is an eigenvector, expressed as:

\[ \{ \phi \} = (\phi_1 \quad \phi_2 \quad \phi_3 \quad \ldots \quad \phi_M)^T, \]  

(3.120)

with \( M = m \times n \), of the square sparse matrix \( [A] \).

We can calculate the longitudinal propagation constant \( \beta \), as well as the field distribution, by solving the eigenvalue matrix equation (3.119). In the presented, fully-vectorial finite-difference frequency-domain (FV-FDFD) mode-solving technique, the sparse matrix is non-symmetric, having only 13 diagonals. Its bandwidth and dimension are \((4m + 7)\) and \(2mn\), respectively. Once matrix \( [A] \) is defined, the eigenvalues and eigenvectors of its nonzero band can be found by means of the shifted-inverse power iteration method. This method is, further, commonly deployed for the solution of various eigenproblems in commercially available software.

### 3.7.5 Numerical Examples of the FV-FDFD Mode-Solver

The finite-difference expressions presented in this chapter are widely used in computer-aided (CAD) software available for the investigation of arbitrary dielectric heterostructures having 2-D cross sections. Here, we shall concisely discuss exemplary results obtained with the previously described FV-FDFD mode-solver. In all cases, for simplicity, a Dirichlet condition was implemented at the boundaries of our numerical space, except from the symmetry-plane, where we made use of a Neumann condition in order to halve the computational overhead.

Figure 3.5 illustrates a rib waveguide structure, also considered by the COST-216 group using 12 different methods [75], and by Xu et al. in Ref. [58]. The refractive indices for a wavelength 1.55 \( \mu \text{m} \) are also indicated in this figure. First, we examined the convergence of our mode-solving technique for several InP thicknesses, \( t \), by progressively reducing the mesh size, while keeping the mesh size ratio \( \Delta x/\Delta y \) fixed at the value of 2. Figure 3.6 clearly illustrates that the mode-solver converges to,
evidently, more accurate values of the calculated normalized propagation constant, $b$ [70]-[72], in all three cases. Very similar results were also reported in [58].

Figure 3.7 illustrates the calculated transverse electric field components of the first TE-like mode for $t = 0.2 \, \mu m$. The refractive index distribution that was considered in our calculations is, also, shown in this figure. The mesh sizes utilized in these investigations were $\Delta x = 0.05 \, \mu m$ and $\Delta y = 0.025 \, \mu m$. The results in all cases were obtained in approximately 11 seconds on an Inter® Pentium 4® 2.80GHz, personal workstation. One may readily notice that in this case, the mode is rather weakly-guided, since it is not well-confined below the central rib. Moreover, we note that the weak transverse electric field component is mainly concentrated at the dielectric

Figure 3.6: Convergence of the developed finite-difference frequency-domain mode solver for various thicknesses $t$ of the InP layer shown in Fig. 3.5.
Figure 3.7: (a) Refractive index 2-D distribution used in the FV-FDFD mode-solver. Computed (b) dominant $E_x$, and (c) weak $E_y$ field component.
corners, as expected, somewhat resembling a surface wave. The obtained normalized propagation constant is, indeed, considerably small (approximately 0.1336) and compares very favourably with the value 0.1365 calculated in [58] for the same mesh sizes. The minor discrepancy is attributed to the finite, approximately 2 μm rather than infinite, as used in [58], height of the central rib that we selected for our present calculations.

In Fig. 3.8 we report the 2-D spatial distribution of the dominant, $E_x$, electric field component of the fundamental TE-like eigenmode, for $t = 0.4$ and 0.0 μm. For the first case, it is seen that the mode is somewhat better guided and confined, compared to the previous case, as judged by observing the field immediately below the dielectric corners. The discontinuity of the dominant electric component at the two interfaces of the rib with vacuum, is clearly observed in Fig. 3.8. The computed normalized constant is $b = 0.1507$ (Fig. 3.8(a)), which is again found to be in excellent agreement with the corresponding value of 0.1511, reported in Ref. [58]. Very similar values were also found in [75]. Finally, for $t = 0.0$ μm, it is readily inferred that the mode exhibits the weakest guidance compared to both previous cases, since it considerably extends to both sides of the InGaAsP layer, while having

![Figure 3.8: Computed two-dimensional spatial distribution of the $E_x$ field component for (a) $t = 0.4$ μm and (b) $t = 0.0$ μm, for the structure shown in Fig. 3.5.](image-url)
a normalized propagation constant equal to just 0.1215 (the corresponding value in [58] is 0.1228).
4 Conventional and Nonstandard Finite-Difference Time-Domain Method

In the analytic investigations of dielectric waveguides that we have been concerned with so far, we have assumed that the structures were axially uniform (see Eq. (3.36)). Moreover, our analysis involved steady-state solutions of the wave equations, in which the time-derivative $\partial/\partial t$ was replaced by $j\omega$ (see Eqs. (3.13)-(3.15)). For the investigation of transient phenomena, axially varying, irregularly shaped or other types of more involved (e.g. nonlinear) structures, one has to resort to full 3-D, time-domain methodologies. A rigorous and well-established such algorithm, based directly on Maxwell’s equations, is the finite-difference time-domain (FDTD) method [59]-[61]. Upon establishing the necessary nomenclature, we shall here examine in detail several aspects of this algorithm, including its numerical stability and the induced numerical dispersion and anisotropy errors that occur within it, which occasionally can contaminate the simulation outcomes and seriously compromise the overall accuracy of the method. As a means of subduing the aforementioned algorithmic errors we also examine the recently introduced concepts of “exact” and “non-standard” finite-differences [62], their implementation within the context of the FDTD algorithm [63] and the significantly enhanced computational performance, in terms of memory and/or computational time, which can arise from their adoption.

4.1 The Yee Algorithm

As we show in section 3.1, Maxwell’s equations of electromagnetism for a homogeneous, isotropic, lossless, non-dispersive, linear, achiral and stationary medium take the form:

$$-\mu_0 \frac{\partial \mathbf{H}}{\partial t} = \nabla \times \mathbf{E}, \quad (4.1)$$

$$\varepsilon_0 \varepsilon_r \frac{\partial \mathbf{E}}{\partial t} = \nabla \times \mathbf{H}. \quad (4.2)$$
We can rewrite these equations in terms of the six electromagnetic field components, as follows:

\[ -\mu_0 \frac{\partial H_z}{\partial t} = \frac{\partial E_y}{\partial y} - \frac{\partial E_z}{\partial z}, \]  
\[ -\mu_0 \frac{\partial H_y}{\partial t} = \frac{\partial E_z}{\partial z} - \frac{\partial E_x}{\partial x}, \]  
\[ -\mu_0 \frac{\partial H_z}{\partial t} = \frac{\partial E_y}{\partial y} - \frac{\partial E_z}{\partial z}, \]  
\[ \varepsilon_0 \varepsilon_r \frac{\partial E_z}{\partial t} = \frac{\partial H_x}{\partial x} - \frac{\partial H_z}{\partial z}, \]  
\[ \varepsilon_0 \varepsilon_r \frac{\partial E_y}{\partial t} = \frac{\partial H_x}{\partial x} - \frac{\partial H_y}{\partial y}, \]  
\[ \varepsilon_0 \varepsilon_r \frac{\partial E_z}{\partial t} = \frac{\partial H_y}{\partial y} - \frac{\partial H_z}{\partial z}. \]

(4.3)  
(4.4)  
(4.5)  
(4.6)  
(4.7)  
(4.8)

In order to discretise these equations using finite differences, we assume that the electromagnetic field components are located in three dimensions according to the

--- (red) Electric-field component
--- (blue) Magnetic-field component

Figure 4.1: Three-dimensional spatial arrangement of the six electromagnetic field components ("Yee cube") for the finite-difference time-domain (FDTD) numerical solution of Maxwell's equations.
arrangement shown in Fig. 4.1, known also as the “Yee cube” after the name of the scientist who first proposed the present algorithm [59]. Further, we assume that the electric field components are calculated at integer time increments $\Delta t$, i.e. at $n\Delta t$ ($n = 0, 1, 2 \ldots$), whereas the magnetic field components are calculated at half-integer time intervals, $(n + 1/2)\Delta t$. With these assumptions in place, we may self-consistently discretise each one of Eqs. (4.3)–(4.8). For instance, the left-hand side of Eq. (4.3) becomes:

$$\frac{-\mu_0}{\Delta t}[H_x^{n+1/2}(i, j + 1/2, k + 1/2) - H_x^{n-1/2}(i, j + 1/2, k + 1/2)],$$  \hfill (4.9)

whereas for the right-hand side one obtains:

$$\frac{1}{\Delta y}[E_x^n(i, j + 1, k + 1/2) - E_x^n(i, j, k + 1/2)] - \frac{1}{\Delta z}[E_y^n(i, j + 1/2, k + 1) - E_y^n(i, j + 1/2, k)].$$  \hfill (4.10)

We, thus, arrive at the following updating equation for the $H_x$ field component:

$$H_x^{n+1/2}(i, j + 1/2, k + 1/2) = H_x^{n-1/2}(i, j + 1/2, k + 1/2)$$
$$- \frac{\Delta t}{\mu_0} \left\{ \frac{1}{\Delta y}[E_x^n(i, j + 1, k + 1/2) - E_x^n(i, j, k + 1/2)] \right\}$$
$$- \frac{1}{\Delta z}[E_y^n(i, j + 1/2, k + 1) - E_y^n(i, j + 1/2, k)].$$ \hfill (4.11)

Similar updating equations are obtained for the remaining five electromagnetic field components. We, hence, have an algorithm for calculating all components, at every discrete point and at each time step. Starting from the electric field components, which at $n = 0$ are assumed to be zero throughout the computational space, we calculate the magnetic field components (e.g., for the new value of the $H_x$ component, we use Eq. (4.11)); then, based on the new values of the $H$-field, we find the new values of the $E$-field components, and so on. This process enables the calculation of the full time evolution of the electromagnetic fields in the defined computational space. We note that, in contrast to the mode-solving techniques analysed in the
previous chapter (see Eq. (3.106)), the FDTD algorithm performs a direct solution of Maxwell’s equations, i.e. it is not based on the wave equations and it does not involve complex computations of eigenvalues/vectors of sparse matrices.

### 4.2 Stability of the Yee Algorithm

A point of attention in the correct implementation of the FDTD algorithm and, in fact, of every explicit algorithm is that the time step, $\Delta t$, cannot be arbitrarily large; it has to be upper-bounded by an explicit figure, which depends solely on the spatial discretization steps and the defined dielectric structure; otherwise the algorithm will become unstable (i.e., the computed fields will unphysically start increasing with time, tending to infinity). In this section we show how one may analytically calculate the aforementioned upper limit of the time step, known also as the “Courant-Friedrichs-Lewy (CFL) limit” [61], [73].

The algebraic calculations are greatly simplified if we consider the (equivalent) scalar wave equation. For the sake of clarity we shall examine the 1D case. We, thus, start with the following equation:

$$\frac{\partial^2 \varphi}{\partial x^2} - \varepsilon \frac{\partial^2 \varphi}{\partial t^2} = 0,$$

where $\varphi$ is an electric or magnetic field component. Assuming without losing generality that the oscillation amplitude of the component is unity and that the propagation constant along the $x$-direction is $\beta_n$, it is a simple matter to see that the first part (a) of the following expression for $\varphi$ satisfies Eq. (4.12):

$$\varphi(x, t) = \exp(j \beta_n x) \exp(at) = \exp(j \beta_n l \Delta x) \exp(an\Delta t) = \exp(j \beta_n l \Delta x) \xi^n,$$

where $\xi = \exp(an\Delta t)$, $l$, $n = 0$, 1, 2 ..., and $\Delta x$ and $\Delta t$ being the spatial and temporal increments, respectively.

We note from Eq. (4.13) that if the solution $\varphi$ is to be stable, that is not growing unphysically with time, then the variable $\xi$ must satisfy the restriction:
We mentioned above that Eq. (4.13a) satisfies Eq. (4.12). Obviously, the same must hold true for Eq. (4.13c). Upon substituting it into Eq. (4.12) we, thus, obtain:

\[
\frac{1}{\Delta x^2} \{ \exp[j \beta_x (l + 1)\Delta x] \xi^n - 2 \exp(j \beta_x l\Delta x) \xi^n + \exp[j \beta_x (l - 1)\Delta x] \xi^n \} 
- \frac{\epsilon \mu}{\Delta t^2} \{ \exp(j \beta_x l\Delta x) \xi^{n+1} - 2 \exp(j \beta_x l\Delta x) \xi^n + \exp(j \beta_x l\Delta x) \xi^{n-1} \} = 0.
\]

(4.15)

Upon eliminating the common \( \exp(j \beta_x l\Delta x) \xi^n \) term and after performing some further algebraic manipulations we arrive at:

\[
\xi^2 - 2\xi + 1 - \frac{\Delta t^2}{\epsilon \mu} \left[ -\frac{4}{\Delta x^2} \sin^2 \left( \frac{\beta_x \Delta x}{2} \right) \right] \xi = 0,
\]

(4.16)

which can be written as:

\[
\xi^2 - 2U\xi + 1 = 0, \quad (4.17)
\]

with \( U \) being:

\[
U = -\frac{2 \Delta t^2}{\epsilon \mu} \frac{1}{\Delta x^2} \sin^2 \left( \frac{\beta_x \Delta x}{2} \right) + 1.
\]

(4.18)

From the two roots of Eq. (4.17), \( \xi_{1,2} = U \pm \sqrt{U^2 - 1} \), because of the restriction of Eq. (4.14) and the fact that it is always \( \sin^2(\ldots) \geq 0 \), we obtain the following relation:

\[
U = -\frac{2 \Delta t^2}{\epsilon \mu} \frac{1}{\Delta x^2} \sin^2 \left( \frac{\beta_x \Delta x}{2} \right) + 1 \leq 1. \quad (4.19)
\]
We, now, consider the following two cases: First, the case where $U > 1$. Then, we immediately see that $|\xi_2| > 1$, and hence the algorithm will run unstably. The second remaining case, given Eq. (4.19), is that where:

$$-1 \leq U \leq 1.$$  \hfill (4.20)

We may promptly find that in this case, $\xi_{1,2}$ (which are, now, complex numbers) satisfy: $|\xi_1| = |\xi_2| = 1$; hence, the algorithm will be stable.

We conclude that the finite-difference algorithm will run stably only when Eq. (4.20) is satisfied. In the light of Eq. (4.19) we see that the right-hand side of this relation is always satisfied. For its left-hand side we have:

$$-1 \leq -\frac{2\Delta t^2}{\varepsilon \mu} \frac{1}{\Delta x^2} \sin^2 \left( \frac{\beta_x \Delta x}{2} \right) + 1.$$  \hfill (4.21)

It should, therefore, be enough to identify the range of values for $\Delta t$, for which the following relation holds true:

$$1 \geq \frac{2\Delta t^2}{\varepsilon \mu} \frac{1}{\Delta x^2} = 1,$$  \hfill (4.22)

from whence we immediately find:

$$\Delta t \leq \sqrt{\varepsilon \mu} \left( \frac{1}{\Delta x^2} \right)^{-1/2} = \sqrt{\varepsilon \mu} \left( \frac{1}{\Delta x^2} \right)^{-1/2} = \frac{1}{\nu} \left( \frac{1}{\Delta x^2} \right)^{-1/2},$$  \hfill (4.23)

where $\nu$ is the speed of the electromagnetic wave in the medium. This equation defines the upper bound in the range of values that the time step of our one-dimensional finite-difference algorithm is allowed to take. Following a similar course of analysis one finds that in three dimensions the corresponding expression for $\Delta t$ is:

$$\Delta t \leq \frac{1}{\nu} \left( \frac{1}{\Delta x^2} + \frac{1}{\Delta y^2} + \frac{1}{\Delta z^2} \right)^{-1/2},$$  \hfill (4.24)
Although this criterion was derived from a finite-difference discretization of the wave equation, it swiftly turns out (as one may intuitively expect) that exactly the same restriction is obeyed by the time step in the Maxwell-equations' based Yee (FDTD) algorithm.

4.3 ‘Exact’ and ‘Nonstandard’ Finite-Differences

To introduce the concepts of ‘exact’ [76]-[79] and ‘nonstandard’ [62], [80] finite-differences and how these are applied within FDTD’s time-marching algorithm, let us once more start by examining the finite-difference discretization of the one-dimensional wave Eq. (4.12). Following the example of Eqs. (4.9) – (4.10), it is an easy task to see that a second-order time and space discretization 0[A^2, Ax2] of this (second-order) equation leads to:

$$\left(\frac{1}{\Delta t^2} \partial_t^2 - v(x)^2 \frac{1}{\Delta x^2} \partial_x^2\right)\phi(x,t) = 0,$$

(4.25)

where $\partial_t^2 \phi(\ell) = \phi(\ell + \Delta\ell) - 2\phi(\ell) + \phi(\ell - \Delta\ell)$, $\ell = t,x$, with $\phi$ and $v$ denoting the numerical field component and phase velocity, respectively.

Let us, now, see how we can construct a more accurate compared to Eq. (4.25) finite-difference approximation to Eq. (4.12), without directly resorting to higher-order implementations. To this end, we start by considering an exact solution to the wave Eq. (4.12). Such a solution can, e.g., be:

$$\phi(x, t) = ae^{i(kx - \omega t)} + be^{i(kx + \omega t)},$$

(4.26)

where $k = 2\pi/\lambda$ is the wavenumber and $\omega$ the wave angular frequency. Assuming that $\phi$ in this expression is also a solution to the numerical, finite-difference, algorithm, we may calculate the first-order central finite difference $\hat{d}_x e^{ikx}$, where $\hat{d}_x \phi(x) = \phi(x + \Delta x/2) - \phi(x - \Delta x/2)$, as:
\[ d_x e^{ikx} = 2ie^{ikx} \sin(k\Delta x/2). \] (4.27)

We, now, note that for \( \phi(x) = e^{ikx} \), by defining:

\[ d_x \phi(x) = \frac{d_x \phi(x)}{s(\Delta x)}, \] (4.28)

with \( s(k, \Delta x) = 2\sin(k\Delta x/2)/k \), we have exactly: \( d_x \phi(x) = \phi'(x) \), for \( \phi \in \{ \sin(kx), \cos(kx), e^{\pm ikx} \} \). Thus, by using the ‘nonstandard’ stipulation \( s(\Delta x) \), instead of the usual \( \Delta x \), in the denominator of Eq. (4.28), we have obtained an exact finite-difference approximation to \( \phi'(x) \). In order to obtain an ‘exact’ finite-difference algorithm for Eq. (4.12), we only need to make the replacements \( \Delta t \to s(\omega, \Delta t) \) and \( \Delta x \to s(k, \Delta x) \) in Eq. (4.25):

\[ \left[ \frac{1}{4\sin^2(\omega\Delta t/2)/\omega^2} \frac{\partial^2}{\partial t^2} - \frac{1}{4\sin^2(k\Delta x/2)/k^2} \frac{\partial^2}{\partial x^2} \right] \phi(x,t) = 0. \] (4.29)

It should be herein noted that even with these ‘nonstandard’ stipulations used in Eq. (4.29), the algorithm may still lead to numerical errors, even at the ‘targeted’ frequency \( \omega \). This is because in our finite (temporally and spatially) computational domain we can never have a strictly monochromatic wave, i.e. it is not possible to define a source that will contain and excite only a single \( \omega \) frequency and wavelength \( \lambda \). Moreover, the phase velocity of the numerical wave will not be equal to the phase velocity of the actual wave, owing to the presence of numerical dispersion. As a result, both, at the frequency \( \omega \) and in a region around it numerical errors will arise. However, as we shall see later by means of numerical examples, these errors are still significantly smaller than those that arise within the corresponding conventional algorithm.
4.4 Generalization to Three Dimensions

4.4.1 Standard Wave-Equation Finite-Difference Algorithm

We, now, consider how the above concepts can be applied to the wave equation in three dimensions. First, we examine the standard three-dimensional finite-difference approximation to the wave equation:

\[(\partial_{tt} - v(\mathbf{r})^2 \nabla^2)\varphi(\mathbf{r},t) = 0,\]

(4.30)

where \(\mathbf{r} = (x, y, z)\) is the position vector and \(\nabla^2 = \partial_{xx} + \partial_{yy} + \partial_{zz}\) the 3D Laplacian operator. Extending Eq. (4.25) to three dimensions, it is an easy task to see that we arrive at the following relation:

\[\left(\frac{1}{\Delta t^2} \hat{D}^2_t - v(\mathbf{r})^2 \frac{1}{\Delta x^2} \hat{D}^2_t\right)\varphi(x,t) = 0,\]

(4.31)

where \(\hat{D}^2_t = \hat{D}^2_x + \hat{D}^2_y + \hat{D}^2_z\), and for simplicity we have assumed that \(\Delta x = \Delta y = \Delta z\), i.e. a cubic spatial cell.

In order to work out the inherent numerical error of the finite-difference Eq. (4.31), we may insert a known, analytic solution of the actual wave Eq. (4.30) to Eq. (4.31) and calculate the resulting residual term(s). Such a solution is, e.g., \(\varphi_0(\mathbf{r}, t) = e^{i(k\cdot\mathbf{r} - \omega t)}\), where \(k = (k_x, k_y, k_z)\), and upon inserting it into Eq. (4.31) we obtain:

\[\left(\frac{1}{\Delta t^2} \hat{D}^2_t - v(\mathbf{r})^2 \frac{1}{\Delta x^2} \hat{D}^2_t\right)\varphi_0(\mathbf{r},t) = \]

\[2 \left[ (\cos(\omega \Delta t) - 1) - v(\mathbf{r}) \frac{\Delta t^2}{\Delta x^2} \sum \left(\cos(k_i \Delta x) - 1\right) \right] \varphi_0(\mathbf{r},t) \]

algorithmic error, \(\tilde{e}\)

(4.32)

Ideally, we would like the error \(\tilde{e}\) to be zero for every direction of wave propagation, i.e. for every \((k_x, k_y, k_z)\) triad, but inspection of Eq. (4.32) reveals that this is never
possible. Thus, an anisotropy error (in addition to the induced numerical dispersion error), originating from the particular finite-difference approximation, $\hat{D}_i^2$, to the 3D Laplacian operator $\nabla^2$ will always be present in Eq. (4.31). In the following section we investigate a method for constructing finite-difference approximations to $\nabla^2$ with significantly reduced anisotropy error. This method is, then, used in the formation of high-accuracy finite-difference operators that can be used within FDTD's time-marching algorithm.

### 4.4.2 Non-Standard Finite-Difference Time-Domain Algorithm

Critical for the NS-FDTD formulae [63], [81] is the introduction of generalised 3D Laplacian operators, $D_i^2$ ($i = 1..3$), shown in Fig. 4.2(a). Focusing on the derivation of $D_1^2$, we start by introducing the parameters: $\xi = h \frac{\partial}{\partial x}$, $\eta = h \frac{\partial}{\partial y}$ and $\mu = h \frac{\partial}{\partial z}$, where $h$ is a real number representing the cell size and:

$$\left(\xi^2 + \eta^2 + \mu^2\right)u(\vec{r}) = h^2\nabla^2 u(\vec{r}), \quad (4.33)$$

with $u(\vec{r})$ being an electromagnetic field component in 3D space. Denoting with $u_i$ ($i = 1..8$) the values of the function $u(\vec{r})$ at the points shown in Fig. 4.2(a), we have:

$$u_i = e^{(\xi + \eta + \mu)}u_0 = \left[1 + (\xi + \eta + \mu) + \frac{(\xi + \eta + \mu)^2}{2} + \ldots\right]u_0. \quad (4.34)$$

Similarly: $u_2 = e^{(-\xi - \eta - \mu)}u_0$, $u_3 = e^{(\xi - \eta - \mu)}u_0$, with analogous expressions for the remaining points. Using series expansions like the one in (4.34) and the identity in (4.33), we obtain:

$$S_2 = \sum_{i=1}^{8} u_i = 8u_0 + 4h^2\nabla^2 u_0 + 0[h^4], \quad (4.35)$$
Figure 4.2: (a) Spatial arrangement of the generalized 3D Laplacian operators. (b) The elements of the nonstandard difference operator $\tilde{d}_x$. A suitable weighting component is shown for the filled points in each side of the element.

from which we can define an alternative second-order Laplacian $D_2^0$ as:

$$D_2^0 u_0 = \frac{S_2/4 - 2u_0}{h^2}. \quad (4.36)$$

Likewise, the second-order operators $D_1^2$ and $D_3^2$ of Fig. 4.2(a) can be derived, as well as higher-order Laplacians. We can efficiently combine $D_i^2$ ($i = 1..3$) to create an optimal Laplacian operator $D_0^2$, which equals a suitably weighted sum of $D_i^2$ and is sixth-order accurate in space for a reference frequency [82]. Decomposing $D_0^2$ into finite-difference operator products and employing nonstandard stipulations, yields the following spatial formulae that preserves the $O[h^6]$ accuracy (Fig. 4.2(b)):
\[ \tilde{d}_x = \frac{1}{S_k(h)} \left( \zeta_1 d_x^{(1)} + \zeta_2 d_x^{(2)} + \zeta_3 d_x^{(3)} \right) \]  

(4.37)

with \( S_k(h) = 2 \sin(kh/2)/k \) [62], \( k \) being the wavevector that corresponds to the central frequency of the input source and \( \zeta_i \) \( (i = 1..3) \) positive weighting components fulfilling \( \zeta_1 + \zeta_2 + \zeta_3 = 1 \) for numerical consistency. For the other two directions analogous difference operators can be formed which, when inserted into the \( E \) or \( H \)-field updating equations, result in a significantly ameliorated FDTD scheme with regard to numerical dispersion and anisotropy.

Very crucial for the simulations herein is the formulation of the nonstandard uniaxial perfectly matched layer (NS-UPML) [61], [82]-[83] inside which extend the waveguide layers. We make use of the UPML in order to suitably terminate our (finite) computational space. The numerical waves excited by our numerical source reach the UPML, enter it with practically zero reflection, and are progressively absorbed while they propagate inside it. For the sake of brevity, only the expression for the advancement of the \( B_y \) component (in Heaviside-Lorentz units) is given below:

\[ B_y = \frac{\kappa_z - S_\omega(\Delta t) \sigma_z/(2\varepsilon_0)}{\kappa_z + S_\omega(\Delta t) \sigma_z/(2\varepsilon_0)} B_y + c_0 \frac{S_\omega(\Delta t)}{\kappa_z + S_\omega(\Delta t) \sigma_z/(2\varepsilon_0)} \left( \tilde{d}_y E_z - \tilde{d}_z E_x \right) \]  

(4.38)

with \( S_\omega(\Delta t) = 2 \sin(\omega\Delta t/2)/\omega \), \( c_0 \) the velocity of light in vacuum, and both finite-difference operators, now, involving eighteen instead of two points in a conventional formulation [82], [83].

For the acquisition of the 2D excitation profiles we recall from Ch. 3 that in rectangular dielectric waveguides there are no precise closed-form expressions for the cross-sectional shape of the modes, owing to the peculiar behaviour of the electromagnetic fields near the dielectric corners where they diverge to a small degree [84]. As we show in the previous chapter, a sufficiently accurate numerical way to overcome this limitation is via computing the \( E \) or \( H \)-field eigenvectors of the vectorial wave equation that takes into account the polarization, vector properties and discontinuity of the guided modes at the dielectric interfaces, including the corner
regions [58]. Upon obtaining the 2D field distributions in this manner, a proper spatial interpolation is necessary to match the Yee-grid. Such an approach is essential, particularly for FDTD in-coupling studies, and an exemplary result for the dominant

![Figure 4.3: Schematic illustration of a high-index contrast (strongly guiding) planar dielectric heterostructure. Also shown is the pre-calculated excitation profile for the generation of the $E_{12}$ mode inside the rectangular waveguide.](image)

component of the $E_{12}$ eigenmode is illustrated in Fig. 4.3. The associated propagation constant is obtained from the eigenvalue of the wave equation and, for sufficiently small mesh size, the error in its estimation can be made very small (e.g. less than 0.01%), as we show in discussing Figs. 3.5–3.8 in the previous chapter [58].

### 4.5 Application of the NS-FDTD Method in the Simulation of a Three-Dimensional Dielectric Waveguide

The structure on which the overall methodology was tested is a symmetric, multimode rectangular waveguide having a core refractive index $n_c = 3.41$, cladding refractive indices $n_s = n_e = 1.5$, width $w = 1.17 \, \mu\text{m}$ and very small core thickness ($t = 180 \, \text{nm}$), as illustrated in Fig. 4.4 [83]. Our choice is motivated by the use of high index-contrast waveguides in existing photonic devices involving waveguide coupling to
planar photonic crystals or various travelling-wave devices (e.g., travelling-wave photodetectors, phototransistors, etc). The length was chosen to be sufficiently large, \( l = 21 \mu m \), to check the precision in the single-mode excitation at the far end of the waveguide. We have examined the dynamical propagation of the first three \( E^\gamma \) modes, generated with the appropriate initial profiles. The investigated spectral range is sufficiently above the cut-off frequency of the third eigenmode (to ensure accurate computation of the propagation constant and mode profile with the mode-solver). Using the effective-index method (section 3.5) we have calculated this frequency to be \( f_{cd} \approx 142 \text{ THz} \). For excitation we used Gaussian pulses modulating a sinusoidal carrier of frequency \( f_0 = 193.5 \text{ THz} \), which coincided with the reference frequency in the nonstandard difference operators. The computed values of \( \xi_i \) in this study were \( \xi_1 = 0.7093, \xi_2 = 0.00449, \xi_3 = 0.2458 \), and an exemplary part of the actual computational code is presented below:

\[
\ldots \text{Hx}(i,j,k) = \text{Hx}(i,j,k) - co^*ux(i,j,k)^*(a1^*( Ez(i+1,j+1,k) - Ez(i,j,k) ) + a2^*( Ez(i+1,j+1,k+1) + Ez(i+1,j+1,k-1) + Ez(i+1,j+1,k-1) + Ez(i+1,j+1,k-1) - Ez(i-1,j+1,k+1) - Ez(i+1,j+1,k+1) - Ez(i+1,j+1,k-1) - Ez(i+1,j+1,k-1) )/4 + a3^*( Ez(i+1,j+1,k+1) + Ez(i-1,j+1,k+1) + Ez(i+1,j+1,k-1) + Ez(i-1,j+1,k-1) - Ez(i+1,j+1,k) + Ez(i+1,j+1,k) - Ez(i+1,j+1,k) + Ez(i+1,j+1,k) + Ez(i+1,j+1,k) - Ez(i-1,j+1,k) - Ez(i-1,j+1,k) - Ez(i-1,j+1,k) - Ez(i-1,j+1,k) + Ez(i+1,j+1,k) - Ez(i+1,j+1,k) - Ez(i+1,j+1,k) - Ez(i+1,j+1,k) - Ez(i+1,j+1,k) - Ez(i+1,j+1,k) + Ez(i+1,j+1,k) - Ez(i+1,j+1,k) - Ez(i+1,j+1,k) - Ez(i+1,j+1,k) + Ez(i+1,j+1,k) - Ez(i+1,j+1,k) - Ez(i+1,j+1,k) - Ez(i+1,j+1,k) )/4 - a1^*( Ey(i,j+1,k+1) - Ey(i,j,k) ) - a2^*( Ey(i+1,j+1,k+1) + Ey(i+1,j+1,k+1) + Ey(i+1,j+1,k+1) + Ey(i+1,j+1,k+1) + Ey(i+1,j+1,k+1) + Ey(i+1,j+1,k+1) + Ey(i+1,j+1,k+1) + Ey(i+1,j+1,k+1) + Ey(i+1,j+1,k+1) - Ey(i+1,j+1,k+1) - Ey(i+1,j+1,k+1) - Ey(i+1,j+1,k+1) - Ey(i+1,j+1,k+1) - Ey(i+1,j+1,k+1) - Ey(i+1,j+1,k+1) - Ey(i+1,j+1,k+1) - Ey(i+1,j+1,k+1) - Ey(i+1,j+1,k+1) + Ey(i+1,j+1,k+1) + Ey(i+1,j+1,k+1) + Ey(i+1,j+1,k+1) + Ey(i+1,j+1,k+1) + Ey(i+1,j+1,k+1) + Ey(i+1,j+1,k+1) - Ey(i+1,j+1,k+1) - Ey(i+1,j+1,k+1) - Ey(i+1,j+1,k+1) - Ey(i+1,j+1,k+1) - Ey(i+1,j+1,k+1) - Ey(i+1,j+1,k+1) - Ey(i+1,j+1,k+1) - Ey(i+1,j+1,k+1) - Ey(i+1,j+1,k+1) )/4 ), \ldots 
\]

where 'co' denotes the vacuum light speed, \( a_i = \xi_i \) \( (i = 1, 2, 3) \), and \( ux = (\sin(\omega \Delta t/2)/\omega)/(\sin(k \Delta x/2)/k) \), with \( ux \) (central film) = \( 1.746516 \times 10^{-9} \), \( ux \) (cover) = \( 1.743246 \times 10^{-9} \), \( ux \) (substrate) = \( 1.745135 \times 10^{-9} \), and \( ux \) (background) = \( 1.732883 \times 10^{-9} \). The corresponding piece of code for the conventional FDTD method is:

\[
\text{Hx}(i,j,k) = \text{Hx}(i,j,k) - co^*dt^*(( Ez(i+1,j,k) - Ez(i,j,k)/dy - ( Ey(i,j,k+1) - Ey(i,j,k) ) /dz ),
\]

where \( dt \) denotes the time step, and \( dy, dz \) the dimensions of the unit cell along the \( y \) and \( z \)-axis, respectively. Note that the updating equations for the NS-FDTD method contain considerably more terms compared with the updating equations of the
Figure 4.4: Schematic illustration of the three-dimensional dielectric waveguide simulated with the NS-FDTD method. The waveguide is laterally surrounded by air, while its top and bottom claddings extend inside the upper and bottom UPML layers, respectively. All three dielectric layers also longitudinally extend inside, both, the front and back UPML.

conventional algorithm, owing to the use of the nonstandard operators $d_m^{(i)}$, $i = 2, 3$ and $m = x, y, z$, shown in Fig. 4.2, which “sample” additional neighbouring points in the calculations of the partial spatial derivatives compared with the conventional $d_m^{(1)}$ operator. As a result, the number of calculations per each time step does considerably increase with the implementation of the NS-FDTD methodology. However, the computational accuracy of the new algorithm increases dramatically, to the point which allows one to use much larger unit cells compared with the conventional algorithm, and still maintain the same (or better) levels of accuracy of the conventional algorithm. Thus, as we demonstrate in the following numerical
simulations, improved accuracy or significant computational savings are achieved when the NS-FDTD strategy is deployed in the analysis of photonic structures.

Figure 4.5 presents a few snapshots from the propagation of the $H_z$-field component along the waveguide of Fig. 4.4. Note that at the excitation plane (at around the fiftieth unit cell in the longitudinal direction) there are, as expected, two pulses generated that travel in opposite directions. The left pulse is completely "absorbed" inside the left UPML and never enters again the main computational domain. By contrast, the right pulse is guided along the dielectric heterostructure until it reaches the right UMPL (not shown here) wherein it is also absorbed. Note, also, that the use of the proper excitation profile has allowed most of the exciting pulse's energy to be launched into the desired mode (here, into the $E_{11}^y$ mode); otherwise, we would have observed a much more involved situation, with more than one mode being excited and propagating (with different speeds) along the waveguide.

Figure 4.6 shows the extracted longitudinal propagation constants, $\beta$, over the bandwidth of the corresponding excitation pulse, which were obtained by dividing the fast Fourier transforms (FFTs) of the pulse's time history at two fixed observation points, shown in Fig. 4.5(b), along the core [61, Chs. 15-16], i.e. we used the relationship:

$$\beta(\omega) = \text{Im} \left\{ \frac{1}{d} \ln \left[ \frac{\text{FT}[H_z(t,x)]}{\text{FT}[H_z(t,x)]} \right] \right\}.$$  

Note the excellent agreement between the values predicted by the NS-FDTD and the fully vectorial mode solvers for all three eigenmodes that confirms the precision of their excitation. Figures 4.7(a)-(c) illustrate the relative error, with respect to the mode-solver, of the conventional and nonstandard FDTD. In each case it is found that the classical Yee scheme only attains the same levels of accuracy, especially in the high-frequency range, when the number of spatial grid points is increased by a factor of 6.5, with a corresponding significant increase in the computational time, as expected. The increase in the computational time for the NS-FDTD, however, is small due to the use of a 30% larger time-step than the maximum allowed one in the standard algorithm [61], [82]. It is also verified that optimum performance for this form of the NS-FDTD is achieved within a narrow region around $f_0$ [63], [82] where it is always found to be more accurate than the usual Yee formulation.

These results allow one to conclude that, owing to the considerable reduction of the numerical dispersion and anisotropy errors, significantly enhanced overall
Figure 4.5: Snapshots from the propagation of the \( H_z \)-field component during the NS-FDTD simulation of the waveguide illustrated in Fig. 4.4. The plane on which the snapshots were obtained is in the middle of the waveguide core layer, lying parallel to the waveguide base. Shown here is the two-dimensional distribution of the \( H_z \)-field component on the aforementioned plane, at the time step: (a) \( n = 1000 \), (b) \( n = 2000 \), (c) \( n = 3000 \), and (d) \( n = 4500 \).
computational performance is obtained when the NS-FDTD formulae is deployed. In particular, either improved accuracy or significant computational savings can be achieved when the nonstandard finite-difference concepts are incorporated within the FDTD modelling methodology of electromagnetic devices and structures.

**Figure 4.6:** Longitudinal propagation constants of the first three $E^y$ eigenmodes: Comparison between fully vectorial mode-solver (dashed line) and NS-FDTD (symbols).
Figure 4.7: (a) Relative error in the NS-FDTD and FDTD-computed propagation constant of the fundamental $E_{11}$ eigenmode. (b) The same calculations as in (a), but for the $E_{12}$ eigenmode. (c) The corresponding calculations for the $E_{13}$ mode supported by the same structure as in both previous cases.
5 Light Propagation in Negative-Refractive-Index Metamaterials and Waveguides

In the previous two chapters we investigated the propagation of light inside conventional dielectric bulk materials and waveguides. The totality of the structures that we therein examined were characterised by constitutive electromagnetic parameters (relative electric permittivity, \( \varepsilon_r \), and magnetic permeability, \( \mu_r \)) that were greater than or equal to unity, i.e. \( \varepsilon_r \geq 1 \) and \( \mu_r \geq 1 \).

The theory describing the interaction of light with and propagation within such structures is well-established, since until very recently media having \( \varepsilon_r > 1 \) and \( \mu_r > 1 \) were, indeed, considered to be the only ones that possess interesting and/or useful electromagnetic properties. For several decades, nonetheless, scientists were well-aware that more ‘exotic’ regions for the electromagnetic parameters of a medium can exist, such as e.g. in metals at optical frequencies, whereupon \( \text{Re}\{\varepsilon_r\} \leq 0 \). However, the wider consensus was that – at least at the time – there was little to benefit from attempting to exploit such ‘exotic’ materials, not least because usually they do not occur naturally and, also, because they were not thought of as offering any particular advantage over their common dielectric counterparts. This view has, now, been revised and overturned, owing primarily to pioneering insights concerning the fundamental properties of such ‘exotic’ materials, but also owing to technological progress that allowed for their construction and characterisation, all the way from the radio up to the optical regime.

In this chapter we will study in some detail the electromagnetic properties of materials possessing simultaneously negative electric permittivity and magnetic permeability, i.e. \( \varepsilon_r \leq 0 \) and \( \mu_r \leq 0 \). These materials have come to be known as metamaterials – although, at present, the same terminology is also used for a number of related materials, such as materials with only \( \text{Re}\{\varepsilon_r\} \) (\( \text{Re}\{\mu_r\} \)) \leq 0 \) or \( \text{Re}\{\varepsilon_r\} \) (\( \text{Re}\{\mu_r\} \)) \approx 0 \), photonic crystals, etc – due to the fact that they allow for overcoming a number of common limitations associated with the electromagnetic properties of naturally occurring materials. Upon concisely reviewing the history behind these materials, we shall examine some of their fundamental properties, including the sign of their refractive index, Snell’s law of refraction and Fresnel’s formulae for the reflection and transmission coefficients of a plane wave impinging at the interface.
between a dielectric material and a metamaterial, as well as the correct new expression for the electromagnetic energy density. We shall then study one of the most remarkable properties of such materials, namely the fact that a planar slab made of a metamaterial with \( \varepsilon_r \leq 0 \) and \( \mu_r \leq 0 \) can, in principle, act as a ‘perfect’ lens, i.e. a lens capable of aberration-free, deep (in theory, infinite) subwavelength resolution. We shall see that this property arises ultimately from the presence of, so called, surface wave modes that are localised at and guided along the interfaces of a metamaterial with a dielectric material. We will conclude the chapter by presenting a detailed analysis into the various types of oscillatory and surface wave modes that metamaterial heterostructures can support and we will highlight some of their intriguing properties that will draw our attention in more detail on chapter 7.

5.1 A Brief History of Metamaterials

The history of metamaterials appears to date back to the pioneering work of Kock [22] in the late 40’s. While working at Bell Labs with Sergei Schelkunoff, renowned for his “field equivalence principles” and for his work on antenna theory, Kock published a series of works [23] wherein he proposed numerous ideas for constructing lightweight and small-volume “artificial dielectrics”, used as microwave lenses in antenna systems. Amongst others, he studied the response to an incident quasi-static electromagnetic radiation of isolated or regularly-arrayed metallic particles of various shapes, such as spheres, discs, ellipsoids and prolate or oblate spheroids. He concluded that such structures effectively behave as a dielectric medium, whose permittivity \( \varepsilon \) and permeability \( \mu \) can be intentionally tuned (but not independently of each other) to an arbitrarily large or small, even negative, value by properly arranging the particles in three dimensions, i.e. the optical properties of the medium depended solely on the particles’ geometrical set up, rather than on their own intrinsic behaviour. Kock also showed that a specially-designed structure, which recently has come to be known as “split-ring resonator” (SRR), can be used to independently increase the effective magnetic permeability \( \mu \), such that one can reduce or altogether eliminate the diamagnetic nature of the aforementioned composite structures. His work attracted considerable interest within the engineering community of the time,
with a number of works extending or elaborating on his ideas. Since then, it has been
the subject of detailed coverage in standard engineering textbooks [23].

More than a decade later, Rotman [85] also considered the quasi-static response of
an array of thin conducting wires, and he showed that such a structure closely
resembles, on the macroscopic level, a plasma medium. In particular, he proved that
the electric permittivity $\varepsilon$ of this artificial dielectric medium varies with frequency
following a Drude-type law. Consequently, below a certain “cutoff” frequency no
incident electromagnetic radiation could penetrate it. Critically, however, neither
Kock nor Rotman nor, indeed, any of the early contributors investigated the properties
of media exhibiting concurrently negative $\varepsilon$ and $\mu$. Partly, that was because the main
motivation behind similar works at that time was to design plasma media at RF or
microwave frequencies that would closely simulate the ionosphere, prompted by
NASA’s desire to secure the safe re-entrance of space-capsules into the earth’s
atmosphere.

Veselago [1], [2] was evidently the first to systematically consider, in the late 60’s,
the possibility and some properties of “double-negative materials” (DNGMs). In
particular, he showed that a negative electric permittivity $\varepsilon$ and magnetic permeability
$\mu$ would imply a reversal of almost all known electromagnetic phenomena, including
the angle of refraction inside a DNGM (e.g., $\theta_r = -20^\circ$), the Doppler effect, the sign
of the refractive index $n$ (e.g., $n = -1$) and the right-handedness of the $\textbf{E}$, $\textbf{H}$ and $\textbf{k}$
vector-triad, from where the designation of such materials as “left-handed” origins.
Moreover, Veselago also proved that the phase velocity ($v_{ph}$) of a plane wave
propagating inside a left-handed metamaterial (LH-MM) has a direction opposite
to that of the group ($v_g$) and energy-flow velocities ($v_E$). However, in spite of his
noteworthy findings, and apparently unaware of Kock’s and other scientists’ research
in the same field, Veselago did not go on to materialise his theoretical conclusions.
Even so, his work did not go unnoticed and he was invited several times to highlight
his research at major international scientific conferences [21].

At present, the realm of “artificial dielectrics” or “meta-materials” (from the Greek
word “meta”, which here means “beyond”) enjoys a breadth of scientific activity and
exploration, having established a sound and coherent mathematical formalism [16]-
[19], the predictions of which have been verified by numerous experimental and
numerical-simulation works. This revived interest followed from a series of works by
Professor Sir John Pendry, wherein he proposed practical means for realizing LH-MMs experimentally [3], [4]. Moreover, building on Veselago’s work, Pendry showed that a slab constructed by a LH-MM having refractive index \( n = -1 \) could, ideally, act as a “perfect lens”, overcoming the well-known diffraction limitations. After these key insights, the physical construction of a composite LH-MM structure has been demonstrated by Shelby et al. [6], [7] and the possibility of achieving subwavelength resolution of an object with the same structure has been demonstrated with a series of further theoretical [30], [31] and experimental [8], [9] works.

5.2 Sign of the Refractive Index and Energy Density Expression in Passive ‘Double Negative’ Metamaterials

Let us consider a passive medium that is characterised by simultaneously negative (effective) electric permittivity, \( \varepsilon \), and magnetic permeability, \( \mu \), e.g. \( \varepsilon = \mu = -1 \). Such media do not normally occur in nature but, as we shall see in the next chapter, they can indeed be constructed in the lab by suitably engineering the electromagnetic ‘molecules’ of a medium. At first sight, one may think that the refractive index of such a ‘double negative’ medium, which is defined as \( n = \sqrt{\varepsilon\mu/(\varepsilon_0\mu_0)} \), should still be positive since the product ‘\( \varepsilon\mu \)’ appearing in the expression for the refractive index remains positive. However, when one also considers the requirement that the ‘double negative’ medium should, in general, be passive, the sign of the square root of the product ‘\( \varepsilon\mu \)’ actually turns out to be negative [86], [87], [138].

Indeed, in general, the relative electric permittivity \( \varepsilon_r \) and magnetic permeability \( \mu_r \) of a medium are complex functions of frequency, i.e., in an obvious notation, they are of the form:

\[
\varepsilon_r(\omega) = \varepsilon_r'(\omega) + i\varepsilon_r''(\omega) = \rho_\varepsilon(\omega)e^{i\theta(\omega)}, \quad \mu_r(\omega) = \mu_r'(\omega) + i\mu_r''(\omega) = \rho_\mu(\omega)e^{i\phi(\omega)},
\]

and therefore:

\[
n = \sqrt{\varepsilon\mu/(\varepsilon_0\mu_0)} = \sqrt{\rho_\varepsilon\rho_\mu}e^{i(\theta+\phi)/2}.
\]
From Eq. (5.2) one may immediately recognize that if the medium is to be passive, i.e. having a positive imaginary part (for an assumed $e^{-\omega t}$ time dependence), it should be:

$$0 \leq (\theta + \varphi)/2 < \pi.$$  \hspace{1cm} (5.3)

On the other hand, in a double negative metamaterial the real parts of the (effective) permittivity and permeability are negative, which based on Eq. (5.1) implies that:

$$\frac{\pi}{2} < \theta < \frac{3\pi}{2}, \text{ and: } \frac{\pi}{2} < \varphi < 3\frac{\pi}{2},$$  \hspace{1cm} (5.4)

or, by adding the two parts of Eq. (5.4):

$$\frac{\pi}{2} < (\theta + \varphi)/2 < 3\frac{\pi}{2}.$$  \hspace{1cm} (5.5)

Comparison of Eqs. (5.3) and (5.4) leads to the following relation:

$$\frac{\pi}{2} < (\theta + \varphi)/2 < \pi,$$  \hspace{1cm} (5.6)

from whence it is unambiguously concluded that the real part of the medium’s refractive index (obtained from Eq. (5.2)) will be negative. Thus, in a medium that has simultaneously (in the same frequency region) negative real parts in its electric permittivity and magnetic permeability, the real part of the refractive index will also be negative. This fact has immediate consequences on the way an electromagnetic wave refracts inside such a medium, but also on the three-dimensional spatial arrangement of the electric field vector, $\mathbf{E}$, magnetic field vector, $\mathbf{H}$, and the wavevector, $\mathbf{k}$, as is explained in the next section.

It is interesting to note that any medium having simultaneously negative $\varepsilon$ and $\mu$ must necessarily be dispersive or the energy density will be negative. Indeed, if the medium was not dispersive, then the cycle-averaged energy density of an electromagnetic wave propagating inside it would have been:

$$W = (\varepsilon |\mathbf{E}|^2 + \mu |\mathbf{H}|^2)/4,$$  \hspace{1cm} (5.7)
which, in view of the fact that \( \epsilon < 0 \) and \( \mu < 0 \), would have obviously been negative. Thus, for a double negative metamaterial the correct expression for the cycle-averaged energy density is that of a dispersive medium [1], [2], i.e.:

\[
W = \frac{1}{4} \left[ \frac{d(\varepsilon \omega)}{d\omega} |E|^2 + \frac{d(\mu \omega)}{d\omega} |H|^2 \right],
\]

(5.8)

which turns out [ ] to be positive for every frequency, since causality and Kramers-Kronig relations for such a metamaterial demand [19] that \( \frac{d(\varepsilon \omega)}{d\omega} > 0 \) and \( \frac{d(\mu \omega)}{d\omega} > 0 \).

5.3 Refraction and E-H-k Vector Triad Inside a ‘Double Negative’ Metamaterial

As we remarked in the previous section, the fact that the real part of the refractive index of a DNGM is negative has some immediate consequences, most notably, on the way a plane wave refracts at the interface between a dielectric medium and a DNGM.

![Figure 5.1](image)

**Figure 5.1**: Reflection and refraction of a plane wave at the interface of a dielectric medium of refractive index \( n_1 \) with (a) a dielectric medium of refractive index \( n_2 \), and (b) a double negative metamaterial with real part of refractive index \( n_2 = -|n_2| \). In both cases the solid black arrows denote the direction of power flow, while the solid red arrows reveal the direction of phase propagation inside the second medium.
Indeed, by a straightforward application of boundary conditions (section 3.3) at the interface between two media with (real parts of) refractive indices \( n_1 \) and \( n_2 \), we obtain the following relations for the angles of incidence, \( \theta_{\text{inc}} \), reflection, \( \theta_{\text{refl}} \), and refraction, \( \theta_{\text{refr}} \):

\[
\theta_{\text{refl}} = \theta_{\text{inc}}, \quad \text{and:} \quad \theta_{\text{refr}} = \text{sgn}(n_2)\sin^{-1}\left(\frac{n_1}{n_2} \sin\theta_{\text{inc}}\right).
\]  

(5.9)

From the second part of Eq. (5.9) one sees that when the second medium has negative (real part of) refractive index, then the angle with which a plane wave refracts inside this medium is ‘negative’, i.e. the wave refracts at the same side of the normal to the media interface as the incident wave (see Fig. 5.1). Hence, remarkably, the refraction of a plane wave, incident from a conventional dielectric medium to an isotropic DNGM, is ‘opposite’ to that of the usual case, wherein refraction occurs inside another (positive index) dielectric medium.

The fact that the refraction of a plane wave inside an isotropic DNGM is ‘negative’ leads to the notable possibility of using a planar slab made of such a material to bring light to a double focus, as illustrated in Fig. 5.2. Indeed, one may readily infer that a slab of thickness \( d \) and refractive index of, e.g., \( n = -1 \), surrounded by air, will bring all rays emanating from a source to a double focus: First, at a point inside the DNGM

![Figure 5.2: Double focusing of a source with a planar double negative (left-handed) metamaterials slab.](image)

66
slab, at a distance \( s = l < d \), where \( l \) is the distance of the source from the slab, and second at a point outside the slab, at a distance \((d-l)\). Hence, such a slab acts like a lens, and is able to bring the rays (corresponding to propagating waves) radiated by a source to a focus outside the slab, without reflections occurring at the media interfaces because the \( n = -1 + i0 \) slab is impedance-matched to free space (see section 5.4). This result can be further verified by means of direct FDTD (numerical) solutions of Maxwell’s equations. An exemplary result of such simulations is presented in Fig. 5.3 below, where we can see a snapshot from the interaction of an electromagnetic pulse with a DNGM slab, impedance matched to the surrounding medium (which is air). Note the formation of a double mirror image of the source, once inside the DNGM slab and a second time outside the slab. We can also observe that there are no reflections occurring at the two media interfaces; the only reflections observed are those which occur close to the (imperfect for DNGM) absorbing boundary conditions that are used to terminate our computational space. The small reflections that we observe in this figure, as well as the departure of the numerical wavefronts from a perfectly cylindrical shape, are entirely owing to the use of imperfect boundary conditions to simulate the extension of the DNGM slab to infinity.

**Figure 5.3:** Snapshot from an FDTD simulation of an electromagnetic pulse radiated by a point source and incident on a DNGM slab \((n = -1)\) that is surrounded by air.

It is to be noted that the aforementioned DNGM slab configuration not only is it capable of bringing outgoing light rays to a focus, but it does so without suffering from the primary (monochromatic or Seidel) aberrations (spherical, coma,
astigmatism, field curvature and distortion) that usually characterise conventional (e.g., spherical) lenses [66]. This is because the DNGM lens, being planar, may extend ‘infinitely’ in the transverse plane, capturing all rays and focusing them at a point exactly ‘in front of’ each object source point, as was shown in Fig. 5.2. Thus, there will be no distortions of the object at the image/focusing plane. Moreover, one may, in principle, suitably engineer an isotropic DNGM that has a refractive index of $n = -1 + i\theta$ over a continuous, but finite, range of frequencies. If such a material, surrounded by air, is used (in the frequency region where $n = -1 + i\theta$) as a lens, then this lens will also be free of chromatic aberrations. Even more remarkable, however, is the fact that such a lens is not limited even by diffraction, i.e. it is capable of reconstructing at the image plane even the highly evanescent components ‘radiated’ off by an object source, which decay rapidly away from the source and are normally completely lost. Therefore, such a configuration acts like a ‘perfect’ lens, and will be the subject of a closer examination in section 5.5.

Finally, one should note yet another interesting property of an isotropic DNGM, pertaining to the three-dimensional arrangement of the vectors $E$, $H$ and $k$ inside such a material. From the first two of Maxwell’s equations (Eqs. (3.8) and (3.9)), as well as from the constitutive relation given by Eqs. (3.1) and (3.2), it can readily be shown that inside every isotropic (possibly dispersive) medium, it is:

\[ k \times E = \omega \mu(\omega)H, \quad \text{and} \quad k \times H = \omega e(\omega)E. \]  

(5.10)

We see from Eq. (5.10) that for a conventional dielectric medium with $\varepsilon, \mu > 0$ (in a given frequency region), the vectors $E$, $H$ and $k$ form a right-handed triad. By contrast, inside a DNGM where $\varepsilon, \mu < 0$, the vectors $E$, $H$ and $k$ form a left-handed triad, from where the designation of such media as ‘left-handed’ (LH) arises. It should also be pointed out that the definition of the Poynting vector, $S = E \times H$, remains the same for both classes of media (right- and left-handed). As a result, whereas in a right-handed medium the direction of the power flow (given by the direction of $S$) is the same as the direction of the phase evolution (given by the direction of the wavevector $k$), in a left-handed medium the vectors $k$ and $S$ are antiparallel, i.e. the phase evolves/‘propagates’ in the opposite direction compared with the direction of the flow of power, as illustrated in Fig. 5.1.
5.4 Fresnel’s Formulae for Plane Wave Incidence at a Planar RH/LH Media Interface

It is interesting to examine how the well-known Fresnel’s formulae for the reflection and transmission coefficient of a monochromatic plane wave incident at the interface between two dielectric media are modified when the second medium is assumed to be a DNGM one.

To this end, let us assume that the incident, reflected and refracted electric field is of the form:

\[
E_{\text{inc}} = E_{\text{0inc}} e^{-i(\omega t - k_{\text{inc}} z)}, \quad E_{\text{refl}} = E_{\text{0refl}} e^{-i(\omega t - k_{\text{refl}} z)}, \quad E_{\text{refr}} = E_{\text{0refr}} e^{-i(\omega t - k_{\text{refr}} z)},
\]

respectively, where both media were assumed to be conventional dielectrics — otherwise the direction of the wavevector \(k_{\text{refr}}\) inside the second, DNGM, medium would have been reversed. Then, from Eq. (5.10), the incident, reflected and refracted B-field will, respectively, be of the form:
In the case where the second medium is a DNGM one, we proceed similarly to the case where both media are normal dielectrics, by implementing the standard boundary conditions (section 3.3) for the tangential field and wavevector components. Following such an analysis for the incident and reflected fields, with the electric field \( \mathbf{E} \) assumed perpendicular to the plane of incidence (Fig. 5.4), yields Snell's law of reflection (Eq. (5.9)) that was highlighted above, as well as the following reflection coefficient [19]:

\[
R = \frac{E_{\text{refl}}}{E_{\text{inc}}} = \frac{(n_1/\mu_1)\cos\theta_{\text{inc}} - (n_2/\mu_2)\sqrt{1 - (n_2^2/n_1^2)\sin^2\theta_{\text{inc}}}}{(n_1/\mu_1)\cos\theta_{\text{inc}} + (n_2/\mu_2)\sqrt{1 - (n_2^2/n_1^2)\sin^2\theta_{\text{inc}}}}. \tag{5.13}
\]

Likewise, one can obtain the following expression for the transmission coefficient of a monochromatic plane wave incident at the interface between a dielectric medium and a DNGM:

\[
T = \frac{E_{\text{trans}}}{E_{\text{inc}}} = \frac{2(n_1/\mu_1)\cos\theta_{\text{inc}}}{(n_1/\mu_1)\cos\theta_{\text{inc}} + (n_2/\mu_2)\sqrt{1 - (n_2^2/n_1^2)\sin^2\theta_{\text{inc}}}}. \tag{5.14}
\]

Similar expressions to Eq. (5.13) and (5.14) are obtained for the case where the electric field is parallel to the plane of incidence. As expected, when the second medium is characterised by \( \varepsilon_2 = -\varepsilon_1 \) and \( \mu_2 = -\mu_1 \) (\( \Rightarrow n_2 = -n_1 \)), then we immediately obtain from Eq. (5.13) that \( R = 0 \), and from Eq. (5.14) \( T = 1 \), i.e. in this case the DNGM is perfectly matched to the dielectric medium-1 and no reflection occurs.

It is interesting to note that Eqs. (5.13) and (5.14) are valid, both, when the second medium is a normal dielectric or when it is a left-handed metamaterial. This is, of course, owing to the fact that these equations contain the absolute values of the electromagnetic parameters of the second medium, and as such they do not lead to different results when the handness of the medium is reversed. However, it is important to realise that these equations were derived for waves that were incident to
the media interface with angle of incidence $\theta_{\text{inc}} < \theta_c$, where $\theta_c$ is the critical angle (given by: $\sin\theta_c = n_2 / n_1$, $n_2 < n_1$), i.e. for propagating waves.

When the incident wave impinges on the two media interface with an angle $\theta_{\text{inc}} > \theta_c$, one needs to suitably modify the expression giving the term: $\cos\theta_{\text{inc}} = [1 - (n_1/n_2)^2 \sin^2 \theta_{\text{inc}}]^{1/2}$, which appears in Eqs. (5.13) and (5.14). This term then becomes: $\cos\theta_{\text{inc}} = \pm \left[ i(n_1/n_2)^2 \sin^2 \theta_{\text{inc}} - 1 \right]^{1/2}$, and since the refracted (transmitted) wave propagates in the $+y$ direction (see Fig. 5.4), with a spatial dependence of the form $\exp[i(\omega y/c)n_2 \cos\theta_{\text{ref}}]$, with $n_2 < 0$, we should choose the minus ("-"") sign in the aforementioned expression for $\cos\theta_{\text{inc}}$, so that the refracted wave will not diverge at infinity. By direct substitution of this expression for $\cos\theta_{\text{inc}}$ into Eq. (5.13) we obtain:

$$r|_{\text{evanescent}} = \frac{\cos\theta_{\text{inc}} + i(\mu_1/\mu_2) \sqrt{\sin^2 \theta_{\text{inc}} - n_2^2/n_1^2}}{\cos\theta_{\text{inc}} - i(\mu_1/\mu_2) \sqrt{\sin^2 \theta_{\text{inc}} - n_2^2/n_1^2}} = r|_{\text{evanescent}}.$$

In the above equation, the term: $\varphi = 2\tan^{-1}\{((\mu_1/\mu_2) \cdot [\sin^2 \theta_{\text{inc}} - (n_2/n_1)^2]^{1/2} / \cos\theta_{\text{inc}}\}$ is the, so called, Goos-Hänchen phase shift that an incident ray experiences upon hitting the two media interface with an angle greater than the critical one, $\theta_c$. We shall again return to this term later, in Chapter 7, when we will study the zigzag ray propagation inside a "slow-light" DNGM waveguide. Here, though, it is interesting to note from Eqs. (5.11) and (5.15) that, in the present case, the incident ray will be spatially-shifted in the $-z$ direction (see Fig. 5.4) upon hitting the media interface with an angle $\theta_{\text{inc}} > \theta_c$, contrary to the case where both media are normal dielectrics whereupon the rays are always shifted in the $+z$ direction.

### 5.5 Metamaterial-Enabled 'Perfect' Lens

We will now turn our attention to a closer examination of one of the most remarkable properties of a DNGM slab, on which we briefly remarked on section 5.3, namely the fact that such a slab may, in principle, work as a 'perfect' lens, enabling super-resolution of an object at the image plane [25].

To this end, let us start by calculating the reflection and transmission coefficient of an evanescent wave (not a plane wave, as in the previous section 5.4) incident at a planar RH/LH interface located at $z = 0$, with the wave initially being in the RH
medium (air). Assuming, without loss of generality, that the electric field is polarised along the x-axis (see Fig. 5.4), i.e. perpendicularly to the plane of incidence, we have the following expressions for the incident and transmitted (electric) field, respectively:

\[
\mathbf{E}_{\text{inc}} = \mathbf{E}_{0} \text{inc} e^{i(k_{z} \gamma + k_{\text{inc}} \leftarrow \text{air})} \hat{x}, \quad \text{with:} \quad k_{z}^{\text{inc}} = i\sqrt{k_{\gamma}^{2} - (\omega/c)^{2}} \tag{5.16}
\]

\[
\mathbf{E}_{\text{trans}} = \mathbf{E}_{\text{trans}} e^{i(k_{z} \gamma + k_{\text{trans}} \leftarrow \text{air})} \hat{x}, \quad \text{with:} \quad k_{z}^{\text{trans}} = i\sqrt{k_{\gamma}^{2} - \epsilon \mu \omega^{2}}, \tag{5.17}
\]

where \( \epsilon \rightarrow -\epsilon_{0}, \mu \rightarrow -\mu_{0} \) are respectively the (negative) permittivity and permeability of the LH medium. Applying the standard boundary conditions for the tangential field components at the media interface, results in the following expressions for the reflection and transmission coefficients:

\[
t = \frac{2\mu k_{z}^{\text{inc}}}{\mu k_{z}^{\text{inc}} + \mu_{0} k_{z}^{\text{trans}}}, \tag{5.18}
\]

\[
r = \frac{\mu k_{z}^{\text{inc}} - \mu_{0} k_{z}^{\text{trans}}}{\mu k_{z}^{\text{inc}} + \mu_{0} k_{z}^{\text{trans}}}. \tag{5.19}
\]

If the incident evanescent wave is initially inside the LH medium, then the new reflection and transmission coefficients are obtained, simply, by interchanging the \( \mu k_{z}^{\text{inc}} \) and \( \mu_{0} k_{z}^{\text{trans}} \) terms in Eqs. (5.18) and (5.19). Furthermore, if the LH medium is assumed to be lossy, i.e. if it has electromagnetic parameters of the form: \( \epsilon = -\epsilon_{0}(1 - i\zeta) \) and: \( \mu = -\mu_{0}(1 - i\zeta), \zeta > 0 \), then (in the limit of \( \zeta \rightarrow 0 \)) it is straightforward to show that Eqs. (5.18) and (5.19) take the form:

\[
t = r = \frac{2i}{\zeta} \left( 1 - \frac{\omega^{2}}{c^{2}k_{\gamma}^{2}} \right). \tag{5.20}
\]

When two RH/LH interfaces are present, as is the case with a LH slab surrounded by air, an evanescent wave incident (from the air) to the slab will be reflected and transmitted multiple times at the media interfaces. A detailed analysis [31] then
reveals that, in this case, the overall transmission coefficient associated with the (evanescent) wave exiting the slab is given by:

\[
I_{\text{total}} = \left[ \frac{1}{2} + \frac{1}{4} \left( \frac{\mu_0 k_{\text{trans}}}{\mu k_{\text{inc}}} + \frac{\mu k_{\text{inc}}}{\mu_0 k_{\text{trans}}} \right) \right] e^{ik_{\text{trans}}d} + \left[ \frac{1}{2} - \frac{1}{4} \left( \frac{\mu_0 k_{\text{trans}}}{\mu k_{\text{inc}}} + \frac{\mu k_{\text{inc}}}{\mu_0 k_{\text{trans}}} \right) \right] e^{-ik_{\text{trans}}d} \right]^{-1},
\]

(5.21)

where \(d\) is the distance between the two interfaces, i.e. the slab thickness.

With the aid of Eqs. (5.16) and (5.17) one can, now, directly infer from Eq. (5.21) that when \(\varepsilon \rightarrow -\varepsilon_0, \mu \rightarrow -\mu_0\) the first term in Eq. (5.21) becomes equal to zero and we, therefore, obtain:

\[
I_{\text{total}} = e^{ik_{\text{trans}}d}.
\]

(5.22)

Thus, in this case, the evanescent field that impinges upon the LH slab does not decay further inside the slab, as it usually occurs with dielectric or metallic slabs/films, but is instead ‘amplified’ inside the LH slab. Realizing that evanescent waves are associated with the high spatial frequencies of the field ‘radiated’ by an object source (see Eqs. (5.16) and (5.17)), i.e. they are the part of the field that ‘carries’ the fine subwavelength features of an object source, Eq. (5.22) implies that the subwavelength features of an object can be recovered at the image plane. Therefore, a DNGM slab can, in principle, enable one to obtain the image of an object with ‘perfect’ resolution, containing all the subwavelength features of an object and overcoming the usual diffraction limitations that characterise conventional lenses.

When losses are present in the DNGM slab then, as one may intuitively expect, ‘perfect’ resolution can not be attained any more. In this case, a straightforward analysis based on Eqs. (5.20) and (5.21) reveals that the smallest feature that can be resolved at the image plane has a size: \(\Delta = (2\pi d)/|\ln\zeta|\) [19], [29], where \(\zeta\) is the imaginary part of the DNGM’s effective permittivity and permeability (see Eq. (5.20)). Thus, even in this case, subwavelength resolution, i.e. resolution \(\Delta < \lambda\), of an image can still be obtained, so long as the losses are smaller than an upper limit: \(\zeta < e^{-2\pi d/\lambda}\). This remarkable feature of a DNGM slab is, ultimately, owing to the fact that the slab supports a special class of waves at its boundaries, known as ‘surface plasmon polaritons’, to the study of which we now turn our attention.
5.6 Surface Plasmon Polaritons in Asymmetric DNGM Slab Heterostructures

As we showed in the previous section, the "perfect lens" action relies critically on the amplification of an object's near field in a surface wave (SW)-like manner inside a LH slab. Due to momentum mismatch, radiative waves cannot couple directly to the formed SWs (or surface polaritons) at the interfaces of the slab with positive-index dielectrics. Pendry, however, showed that the near field of an object, which describes its finest features and decays exponentially away from the source (evanescent), can couple to an exponentially increasing field inside the LH slab that decays similarly on the other side. The whole field pattern resembles that of a surface polariton, although the proof of its existence for the particular 'perfect lens' structure ($\varepsilon = \mu = -1$) was not further elaborated in [25], as well as in other more detailed analyses [88].

Generalizations of such analyses were investigations of asymmetric DNGM slab configurations for lensing [89]-[91], sensing and directional coupling [92] applications. In both cases, the role of the coupled surface polaritons at the two interfaces of the slab waveguide was shown to be of crucial importance. For the first class of applications, the asymmetry helped to improve the limit imposed on the image resolution by the losses of the core. For the second class it improved the amplification of the evanescent waves in the device-working region, leading to enhanced performance.

In this section we shall identify and classify in detail all surface plasmon polariton (SPP) eigenmodes supported by generalized asymmetric slab heterostructures. To this end, a rigorous analytical study will be pursued, which will prove that a total of 30 solutions to the involved characteristic equation giving the SPP eigenmodes can exist for all choices of the refractive index distribution, constitutive parameters $\varepsilon$ and $\mu$ and the thickness of the core [64]. Such an approach is essential [93], particularly for the investigation of asymmetric slab configurations, because the graphical methodologies that have been proposed in the past for the modal analysis of LH waveguides [88]-[92] did not reveal all SPP eigenmodes. We will see that a suitably modified form of the associated transcendental equation, derived from macroscopic electrodynamics using the well-known boundary conditions for the tangential electric and magnetic field components (section 3.3), can obviate this limitation and allow for
an analytical, unified treatment. We will confine ourselves to the discussion of the geometric dispersion (SPP effective index vs. reduced slab thickness) since, as we shall see in the following chapter, negative material parameters occur near resonances; hence most experimental realizations of LH materials considered so far were for narrow bands. In addition, all transmission (lensing) analyses of LH slab heterostructures, as well as investigations of asymmetric LH [92] and metallic [93], [94] films in the past, involved mainly monochromatic waves. All the important modal features, such as the number and classification of modes, number and kind of cutoffs, field enhancement, phase reversal and possible double mode-degeneracy occurrence, can be derived in a clear and conclusive way following this methodology.

The organization of this section is the following. Section 5.6.1 will make some introductory remarks regarding SPP waves at a single interface between a right-handed (RH) and a LH material. Emphasis is given on the conditions for the existence of such waves, as these are used later when the effects of retardation are taken into account. Section 5.6.2 is devoted to the discussion of the SPP eigenmodes supported by an asymmetric slab waveguide with a negative refractive index core. Following a macroscopic analysis, the DNGM waveguide is treated as a boundary value problem. Special solutions of the scalar wave equation are sought, subject to boundary conditions, to obtain the characteristic equation of the SPP eigenmodes. From the restrictions inherent in this equation, which depend on the refractive index distribution, we identify all supported SPP eigenmodes and classify them as forward or backward propagating via a closed-form expression for the total power flow $P$ in the guide. Finally, in section 5.6.3 we will summarize the findings of this section and we will present the main conclusions of this study.

### 5.6.1 Surface Plasmon Polaritons at a Planar LH/RH Interface

The negative permittivity and permeability of a LH medium allows for the existence of surface waves (SWs), also called surface polaritons (SPs), at the interface with a conventional dielectric. In order to investigate these solutions and derive the conditions for their existence we consider the geometry illustrated in Fig. 5.5. Here, both media are considered to be semi-infinite, homogeneous [88] and isotropic [6]. Medium 1 has negative relative permittivity $\varepsilon_{r1} = -\varepsilon_{r1}\mu < 0$ and permeability...
\( \mu_{r1} = -\mu_{r1p} < 0 \), whereas in medium 2 we assume \( \varepsilon_{r2} > 0 \) and \( \mu_{r2} > 0 \). The coordinate axes are chosen so that the z-axis is directed along the SP propagation and the x-axis is perpendicular to the media interface.

![Diagram](image)

**Figure 5.5:** Isolated interface between a left-handed (LH) material and a right-handed (RH) material. The change in sign of the permittivity allows a \( p \)-polarised surface polariton (SP) to exist at this interface.

In what follows, we will examine \( p \)-polarised (transverse magnetic, TM) SP waves that exist due to the change in the sign of the permittivities. Analogous results can be obtained for \( s \)-polarised (TE) waves, following a dual analysis. The following relations describe the field components in both dielectrics (in SI units):

\[
\begin{align*}
\frac{d^2 H_y}{dx^2} + (\varepsilon_r \mu_r k_0^2 - \beta^2) H_y &= 0, \\
E_x &= \frac{\beta}{\omega \varepsilon} H_y, \\
E_z &= -\frac{j}{\omega \varepsilon} \frac{\partial H_y}{\partial x}.
\end{align*}
\]

(5.23a)

(5.23b)

(5.23c)

Inside the LH medium, Eq. (5.23a) becomes:
where \( \beta \) is the longitudinal propagation constant of the SP wave. Assuming 
\[ \beta^2 > \max \{ \epsilon_{r1p}\mu_{r1p}k_0^2, \epsilon_{r2}\mu_{r2}k_0^2 \}, \]
the \( H_y \)-field component in medium 1 will be of the form 
\[ H_y = Ae^{\kappa x}, \]
where \( \kappa = \sqrt{\beta^2 - \epsilon_{r1p}\mu_{r1p}k_0^2} \) and \( A \) is an arbitrary constant. From 
Eq. (5.23c) it follows that the \( E_z \) component will be 
\[ E_z = (j\kappa/\omega\epsilon_0\epsilon_{r1p})e^{\kappa x}. \]
Note that the previously mentioned assumption implies a momentum mismatch between the 
supported SP and a radiative electromagnetic wave in the second dielectric, hence 
radiative waves cannot directly excite an SP at this interface.

For a bound wave to be supported by the interface, we seek \( H_y \)-solutions that 
decay exponentially as \( x \to \pm \infty \). Therefore, we seek a solution for medium 2 of the 
form \( H_y = Ce^{-\gamma x} \). By direct substitution of this expression into the scalar wave 
equation for \( H_y \) in Eq. (5.23a) we obtain: 
\[ \gamma = \sqrt{\beta^2 - \epsilon_{r2}\mu_{r2}k_0^2}, \]
with the \( E_z \) component given by 
\[ E_z = (j\gamma/\omega\epsilon_0\epsilon_{r2})e^{-\gamma x}. \]
Applying the boundary conditions associated with the tangential \( H_y \) and \( E_z \) fields at \( x = 0 \), yields a characteristic or 
eigenvalue equation for the formed SP at the plane interface that allows us to 
determine the conditions for its existence. In terms of the eigenmode's effective 
index, \( n_{\text{eff}} = \beta/k_0 \), the aforementioned equation takes the form:

\[ n_{\text{eff}} = \left[ \frac{\epsilon_{r1p}\epsilon_{r2}(\mu_{r2}\epsilon_{r2} - \mu_{r2}\epsilon_{r1p})}{\epsilon_{r2}^2 - \epsilon_{r1p}^2} \right]^{1/2}. \]  

(5.25)

It is convenient for the subsequent discussions to rewrite Eq. (5.25) using the ratios 
of the permittivities, \( \rho_{\epsilon} = \epsilon_{r2}/\epsilon_{r1p} \), and permeabilities, \( \rho_{\mu} = \mu_{r2}/\mu_{r1p} \), of the two 
media:

\[ n_{\text{eff}} = \left| n_1 \left[ \frac{\rho_{\epsilon}(\rho_{\epsilon} - \rho_{\mu})}{\rho_{\epsilon}^2 - 1} \right]^{1/2} \right| \]  

(5.26)
Since \( \varepsilon_e \) is a positive quantity and (always) \( n_{\text{eff}} > |n_1| \), we conclude from Eq. (5.26) that a SP at a LH/RH interface can only exist if:

\[
\{ \varepsilon_e \mu_p < 1 \ \text{and} \ : \ (\varepsilon_e > 1, \ \varepsilon_e > \mu_p) \},
\]

(5.27a)

or:

\[
\{ \varepsilon_e \mu_p < 1 \ \text{and} \ : \ (\varepsilon_e < 1, \ \varepsilon_e < \mu_p) \}.
\]

(5.27b)

Before closing this subsection, we wish to emphasise that these restrictions concern uncoupled (“unretarded”) SPs existing at isolated LH/RH interfaces. We demonstrate in the next section that an SP eigenmode violating these constraints may exist if the interface that supports it, is brought sufficiently close to another LH/RH interface creating a new ‘supermode’, which is not obliged to obey the two different cases in Eq. (5.27).

### 5.6.2 Surface Plasmon Polaritons in Asymmetric LH Slab Waveguides

In the following we study surface plasmon polaritons propagating along a homogeneous isotropic slab of negative permittivity and permeability bounded asymmetrically by two dielectric media, as illustrated in Fig. 5.6. The SPP eigenmodes in the slab waveguide will be travelling along the \( z \) direction. There is no variation in the guide geometry in the \( z \) direction and by symmetry no variation in the

![Figure 5.6: Schematic representation of the asymmetric left-handed slab heterostructure. The core is a medium with negative refractive index \( n_1 \) and thickness \( 2a \). Also shown are a possible field pattern of a supported SP and the direction of the longitudinal propagation constant \( \beta \).](image-url)

78
field distributions in the $y$ direction. The thickness of the slab is $2\alpha$ and in all the subsequent discussions we assume, without loss of generality, that $n_2 > n_3$.

In the analysis of planar dielectric waveguides [57] the solution ansatz to the master equation is a monochromatic plane wave of frequency $\omega$ with a functional expression that can be symbolically written as:

$$\Psi(x, z, t) = \Phi(x) e^{-j\beta z} e^{j\omega t},$$

(5.28)

where $\Psi$ represents an electric or magnetic field component, $\Phi$ describes its amplitude in the $x$-axis and $\beta$ is the longitudinal component of the wavevector $k$ in the slab. For TM SPPs, where the three existing field components are given by Eq. (5.23), we are seeking $H_y$-solutions in the three media, of the form:

$$H_y(x) = \begin{cases} 
A e^{j\beta x}, & x \leq 0 \\
B \cosh(\kappa x) + C \sinh(\kappa x), & 0 \leq x \leq 2\alpha, \\
D e^{-(x-2\alpha)/2}, & x \geq 2\alpha
\end{cases}$$

(5.29)

with $\kappa = \sqrt{\beta^2 - \varepsilon_{r1} \mu_{r1} k_0^2}$, $\gamma_2 = \sqrt{\beta^2 - \varepsilon_{r2} \mu_{r2} k_0^2}$ and $\gamma_3 = \sqrt{\beta^2 - \varepsilon_{r3} \mu_{r3} k_0^2}$, as in Section 5.6.1. Similarly to the single interface case we require $n_{eff} > \max\{|n_1|, n_2, n_3\}$, from which we make the ansatz to Eq. (5.29). By matching the tangential components at $x = 0$ and $x = 2\alpha$, we find:

$$B = A,$$

(5.30a)

$$C = -\frac{\varepsilon_{r1} \gamma_3}{\varepsilon_{r3} \kappa} A,$$

(5.30b)

$$D = \left[ \cosh(2\alpha \kappa) - \frac{\varepsilon_{r1} \gamma_3}{\varepsilon_{r3} \kappa} \sinh(2\alpha \kappa) \right] A,$$

(5.30c)

and the following SPP characteristic equation is obtained:
\[
\tanh(2\alpha K) = \frac{\varepsilon_{r1p} \kappa (\varepsilon_{r3}\gamma'_2 + \varepsilon_{r2}\gamma'_3)}{\varepsilon_{r2}\varepsilon_{r3}\kappa^2 + \varepsilon_{r1p}^2 \gamma'_2 \gamma'_3}.
\]  

(5.31)

As in classical fiber theory [57], it is advantageous to introduce the following reduced, dimensionless, modal parameters:

\[
U = \alpha K = \alpha k_0 \sqrt{n_{\text{eff}}^2 - \varepsilon_{r1p} \mu_{r1p}}, \quad (5.32a)
\]

\[
W_2 = \alpha \gamma_2 = \alpha k_0 \sqrt{n_{\text{eff}}^2 - \varepsilon_{r2} \mu_{r2}}, \quad (5.32b)
\]

\[
W_3 = \alpha \gamma_3 = \alpha k_0 \sqrt{n_{\text{eff}}^2 - \varepsilon_{r3} \mu_{r3}}. \quad (5.32c)
\]

With these definitions, Eq. (5.31) takes the form:

\[
\tanh(2U) = \frac{\varepsilon_{r1p} U (\varepsilon_{r3} W_2 + \varepsilon_{r2} W_3)}{\varepsilon_{r2} \varepsilon_{r3} U^2 + \varepsilon_{r1p}^2 W_2 W_3}.
\]  

(5.33)

In order to determine the power propagation direction, we calculate the cycle-averaged power flow in the slab heterostructure, obtained by the integral over the guide’s cross-section of the z component of the complex Poynting vector \(S_z\):

\[
P = \int_{-\infty}^{\infty} S_z dx = \frac{1}{2} \int_{-\infty}^{\infty} \text{Re}(\mathbf{E} \times \mathbf{H}^*) dx.
\]  

(5.34)

For p-polarized SPP eigenmodes, \(S_z\) is given by:

\[
S_z = \frac{1}{2} H_y^* E_z = \frac{\beta}{2 \omega \varepsilon_0 \varepsilon_r} |H_y|^2 \quad (t = 1, 2, 3). \quad (5.35)
\]

From Eqs. (5.29)-(5.31) we find the power, \(P_i\), confined in each region to be:

\[
P_i = \left( \frac{A^2}{4 \omega \varepsilon_0} \right) \frac{1}{\varepsilon_{r1p} \sigma_y} \frac{\beta}{\varepsilon_{r1p} \sigma_y},
\]  

(5.36a)
\[ P_1 = \left( \frac{A^2}{4 \omega \varepsilon_0} \right) \frac{\beta \sigma^2_2 \kappa^2 - \gamma_3^2}{\sigma^2_2 \kappa^2} \left[ 2 \alpha + \frac{\rho_2 \gamma_2}{\rho^2_2 \kappa^2 - \gamma_2^2} + \frac{\sigma_3 \gamma_3}{\sigma^2_2 \kappa^2 - \gamma_3^2} \right], \quad (5.36b) \]
\[ P_2 = \left( \frac{A^2}{4 \omega \varepsilon_0} \right) \frac{\beta \rho_2^2 \sigma^2_2 \kappa^2 - \gamma_3^2}{\rho^2_2 \sigma^2_2 \kappa^2 - \gamma_2^2}, \quad (5.36c) \]

where \( \sigma_\varepsilon = \varepsilon_{r1}/\varepsilon_{r1p} \) and \( \rho_\varepsilon = \varepsilon_{r2}/\varepsilon_{r1p} \). From Eq. (5.36) we can derive a closed-form expression for the total power \( P_{\text{tot}} = \sum_{i=1}^{3} P_i \) in terms of the dimensionless parameters defined in Eq. (5.32) and the reduced slab thickness, \( a_k \):

\[ P_{\text{tot}} = \left( \frac{A^2}{4 \omega \varepsilon_0} \right) \frac{W_3^2 - \sigma_\varepsilon^2 U^2}{\sigma^2_2 U^2} \frac{U^2 + \varepsilon_{r1p} \mu_{r1p} (a_k)^2}{\varepsilon_{r1p}} \left[ 2 + \frac{P_\varepsilon}{W_2} \frac{U^2 - W_2^2}{W_2^2 - \rho_\varepsilon^2 U^2} + \frac{\sigma_\varepsilon}{W_3} \frac{U^2 - W_3^2}{W_3^2 - \sigma_\varepsilon^2 U^2} \right]. \quad (5.37) \]

The central task at this point is the determination of the solutions to Eq. (5.33). Their existence is identified following an analytical methodology. In what follows, we discuss the features and dependence of these solutions on the thickness of the inner layer, for the various cases of the refractive index distribution. A summary of the results for each case is found in Section 5.6.3, in Tables I-III.

A. Case I: \(|n_1| > n_2 > n_3\)

For this case, we start by defining the following two \( V \)-parameters:

\[ V_2(a_k) = (W_2^2 - U^2)^{1/2} = a_k \left( \varepsilon_{r1p} \mu_{r1p} - \varepsilon_{r2} \mu_{r2} \right)^{1/2}, \quad (5.38a) \]
\[ V_3(a_k) = (W_3^2 - U^2)^{1/2} = a_k \left( \varepsilon_{r1p} \mu_{r1p} - \varepsilon_{r3} \mu_{r3} \right)^{1/2}, \quad (5.38b) \]

where it is seen that the usual notation of the \( V \)-number found in fiber theory [57] has been properly modified to accommodate the changes in the refractive index.
distribution and that both parameters are functions of \( \alpha k_0 \). We also introduce the following ratios:

\[
b(n_{\text{eff}}) = \frac{U}{V_2} = \left[ \frac{(n_{\text{eff}}/n_i)^2 - 1}{1 - \rho_e \rho_p} \right]^{1/2},
\]

\[
t = \frac{V_3}{V_2} = \left( \frac{1 - \sigma \sigma_p}{1 - \rho_e \rho_p} \right)^{1/2},
\]

(5.39)

(5.40)

obeying the restrictions \( b > 0 \) and \( t > 1 \), where \( \rho_p = \mu_r / \mu_{r1p} \) and \( \sigma_p = \mu_r / \mu_{r1p} \).

Note that similar ratios to \( b \) and \( t \) are utilized in the analysis of conventional slab waveguides to denote the "normalized guide index" and the "asymmetry measure", respectively [95] and that \( b \) is a function of the SPP eigenmode's effective index.

For the explicit acquisition of the dispersion diagrams and the derivation of the analytical restrictions inherent in Eq. (5.33), a common strategy is to produce an inverted version of the associated characteristic equation [94], [95]. First, we note from Eqs. (5.38) and (5.39) that

\[
U = b V_2, \quad W_2 = V_2 \left( b^2 + 1 \right)^{1/2} \text{ and } W_3 = V_2 \left( b^2 + t^2 \right)^{1/2}.
\]

With these observations, Eq. (5.33) can be rewritten in the form:

\[
V_2 = \frac{1}{4b} \ln \left[ \frac{(X + 1)(Y + 1)}{(X - 1)(Y - 1)} \right],
\]

(5.41)

where \( X(b) = \sigma_e b / \left( b^2 + t^2 \right)^{1/2} \) and \( Y(b) = \rho_e b / \left( b^2 + 1 \right)^{1/2} \). Since \( V_2 \), given in Eq. (5.38a), is a real number, we immediately see that Eq. (5.41) only has solutions, when the argument of the logarithm is positive, i.e. for:

\[
\{ X > 1 \text{ and } Y > 1 \}, \quad (5.42a)
\]

or:

\[
\{ 0 < X < 1 \text{ and } 0 < Y < 1 \}. \quad (5.42b)
\]
We now examine in detail the consequences of these restrictions on $X(b)$ and $Y(b)$, based on which the existing SPP eigenmodes are rigorously identified. In particular, we derive analytically the allowable range of values that $b$, hence the eigenmodes’ effective index, can take. In pursuing this analysis, we find that it is necessary to distinguish between the following four situations that describe the possible variations in the permittivity distribution. In all four of them, it is implied that $\sigma_\varepsilon \sigma_\mu < \rho_\varepsilon \rho_\mu < 1$ from the initial assumption $|n_1| > n_2 > n_3$.

The first situation occurs for $\{\sigma_\varepsilon > 1 \text{ and } \rho_\varepsilon > 1\}$. When $\rho_\varepsilon > \sigma_\varepsilon$, inspection of Eq. (5.42) results in the allowable values for $b$; that is $0 < b < b_1$ or $b > b_2$, where:

$$b_1 = \left(\frac{1}{\rho_\varepsilon^2 - 1}\right)^{1/2}, \quad (5.43)$$

and:

$$b_2 = \left[\frac{1 - \sigma_\varepsilon \sigma_\mu}{(\sigma_\varepsilon^2 - 1)(1 - \rho_\varepsilon \rho_\mu)}\right]^{1/2}. \quad (5.44)$$

The corresponding geometric dispersion diagram is shown in Fig. 5.7(a), and the variation of the normalised power $P = P_{\text{tot}}/(|P_1| + |P_2| + |P_3|)$ with the reduced slab thickness for each solution is given in Fig. 5.7(b). Three SPP eigenmodes exist in this case. The first has a lower cutoff and an upper one at $b = 0$, is forward propagating, having positive total power $P$, and the field intensity has a node in the core region. At the lower cutoff point this solution degenerates into the second eigenmode. This SPP has a low cutoff and no upper cutoff and is backward propagating, having negative energy velocity and, since the dielectric media are non-absorbing, negative group velocity, as well [96], [97]. After a finite gap a third eigenmode appears, which has no cutoff and is also backward propagating having negative $P$ for every core thickness. The corresponding fields have no node in the core region.

It is interesting to discuss the behaviour of these solutions in the extreme cases of $b \rightarrow b_1$ and $b \rightarrow b_2$. First, by letting $b$ become equal to $b_1$, we recover asymptotically
Effective index $n_{ef}$

Reduced slab thickness $\alpha k_0$

(a)

(b)

Reduced slab thickness $\alpha k_0$

(c)

(d)

Reduced slab thickness $\alpha k_0$

(e)

(f)
Figure 5.7: Variation of an SP eigenmode’s effective index $n_{\text{eff}}$ and normalised power $P$ with the reduced slab thickness $\alpha k_0$ in a generalised LH slab waveguide (case I: $|n_1| > n_2 > n_3$, as indicated by the shaded background). In all cases it is assumed that $\varepsilon_r = 2$, $\mu_r = 1.2$. (a), (b) $n_{\text{eff}}$ and $P$ variation for $\sigma_\varepsilon = 1.1$, $\sigma_\mu = 0.5$, $\rho_\varepsilon = 1.15$, $\rho_\mu = 0.6$. (c), (d) Variations for $\sigma_\varepsilon = 1.8$, $\sigma_\mu = 0.5$, $\rho_\varepsilon = 1.55$, $\rho_\mu = 0.6$. (e), (f) Variations for $\sigma_\varepsilon = 1.1$, $\sigma_\mu = 0.5$, $\rho_\varepsilon = 0.95$, $\rho_\mu = 0.8$. (g), (h) Variations for $\sigma_\varepsilon = 1.1$, $\sigma_\mu = 0.2$, $\rho_\varepsilon = 1.05$, $\rho_\mu = 0.2$. (i), (j) Variations for $\sigma_\varepsilon = 0.8$, $\sigma_\mu = 1$, $\rho_\varepsilon = 0.9$, $\rho_\mu = 1.1$.

Eq. (5.26), which is the SPP characteristic equation at the 1-2 media interface. Such a result is expected, since from Eqs. (5.38a) and (5.41) and from Fig. 5.7(a) we note
that for $b \to b_1$ it is $ak_0 \to \infty$, hence the two slab interfaces decouple and the SPP at an isolated interface should be recovered. For a relatively large value of $ak_0$, the result of plotting this SPP eigenmode is shown in the middle right inset of Fig. 5.7(a), reflecting the previous conclusions. Then, assuming that $b = b_2$, we obtain:

$$n_{\text{eff}3} = n_1 \sqrt{\frac{\sigma_e (\sigma_e - \sigma_\mu)}{\sigma_e^2 - 1}}^{1/2},$$

and the SPP eigenmode at the 1-3 interface is recovered asymptotically, shown in the top inset.

When $\rho_e < \sigma_e$, it proves necessary to examine the intervals that the ratio $\rho_e$ belongs to. If $\rho_e \in \Delta$, where $\Delta = \{(\rho_{e,1} < \rho_e < \rho_{e,2}) \cap (\rho_e > 1)\}$ and $\rho_{e,1}$, $\rho_{e,2}$ are the two roots of the polynomial:

$$\Pi(\rho_e) = (1 - \sigma_e \sigma_\mu) \rho_e^2 + \rho_\mu (\sigma_e^2 - 1) \rho_e - [(1 - \sigma_e \sigma_\mu) + (\sigma_e^2 - 1)],$$

it is $b_1 > b_2$, and we see from Fig. 5.7(c) that two eigenmodes exist, both of which are backward propagating. The first SPP has only a lower cutoff at $b = 0$ and, contrary to the previous case, it concentrates asymptotically at the 1-3 interface while exhibiting a phase reversal. The second (upper) SPP shows no cutoff, has positive $H_y$-field component throughout the slab and concentrates at the 1-2 media interface for large core thickness. For $\rho_e \in \Delta'$, where $\Delta' = \{(\rho_e < \rho_{e,1}) \cup (\rho_e > \rho_{e,2})\} \cap (\rho_e > 1)$, we have $b_1 < b_2$ and the results of the analysis hold the same as in the case $\rho_e > \sigma_e$, apart from the disappearance of the first SPP eigenmode that was forward propagating.

The second situation occurs for $\{\sigma_e > 1 \text{ and } \rho_e \leq 1\}$. From the conditions in Eq. (5.42), we find that $b$ ranges from 0 to $b_2$ and the corresponding dispersion diagram is illustrated in Fig. 5.7(e). In this case two eigenmodes are shown to exist; for both, the $H_y$ component has a node in the core region. The first has a low cutoff and an upper one occurring at $b = 0$, and is forward propagating. The second SPP has only a low cutoff and it concentrates asymptotically at the 1-3 interface. As in all the encountered
cases that contain degeneracy, the total power $P$ at this point equals zero, corresponding to zero group velocity. Waves with this feature are of great practical interest for optical communication and data storage applications [19]. It should be noted that, for the particular values of the permittivity and permeability ratios shown in Fig. 5.7(e), the 1-2 single interface does not support an SPP because the constraints in Eq. (5.27) are violated. However, the coupled SPPs at the interfaces of the slab overcome this limitation, creating the two new SPP ‘supermodes’ for relatively large slab thickness.

In the third case, which exists for $\{\sigma_e \leq 1$ and $\rho_e > 1\}$, it is found that the range of values for $b$ is between 0 and $b_1$. The variation of the eigenmodes’ effective index and power with the reduced guide thickness is shown in Figs. 5.7(g) and (h). The conclusions for the supported SPPs are similar with those in the previous case, with the difference that the second (upper) eigenmode concentrates asymptotically at the 1-2 interface while taking negative values. In this case also, it is the 1-3 interface that violates the SPP existence conditions and, if isolated, would not support a bound wave.

The final situation occurs for $\{\sigma_e < 1$ and $\rho_e < 1\}$. In this case there are no additional restrictions on $b$ other than $b > 0$. From Figs. 5.7(i) and (j) we see that only one SPP eigenmode exists, which is forward propagating with only a higher cutoff at $b = 0$ and has opposite sign at the two interfaces of the slab.

B. Case II: $n_2 > n_3 > |n_1|$

For the refractive index distribution considered here, we use the following definitions for the $V$-numbers:

$$V_2(ak_0) = (U^2 - W_2^2)^{1/2} = ak_0(e_{r2}e_{r2} - e_{r1p}e_{r1p})^{1/2}, \quad (5.47a)$$
$$V_3(ak_0) = (U^2 - W_3^2)^{1/2} = ak_0(e_{r3}e_{r3} - e_{r1p}e_{r1p})^{1/2}, \quad (5.47b)$$

with $\rho_e\rho_p > \sigma_e\sigma_p > 1$ used throughout the following analysis. Then, the previously introduced $b$ and $t$ parameters take the form: 87
obeying the restrictions $b > 1$ and $t < 1$. The general form of Eq. (5.41) and the solution conditions of Eq. (5.42) remain the same, but now we have $W_2 = V_2 \left( b^2 - 1 \right)^{1/2}$, $W_3 = V_2 \left( b^2 - t^2 \right)^{1/2}$, $X(b) = \sigma_x b / \left( b^2 - t^2 \right)^{1/2}$ and $Y(b) = \rho_x b / \left( b^2 - 1 \right)^{1/2}$. By letting $X(b)$ and $Y(b)$ fulfill these conditions, we again find that four distinct situations arise depending on the permittivity profile.

The first situation occurs for $\{\sigma_x \geq 1 \text{ and } \rho_x \geq 1\}$, and does not contain further restrictions on $b$. The eigenmodes' effective index and $P$ dispersion diagrams are shown in Figs. 5.8(a) and (b). We see that two SPPs exist; both have no node in the middle layer. The first, which is forward propagating, has a lower cutoff at $b = 1$ and also an upper cutoff point, where it degenerates into the second eigenmode. The second SPP has only a high cutoff and is backward propagating having negative total power $P$.

The second situation occurs for $\{\sigma_x \geq 1 \text{ and } \rho_x < 1\}$ and it can be shown that $b$ is in the range $1 < b < b_1$, where now:

$$b_1 = \left( \frac{1}{1 - \rho_x^2} \right)^{1/2}. \quad (5.50)$$

In this case a single SPP eigenmode is shown to exist, having only a low cutoff at $b = 1$. This SPP is forward propagating and it concentrates at the 1-2 media interface for large core thickness, as illustrated in the inset of Fig. 5.8(c). This result is also verified by letting $b = b_1$, where we obtain the SPP characteristic Eq. (5.26).

The third case exists for $\{\sigma_x < 1 \text{ and } \rho_x \geq 1\}$ and it is found that $b$ ranges from 1 to $b_2$, where now:
Figure 5.8: Variation of an SP eigenmode’s effective index $n_{\text{eff}}$ and reduced power $P$ with the reduced slab thickness $ak_0$ in a generalised LH slab waveguide (case II: $n_2 > n_3 > |n_1|$, as indicated by the shaded background). In all cases it is assumed that $\varepsilon_r = 2, \mu_r = 1.2$. (a), (b) $n_{\text{eff}}$ and $P$ variation for $\sigma_\varepsilon = 1.7, \sigma_\mu = 1.1, \rho_\varepsilon = 1.8, \rho_\mu = 1.2$. (c), (d) Variations for $\sigma_\varepsilon = 1.2, \sigma_\mu = 0.9, \rho_\varepsilon = 0.8, \rho_\mu = 1.6$. (e), (f) Variations for $\sigma_\varepsilon = 0.9, \sigma_\mu = 1.2, \rho_\varepsilon = 1.1, \rho_\mu = 1.1$. (g), (h) Variations for $\sigma_\varepsilon = 0.7, \sigma_\mu = 1.6, \rho_\varepsilon = 0.71, \rho_\mu = 1.7$. (i), (j) Variations for $\sigma_\varepsilon = 0.7, \sigma_\mu = 2.7, \rho_\varepsilon = 0.4, \rho_\mu = 5.8$. (k), (l) Variations for $\sigma_\varepsilon = 0.7, \sigma_\mu = 2.7, \rho_\varepsilon = 0.85, \rho_\mu = 3.5$.

$$b_2 = \left[ \frac{\sigma_\varepsilon \sigma_\mu - 1}{(1-\sigma_\varepsilon^2)(\rho_\varepsilon \rho_\mu - 1)} \right]^{1/2}$$

(5.51)

with the constraint $\rho_\varepsilon \rho_\mu < (\sigma_\varepsilon \sigma_\mu - \sigma_\varepsilon^2)/(1-\sigma_\varepsilon^2)$. The variation of the effective index and total power with the reduced guide thickness are shown in Figs. 5.8(e) and (f). The conclusions are similar with the previous case, with the difference that the supported SPP concentrates asymptotically at the 1-3 interface. For the values of $\rho_\varepsilon$ and $\rho_\mu$ shown in Fig. 5.8(e), the isolated 1-2 interface would not support an SPP.
The final situation occurs for \( \{\sigma_e < 1 \text{ and } \rho_e < 1\} \). If \( \rho_e \rho_p < \left( \sigma_e \sigma_p - \sigma_e^2 \right) / (1 - \sigma_e^2) \), it proves necessary to examine the intervals that the ratio \( \rho_e \) belongs to. When \( \rho_e \in A \), with \( A = \{(\rho_e < \rho_{e,1}) \cup (\rho_e > \rho_{e,2})\} \cap (0 < \rho_e < 1) \) and \( \rho_{e,1}, \rho_{e,2} \) being the two roots of the polynomial:

\[
H(\rho_e) = (\sigma_e \sigma_p - 1) \rho_e^2 + \rho_p (1 - \sigma_e^2) \rho_e - \left[ (1 - \sigma_e^2) + (\sigma_e \sigma_p - 1) \right],
\]

it is \( b_1 > b_2 \), and we see from Fig. 5.8(g) that two eigenmodes exist, both of which are forward propagating. The first SPP has only a low cutoff at \( b = 1 \) and concentrates asymptotically at the 1-3 interface; the corresponding field intensity has no node in the core region. The second SPP has no cutoff, exhibits phase reversal and concentrates at the 1-2 media interface for large core thickness. For \( \rho_e \in A' \), where \( A' = \{ (\rho_{e,1} < \rho_e < \rho_{e,2}) \cap (0 < \rho_e < 1) \} \), we have \( b_1 < b_2 \) and again two SPP eigenmodes are shown to exist in Figs. 5.8(i) and (j). The first has positive \( H_y \) field amplitude across the slab. Compared with the previous two eigenmodes, the cutoff characteristics remain the same, but the order of the interfaces to which these SPPs concentrate is reversed. Finally, if \( \rho_e \rho_p > \left( \sigma_e \sigma_p - \sigma_e^2 \right) / (1 - \sigma_e^2) \), the constraint \( b > b_1 \) becomes mandatory. The dispersion diagrams for this case are shown in Figs. 5.8(k) and (l). We see that a single SPP exists, which shows no cutoff, exhibits phase reversal, concentrates asymptotically at the 1-2 interface and has positive total power for all core thicknesses.

C. Case III: \( n_2 > |n_1| > n_3 \)

To reflect the refractive index distribution considered here, the two \( V \)-numbers are defined as:

\[
V_2(ak_0) = (U^2 - W_2^2)^{1/2} = ak_0 \left( \varepsilon_{r,2} \mu_{r,2} - \varepsilon_{r,1,3} \mu_{r,1,3} \right)^{1/2},
\]

\[
V_3(ak_0) = (W_3^2 - U^2)^{1/2} = ak_0 \left( \varepsilon_{r,1,3} \mu_{r,1,3} - \varepsilon_{r,3} \mu_{r,3} \right)^{1/2},
\]

(5.53a)

(5.53b)
with \( \rho_e \rho_\mu > 1 > \sigma_e \sigma_\mu \) used in the remaining analysis. Accordingly, the \( b \) and \( t \) ratios take the form:

\[
b(n_{\text{eff}}) = \frac{U}{V_2} = \left[ \frac{(n_{\text{eff}}/n_1)^2 - 1}{\rho_e \rho_\mu - 1} \right]^{1/2},
\]

(5.54)

\[
t = \frac{V_1}{V_2} = \left( \frac{1 - \sigma_e \sigma_\mu}{\rho_e \rho_\mu - 1} \right)^{1/2},
\]

(5.55)

obeying the restrictions \( b > 1 \) and \( t > 0 \). Once more, the general form of Eq. (5.41) and the solution conditions of Eq. (5.42) remain the same, but now

\[
W_2 = V_2 \left( b^2 - 1 \right)^{1/2},
\]

\[
W_3 = V_2 \left( b^2 + t^2 \right)^{1/2}, \quad X(b) = \sigma_e b / \left( b^2 + t^2 \right)^{1/2} \quad \text{and} \quad Y(b) = \rho_e b / \left( b^2 - 1 \right)^{1/2}.
\]

By letting \( X(b) \) and \( Y(b) \) fulfil the aforesaid conditions, the following four distinct situations arise depending on the permittivity profile.

The first situation occurs for \( \{ \sigma_e > 1 \text{ and } \rho_e \geq 1 \} \). If \( \rho_e \rho_\mu < \left( \sigma_e^2 - \sigma_e \sigma_\mu \right) / \left( \sigma_e^2 - 1 \right) \), the allowable range of values for \( b \) is \( b > b_2 \), where:

\[
b_2 = \left[ \frac{1 - \sigma_e \sigma_\mu}{(\rho_e \rho_\mu - 1)(\sigma_e^2 - 1)} \right]^{1/2}.
\]

(5.56)

We see from Figs. 5.9(a) and (b) that a single eigenmode exists, which is backward propagating and has no cutoff. For large core thickness, the SPP characteristic equation at the 1-3 interface is recovered. The other interface would not support a SPP, if isolated. In the case where \( \rho_e \rho_\mu > \left( \sigma_e^2 - \sigma_e \sigma_\mu \right) / \left( \sigma_e^2 - 1 \right) \), it is seen from Figs. 5.9(c) and (d) that two SPP eigenmodes can exist; both have no node in the inner layer. The first one, which is forward propagating, has a low cutoff point at \( b = 1 \) and a high cutoff, where it degenerates into the second SPP. This eigenmode is backward propagating having only an upper cutoff.

The second situation is described by \( \{ \sigma_e > 1 \text{ and } \rho_e < 1 \} \). For \( \rho_e \rho_\mu < \left( \sigma_e^2 - \sigma_e \sigma_\mu \right) / \left( \sigma_e^2 - 1 \right) \), it proves necessary to examine the intervals that the ratio
Figure 5.9: Variation of an SP eigenmode’s effective index $n_{\text{eff}}$ and reduced power $P$ with the reduced slab thickness $a k_0$ in a generalised LH slab waveguide (case III: $n_2 > |n_1| > n_3$, as indicated by the shaded background). In all cases it is assumed that $\varepsilon_r = 2$, $\mu_r = 1.2$. (a), (b) $n_{\text{eff}}$ and $P$ variation for $\sigma_\varepsilon = 1.1$, $\sigma_\mu = 0.7$, $\rho_\varepsilon = 1.38$, $\rho_\mu = 1.5$. (c), (d) Variations for $\sigma_\varepsilon = 1.1$, $\sigma_\mu = 0.7$, $\sigma_\varepsilon = 1.1$, $\sigma_\mu = 0.7$, $\rho_\varepsilon = 1.95$, $\rho_\mu = 1.5$. (e), (f) Variations $\sigma_\varepsilon = 1.1$, $\sigma_\mu = 0.7$, $\sigma_\varepsilon = 1.1$, $\sigma_\mu = 0.7$, $\rho_\varepsilon = 0.75$, $\rho_\mu = 2.5$. (g), (h) Variations for $\sigma_\varepsilon = 1.1$, $\sigma_\mu = 0.7$, $\rho_\varepsilon = 0.5$, $\rho_\mu = 2.5$. (i), (j) Variations for $\sigma_\varepsilon = 1.1$, $\sigma_\mu = 0.7$, $\rho_\varepsilon = 0.92$, $\rho_\mu = 2.5$. (k), (l) Variations for $\sigma_\varepsilon = 0.9$, $\sigma_\mu = 0.9$, $\rho_\varepsilon = 0.95$, $\rho_\mu = 1.2$. 

$\rho_\varepsilon$ belong to. When $\rho_\varepsilon \in \Delta$, with $\Delta = \{(\rho_\varepsilon < \rho_{\varepsilon,1}) \cup (\rho_\varepsilon > \rho_{\varepsilon,2})\} \cap (0 < \rho_\varepsilon < 1)$ and $\rho_{\varepsilon,1}$, $\rho_{\varepsilon,2}$ being the two roots of the polynomial:

$$
\Pi(\rho_\varepsilon) = (1 - \sigma_\varepsilon \sigma_\mu) \rho_\varepsilon^2 + \rho_\mu (\sigma_\varepsilon^2 - i) \rho_\varepsilon - \left[(1 - \sigma_\varepsilon^2) + (1 - \sigma_\varepsilon \sigma_\mu)\right],
$$

(5.57)

it is $b_1 > b_2$, where:
\[ b_1 = \left( \frac{1}{1-\rho_e^2} \right)^{1/2}, \]  

(5.58)

and it can be shown that \( b \) is in the range \( b_2 < b < b_1 \). The corresponding dispersion diagrams are illustrated in Figs. 5.9(e) and (f). We see that two node-less eigenmodes exist. The first one has only a low cutoff, is backward propagating and concentrates asymptotically at the 1-3 interface. The second SPP has only a low cutoff, is forward propagating and concentrates at the 1-2 interface for relatively large core thicknesses. For \( \rho_e \in \mathcal{A}' \), where \( \mathcal{A}' = \{(\rho_{e,1} < \rho_e < \rho_{e,2}) \cap (0 < \rho_e < 1)\} \), it is \( b_1 < b_2 \) and again two SPP eigenmodes are shown to exist in Figs. 5.9(g) and (h). Compared with the previous two eigenmodes, the cutoff characteristics and the classification remain the same, but the order of the interfaces to which these SPPs concentrate is reversed. Also, both SPPs now exhibit phase reversal. For \( \rho_e \rho_\mu > (\sigma_{\epsilon}^2 - \sigma_\mu \sigma_\mu)/(\sigma_\epsilon^2 - 1) \), we have \( 1 < b < b_1 \) and the dispersion diagrams are shown in Figs. 5.9(i) and (j). In this case, three distinct solutions of Eq. (5.33) are found; the \( H_y \) field intensity for all of them has no node in the core region. The first one, which corresponds to positive total power \( P \), has a low cutoff point at \( b = 1 \) and an upper cutoff, where it degenerates into the second SPP. This eigenmode has also a lower and a higher cutoff, but is backward propagating. The third branch shows only a low cutoff, corresponds to positive \( P \) and, asymptotically, degenerates into the isolated 1-2 interface solution. This situation is the only one where two double mode-degeneracy points occur.

In the third situation, which occurs for \( \sigma_\epsilon < 1 \) and \( \rho_e > 1 \), there are no solutions to Eq. (5.41) for any \( b > 1 \), hence the slab waveguide does not support SPPs.

The final situation occurs for \( \sigma_\epsilon \leq 1 \) and \( \rho_e < 1 \) and the corresponding dispersion diagrams are shown in Figs. 5.9(k) and (l). From the solution constraints of Eq. (5.42), we find that \( b > b_1 \). A single SPP exists, which does not have cutoff, is forward propagating, exhibits phase reversal and concentrates asymptotically at the 1-2 media interface. It should be noted that, for the chosen values of \( \sigma_\epsilon \) and \( \sigma_\mu \), the isolated 1-3 interface does not support an SPP.
5.6.3 Summary of the Identified SPP Solutions

We have presented a systematic investigation of all solutions of the involved characteristic equation giving the $p$-polarised surface plasmon polariton eigenmodes in generalised slab waveguides, comprised of a negative refractive index core that is bounded by two different positive-index dielectrics. Following an analytical methodology, all SPPs are classified as forward or backward propagating, depending on the sign of the associated power flow. Contrary to the case of thin metallic films

### TABLE I. Summary of the discussion for the case $|n_1| > n_2 > n_3$.

<table>
<thead>
<tr>
<th>$\rho_e \geq 1, \rho_e &gt; 1$</th>
<th>$\sigma_e &gt; 1, \rho_e &gt; 1$</th>
<th>$\sigma_e &gt; 1, \rho_e \leq 1$</th>
<th>$\sigma_e \leq 1, \rho_e &gt; 1$</th>
<th>$\sigma_e \leq 1, \rho_e \leq 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_e = \sigma_e$</td>
<td>$\rho_e &gt; \sigma_e$</td>
<td>$\rho_e \in \mathcal{D}$</td>
<td>$\rho_e \in \mathcal{D}$</td>
<td>$\rho_e \in \mathcal{D}$</td>
</tr>
<tr>
<td>2nd: backward, low cutoff.</td>
<td>2nd: backward, low cutoff.</td>
<td>2nd: backward, no cutoff.</td>
<td>2nd: backward, low cutoff.</td>
<td>2nd: backward, low cutoff.</td>
</tr>
</tbody>
</table>

### TABLE II. Summary of the discussion for the case $n_2 > n_3 > |n_1|$.

<table>
<thead>
<tr>
<th>$\sigma_e \geq 1, \rho_e \geq 1$</th>
<th>$\sigma_e \geq 1, \rho_e &lt; 1$</th>
<th>$\sigma_e \leq 1, \rho_e \geq 1$</th>
<th>$\sigma_e &lt; 1, \rho_e &lt; 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_e \in \mathcal{D}$</td>
<td>$\rho_e \in \mathcal{D}'$</td>
<td>$\rho_e \in \mathcal{D}'$</td>
<td>$\rho_e \in \mathcal{D}'$</td>
</tr>
<tr>
<td>1st: forward, 2 cutoffs.</td>
<td>1st: forward, low cutoff.</td>
<td>1st: forward, low cutoff.</td>
<td>1st: forward, low cutoff.</td>
</tr>
<tr>
<td>2nd: backward, upper cutoff.</td>
<td>2nd: forward, no cutoff.</td>
<td>2nd: forward, no cutoff.</td>
<td>2nd: forward, no cutoff.</td>
</tr>
</tbody>
</table>

### TABLE III. Summary of the discussion for the case $n_2 > |n_1| > n_3$.

<table>
<thead>
<tr>
<th>$\sigma_e &gt; 1, \rho_e &gt; 1$</th>
<th>$\sigma_e &gt; 1, \rho_e &lt; 1$</th>
<th>$\sigma_e \leq 1, \rho_e \leq 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_e \rho - \sigma_e^2 \rho - \sigma_e^2 \rho \in \mathcal{D}$</td>
<td>$\rho_e \rho - \sigma_e^2 \rho - \sigma_e^2 \rho \in \mathcal{D}'$</td>
<td>$\rho_e \rho - \sigma_e^2 \rho - \sigma_e^2 \rho \in \mathcal{D}'$</td>
</tr>
<tr>
<td>1st: forward, no cutoff.</td>
<td>1st: forward, 2 cutoffs.</td>
<td>1st: forward, 2 cutoffs.</td>
</tr>
<tr>
<td>2nd: backward, upper cutoff.</td>
<td>2nd: backward, low cutoff.</td>
<td>2nd: backward, low cutoff.</td>
</tr>
<tr>
<td>3rd: forward, low cutoff.</td>
<td>3rd: forward, low cutoff.</td>
<td>3rd: forward, low cutoff.</td>
</tr>
</tbody>
</table>

98
surrounded by two different dielectrics, where the SPPs depend solely on the permittivity distribution, we have shown analytically that, for the corresponding left-handed structures considered here, the refractive index distribution must also be included in the analysis.

In particular, when the absolute value of the core refractive index is greater than those of the surrounding dielectrics we identified ten TM SPP eigenmodes (see Table I); six of these solutions are backward propagating. Nine solutions existed if the core index has the smallest absolute value (Table II); one of these SPPs has antiparallel phase and group velocities. In the final case, where the core index value is between those of the claddings, eleven SPP eigenmodes were identified (Table III); four of these solutions have backwards power flow in relation to the direction of the phase velocity. The total number of 30 SPP eigenmodes, which is significantly higher than the 8 SPPs that exist in similar metallic film geometries [93], is a direct result of the presence of negative magnetic permeability in the LH structures that provides additional degrees of freedom in the definition of the \( V \)-numbers.

The analytical treatment revealed features not observed before, such as the occurrence of 'supermodes' in the case of a violation of the isolated interface condition, strong field enhancement and opening of gaps in the geometric dispersion diagrams owing to the asymmetry, and coexistence of three eigenmodes with double mode-degeneracy points occurring twice. We also demonstrated that the group velocity of the investigated slow waves could be decreased down to zero or become negative, by adjusting suitably the thickness of the core.

5.7 Oscillatory Guided Modes Supported by Asymmetric DNGM Slab Heterostructures

In the previous section we have seen that important potential applications of LH materials in optics and microwaves, such as the possibility to create super-resolving lenses or to improve the performance of biosensing devices, prompted detailed investigations into the properties of specific LH heterostructures [88]-[92]. From this research it has been shown that LH waveguides support two classes of waves: Oscillatory or waveguide modes (OM/fast waves), which are also found in regular
dielectric slab structures, and a rich variety of surface plasmon polariton modes (SPP/slow waves). We showed that up to 30 bound SPPs can exist in a generalized LH slab waveguide for all choices of opto-geometrical parameters. Although approximately a third of the SPP solutions allow for attaining zero group velocity – a notable characteristic of guided modes in such waveguides, to which we shall turn our attention later, in Chapter 7 – their inherent sensitivity to small variations of the media interfaces [64] may limit their practical use. For this reason, oscillatory modes, which generally have their maxima inside the waveguide core, are more suitable for most conceived applications, including slow-light. Nonetheless, in most relevant studies, these modes have been described only qualitatively, following ad hoc graphic solution approaches [88], [90]. In this manner, relevant modal properties were either not revealed or not conclusively proven and, importantly, explicit expressions for the cutoff conditions of each mode have not been established yet.

In this section we shall report on an extension of waveguide mode theory that considers an asymmetric three-layer slab heterostructure in which either the guiding region or the cladding region may have a negative refractive index. We derive dimensionless, closed-form expressions for the cutoff(s) of the investigated oscillatory modes that allow for the identification of their existence regions. The modes with negative energy flux that give rise to negative group velocity are identified via an explicit expression for the cycle-averaged total power flow $P_{\text{tot}}$ in the guide. Based on this treatment we prove rigorously that, with judicious choice of parameters, there is a frequency region where the second mode can simultaneously exist alone and attain zero group velocity. Moreover, we show that the inverted (LH-RH-LH) arrangement supports similar modes as the RH-LH-RH heterostructure but with opposite power flux. To our knowledge, these are the only slab waveguide structures utilizing homogeneous, isotropic [98] media that have the potential for single-mode operation in the “slow-light regime.”

In the analysis that follows we consider the geometry illustrated in Fig. 5.10, where all media are assumed to be lossless [99], homogeneous and isotropic. We shall be concerned primarily with $p$-polarized (TM) waves, in which the magnetic field is directed along the $y$-axis. First, we investigate the case where the core is LH with thickness $2\alpha$, bounded asymmetrically by two RH media that satisfy $n_2 > n_3$. The fields in the guide consist of counterpropagating forward and backward modes.
trapped within the core by total-internal reflection. For wave guidance to occur the wave components of both types of modes must satisfy the total-internal reflection condition. This condition can be stated as \(|n_1 > n_{\text{eff}} > n_2\), where \(n_{\text{eff}}\) is the effective index of the mode. We point out here that there is a marked contrast between the present case and that of SPP waves that we studied in the previous section: For the latter case, we required \(n_{\text{eff}} > \max\{|n_1|, n_2\}\) (see remarks after Eq. (5.29)) as a result of the somewhat different solution ansatz (see Eq. (5.29)) that we therein made to the wave equation. Matching the \(E_z\) and \(H_y\) fields at the boundaries yields the following dimensionless characteristic equation for oscillatory modes in generalized LH slab waveguides:

\[
\alpha k_0 = \frac{1}{2|n_1|\sqrt{1-\rho_\sigma \rho_\mu}} \left[ \arctan \left( \frac{\rho_\sigma b}{\sqrt{1-b^2}} \right) + \arctan \left( \frac{\sigma_\sigma b}{\sqrt{t^2-b^2}} \right) + m\pi \right],
\]

where, similarly to the previous section, the following ratios are introduced:

\[
b(n_{\text{eff}}) = \frac{U}{V_2} = \left[ \frac{1 - (n_{\text{eff}}/n_1)^2}{1 - \rho_\sigma \rho_\mu} \right]^{1/2},
\]

\[
t = \frac{V_3}{V_2} = \left( \frac{1 - \sigma_\gamma \sigma_\mu}{1 - \rho_\sigma \rho_\mu} \right)^{1/2},
\]

with \(\rho_\sigma = e_{r_2}/|e_{r_1}|\), \(\rho_\mu = \mu_{r_2}/|\mu_{r_1}|\), \(\sigma_\gamma = e_{r_3}/|e_{r_1}|\), \(\sigma_\mu = \mu_{r_3}/|\mu_{r_1}|\) and \(k_0\) being the free-space wavevector. In this section, the two \(V\)-parameters are defined as
$V_2(ak_0) = ak_0(e_{r1}\mu_{r1} - e_{r2}\mu_{r2})^{1/2}$ and $V_3(ak_0) = ak_0(e_{r1}\mu_{r1} - e_{r3}\mu_{r3})^{1/2}$, whereas $U = ak_0 e_{r1}\mu_{r1} - n^2_{eff}^{1/2}$, with $\kappa$ standing for the transverse component of the wavevector in the core. It is noteworthy that similar dispersion expressions, cast in an “inverted” form that is obtained following zigzag-ray model analysis, describe oscillatory modes in standard asymmetric optical waveguides [95]. The parameters used in Eqs. (5.60) and (5.61) here, though, obey the restrictions $0 < b < 1$ and $t > 1$. The equalities for $b$ hold at the mode-cutoff points. For each mode ($m = 0, 1, ...$) and refractive index distribution, the dispersion diagrams, $n_{eff}$ versus reduced slab thickness $ak_0$, may be directly obtained by noting that $b$ in Eq. (5.59) is a function of $n_{eff}$, which in turn increases monotonically from $n_2$ to $|n_1|$.

In discussing the cutoff conditions, it is important to recognize the fact that several double mode-degeneracy points can, in principle, appear in the dispersion diagrams since some of the modes in the LH heterostructure will be backward propagating, having antiparallel phase ($\nu_p$) and group ($\nu_g$) velocities. The cycle-averaged total power flow $P_{tot}$ at these (cutoff) points will vanish. It is therefore useful to derive here a general closed-form expression for $P_{tot}$. Integrating the $z$ component of the complex Poynting vector over the guide’s cross-section [55], [57], considering Eq. (5.59), we obtain after some algebraic manipulations:

$$P_3 = C \frac{1}{|e_{r1}|} \frac{\alpha\beta}{\sigma_3 W_3}, \quad (5.62a)$$

$$P_1 = C \frac{\alpha\beta}{|e_{r1}|} \frac{W_2^2 + \sigma_2^2 U_2^2}{\sigma_2^2 U_2^2} \left( \frac{\rho_2 W_2}{W_2^2 + \sigma_2^2 U_2^2} + \frac{\sigma_3 W_3}{W_3^2 + \sigma_3^2 U_2^2} - 2 \right), \quad (5.62b)$$

$$P_2 = C \frac{\alpha\beta}{|e_{r1}|} \frac{\rho_2^2 W_2^2 + \sigma_2^2 U_2^2}{\sigma_2^2 W_2^2 + \rho_2^2 U_2^2}, \quad (5.62c)$$

where $P_i$ ($i = 1, 2, 3$) is the power confined in the waveguide $i$-layer (Fig. 5.10), $\beta$, $W_j$ ($j = 2, 3$) the mode longitudinal propagation and decay constants, respectively, and $C$ an arbitrary positive constant. It is inferred from Eqs. (5.62a-c) that the net power flow can become negative in the core layer, but remains positive in the cladding region. Calculating the total power $P_{tot} = \sum_{i=1}^{3} P_i$, in terms of the dimensionless parameters defined before and the reduced slab thickness, we arrive at:
To facilitate the discussion of the mode cutoff relationships, we first note from Eq. (5.59) that the upper branches \((b \to 0)\) for all modes other than the fundamental \((m \neq 0)\), show no upper cutoff thickness, since \(b \to 0\) in this case yields \(a k_0 \to \infty\). The situation is reminiscent of OMs in normal dielectric slab structures and is illustrated in Fig. 5.11(b). However, for the fundamental mode \((m = 0)\) inspection of Eq. (5.59) reveals a somewhat unusual behaviour and a low cutoff point given by:

\[
(a k_0)_{\text{LOW}}^{m=0} = \frac{1}{2 |n_1|} \left( \frac{\rho_\varepsilon}{\sqrt{1 - \rho_\varepsilon^2 \rho_\mu^2}} + \frac{\sigma_\varepsilon}{\sqrt{1 - \sigma_\varepsilon^2 \sigma_\mu^2}} \right). \tag{5.64}
\]

At this transition point, depicted in Fig. 5.11(a), the SPP-1 mode \([64]\) transforms continuously to the \(m = 0\) oscillatory mode, whose field pattern resembles closely that of a surface wave. In Ref. \([100]\), the existence of such cutoff point for the \(m = 0\) oscillatory mode in asymmetric three-layer LH heterostructures has been associated

**Figure 5.11:** (a) Variation of SPP and fundamental oscillatory mode (OM) effective index \(n_{\text{eff}}\) with the reduced slab thickness \(a k_0\), for \(\sigma_\varepsilon = 1.1, \sigma_\mu = 0.5, \rho_\varepsilon = 1.15, \rho_\mu = 0.6, \varepsilon_r = 2, \mu_r = 1.2\). (b) Variation of oscillatory modes effective index with reduced slab thickness, for \(\sigma_\varepsilon = \sigma_\mu = 0.05, \rho_\varepsilon = \rho_\mu = 0.08, \varepsilon_r = \mu_r = 6\). (c) Variation of oscillatory modes normalized power flow \(P\) with reduced slab thickness for the parameters defined in (b).
with the appearance of a complete 3D photonic bandgap. Critically, the fundamental mode also exhibits an upper cutoff that can be calculated as:

\[
(a k_0)_{\text{UPPER}}^{m=0} = \frac{1}{2 n_1 \sqrt{1-\rho_\varepsilon \rho_\mu}} \left[ \frac{\text{atan}}{\sqrt{r^2-1}} + \pi (m + \frac{1}{2}) \right].
\]  

(5.65)

This relationship also describes the upper cutoff point of all lower branches in the mode-dispersion diagrams. An exemplary result of such curves for a "slow-light oriented" choice of optogeometric parameters is shown in Fig. 5.11(b), with the insets illustrating the corresponding field-profiles. For the sake of clarity, the dispersion curve of the fundamental TM\(_1\) mode is not shown here, since the upper cutoff point of this mode is \((a k_0)_{\text{UPPER}}^{m=0} \approx 0.18763\), i.e. well below the “degeneracy” point of the TM\(_2\) branch, which occurs at \((a k_0)_{\text{DEG}}^{m=0} \approx 0.31047\). The proposed nomenclature for recognizing the oscillatory modes, consists of a pair of letters to identify the polarization, followed by a subscript to track the number of nodes in the core region, and a superscript to designate that the mode is forward (f) or backward (b), i.e. TM\(_{m+1}^{f/b}\) for the \(m\)\(^{th}\)-order mode.

Figure 5.11(c) reports the variation of the normalized power flow
\[ P = \frac{P_{\text{tot}}}{(|P_1| + |P_2| + |P_3|)} \]  
[55], [88] with reduced slab thickness for the previous solutions. One may notice that the two branches of the TM\(_2\) mode merge at a critical slab thickness. At this (cutoff) point the total power, \(P_{\text{tot}}\), of the resulting degenerate mode vanishes and the group velocity reduces to zero [96], [97]. An exact expression for this cutoff point cannot be given except in the somewhat simplified form of \(P_{\text{tot}} = 0\). However, for every mode \((m = 1, 2 \ldots)\) this point can be swiftly calculated following the steps for the derivation of the dispersion diagrams outlined above and requiring \(ak_0(n_{\text{eff}}) \rightarrow \text{min}\).

Since the TM\(_1\) and TM\(_2\) modes always show an upper cutoff, given by Eq. (5.65), we conclude that there is a unique and experimentally intriguing region, highlighted by the gray area in Fig. 5.11(c), where the backward TM\(_2\) mode can simultaneously exist alone and allow for attaining very small or zero group velocity via adiabatically tapering to the degeneracy point. Under weaker guidance conditions,
i.e. $\rho_r \rho_i, \sigma_r \sigma_i \to 0$, it is further possible to altogether suppress the TM$_2$ branch, thereby increasing the operable width of the highlighted region from approximately 0.134 to 0.274, for the case shown in Fig. 5.11(c).

It is interesting to note that similar modes, but with reversed power flow, are supported by the inverted heterostructure, i.e. one with a RH core (e.g. air) and LH claddings. Indeed, by again solving the wave equation in each waveguide region and requiring continuity of the tangential fields at the boundaries, we find that the characteristic Eq. (5.59) remains unchanged, and so does the expression for the sole magnetic field component, $H_y$. In a similar vein, it is found that for the new expressions of the cycle-averaged power flow in each layer, one only needs to replace $|\varepsilon_{r1}|$ in Eqs. (5.62a-c) with $-|\varepsilon_{r1}|$. Based on the explicit form of Eq. (5.59) and since the $V$-parameters defined above are independent of the refractive index sign distribution, it is inferred that all modal properties of the RH-LH-RH arrangement, previously analysed, are replicated by its “dual” counterpart. It appears at this time that this is the only pair of planar waveguides constructed from isotropic media that can do so. This property also provides increased flexibility in slow-light waveguiding design utilizing LH materials.

In summary, on the basis of an exact, analytic appraisal, it has been shown that an asymmetric planar waveguide utilizing LH media in either the core or the cladding can support single-mode operation in the slow-light regime. The investigated scheme, relying solely on sufficient decrease of slab thickness, combines the remarkably simple approach for slowing down light suggested in Ref. [54], with the use of efficiently excitable waveguide modes used in Refs. [50], [51], since the profile of the TM$_2$ solution here closely matches that of a single-mode fibre. Moreover, the heterostructures investigated here can be designed to be monomode in the desired frequency range [101]. The control of the group velocity is achieved solely by varying the core thickness rather than by varying the temperature or field intensity [50], [102]. The same is true for the light in- and out-coupling, which may be satisfactorily adjusted by adiabatically tapering the size of the waveguide core. It should be stressed that the mechanism for decelerating light here does not directly rely on refractive index resonances but merely on the exchange of power between the core and cladding regions, as indicated by Eqs. (5.62)-(5.63). Hence, broadband slow light can be
obtained provided that the negative material parameters are designed to exist over relatively large bandwidths [103] at optical frequencies [104]. This intriguing possibility is further analysed in detail in Chapter 7.
6 Design Methods for Constructing Metamaterials

In the previous chapter we examined in detail some of the intriguing characteristics of light propagation in bulk metamaterials and waveguides. There, we assumed that such materials were available and we were not concerned with how to actually construct them. However, as we remarked in the previous chapter, materials characterised by regions where their electromagnetic parameters are *simultaneously* negative do not occur naturally, and methods to construct them are not immediately apparent. Thus, though there are indeed naturally occurring media, well-known to scientists, that do possess *either* negative permittivity (such as plasmonic media) or negative permeability (such as gyrotropic media) in a given frequency region, double-negative metamaterials do not (under normal conditions) occur in nature and have to be ‘built’ in the laboratory. It was this conception of practical recipes for the construction of man-made metamaterials, followed by their experimental realisation and demonstration of their unique properties, which led to a reappraisal of their usefulness and to a resurgence of interest by the scientific community in these exotic materials.

In this chapter we shall examine how one can construct effective materials whose interaction with light can accurately be described in terms of *negative* effective electromagnetic parameters, which can be engineered to exist in a desired frequency regime. We will start by studying the pioneering proposal towards obtaining materials with negative (effective) electric permittivity in the microwave regime. We will then see how one can also construct an engineered medium exhibiting negative magnetic permeability, again in the microwave regime – thus, opening the way to a man-made double-negative metamaterial, by suitably overlaying the negative-ε and negative-μ media. We will close this chapter by analysing in detail some novel configurations that can give rise to ultra-/zero-loss magnetic metamaterials or, indeed, engineered materials that can exhibit magnetic “gain” over a continuous range of frequencies without having to use gain media to overcome (dissipation, radiation, impedance-mismatch or surface-roughness) losses – solely by a judicious redesign, at the unit-cell level, of the previously studied magnetic metaparticles. These design recipes are, thus, expected to open the way to a multitude of metamaterial-enabled applications, such as the ‘perfect’ lens that we studied in section 5.5, or slow-light metamaterial
waves that we briefly considered in section 5.7 – and to the more detailed study of which we shall focus in the next chapter.

6.1 Metamaterials With Negative Effective Permittivity in the Microwave Regime

It is well-known that metals at optical frequencies are characterised by an electric permittivity that varies with frequency according to the following, so called Drude, relation:

\[ \varepsilon(\omega) = \varepsilon_0 \left[ 1 - \frac{\omega_p^2}{\omega(\omega + i\gamma)} \right], \]

(6.1)

where:

\[ \omega_p^2 = \frac{N e^2}{m c_0} \]

(6.2)

is the plasma frequency, i.e. the frequency with which the collection of free electrons (plasma) oscillates in the presence of an external driving field, \( N, e \) and \( m \) being, respectively, the electronic density, charge and mass, and \( \gamma \) is the rate with which the amplitude of the plasma oscillation decreases. One can directly infer from Eq. (6.1) that, e.g., when \( \gamma = 0 \) and \( \omega < \omega_p \) it is \( \varepsilon < 0 \), i.e. the medium is characterised by a negative electric permittivity. Typical values for \( \omega_p \) are in the ultraviolet regime, while for \( \gamma \), a typical value (e.g., for copper) is \( \gamma \approx 4 \times 10^{13} \) rad/s. Unfortunately, for all frequencies (\( \omega < \omega_p \)) for which \( \varepsilon < 0 \), it is also \( \omega \ll \gamma \), i.e. the dominant term in Eq. (6.1) is the imaginary part of the plasma electric permittivity, which is associated with losses (light absorption). Thus, if one is interested in obtaining a bulk structure characterised by a negative electric permittivity in, e.g., the microwave (GHz) regime – where the construction of an effective medium or device is, of course, considerably less demanding than the optical regime –, simply deploying metals will result in an impractically lossy (mainly evanescent) propagation of light inside such a structure.
A method for overcoming this limitation was first proposed and analysed in detail by Pendry *et al.* [3], based on the observation that the plasma frequency, given by Eq. (6.2), depends critically on the density and mass of the collective electronic motion. They considered the structure illustrated in Fig. 6.1, wherein thin metallic wires (infinite in the vertical direction, $z$) of radius $r$ are periodically arranged on a horizontal plane ($xy$). The unit cell of the periodic structure is a square whose sides have length equal to $a$. If an electric field $\mathbf{E} = E_0 e^{-i(\omega t-kz)} \mathbf{z}$ is incident on the structure, then the (free) electrons inside the wires will be forced to move in the direction of the incident field. If the wavelength of the incident field is considerably larger compared to the side length of the unit cell, $\lambda \gg a$, then the whole structure will appear (to the incident electromagnetic field) as an effective medium whose electrons (confined in the wires) move in the $+z_0$ direction. The crucial observation here is that, since the electrons are confined to move only inside the thin wires, the effective electron density of the *whole* structure (effective medium) is:

$$N_{\text{eff}} = N \frac{\pi r^2}{a^2},$$

(6.3)

with $N$ being the electron density inside each wire. Thus, for sufficiently thin wires

![Figure 6.1: Schematic illustration of a period arrangement of infinitely long thin wires, used in the creation of an effective plasma medium at microwave frequencies.](image)
the effective electron density, \( N_{\text{eff}} \), of the engineered medium can become much smaller compared to \( N \), thereby substantially decreasing the effective plasma frequency, \( \omega_p \), of the engineered medium according to Eq. (6.3). For instance, for a wire radius \( r = 1 \mu m \) and wire spacing \( a = 5 \, \text{mm} \), we find from Eq. (6.3) that: \( N_{\text{eff}} \approx 1.3 \times 10^7 N \), i.e. the effective electronic density of the new medium is reduced by seven orders of magnitude compared to that of the free electron gas inside an isolated wire.

Moreover, it also turns out that the effective mass, \( m_{\text{eff}} \), of the electrons moving inside the engineered medium is considerably larger compared to the free electron mass, \( m \). To prove this result, we start by assuming that the total current flowing in the \(+z\) direction inside the whole engineered medium is \( I \) and that the associated current density is \( J(R) \), \( R \) being the radius of a circle centred at each wire, i.e.:

\[
I = \int_{0}^{\infty} J(R) \, d(\pi R^2), \quad (6.4)
\]

\[
J = \frac{I}{\pi R} z_0. \quad (6.5)
\]

We, then, for the sake of convenience, divide the \( xy \) plane into circles of radius \( R_c \), centred at each wire and having area equal to that of the square unit cell, i.e. \( R_c = a / \sqrt{\pi} \). Furthermore, we assume that the wires are sufficiently apart from each other (‘isolated’) so that the magnetic field inside each circle arises only from the current \( I \) that flows perpendicularly to the centre of the circle, and that the field at the circumference of each circle vanishes, i.e. \( H(R_c) = 0 \). It, then, readily follows from Eq. (3.9) that the magnetic field intensity at a distance \( R \) from each wire is given by:

\[
H = \frac{I}{2\pi R} \left(1 - \frac{R^2}{R_c^2}\right) \varphi_0. \quad (6.6)
\]

Realising that because of Eq. (3.10) the magnetic field \( \mathbf{H} \) can be associated with a vector potential \( \mathbf{A} \) according to the relation: \( \mathbf{H} = \mu_0^{-1} \nabla \times \mathbf{A} \), yields the following expression for the vector potential \( \mathbf{A} \):
where, again, it has been assumed that $A(R \geq R_c) = 0$. For distances very close to the wires, i.e. for $R \rightarrow r \ll R_c$, Eq. (6.7) takes, to a great accuracy, the following form:

$$A \approx \frac{\mu_0 I}{2\pi} \ln(a/r)z_0. \quad (6.8)$$

We may, now, recall that for any electromagnetic field we can write $B = \nabla \times A$ (as we mentioned above) and $E = -\nabla \varphi - \partial A/\partial t$, the latter relation stemming directly from Eq. (3.11). Thus, the equation of motion of a moving electron will be:

$$d(mv)/dt = q[E + v \times B] \rightarrow \frac{d}{dt}[mv + qA] = -q\nabla(\varphi - v \cdot A), \quad (6.9)$$

with $q = -e$ and $v$ being the (absolute value of the) electronic charge, and the (magnitude of the) electron velocity, respectively. Since, both, $v$ and $A$ point along $z$, the RHS of the second part of Eq. (6.9) will vanish (for a given $R$), i.e. the, so called, “conjugate momentum” $(mv + qA)$ of an electron will be conserved along the $z$-direction. As a result, an electron will be moving in the above engineered medium with an effective mass, $m_{\text{eff}} = qA(r)/v$. Realising that the current $I$ can be simply re-expressed as: $I = -(\pi r^2)(Nev)$, we finally obtain using Eq. (6.8) the effective mass of a moving electron inside our effective medium:

$$m_{\text{eff}} = 0.5 \times \mu_0 Ne^2 r^2 \ln(a/r). \quad (6.10)$$

Thus, for copper wires of radius $r = 1 \mu m$, being separated by $a = 5 mm$, we obtain: $m_{\text{eff}} \approx 1.3 \times 10^4 m$, i.e. the effective mass of an electron in our engineered medium is increased by more than four orders of magnitude. This, combined with the fact that (as we show before) the effective electron density is reduced by approximately seven orders of magnitude, leads to an effective plasma frequency that is in the microwave regime:
It should be noted that, based on Eq. (6.11), the calculated wavelength, \( \lambda_p = \frac{c}{f_p} \), which corresponds to our medium's effective plasma frequency turns out to be considerably larger compared to the periodicity of the structure (\( \lambda_p \approx 7a \)), justifying the description of the periodic structure as an effective medium. Therefore, with the herein presented methodology, we are indeed able to construct an engineered medium that can exhibit a negative electric permittivity in the microwave regime (with reasonably low losses and high field penetration inside the structure), thereby mimicking the interaction of light with real metals in the optical regime.

\[
\omega_p^2 = \frac{N_{\text{eff}}}{m_{\text{eff}}} = 5.1 \times 10^{10} \text{[rad/s]}^2 \rightarrow f_p = \frac{\omega_p}{2\pi} = 8.2 \text{ GHz.}
\]  

(6.11)

6.2 Metamaterials With Negative Effective Permeability in the Microwave Regime

In the previous section we examined how we can construct a metamaterial exhibiting negative electric permittivity (\( \varepsilon \)) in the microwave regime. However, harnessing the remarkable properties of left-handed metamaterials also requires, as we saw in the previous chapter, a design strategy for obtaining negative (effective) magnetic permeability (\( \mu \)) at the same frequency region with the negative permittivity. Unfortunately, with the exception of some magnetic gyrotropic materials, media exhibiting negative \( \mu \) do not occur naturally and should, thus, be built in the lab. In

**Figure 6.2:** Schematic illustration of two concentric split-ring resonators (SRRs) for obtaining negative effective permeability in the microwave regime.
this section we shall see how one can construct such magnetic metamaterials in the microwave regime using entirely non-magnetic structured metallic elements [4], which act as the magnetic ‘molecules’ of the engineered medium.

To this end, let us consider a three-dimensional periodic repetition of the external (larger) ring, shown in Fig. 6.2. The radius of the ring is \( r \), and the whole arrangement is assumed to be immersed in air. This ‘split ring resonator’ (SRR) is equivalent to a simple \( R-L-C \) circuit, \( R \) being the resistance of the metallic ring, \( L \) its inductance and \( C \) (primarily) the capacitance between its unconnected ends. The rings residing on a given \( z = z_i \) plane have the same axis (i.e., they are ‘concentric’) with the corresponding rings on the \( z \)-planes below and above them. The side of the square unit cell on an \( xy \) plane is equal to \( a \).

Assuming that a magnetic field \( \mathbf{H} = H_0 e^{i(oH - k\rho)} \mathbf{z}_0 \) is incident on the structure, the induced (electromotive) source will (from Eq. (3.8)) be: \( U = i\omega\mu_0\pi^2 H_0 \), generating an electric current \( I \) that circulates each ring (see Fig. 6.2). If the rings sitting on successive \( z \)-planes are close together (‘solenoid’ approximation) there will be negligible ‘loss’ of magnetic flux between the rings in each column, and therefore the magnetic flux will be: \( \Phi = \mu_0\pi^2 l/l \), \( l \) being the distance between corresponding SRRs lying on successive \( xy \) planes. Accordingly, the inductance \( L \) (in Henry) of each SRR will be: \( L = \Phi/I = \mu_0\pi^2 l/l \). One may, further, assume that the depolarizing magnetic flux lines generated by all rings are uniformly spread on a given \( xy \) plane, which results in a mutual inductance between two SRRs given simply by: \( M = (\pi^2/a^2) L = FL \), \( F \) being the fractional volume within a unit cell occupied by an SRR.

We may, now, apply Ohm’s second law across a closed SRR ‘circuit’ to obtain:

\[
U = [R + i l/(\omega C) - i \omega L + i \omega M] I, \quad (6.12)
\]

where \( R = 2\pi r\sigma \) is the (ohmic) resistance of each ring, \( \sigma \) being the resistance per unit length. Thus, the induced magnetic dipole moment per unit volume, \( M_d \), will be:

\[
M_d = I(\pi r^2)/(a^2 l), \quad (6.13)
\]

with the current \( I \) inferred from Eq. (6.12) to be:
\[ I = -\frac{H_0 I}{(1 - F) - 1/(\omega^2 LC) + iR/(\omega L)}. \]  

(6.14)

As a result, the (relative) effective magnetic permeability associated with this medium will be (in the direction, z, that the incident magnetic field is polarised):

\[ \mu = \frac{B/\mu_0}{B/\mu_0 - M_d} = 1 - \frac{F}{1 - 1/(\omega^2 LC) + iR/(\omega L)}. \]  

(6.15)

From Eq. (6.15) we can see that \( \mu \) assumes negative values in the range: \( 1/\sqrt{LC} < \omega_p < 1/\sqrt{LC(1 - F)} \), where \( \omega_p = 1/\sqrt{LC} \) is the resonance frequency of the Lorentzian variation of the medium's magnetic permeability, and \( \omega_{mp} = 1/\sqrt{LC(1 - F)} \) is the corresponding plasma frequency (where \( \text{Re}\{\mu\} = 0 \)).

Crucially, we note that the resonant wavelength \( (\lambda_{m0}) \) of the structure depends entirely on the rings' effective inductance \( (L) \) and capacitance \( (C) \), and can therefore be made considerably larger that the periodicity \( (a) \) of the structure, thereby fully justifying its description as an effective medium. Had we placed the SRRs on the other two planes (yz and xz) we would have, similarly, obtained negative effective permeabilities in the other two directions, x and y, as well, and the variation with frequency of these permeabilities would have been given by an expression similar to Eq. (6.15). Thus, with the present methodology we are able to construct a three-dimensional, isotropic metamaterial exhibiting negative effective permeability in a specified frequency region. The same effect can be obtained all the way from the radio up to the optical frequencies, simply by accordingly scaling the size of the SRR metaparticles.

### 6.3 Intrinsically Lossless Magnetic Metamaterials

As we have seen previously, it is by now well-established that not only do negative-\( \mu \) and negative-index materials exist, but that they can even be constructed to exhibit broadband behaviour – in fact, they may even posses infinite bandwidth – [105], scaled down to optical frequencies and be built in three dimensions [34], as well as allow for efficient and fast tuning and switching [35]. Unfortunately, however, at
present their performance is, as we remarked in section 5.5, limited by the occurrence of losses (light absorption), which occasionally may reach levels of up to tens of dB/wavelength [106]. Clearly, if metamaterials are to find their way towards practical applications, this issue has to be adequately addressed.

To this end, a number of diverse strategies have been proposed [37]. A first approach relies on the use of gain media to compensate for the losses that originate from the presence of metallic elements in the ‘meta-molecules’ [38]. This approach is very promising and it has, in fact, been theoretically shown that it can lead to zero-loss metamaterials [107], even over a broad but finite bandwidth [39]. However, it may not always be possible to find a suitable gain medium to provide the necessary gain at the desired frequency regime. Another method for reducing losses in metamaterials relies on a judicious optimisation of the geometry of the ‘meta-molecules’, whereby increasing the effective inductance to capacitance ratio, \( L/C \), increases the quality factor, \( Q = \frac{1}{2R \sqrt{\frac{L}{C}}} \), leading to reduction in the overall losses [108]. This approach results in metamaterial losses being substantially reduced, but normally by less than an order of magnitude, without being fully overcome, i.e. the resulting structures are never lossless. A third, recently proposed, interesting scheme makes use of negatively reflecting/refractive interfaces to reproduce the features of bulk negative-index metamaterials [109]. Upon being negatively refracted at the engineered interfaces, light is allowed to propagate through lossless dielectric materials, such as air, thereby avoiding high-attenuation regions. This strategy relies on the use of nonlinear media and therefore requires intense incident light.

6.3.1 Case of ‘Isolated’ Unit Cells

To arrive at a design for an intrinsically lossless resonant-type magnetic metaparticle let us start by noting that, in general, the magnetic metaparticles proposed so far correspond to equivalent electrical circuit models with only one mesh. For instance, as we show in the previous section, by determining the equivalent resistance \( (R) \), inductance \( (L) \) and capacitance \( (C) \) of a split-ring resonator (SRR) particle, one can accurately describe it in terms of a single \( RLC \) mesh, and then extract the effective magnetic permeability of a periodic array of such particles [110], [111]. Conversely,
by connecting discrete $R$, $L$ and $C$ elements in a single loop, and for sufficiently large incident wavelengths to allow for being well within the ‘effective medium’ regime, one can precisely reconstruct the magnetic behaviour of an SRR. Observing this common feature in the existing designs of magnetic meta-molecules, let us in the following analyse the quasi-static magnetic response of a pair of electrically connected $RLC$ meshes, placed at the centre of a unit cell of volume $V = a^2 \times l$ which is periodically repeated in three dimensions. It is to be noted that electrical connection (‘tight coupling’) of $RLC$ meshes in an infinite 1D array was recently considered in [105] in order to elucidate the broadband and low-loss behaviour of transmission-line metamaterials. It has to be emphasized that in the designs introduced herein, each pair of electrically connected $RLC$ meshes resides at the centre of its own unit cell and is not connected to its neighbouring pairs. Furthermore, the wavelength of the incident light is considerably larger compared to the dimensions of the unit cell. For these reasons, the resulting structure is neither a backward transmission line [105] nor a magnetoinductive waveguide [112], but a resonant-type effective magnetic medium.

Consider a plane wave with magnetic field component $\mathbf{H} = H_0 \exp(i \omega t - ik \cdot \mathbf{r}) \hat{e}_0$, where $k$ is the wavevector, $\mathbf{r}$ is the vector along the direction of the wave propagation, $\omega$ is the angular frequency, $t$ is the time and $\hat{e}_0$ the unit polarization vector. We assume that the magnetic field is incident on a pair of the electrically connected meshes (see Fig. 6.3) such that $\hat{e}_0$ is parallel to $\hat{a}$, where $\hat{a}$ is the unit vector normal to the plane of the two meshes. From Faraday’s law, the electromotive sources (voltages) induced in the two closed meshes by the $\mathbf{H}$-field will be $V_m = -i \omega \mu_0 S_m H_0$, $m = 1, 2$, with $S_m$ being the surface of each mesh. From the well-known ‘mesh current

![Figure 6.3: Schematic illustration of two electrically connected $RLC$ meshes, each corresponding to a split-ring resonator (SRR). In the present case, both meshes reside on the same plane and are excited by the same incident magnetic field.](image-url)
method' [139] one may determine the currents circulating in each loop in Fig. 1(b), as

\[ I_m = \frac{\text{det}(G_m)}{\text{det}(G)}, \]

where:

\[ G = \begin{bmatrix} R_1 + i\omega L_1 - i/\omega C_1 - i/\omega C & i/\omega C \\ i/\omega C & R_2 + i\omega L_2 - i/\omega C_2 - i/\omega C \end{bmatrix}, \]

\[ G_1 = \begin{bmatrix} V_1 & i/\omega C \\ V_2 & R_2 + i\omega L_2 - i/\omega C_2 - i/\omega C \end{bmatrix}, \quad G_2 = \begin{bmatrix} R_1 + i\omega L_1 - i/\omega C_1 - i/\omega C & V_1 \\ -i/\omega C & V_2 \end{bmatrix} \]

(6.16)

with \( R_m, L_m, C_m \ (m = 1, 2) \) being respectively the resistance, inductance and capacitance of the \( m^{th} \) mesh, \( C \) the capacitance of their common branch, and 'det' designating the matrix determinant. Here, the resulting matrices are two-dimensional and the structure of the matrix \( G \) is, as we explain later in section 6.3.6, typical of what is often called a 'two-degrees-of-freedom' (2-DEG) system [113]. Had we considered a cube with \( m \) electrically connected meshes, we would have obtained an \( M \)-DEG system described by \( m \)-dimensional matrices. Such systems are, indeed, well-known in diverse realms of science [113], with some typical examples being highlighted in section 6.3.6.

Having calculated the currents \( I_m \) circulating in each mesh, we can determine the total (emanating from both meshes) magnetic dipole moment per unit volume (magnetisation), \( M \), as \( M = (S_1 I_1 + S_2 I_2)/V \). Let us start by investigating the crucial (for the points to be made later) case where the meshes at the centres of neighbouring cubes are sufficiently apart from each other on the horizontal plane, i.e. the filling factor \( F \) – defined as the ratio between the total area of the two meshes and the area of the basis of a unit cell [111] – is sufficiently small, so that the cells can be regarded as weakly interacting ('isolated'). This is typical of low-density (dilute) gases, but note that contrary to the case of gases wherein the weak interaction of the molecules leads to the absence of any magnetic response and to \( n \approx 1 \), metamaterials can be designed to exhibit strong magnetic response [4], even for low densities, simply by increasing the \( Q \)-factor (e.g., by decreasing the value of the resistance \( R \)) of the resonant \( RLC \) meshes. We examine this crucial case of 'isolated' cells in order to establish that the observed effects are intrinsic, i.e. that they occur at the unit-cell level and are not an artefact of the periodicity of the structure (see section 6.3.3 for a detailed discussion.)
of the latter situation). The usual case of ‘tightly coupled’ meshes on the horizontal plane does not fundamentally differ in any of its conclusions with the case presented herein, as is explained in detail in section 6.3.3. Note that in order to illuminate the subtle distinctions and facilitate a direct comparison between the two cases, in the numerical computations we used the same filling factors for both cases.

Under the aforementioned conditions, higher-order multipole terms are, indeed, negligible [114] and the local fields in each cube are just $H_{loc} = H_0$ and $B_{loc} = B_0$, where $B$ is the magnetic flux density, so that we may define $\mu_{r,eff} = 1 + M/H_0$. Note that in this case the effective (relative) permeability is directly proportional to the magnetization $M$. If the neighbouring cubes were considered ‘tightly coupled’ then

$$\mu_{r,eff} = \frac{1+rac{2}{3} \frac{M}{H_0}}{1-rac{2}{3} \frac{M}{H_0}} \quad \text{(see section 6.3.3)}.$$

In that case, $\mu_{r,eff}$ is not any more simply proportional to neither the magnetisation $M$ nor to the total active power (see section 6.3.2) $P = P_1 + P_2$, $P_m = \text{Re}\{V_m I_m^*\}$ ($m = 1, 2$) of the pair of meshes, and the reported effects become slightly more intricate, but do not fundamentally nor qualitatively differ from those presented below. It is useful to also note that, as we show in the previous section (Eq. (6.12)), in this case and for meshes that are stacked up closely together (solenoid approximation), the uniform depolarization magnetic field [4] only modifies the value of the inductances $L_m$ in each cube [111], i.e. merely reducing the value of $L_m$ in each mesh is sufficient to incorporate the effect of the depolarisation field in the analysis.

Figure 6.4 furnishes exemplary results for the real and imaginary parts, as well as for the corresponding figure of merit $\text{FOM} = - \text{Re}\{\mu\}/\text{Im}\{\mu\}$, of the (relative) effective permeability $\mu$ in structures where the area $S_2$ of the second mesh in each unit cell is much smaller compared to the area $S_1$ of the first mesh. Let us start with the case (shown in Figs. 6.4(a) and (b)) where the second mesh is also open-circuited. As expected, one observes a typical Lorentzian absorption pattern for the imaginary part of $\mu$, as well as for the total (absorbed) active power $P$ (see section 6.3.2). The specific variation of the real part of $\mu$ with frequency stems from the fact that the engineered medium, since it is physically realizable, is causal with $\text{Re}\{\mu\}$ and $\text{Im}\{\mu\}$
Figure of merit Magnetic permeability $\mu$

- Frequency (GHz)
- Magnetic permeability $\mu$
  - Real
  - Imaginary

(c)

Figure of merit

- Frequency (GHz)
  - Figure of merit

(d)
Figure of merit: Magnetic permeability $\mu$

(e)

Frequency (GHz)

Real

Imaginary

(f)

Frequency (GHz)

Figure of merit
Figure 6.4: Calculated effective (relative) magnetic permeability ($\mu$) and figure of merit (FOM = $-\text{Re}\{\mu\}/\text{Im}\{\mu\}$) for the case where both loops lie on the $yz$ plane. A perpendicularly incident magnetic field $H_0 = 1$ A/m induces two electromotive voltage sources, $V_1$ and $V_2$ (see Fig. 1). In all cases, the values of the lumped elements were: $R_1 = 50$ Ohms, $R_2 = 0.1$ Ohms, $C_1 = C_2 = 10^{-14}$ F, $L_1 = L_2 = 1.6\times10^{-8}$ H. The surface of the first loop is $S_1 = \pi\times(2\times10^{-3})^2$ m$^2$ and both loops are located at the centres of unit cells of volume $(5\times10^{-3})\times(5\times10^{-3})\times(1\times10^{-3})$ m$^3$. (a) Real and imaginary part of effective $\mu$ for the case where the second mesh is open circuited and the value of the coupling capacitance is $C = 0.1\times10^{-12}$ F. (b) Corresponding figure of merit for the configuration with the parameters of (a). (c) Real and imaginary part of $\mu$ for the case where the second mesh is included, but has surface $S_2 \to 0$. The value of the coupling capacitance is, now, $C = 1\times10^{-12}$ F. (d) Corresponding figure of merit for the configuration with the parameters of (c). (e) Same as in (c), but now with $C = 0.3\times10^{-12}$ F. (f) Corresponding figure of merit for the configuration with the parameters of (e). (g) Same as in (c), but now with $S_2 = S_1/10$ and $C = 1\times10^{-12}$ F. (h) Corresponding figure of merit for the configuration with the parameters of (g).

strictly adhering to Kramers-Kronig relations. Figure 6.4(b) presents the FOM for this case, which can assume appreciable values only in regions where Re{$\mu$} is either positive or very close to zero. Figures 6.4(c) and (d) report the corresponding results for the case where the second mesh is included in each cube, but has area $S_2 \ll S_1$ (with: $S_2 \to 0$). Note the abrupt dip in the spectra of, both, the imaginary part of $\mu$ and
the total active power $P$ (the latter shown in section 6.3.2), accompanied by a sudden small jump in the FOM, shown in Fig. 6.4(d). This behaviour is, indeed, expected since it is well-known that macroscopic coupled resonators (in this case, electrical resonators) can mimic the electromagnetically induced transparency effect that occurs in atomic gases [44], [48] leading to a so-called, 'coupled resonator induced transparency' [115]-[118]. The associated very strong dispersion in a close region around the dip can be used for decelerating light [44], although over only a very narrow bandwidth. Figures 6.4(e) and (f) present similar results to the previous case, but for a coupling capacitance reduced tenfold. One can clearly observe the evolution of the spectra of $P$ and the imaginary part of $\mu$ according to the well-known Autler-Townes effect [115]. The calculated FOM in this case becomes large, assuming a maximum value of approximately 60, but mainly in regions where $\text{Re}\{\mu\} > 0$. These low-loss regions of $\mu$ (with $0 < \text{Re}\{\mu\} < 1$) could be useful for good-quality diamagnetic metamaterials, as well as in the design of 'invisibility' cloaks [42], [43]. For instance, at $f = 12.58$ GHz, one obtains $\mu \approx 0.9346 - i0.0355$, with a FOM $\approx 26$.

To complete the presentation of the results for this general case, Figs. 6.4(g) and (h) report the corresponding calculations to Figs. 6.4(c) and (d), but now with $S_2 = S_1/10$. As shown in section 6.3.2, this is the first instance where $P_2$ becomes nonzero, and it does so in an intriguing way: There is a narrow region wherein $P_2$ actually becomes negative, which suggests that in this region the second mesh does not absorb but, instead, remits active power to the first mesh. We will return to this important observation later when we discuss the design of an intrinsically lossless magnetic metamaterial, but at the moment let us note that in this region the total power $P$ absorbed by the pair of meshes is reduced, owing to the negative contribution from $P_2$, and consequently the FOM exhibits an increased jump — within the $\text{Re}\{\mu\} < 0$ region — compared with that of Fig. 6.4(d).

Let us, now, turn our attention to the important case where the areas of the two meshes inside each unit cell are equal, i.e. $S_2 = S_1$. Figures 6.5(a)-(e) report the real/imaginary parts of $\mu$, the corresponding FOM and the absorbed powers $P_1$, $P_2$ and $P = P_1 + P_2$. The exchange of active power between the two meshes (electrical oscillators) is readily inferred from Figs. 6.5(c)-(e). There are, indeed, regions wherein $P_1$ or $P_2$ become negative, but never simultaneously (which would violate the conservation of energy). In the region where $P_1 < 0$ the second mesh receives active
Figure 6.5: In all cases the electromagnetic and geometric parameters are those of Fig. 2, but now with $S_2 = S_1$ and $C = 0.1 \times 10^{12}$ F. Variation with frequency of the: (a) Real and imaginary part of effective $\mu$. (b) Corresponding figure of merit. (c) Active power $P_1$. (d) Active power $P_2$. (e) Total active power $P = P_1 + P_2$.

power from the first mesh, with the opposite being true for the region where $P_2 < 0$. The fact that $P_1$ and $P_2$ are never simultaneously negative is evident from the fact that $P > 0$ throughout, as required from the passivity of the structure. Interestingly, one can observe a region where the total absorbed power $P$ reduces abruptly towards zero,
i.e. it loses its Lorentzian shape. This is also reflected in the imaginary part of $\mu$, shown in Fig. 6.5(a), and as a result the FOM at precisely this region increases dramatically to around 987; for instance, at $f = 13.196$ GHz, we obtain $\mu \approx 4.4575 - i0.0045165$, i.e. we achieve a substantially negative $\text{Re}\{\mu\}$ with simultaneously excellent FOM.

Until now we have assumed that the two electrically connected meshes lie on the same plane inside each unit cell, but this restriction is not necessary. Consider the case where the first mesh is placed on the $xz$ plane and the second mesh on the $yz$ plane, as illustrated in Fig. 6.6. Let us, further, assume that two uniform harmonic magnetic fields of equal amplitude, $H_x = H_2 = H_0 \exp\{i(k_1 r_1 - \omega t)\} \hat{x}_r$ and $H_y = H_1 = H_0 \exp\{i(k_2 r_2 - \omega t)\} \hat{y}_r$, are simultaneously incident on the structure, generating the electromotive voltages $V_1$ and $V_2$ shown in Fig. 6.3. In this particular arrangement, the magnetic field $H_1$ generates electric currents circulating in both meshes. As a result, the magnetic field $H_y = H_1$ generates magnetic moments on both $xz$ and $yz$ planes, i.e. it is responsible for the generation of $\mu_{yx}$ and $\mu_{xy}$. The corresponding is also true for the magnetic field $H_x = H_2$. The ‘total’ $\mu_x = \mu_2$ that the $x$-component of the magnetic flux density, $B_x$, experiences $B_x = \mu_{xx} H_x + \mu_{xy} H_y = (\mu_{xx} + \mu_{xy}) H_0 = \mu_x H_0$ - can, now, be calculated by finding the ‘total’ current $I_2$ shown in Fig. 6.3 and following the methodology for the calculation of $\mu_{r,\text{eff}}$ that was outlined in the previous examples. Figures 6.7(a) and (b) report the results of such a calculation of $\mu_x = \mu_2$ and $\mu_y = \mu_1$ for the case where the coupling capacitance, $C$, of the two meshes is $C = 0.4pF$. Note that the imaginary parts of $\mu_1$ and $\mu_2$ follow closely the variation with frequency of the active powers $P_1$ and $P_2$, respectively, which were previously studied in Figs. 6.5(c)-(e), and as such there are now regions where either $\mu_1 (\mu_2)$ or $\mu_2 (\mu_1)$ become zero or even positive. Furthermore, owing to the passivity of the structure, $\mu_1$ and $\mu_2$ are never simultaneously positive; in fact, $\text{Im}\{\mu_1\} + \text{Im}\{\mu_2\}$ is exactly equal to the imaginary part of $\mu$ that was shown in Fig. 6.5(a).

Thus, with the present anisotropic design we are, in principle, able to harness zero-loss negative-$\mu$ magnetic metamaterials – in fact, even with artificial magnetic “gain” – over a continuous range of frequencies along one direction, at the cost of having increased magnetic losses in the other direction. For instance, assuming that the imaginary part of $\mu_1$ is positive in a frequency region, one will observe an increase
Figure 6.6: To arrive at a design that can allow for harnessing intrinsically lossless
metamaterial magnetism, we place – in each unit cell – the first mesh on the $xz$ plane
and the second mesh on the $yz$ plane (see top left inset). We then use two separate
light beams of the same incident amplitude: One (shown in green) whose magnetic
field oscillates perpendicularly to the second mesh, i.e. parallel to the $x$-axis, and a
second beam (shown in red) whose magnetic field oscillates perpendicularly to the
first mesh, i.e. parallel to the $y$-axis. Our analysis reveals that there are certain
frequency regions wherein the imaginary part of the structure’s effective magnetic
permeability, $\text{Im}\{\mu\}$, becomes positive (corresponding to magnetic ‘gain’) along the $y$
($x$) direction and negative along the $x$ ($y$) direction. The magnetic losses in the
direction where $\text{Im}\{\mu\} < 0$ are larger compared to the ‘gain’ in the direction where
$\text{Im}\{\mu\} > 0$, so that the conservation of the magnetic energy is satisfied.
Figure 6.7: The electromagnetic and geometric parameters are those of Fig. 3, but now the first mesh lies on the $xz$ plane, while the second mesh on the $yz$ mesh. Variation with frequency of the: (a) Real and imaginary part of the effective relative permeability along the $x$-axis ($\mu_x$). The inset illustrates the variation with frequency of the imaginary part of $\mu_x$ (red) and $\mu_y$ (green), from where one can observe that in the region where $\mu_y$ ($\mu_x$) $> 0$ it is, also, $\mu_x$ ($\mu_y$) $< 0$. (b) Real and imaginary part of the effective relative permeability along the $y$ direction ($\mu_y$). The inset illustrates the variation with frequency of the real part of $\mu_x$ (green) and $\mu_y$ (blue).
(artificial “gain”) in the magnetic density $B_y$ in the direction of its propagation, owing to $\text{Im}\{\mu_1\} > 0$; at the same time, however, the decrease of the magnetic density $B_x$ along the direction to which it propagates, owing to $\text{Im}\{\mu_2\} < 0$ in the same frequency region, will be larger compared with the increase that $B_y$ experienced, so that – at every frequency point – there will be no violation of the conservation of the magnetic energy.

In summary, in this section we have presented a design paradigm that allows for ultra-low or (uniaxially) zero-loss magnetic metamaterials. The proposed blueprint is based on a 2-DEG topology and relies critically on the presence of electrically connected $RLC$ meshes that exchange active power. In certain frequency regions, one mesh is able to naturally ‘pump’ energy to the other mesh, which leads to the imaginary part of $\mu$ in a specified direction becoming exactly zero or even positive. The scheme is, in principle, scalable and realizable at any frequency regime, and does not require nonlinear or gain media to compensate for the losses. This design paradigm may conceivably lead to a new generation of ultralow- or zero-loss metamaterials that could be exploited in the design of ‘perfect’ lenses, ‘invisibility’ cloaks, slow-light waveguides, optical nanocircuits and magnetic resonance imaging systems. In the following subsections we present a number of elaborating remarks on the aforementioned scheme, particularly pertaining to the calculations of the active powers and to the case of ‘tightly coupled’ neighbouring unit cells.

### 6.3.2 Calculations of Active Powers in the Equivalent Electrical Circuits

Consider two circuits, $C_1$ and $C_2$, electrically connected with two conducting wires, as shown in Fig. 6.8. Circuit $C_1$ contains sources and passive elements (resistors, inductors, capacitors), while circuit $C_2$ contains only passive elements. Let us assume that the time-domain (designated with low case symbols) voltage, $v$, and the current, $i$, shown in Fig. 6.8, are sinusoidal functions of time, i.e.: $v(t) = V_0\cos(\omega t)$, and $i(t) = I_0\cos(\omega t - \phi)$, where $\phi$ is the phase difference between the current and the voltage. For the sign conventions shown in Fig. 6.8, a positive time-domain (instant) power, $p(t) = v(t)i(t) > 0$, means that electrical power flows from circuit $C_1$ towards circuit $C_2$. Conversely, if $p(t) < 0$ then power flows from circuit $C_2$ to circuit $C_1$.  

129
It is instructive to write the expression for the instant power $p$ in the form:

$$p(t) = V I \cos \varphi [1 + \cos(2\omega t)] + V I \sin \varphi \sin(2\omega t),$$

(6.17)

where $V = V_0 / \sqrt{2}$ and $I = I_0 / \sqrt{2}$. Since we have assumed that the circuit $C_2$ contains only passive elements, the phase difference $\varphi$ between the current and the voltage will always be: $-\pi/2 < \varphi < \pi/2$, and therefore the first term on the RHS of Eq. (6.17) will be positive, corresponding to power flowing from circuit $C_1$ to circuit $C_2$, where it is consumed (dissipated). For this reason, this power, $p_R(t) = V I \cos \varphi [1 + \cos(2\omega t)]$, is called real or active power. The active power varies between 0 and $2V I \cos \varphi$ in one period, and a metric that is frequently used for it, is its time-averaged value: $\overline{P} = V I \cos \varphi$, which customarily is again referred to as active power. By contrast, the second term in the RHS of Eq. (6.17), $p_X(t) = V I (\sin \varphi) \sin(2\omega t)$, changes sign twice during the period $T = 2\pi/\omega$. During the first half-period, $p_X$ flows from $C_1$ to $C_2$, while in the second half-period the direction of the flow of $p_X$ is reversed, i.e. $p_X$ now flows from $C_2$ to $C_1$, so that the overall effect is zero. Because of the fact that $p_X$ does not (on average) produce any work, it is usually referred to as reactive power. Since the time-averaged value of $p_X$ is zero, a metric that is frequently used for its quantification is its amplitude, $Q = V I \sin \varphi$. Generally, the reactive power refers to the part of the instant power $p$ that is send by a source to the inductors and capacitors of the circuit, stored there temporarily, and then send back to the source.

In many situations it is useful to consider the transformations of the voltage $V$ and current $I$ in the frequency domain, $\tilde{V} = V e^{\jmath \theta_1}$ and $\tilde{I} = I e^{\jmath \theta_2}$, $\theta_1$ and $\theta_2$ being the angles
of the rotating vectors \( \vec{V} \) and \( \vec{I} \) with the real axis, and to work with the so called complex power, \( S \):

\[
S = \vec{V}^* \vec{I} = V I e^{i(\theta_0 - \theta_1)} = V I e^{i\theta}.
\]  

(6.18)

It can, then, be readily observed that the reactive power is simply given by: \( Q = \text{Im}\{S\} \), while the (time-averaged) active power by: \( \bar{P} = \text{Re}\{S\} \), which is the relation that we used in the calculations of the active powers referred to in section 6.3.1.

It should be noted that when the circuit \( C_2 \) does not contain sources, the active power \( \bar{P}_2 \) in \( C_2 \) is strictly positive, i.e. \( C_2 \) absorbs (consumes) real power, which is converted into heat at the resistances. In contrast, if \( C_2 \) also includes sources of electrical energy, then it is possible that \( \bar{P}_2 \) may become negative (for the sign convention shown in Fig. 6.8) in a frequency region, corresponding to active power flowing away from \( C_2 \), i.e. in that region \( C_2 \) acts as a source of real power: It does not absorb real power, but remits it to the neighbouring meshes that it is electrically connected with. However, the total active power, \( P = \sum_i P_i \) \((i = 1, \ldots, n)\), \( n \) being the total number of electrically connected meshes, must still be positive at every frequency point, since overall we have a net consumption of real power. These features are, indeed, precisely what we observe in the analytic calculations (shown below) of the active powers in the meshes of the 2-DEG system that we described in section 6.3.1.

In the following we present the results of the calculations of the active powers for the various mesh configurations that were analysed in section 6.3.1. Note, in each case, the precise correspondence (in their variation with frequency) between the total active power \( P = P_1 + P_2 \) and the imaginary part of the effective magnetic permeability \( (\mu) \) that was presented in section 6.3.1.
Active power $P_1$ (mW) vs. Frequency (GHz)

(c)

Active power $P_1$ (mW) vs. Frequency (GHz)

(d)
Figure 6.9: (a) Total (time-averaged) active power for the configuration with the parameters of Fig. 6.4(a). (b) Same as before, but now with the parameters of Fig. 6.4(c). (c) Same as before, but with the parameters of Fig. 6.4(e). (d) Active power of the first mesh, with the parameters of Fig. 6.4(g). (e) Active power of the second mesh, with the parameters of Fig. 6.4(g). (f) Total (time-averaged) active power for the configuration with the parameters of Fig. 6.4(g).
6.3.3 Case of ‘Tightly Coupled’ Unit Cells

In section 6.3.1 we studied the quasi-static magnetic response of ‘isolated’ pairs of electrically connected meshes, in order to establish the fact that the observed regions of positive imaginary parts of the effective magnetic permeabilities do not arise owing to the periodicity of the structure [119]-[124] but that they are an intrinsic feature of the 2-DEG system, occurring within each individual, ‘isolated’ unit cell. There, we showed that such regions are ultimately the result of the exchange of active power within each pair of meshes.

Having established that the aforementioned observed effects are intrinsic, let us in the following examine the corresponding situations with those of the section 6.3.1, but now for the case of strongly interacting (‘tightly coupled’) cells. For an incident magnetic field of amplitude \( H_0 \), inducing a magnetisation \( M \) per unit cell, a pertinent analysis reveals that, in this case, the effective magnetic permeability may be given by:

\[
\mu_{\text{eff}} = \frac{1 + \frac{2}{3} \frac{M}{H_0}}{1 - \frac{1}{3} \frac{M}{H_0}}
\] (6.19)

with care exercised to accordingly modify the total impedance, \( Z \), in each mesh [110], [111] (see also later in this subsection). It was further shown in [111] that for meshes stacked closely together, one can accurately consider each pile to be a ‘solenoid’. In that approach, the depolarisation magnetic field results in an additional (mutual) inductance, \( M \), which is simply subtracted from the (lumped) inductance \( L \) of each mesh.

To facilitate direct comparison with the corresponding results in section 6.3.1, here, we retain the same value (16 nH) for the inductance of each mesh as in section 6.3.1. This is, of course, equivalent to assuming that the actual inductance, \( L \), in each mesh is slightly larger, so that when \( M \) is subtracted from \( L \) we end up with a total inductance, in each mesh, equal to 16 nH.

All the results concerning the active powers in each mesh are precisely the same as the corresponding ones presented in the previous section or in section 6.3.1.
Accordingly, all the dips or peaks in the imaginary part of the effective permeabilities associated with the exchange of active power between the meshes continue to occur at the same frequency (≈ 12.5 GHz) as in the case of ‘isolated’ cells. However, since the effects of neighbouring cells are now incorporated in the analysis, one does expect to observe a redshift in the collective resonant response of the composite medium. For instance, it is well-known from the classical Lorentz theory for the description of dielectric molecules that the use of the Clausius-Mossotti relation (which is very similar to Eq. (6.19) above) results in a redshift of the resonant frequency from $\omega_0^2$, when the density of the molecules is reasonably low (‘isolated’ unit cells, as in a gas), to $\omega_0^2 - N e^2 / 3\varepsilon_0 m$, where $N$ is the density of the molecules, $e$ the electronic charge, $m$ its mass, and $\varepsilon_0$ the free-space permittivity. This is also what we observe in the results presented below.

Inspection of Figs. 6.10–6.12 reveals that the variation with frequency of the calculated relative effective permeabilities are, as expected, very similar to the ones that was presented in Figs. 6.4, 6.5 and 6.7. In particular, compared with the case of ‘isolated’ unit cells, there is a redshift in the resonant response of the system (e.g., compare Fig. 6.4(a) with Fig. 6.10(a)), as highlighted above. The location of the sudden dips and rises, however, which is due to exchange of active power between the electrically connected meshes in each unit cell, remains at around 12.5 GHz, similarly to Figs. 6.4, 6.5 and 6.7, as one would expect. This is because the influence of the neighbouring cells to the currents circulating the pair of meshes in each cell arises primarily through the mutual inductance $M$, which is subtracted from the (herein assumed slightly increased) inductance $L$ of each mesh, so that $L - M = 16$ nH, equal to the value of $L$ that was used in section 6.3.1. As a result, the currents circulating each pair of meshes inside the ‘tightly’ coupled cells are exactly equal to those of the case of ‘isolated’ cells. Hence, here, the variations of the active powers with frequency, which are responsible for stipulating the location of the dips and rises, are the same with the variations presented in section 6.3.1.

It is interesting to note from Fig. 6.12(c) that using Eq. (6.19) to describe the quasi-static response of a medium with ‘tightly’ coupled unit cells, in which the meshes are located at different planes (anisotropic medium), results in a region where the summation of the imaginary parts of the effective permeabilities along the directions perpendicular to these planes becomes positive, implying the occurrence of ‘net’ gain.
in that region. As explained in section 6.3.5, such an outcome, highlighted by a dashed red line in Fig. 6.12(c), emergences ultimately owing to the violation of the assumptions used in the derivation of Eq. (6.19).
Figure of merit Magnetic permeability

(c)

(d)
Figure of merit: Magnetic permeability $\mu$

- Frequency (GHz)

(e)

- Frequency (GHz)

(f)
Figure 6.10: In all cases, the electromagnetic and geometric parameters are those of Fig. 6.4. (a) Relative effective permeability when the second mesh is open-circuited. (b) Corresponding figure of merit. (c) Relative effective permeability for $S_2 = 0$ and $C = 1 \text{ pF}$. (d) Corresponding figure of merit. (e) Relative effective permeability for $S_2 = 0$ and $C = 0.3 \text{ pF}$. (f) Corresponding figure of merit. (g) Relative effective permeability for $S_2 = S_1/10$ and $C = 1 \text{ pF}$. (h) Corresponding figure of merit.
Figure 6.11: The electromagnetic and geometric parameters are those of Fig. 6.5. (a) Relative effective permeability for $S_2 = S_1$ and $C = 1 \text{ pF}$. (b) Corresponding figure of merit.
Imaginary part of magnetic permeability

Real part of magnetic permeability

Frequency (GHz)

(a)

Imaginary part of magnetic permeability

Frequency (GHz)

(b)
Figure 6.12: The electromagnetic and geometric parameters and arrangements are those of Fig. 6.7. (a) Real parts of the relative effective permeabilities along the x- ($\mu_x$) and the y-directions ($\mu_y$). (b) Corresponding imaginary parts. (c) Sum of the imaginary parts of $\mu_x$ and $\mu_y$.

6.3.4 Remarks on the Working Principle of theMagnetically Lossless 2-DEG Configuration

The presented 2-DEG structure allows for the realisation of ultralow- or zero-loss magnetic metamaterials over a continuous range of frequencies. As was highlighted at the beginning of section 6.3, a number of approaches have previously been proposed for compensating losses [37], [38] or even creating stable gain [36] in metamaterials. Those approaches usually relied on providing optical gain, e.g. in the form of optical parametric amplification [125] or electromagnetically induced chirality [38], or on the use of active, negative-resistance, diode elements, such as Gunn or resonant tunnel diodes [36], which enable ‘cancelling’ the metaparticles’ ohmic losses and even leading to magnetic metamaterials with gain. However, though promising, such
approaches normally require high field intensities and/or are accompanied by nonlinearities in the response of the engineered medium, which may not always be desirable. Moreover, their scaling from radio to visible frequencies (or vice versa) can be challenging. Clearly, a much more desirable and convenient approach would be to judiciously redesign the structure of the metaparticles at the unit-cell level, such that it could open a ‘window’ for harnessing perfectly lossless artificial magnetism over a continuous range of frequencies. This is the route that was followed here.

The approach presented herein is based solely on the exchange of active powers between the two electrically connected meshes inside each unit cell. This results in frequency regions wherein the sign of the active powers becomes negative (the meaning of which was explained in section 6.3.2) and the imaginary parts of the susceptibilities associated with the two meshes become positive (for the: exp(\textit{i}\omega t - \frac{i\omega r}{k} ) spatiotemporal dependence that was assumed herein), but not at the same frequency region. When an RLC mesh related to a permeability with positive imaginary part is placed on a different plane (say \textit{xz}) than the mesh it is electrically connected to, this results in lossless artificial magnetism for a plane wave that has its magnetic field component polarized perpendicularly to that plane (\textbf{H}_y) and propagating along any of the orthogonal axes (x or z) of that plane.

The scheme works by invoking (at the unit-cell level) a mechanism that manages to naturally ‘pump’ magnetic energy in a specified direction – along which we obtain lossless metamaterial magnetism. As expected, the mesh that absorbs (receives) the electrical active power from its pair is associated with increased magnetic losses (in the same frequency region over which its pair is magnetically active), so that the conservation of energy is preserved. The scheme results in a light beam propagating without magnetic losses along a specified direction inside the metamaterial, whilst another beam, propagating perpendicularly to the first, naturally ‘supplies’ the required ‘gain’ to the first beam and is itself experiencing increased magnetic losses. Note the subtle, but important, conceptual difference between the present approach and those that rely on provision of optical or electrical gain: The latter approaches provide gain to cancel (or even reverse) the magnetic losses that are already present in the metamaterial. By contrast, the approach presented herein relies on ‘enforcing’ one of the two meshes not to absorb the incident magnetic energy at all, but remit it to the mesh it is electrically connected with. As a result, one of the two meshes naturally
becomes a ‘source’ of energy, which is the reason behind the occurrence of positive imaginary part for the magnetic susceptibility associated with that mesh.

From the examples presented in section 6.3.1 it should also be clear that the present scheme does not require large intensities for the incident fields – in fact, it works equally well even with low field intensities. Indeed, the results shown in section 6.3.1 (Fig. 6.4) were obtained assuming incident magnetic fields having intensity of just 1A/m. Obviously, similar results can be obtained with even smaller magnetic intensities, so long as they suffice to excite a useful or interesting collective magnetic response of the effective medium.

Furthermore, it turns out that the underlying mechanism (i.e., the exchange of active power between the meshes) that is responsible for enabling lossless artificial magnetism is very robust against the presence of ohmic losses, insofar as the difference in (not the actual values of) the ohmic resistances of the meshes is substantial – in the examples of the section 6.3.1 it was: $R_1 = 50 \, \Omega$ and $R_2 = 0.1 \, \Omega$. Indeed, from Fig. 6.13 one observes that even when the resistance in each mesh increases by an order of magnitude ($R_1 = 500 \, \Omega$, $R_2 = 1 \, \Omega$) there are, as before, regions wherein the imaginary parts of the effective permeabilities become positive. The same holds true for any increase in $R_1$ and $R_2$, to the extend that we do not enter the ‘overdamped’ region wherein there is no effective magnetic oscillation (response) in the effective medium at all. Ultimately, this robustness against losses is, as highlighted before, owing to the fact that the underlying mechanism relies critically on an imbalance in the values of the resistances of the pair of meshes, and not on the actual values of the resistances themselves. As long as such an imbalance is present (and provided that we are not in the ‘overdamped’ regime) there will always be power flowing away from a mesh to its pair, resulting in algebraically negative active power for the mesh that remits this power and in, correspondingly, positive imaginary part of the associated magnetic susceptibility/permeability.

This suggests the remarkable (and counterintuitive) prospect of actually deploying ‘poor’ conductors to create a perfectly lossless magnetic metamaterial in a specified direction (or directions, see section 6.3.6). For instance, let us assume that an 1-DEG system, corresponding to the equivalent electrical circuit of a split ring resonator (SRR), uses a ‘good’ conductors of resistance $0.1 - 1 \, \Omega$, which is typical for conductors in the GHz regime [4]. As we show in section 6.2, a periodic arrangement
Figure 6.13: Imaginary parts of \( \mu_x \) and \( \mu_y \) for the electromagnetic and geometric parameters used in Fig. 6.4, but now with \( R_1 = 500 \, \Omega \) and \( R_2 = 1 \, \Omega \). (a) Case of low-density (weakly interacting) unit cells. (b) Case of ‘tightly coupled’ unit cells.
of the single RLC mesh (or SRR) in this 1-DEG system will, in the quasi-static regime, result in an effective medium exhibiting artificial magnetism [4], but limited by the presence of the ohmic losses. Instead of attempting to further reduce the losses present in this medium (by, e.g., using even better conductors for the single meshes), our analysis shows that one may (or, in fact, should) deploy a ‘poor’ conductor of resistance 50–500 Ω to create a second RLC mesh on a different plane than the first. The imbalance in the values of the meshes’ resistances residing at the two planes will give rise to power being exchanged between the two meshes which, according to what was explained above, will cause positive imaginary parts for the associated permeabilities over certain frequency regions, and therefore in perfectly lossless magnetism in the corresponding directions – in fact, even with the presence of gain.

The approach introduced here for creating lossless magnetic metamaterials can be most conveniently realized experimentally in the radio and microwave frequencies by using discrete lumped resistors, inductors and capacitors, or SRRs placed at different planes and made of different conductors. The scheme is also scalable down to optical frequencies, where magnetic metamaterials made of arrays of SRRs have already been demonstrated. A further method for the construction of the herein proposed structures at optical frequencies could be the use of discrete nanoresistors, nanoinductors and nanocapacitors [126], which have already been studied and were shown to hold promise in connection with the creation of optical nanocircuits [127] and nanoantennas [128].

Finally, it should be noted that the electrical connection of the meshes is a crucial aspect of the proposed mechanism for overcoming metamaterial losses, not only because it allows electrical power to flow and be exchanged more easily between the meshes, but also because it allows each magnetic field component to generate a magnetic moment at, both, the plane to which it is perpendicular and at a plane to which it is parallel. For instance, the \( H_z \)-field component generates currents circulating both meshes residing at the \( xz \) and \( yz \) planes. As a result, the \( H_z \)-field component induces a magnetic moment not only along the \( y \)-direction (perpendicularly to the \( xz \) plane), but also along the \( x \)-direction (perpendicularly to the \( yz \) plane, to which the \( H_z \)-field component is parallel). The corresponding is, of course, also true for the \( H_y \)-field component. This superposition of the magnetic moments generated, at both planes, by both \( \mathbf{H} \)-field components is an essential feature of the present design. Note also that in order for the ‘superposition principle’ to be in
force, our present scheme not only is it not nonlinear, but it actually requires linearity and low or moderate field intensities to function according to its conception.

6.3.5 Issues with the Homogenisation Method for the Case of ‘Tightly Coupled’ Unit Cells

As was noted in Fig. 6.12(c), when the density of the unit cells is high (‘tightly coupled’ unit cells) application of Eq. (6.19) leads to a certain frequency region wherein \( \text{Im}\{\mu_1\} + \text{Im}\{\mu_2\} > 0 \), implying the occurrence of ‘net’ gain in that frequency region, i.e. that the magnetic gain that a beam experiences along, e.g., the \( x \)-direction is larger than the dissipation that a second beam experiences along, e.g., the \( y \)-direction. Such an outcome cannot be physically justified, since in our scheme we do not use an active medium or electronic element to provide gain to the structure. It is interesting to note that similar observations, i.e. occurrence of positive imaginary parts (corresponding to ‘gain’) in the spectra of effective permittivities and permeabilities, have in the past also been reported in a number of theoretical [119]-[124], but also experimental [129], [130] studies. However, in those cases such outcomes were not occurring at the unit-cell basis and were entirely an artefact of the periodicity of the structure, i.e. they were only appearing owing to the inappropriateness of describing a periodic medium in terms of a homogeneous effective medium for wavelengths that are not sufficiently larger compared to the periodicity of the engineered structure. As was noted in [124], in such a pseudo-effective medium, spatial dispersion (i.e., dependence of the effective electromagnetic parameters not only on the frequency, but also on the wavevector), as well as resonant-antiresonant coupling between the electric and magnetic responses of the medium, should not be ignored. Those effects, though substantially diminished, still persisted even for relatively large wavelengths [124].

As was explained in section 6.3.1, as well as in the previous section 6.3.4, in the designs presented herein the presence of positive imaginary parts in the spectra of the effective permeabilities is physically fully justified on the basis of the exchange of active power between the electrically connected meshes. The possibility that the observed effects could arise due to the periodicity of the structure was further eliminated by studying the magnetic moments of the individual unit cells and noting
that they possessed positive imaginary parts. No periodic boundary conditions were used at the edges of the cells, and the analytic, exact circuit calculations in each unit cell revealed the presence of active powers with negative algebraic sign. The herein reported positive imaginary parts in the spectra of the calculated effective permeabilities will clearly continue to occur even in the deep-subwavelength ('true' effective medium) regime, as long as frequency regions wherein $P_i < 0$ ($i = 1, 2$) are suitably engineered. For the same reason, the effects reported here are, also, not related to a resonant-antiresonant coupling between electric and magnetic dipoles induced in the structure.

It should be noted that in our structures frequency regions wherein $\text{Im}\{\mu_1\} + \text{Im}\{\mu_2\} > 0$ never occur for the case of low-density, weakly interacting ('isolated') unit cells. This is simply because in this case the effective (relative) permeability of the structure along a particular direction is directly proportional to the magnetisation $M$ ($\mu_{r,\text{eff}} = 1 + M/H_0$) and therefore, ultimately, directly proportional to the active power of the corresponding mesh, which takes on negative values in a certain frequency region. This observation is further verified in Fig. 6.14(a), which shows the frequency variation of the sum $\text{Im}\{\mu_1\} + \text{Im}\{\mu_2\}$ for the case of 'isolated' unit cells and for the structure that we studied previously in Fig. 6.13(a). Evidently, the aforesaid sum is negative throughout, in accord with what one would expect for a structure to which there is no net supply of gain.

Figure 6.14(b) reports the variation of the same quantity with frequency, but now for the case of strongly interacting cells. Here, because of the damping of the magnetic resonance caused by the increased values of resistances in each mesh, the region wherein $\text{Im}\{\mu_1\} + \text{Im}\{\mu_2\} > 0$ is more clearly seen. Before we proceed into furnishing a plausible explanation for this observation, it is important at this point to reiterate that the mechanism for overcoming losses and producing magnetic gain in our structure has a local origin, i.e. it occurs because we are able to design the meshes in each unit cell so that they do not consume active power in certain frequency regions, but remit it to each other. As a result, each individual unit cell does not absorb but amplifies the corresponding magnetic flux density component. The working principle of our scheme does not rely at all on (destructive) interference amongst the cells to eliminate losses, i.e. it does not have a 'global' origin, but a much stronger 'local' one, occurring at the unit-cell level. It follows that regardless of how
Figure 6.14: Sum of the imaginary parts of $\mu_x$ and $\mu_y$ for the electromagnetic and geometric parameters used in Fig. 6.4, but now with $R_1 = 500$ $\Omega$ and $R_2 = 1$ $\Omega$. (a) Case of low-density (weakly interacting) unit cells. (b) Case of ‘tightly coupled’ unit cells.
strongly or weakly do the unit cells interact with each other or the precise form of
their interaction, the collective (effective) medium must necessarily also exhibit zero
absorption of the magnetic flux density component in a specified direction (and, also,
$\text{Im}\{\mu_1\} + \text{Im}\{\mu_2\} > 0$, so that the conservation of the magnetic energy is honoured).
As a result, the reason for the presence of frequency region wherein $\text{Im}\{\mu_1\} + \text{Im}\{\mu_2\}$
$> 0$ should be traced to the limitations of Eq. (6.19) in describing such a medium.

Such an explanation, indeed, becomes plausible when one considers the precise
methodology that should be followed in assigning bulk electromagnetic parameters in
a medium. The local quasi-static electromagnetic field components will be owing to,
both, the incident field and the field scattered and/or induced by the neighbouring
metaparticles. For the methodology to be self-consistent, one should start by
computing the electric and magnetic multipoles induced by the incident field [114]. In
doing so, one expands the electric and magnetic dipoles in terms of the incident
electromagnetic field components and their derivatives, and the electric quadrupole
tensor in terms of incident electric field. One then proceeds by determining the
multipolar coefficients in the previous expansion by deploying, e.g., time-dependent
quantum perturbation theory. In our case, the determination of the scattered fields
should also take into account the electrical connectivity of the equivalent meshes
which, as we show, results in (active and reactive) power being exchanged between
them. However, the most important point for our discussion here is that, unless the
unit cells are very weakly interacting, the higher-order multipolar terms cannot be
ignored, even in the long-wavelength, quasi-static regime [131]. Moreover, it turns
out that these terms further depend on the origin of our coordinate system, a point that
requires careful treatment before meaningful effective-medium parameters can be
assigned. Such a methodology should –according to what was explained above– result
in the summation $\text{Im}\{\mu_1\} + \text{Im}\{\mu_2\} < 0$ at every frequency point, as is the case with
the ‘isolated’ cells. Here, however, it is the prospect of overcoming losses in
metamaterial that is investigated, leaving the detailed development of a more
appropriate homogenisation methodology to be the subject of a future work.
6.3.6 Systems with \( M \) Degrees of Freedom

The notion of, so called, systems having \( M \) degrees of freedom is a one that is very frequently encountered in diverse realm of science, from civil or mechanical engineering to quantum mechanics. In these situations, the properties (e.g., movement or oscillation) of a multi-degrees-of-freedom system in the real or in the phase space are described with the aid of \( M \) independent parameters, \( u_1-u_M \), known as generalised Lagrange coordinates of the system. The number of these parameters (or “coordinates”) depends upon the particular form or structure of the system, on the way it is excited, as well as on the required accuracy. Generally, increasing the number of the independent parameters in the description of a system also increases the accuracy of the obtained results. As a result, there is only a limited number of cases where, e.g., an infinite-degrees-of-freedom system (also known as a continuous system), such as a transmission line, is described in terms of an 1-DEG equivalent system. By contrast, with the aid of a relatively small number of independent parameters, the description of a continuous system in terms of an \( M \)-DEG system can normally be accommodated with sufficient accuracy.

Figure 6.15: Main modes of oscillation of a 3-DEG vertical rod oscillator.
Figure 6.15 schematically illustrates an example of a 3-DEG system, frequently encountered in the realms of, e.g., civil or mechanical engineering. It shows the three main modes of oscillation of 3-DEG oscillator composed of three masses, $m_1 - m_3$, which are elastically attached into a vertical pole. The latter may, e.g., be simulating a concrete rod in a building or a shaft joining parts of a machine. It turns out [113] that the $3 \times 3$ matrix describing the movement of this system is remarkably similar in its structure to the matrix resulting from the application of the ‘mesh current method’ to an electrical circuit of three coupled $RLC$ meshes. This is simply owing to the well-known analogy between mechanical and electrical oscillators, which is also applicable in the case of infinite-degree-of-freedom (continuous) systems.

The system that we introduced in this work for overcoming losses in metamaterials consisted of two electrically connected meshes; hence, it was described in terms of two independent parameters ($I_1$ and $I_2$) and the resulting matrix was $2 \times 2$, i.e. the system was a 2-DEG one. This arrangement enabled lossless propagation for one component of the magnetic flux density (propagating along any of the allowed two directions, orthogonal to the direction of this component) at a given frequency region – actually, as we show in Figs. 6.7, 6.12 and 6.13, the propagation could, indeed, be made lossless for two magnetic flux density components, but this occurred at different frequency regions for each component.

It is possible, however, to envisage a system with a larger number of meshes being electrically connected, e.g., a system wherein each mesh resides at a separate plane in a cubical cell, i.e. a 6-DEG system. Depending on the design, the structure may, of course, posses even more degrees of freedom (e.g., a 12-DEG system) if more than one meshes are electrically connected in each side of the unit cell. We speculate that with a judicious choice of the electrical ($R$, $L$, $C$) parameters, such a discrete (i.e., non-continuous) $M$-DEG system may enable lossless propagation for two out of the three orthogonal component of the magnetic flux density, in the same frequency region. This can occur as long as there is sufficient imbalance in the values of the resistances of the meshes to allow for active power to flow away from the meshes residing at two orthogonal planes (e.g., the $xz$– and $yz$–planes), towards the meshes residing at the third plane or planes (e.g., the $xy$–plane). In this manner, the meshes at the $xz$– and $yz$–planes will, at the same frequency region, act as ‘sources’ of electrical power, i.e. they will not absorb the incident magnetic energy but they will remit it to the meshes residing at the $xy$–plane. It follows that the imaginary parts of the
magnetic susceptibilities and permeabilities associated with the meshes at the \( xz \)- and \( yz \)-planes will be positive in the aforementioned frequency region (or regions), i.e. the structure will be magnetically active for the magnetic flux density components perpendicular to the \( xz \)- and \( yz \)-planes.

Finally, it should be emphasised that the present scheme requires that the meshes in each unit cell are \textit{not} electrically connected to the meshes of their neighbouring cells, i.e. the system should be discrete at the unit-cell level. If the meshes of neighbouring unit cells are electrically connected then, in the long-wavelength regime, the system becomes a continuous one (such as, e.g., a backward transmission line). These systems have been well-studied in the recent past [16] and, though they allow for the attainment of broadband and relatively low-loss metamaterials, they do not exhibit regions of positive imaginary parts for their effective electromagnetic parameters. As was detailed herein, the latter feat is ultimately accomplished due to the exchange of active power amongst the meshes, which results in lossless magnetism in each of the isolated, discrete unit-cells that ‘built’ the effective medium.
7 ‘Trapped Rainbow’ Stopping of Light in Metamaterials

Light usually propagates inside transparent materials with well-known ways that science students are taught. However, as we saw in the previous chapter, researchers have recently been examining the possibility of taking a normal transparent material and inserting tiny metallic inclusions (meta-molecules/particles) of various shapes and arrangements. As light passes through these structures, oscillating electric currents are set up that generate electromagnetic field moments (see, e.g., Eq. (6.13)), which modify the way the light travels through the material. The effects can be dramatic leading to light propagation or bending in very unusual ways (e.g. negative refraction, section 5.3). As we show in section 5.5, lenses that break traditional diffraction limits are one possibility [25]; an ‘invisibility cloak’ another [42], [43].

Significantly less research has focused on the potential of such structures for slowing, trapping and releasing light signals. In this chapter, we shall show that an axially varying heterostructure with a core of negative refractive index metamaterial can be used to efficiently and coherently bring light to a complete standstill. We will see that in stark contrast to previously proposed schemes for decelerating and storing light [44]-[54], the present one simultaneously allows for high in-coupling efficiencies and broadband, room-temperature, operation. Our analysis reveals that at a critical point the effective thickness of the waveguide reduces to zero, preventing the lightwave to propagate further. At this point, a light ray is found to be permanently trapped, its trajectory forming a double light-cone that we call an ‘optical clepsydra’. Each frequency component of a wave packet is stopped at a different guide thickness, leading to the spatial separation of the packet’s spectrum and the formation of a ‘trapped rainbow’. Our results, thus, bridge the gap between two important contemporary realms of science, that is metamaterials and slow light, and may open a new host of combined investigations. Such macroscopic control of photons may , also, conceivably find applications in optical data processing and storage and in the realisation of quantum optical memories.

Section 7.1 will present the main theory behind the slowing and stopping of light in metamaterial waveguides, with the remaining section 7.2 – 7.5 presenting elaborating remarks on and derivations of the main results described in section 7.1.
7.1 Broadband Stopping of Light in Metamaterial Waveguides

For decades scientists were arguing [132] “optical data cannot be stored statically and must be processed and switched on the fly”. The reason for this conclusion was that stopping and storing an optical signal by dramatically reducing the speed of light itself was thought to be infeasible. Undeniably, the absence of any form of interaction between photons and other elementary particles, as well as their enormous speed, makes confining them to a finite volume by reducing their velocity down to zero excessively difficult. However, in recent years a series of major scientific breakthroughs contributed towards annulling such assertion and proved conclusively that it is, indeed, possible to bring light to a complete standstill. Amongst others, electromagnetically induced transparency (EIT) [44], [48], quantum-dot semiconductor optical amplifiers (QD-SOAs) [49], photonic crystals (PhCs) [50], [51], coherent population oscillations (CPOs) [46], [47], stimulated Brillouin scattering (SBS) [52] and surface plasmon polaritons (SPPs) [53], [54] in metallodielectric waveguides have been proposed as means of producing ‘slow light’. However, so far most of these methods bear inherent limitations that may hinder their practical deployment. For instance, EIT uses ultracold atomic gases and not solid state materials, QD-SOAs usually allow for only modest delays but for potentially ultra-broadband light pulses, CPOs and SBS are very narrowband owing to the narrow transparency window of the former and the narrow Brillouin gain bandwidth (around 30 MHz in standard single-mode optical fibres) of the latter, SPPs are very sensitive to surface roughness and are relatively difficult to excite, while PhCs are normally highly multimodal [51]; this, combined with the strong impedance mismatch in the ‘slow-light regime’ causes launching the incoming light energy to a single, slow mode alone overly difficult [133].

As we saw in chapters 5 and 6, during the same period and in parallel with the above advances a different, wide-ranging realm of contemporary research has also been developing. It largely followed from a sequence of works by Pendry, Smith and co-workers, wherein they proposed practical means for realizing negative refractive index metamaterials [1], [2] and meticulously demonstrated their operation [3]-[7]. These materials offer a completely new perspective to the optical world and provide
additional ‘degrees of freedom’ in the design of photonic devices [42], thereby allowing unprecedented control over the flow of light. A ‘perfect’ lens [25], highly anisotropic electromagnetic ‘cloaks’ that render objects invisible to incident radiation [42], [43], as well as focusing of electron de Broglie waves by sharp p-n junctions in graphene [41] are a few characteristic examples. Moreover, there is by now compelling evidence that, via building on familiar transmission-line concepts borrowed from microwave analysis, such materials can be designed to exhibit broadband negative-index behaviour and relative robustness to losses, all through the radio [134] up to the visible [37] regime.

In this work we bring together the realms of metamaterials and slow light as a result of studying the stimulating physics associated with wave propagation in slowly, spatially varying negative-index heterostructures. To gain an insight into the physics of the problem at hand, let us for a moment imagine a ray of light zigzag propagating along a waveguide with a left-handed core. The ray experiences negative Goos-Hänchen lateral displacements [135] (see Eq. (5.15)) each time it strikes the interfaces of the core with the right-handed claddings. Accordingly, the cross points of the incident and reflected rays will sit inside the LH core and the effective thickness of the guide will be smaller than its natural thickness. It is reasonable to expect that by gradually reducing the core physical thickness, the dotted border lines indicating the effective thickness of the guide will eventually meet, i.e. the effective thickness of the guide, $t_{\text{eff}}$, will vanish. Obviously, beyond that point the ray will not be able to propagate further down, and will effectively be trapped inside the left-handed heterostructure (LHH). For nonradiative trapping of the ray, we hope that the point where the effective thickness vanishes will not occur below the lower cutoff frequency of the guide. If this condition is fulfilled, then one of the frequencies of a guided wave packet will be completely arrested. The arresting points for the other frequency components will be continuously spaced out. Thereby, one will in principle have a means of stopping light in its track and storing it indefinitely, without radiation, over a whole range of frequencies. Such an approach will also be totally immune to atomic decoherence effects [19] and will not require atomic media nor group index resonances.

In the following, we show that efficient, broadband, nonradiative storage of light can be realized inside an LHH. To this endeavour, we consider an axially nonuniform, linearly tapered, planar waveguide with a core of lossless (see section
6.3), isotropic and homogeneous negative-index metamaterial, bounded asymmetrically by two positive-index media, as illustrated in Fig. 7.1. The simultaneous negative sign of $\varepsilon$ and $\mu$ in the left-handed core allows for the existence of efficiently

![Guided electromagnetic wave](image)

**Figure 7.1:** *Trapped rainbow*: Different frequency components of a guided wave packet stop at correspondingly different thicknesses inside a tapered LHH. The upper (green) and bottom (cyan) dielectric layers, generally, vary between the right- and left-handed (RH and LH) waveguides. The thick empty arrows reveal the direction of power flow propagation ($P_{tot}^{\pm}$), while the thin black arrows show the direction of phase propagation ($\beta$). A guided wave packet is efficiently injected from the ordinary waveguide to the LHH, inside which propagates smoothly owing to the slow (adiabatic) reduction in the thickness of the core. Inside the RHH $P_{tot}^{\pm}$ and $\beta$ are parallel, but become oppositely directed in the LHH. The guided wave is altogether halted within a continuous range of 'critical' thicknesses, as designated by the short vertical line inside the LH core. The smallest (red) frequency components of the wave are stopped at the smallest core thicknesses, while the largest (blue) components stop at correspondingly larger core thicknesses. Thereby, the spectrum of the oscillatory field will be spatially decomposed into its frequency constituents, similar to the decomposition of sunlight when it illuminates water droplets and the appearance of a rainbow. The inset shows an example of the dependence of the 'critical' thicknesses of the LHH on the spectrum of the wave packet. The optogeometric parameters used in this example are those of Fig. 7.2
excitable (oscillatory) waveguide modes (see section 5.7) [55], which are not supported by single-negative (SNG) plasmonic structures. We arrange the variation of the core half thickness $a$ with distance $z$ to be appropriately slow [57] (adiabatic), $\frac{da}{dz} < 0.05 \times \min\{(ak_0)\times(n_{\text{eff}2} - n_{\text{eff}3})/(2\pi)\}$, with $ak_0$ being the reduced guide thickness, $k_0$ the free space wavenumber and $n_{\text{eff}2}, n_{\text{eff}3}$ the effective indices of the second and third-order oscillatory modes [24], so that the power of a local mode is conserved along the structure. One may then analytically show (see section 7.2) that the electromagnetic fields at a position $z_t$ are given by:

$$G(x, y, z_t, t) = F(z_t)g(x, y, \beta(z_t))\exp\left[i\sum_{z=0}^{z_t} \beta(z)\Delta z - i\omega t\right].$$ (7.1)

Here, $G = E$ or $H$, $g$ is the solution to the $E$- or $H$-field vector wave equation for $z = z_t$, $\beta$ is the longitudinal propagation constant, $\omega$ is the wave angular frequency and $F$ is an appropriate factor, which carries all the information about the mode energy conservation and is calculated as:

$$F = 2\sigma_e U\left[\frac{\alpha \beta |e_{r1}|P_{\text{tot}}^{\text{c}}}{\alpha \beta (W_3^2 + \sigma_e U^2)}\right]^{1/2},$$ (7.2)

where $\sigma_e = e_{r3}/|e_{r1}|$, $U = \alpha \kappa$, $W_3 = \alpha \gamma_3$ are the reduced transverse and layer-3 decay constants, respectively ($\kappa$ and $\gamma_3$ are the corresponding unnormalized constants), $P_{\text{tot}}^{\text{c}} = |P_1| + |P_2| + |P_3|$ is the conserved total time-averaged power flow, with $P_i$ ($i = 1, 2, 3$) being the time-averaged power flow in layer $i$, propagating in the increasing $z$ direction, and:

$$\Theta = 2 + \frac{\rho_e}{W_2 W_3^2} \frac{U^2 - W_3^2}{W_2^2 + \rho_e U^2} + \frac{\sigma_e}{W_3 W_3^2} \frac{U^2 - W_3^2}{W_3^2 + \sigma_e U^2},$$ (7.3)

with $\rho_e = e_{r2}/|e_{r1}|$ and $W_2 = \alpha \gamma_2$ being the reduced decay constant in the lower layer ($\gamma_2$ is the unnormalized decay constant for layer 2).
One can judiciously choose the optical parameters of the left-handed heterostructure (LHH) to facilitate total suppression of all surface polariton (SP) modes, i.e. the minimum thickness of the core layer can be chosen above the upper cutoff of the parasitic SP mode, so that only oscillatory guided modes may exist [56], [64]-[65].

Figure 7.2: Wave propagation in an axially varying left-handed heterostructure. The optical parameters of the structure are $\varepsilon_{r1}^{LH} = -5$, $\varepsilon_{r2}^{LH} = 2.56$, $\varepsilon_{r3}^{LH} = 2.25$, $\mu_{r1}^{LH} = -5$ and $\mu_{r2}^{LH} = \mu_{r3}^{LH} = 1$. Shown is a snapshot of the propagation of the monochromatic ($f = 1$ THz) $p$-polarised magnetic-field component, which enters the LHH from the wide-thickness end ($ak_0 = 1.7$) and stops at the pre-arranged ‘critical’ thickness of $ak_0 \approx 0.55$. The light signal carries a total power $P_{tot}^{con} = 82.6 \mu W/m^2/sec$, which to an excellent approximation is assumed to be unaffected by the negligibly small reflections from the approaching media interfaces. The light signal propagates inside the LHH with its group ($v_g$) and phase ($v_{ph}$) velocities being antiparallel. The group velocity is parallel to the total power flow $P_{tot}^{zz}$, while the phase velocity is parallel to the longitudinal propagation constant $\beta$. As the lightwave propagates, its amplitude progressively increases; at the critical core thickness it has increased by a factor of 4 compared to the amplitude of the field at the wide-thickness end. Moreover, we find that the exponentially decaying extension of the $H_y$-field inside the lower dielectric layer increases, while the wavelength gradually reduces and the field becomes spatially compressed.
For such a structure, Fig. 7.2 furnishes an example of *ab initio* calculations, following the previously outlined methodology, of the propagating, monochromatic, $p$-polarised magnetic-field component at a time instant $t = 0$. We find that the magnitude of the total time-averaged power flow propagating in the positive $+z$ direction $P_{\text{tot}}^{+z} = \frac{1}{2} \int_{-\infty}^{\infty} \text{Re}(E \times H^*)_z \, dx$, gradually drops off until it totally peters out, even though $P_{\text{tot}}^{\text{con}}$ is conserved along the nonuniform waveguide. Accordingly, one discovers that whilst the guided oscillatory fields propagate along the structure, with their phase ($v_{ph}$) and group ($v_g$) velocities being antiparallel, the group and energy ($v_e$) velocities [96] progressively decrease, eventually becoming zero at a ‘critical’, pre-determined, guide thickness. At this point, the fields are slowly spatially compressed and amassed [56], with their total amplitude increasing by a factor of 4 for the present set of optogeometric parameters. The exponentially decaying extension of the field inside the lower positive-index layer (medium 3) also progressively increases, as anticipated. Note that in a regular dielectric waveguide the monochromatic oscillatory field would have been reflected and radiated off at the guide cut off thickness. In that case, an increase in the total amplitude of the field would have been the result of interference between the two lightwaves propagating in opposite directions. In the present LHH case, however, the wave travels solely in the positive $z$ direction and completely stops upon reaching the ‘critical’ guide thickness; hence, neither reflection nor radiation or interference occurs.

The trapping of light is more intuitively recognized by tracing the trajectory of a light ray inside the core for guide thicknesses nearly equal to the critical one mentioned above. Recall that the variation of the core refractive index with distance $z$ is much smaller than the ray half period $z_{np}$, so that locally the guide appears practically uniform. Let us assume that the light ray arrives at the 1-3 media interface of the LHH with an angle $\theta$ (Figs. 7.3c-e). Following reflection from this interface, the ray experiences a Goos-Hänchen phase shift [135] $\delta_{p13} = -2\tan^{-1}[W_2/(\sigma_c U)]$, which is equal in magnitude but opposite in sign to the shift of the corresponding positive-index case. Detailed calculations then reveal that the distance $x_{p13}$ between the cross point A of the two rays (Figs. 7.3c-e) and the 1-3 media interface is (see section 7.3) $x_{p13} = (a\sigma V_3^2)/[W_3(W_3^2 + \sigma_c^2 U^2)]$, where $V_3^2 = U^2 + W_3^2$. Likewise, the distance $x_{p12}$
Figure 7.3: Ray analysis reveals that the effective LH guide thickness is smaller than the physical thickness and can become zero or even negative. (a) Variations of normalised effective guide thickness $t_{\text{eff}}/a$ (solid blue line), conserved $P_{\text{tot}}^{\text{con}}$ (dotted green line) and forward $P_{\text{tot}}^{+}$ (solid red line) total time-averaged power flow with the reduced guide thickness $a_k$. The forward component (red) of the conserved power flow (green) gradually decreases in magnitude, until it becomes exactly zero at the critical core thickness. At this point, the effective thickness of the LHH also vanishes. For larger guide thicknesses, $t_{\text{eff}}/a$ tends asymptotically to the value of 2. The insets associate characteristic regions of $P_{\text{tot}}^{+}$ with the ray analysis results shown in (c) – (e). In (b) – (e), a ray of light (black), which here signifies power propagation, hits the media interfaces with an angle $\theta$ while propagating down the waveguide, and experiences a Goos-Hänchen (GH) lateral displacement. The black dotted arrows denote the evanescent field power flow from the optically denser core to the rarer claddings. In a regular dielectric waveguide [ , ], the GH phase shifts are positive, the core appears to extend (dotted orange lines) inside the cladding layers and the effective guide thickness is always larger than the physical core thickness $2a$. In the slowly varying LHH case (c) – (e), the thickness of the core remains practically constant over many ray periods owing to the slow variation criterion. Here, the ray experiences a negative, i.e. antiparallel to $P_{\text{tot}}^{+}$, later displacement originating from the reversed power flow from the core to the claddings and the associated negative GH phase shift. In (b), the shifts are relatively small, such that $x_{12} + x_{13} < 2a$. Accordingly, the ray (blue) is effectively confined in the middle region of thickness $t_{\text{eff}}$ by repeatedly bouncing off points A and B. For an appropriate choice of optogeometric parameters (c), the two phase shifts can become such that $x_{12} + x_{13} = 2a$ exactly. In this case, the effective thickness of the guide ($t_{\text{eff}}$) vanishes and the ray becomes permanently trapped, forming a double light cone (‘optical clepsydra’). For even larger, compared to the ray period, later displacements (d) the effective thickness becomes negative; the ray is still guided in the middle region of thickness $|t_{\text{eff}}|$, but now is forward-propagating, i.e. the direction of phase propagation is parallel to the direction of power flow.
between point B and the 1-2 interface is \( x_{p12} = \frac{V_2^2}{\rho_2 V_2^2 + W_2^2 + \rho_2^2 U_2^2} \), where \( V_2^2 = U^2 + W_2^2 \). One is, thus, led to discern (Fig. 7.3c, blue line) that the light ray is effectively altogether confined within the middle region of thickness \( t_{eff} \) wherein it repeatedly bounces off points A and B.

We argue that \( t_{eff} = 2\alpha - x_{p12} - x_{p13} \) is the effective thickness of the left-handed waveguiding heterostructure. This conclusion is strongly supported by the preceding remarks concerning the ray trajectory and can be formally established by noting that the total time-averaged power flow \( P_{tot}^{**} \) can be directly linked to \( t_{eff} \) through the following relation (see section 7.4):

\[
P_{tot}^{**} = \frac{1}{4} E_{x}^{\max} H_{y}^{\max} t_{eff},
\]

where \( E_{x}^{\max} \) and \( H_{y}^{\max} \) are the maximum values of the \( E_x \) and \( H_y \) field components in the guide, respectively. The similarities between Eq. (7.4) and its counterpart in the case of conventional, right-handed heterostructures (RHH) [136], allow one to conclusively infer that \( t_{eff} \), as defined in Eq. (7.4) is, indeed, the effective thickness of the LHH. However, in stark contrast to conventional waveguides, \( t_{eff} \) is here always smaller than the physical thickness of the core and can become zero (Fig. 7.3(d)) or even negative (Fig. 7.3(e)). When \( t_{eff} \) becomes negative, we deduce from Eq. (7.4) that \( P_{tot}^{**} \) and \( \beta \) become parallel; thereby, the corresponding guided mode will be a forward-propagating one. Interestingly, based on Fig. 7.3(a) and Eq. (7.4) we infer that for a particular value of the guide’s physical thickness \( 2\alpha \), the effective thickness vanishes. In this case, the lateral shifts \( x_{p12}, x_{p13} \) experienced by the ray upon reflection from the two interfaces are such that \( x_{p12} + x_{p13} = 2\alpha \) exactly. For the aforementioned ‘critical’ physical thickness, the light ray is permanently trapped inside the LHH, being unable to propagate further down. From Fig. 7.3(d) we see that in this case the trajectory of the ray forms a double light-cone. In view of its characteristic form we will call it the ‘optical clepsydra’.

Following a similar course of analysis we discover that for different excitation frequencies the guided oscillatory fields stop at correspondingly different guide thicknesses. Accordingly, a guided electromagnetic wave packet can be altogether
trapped within a fixed area, spanning a continuous range of guide thicknesses. The leading (trailing) part of the pulse, composed of the smallest (highest) frequencies (Fig. 7.1), stops at the smallest (highest) guide thicknesses. Thereby, in the small-intensity, linear case wherein the propagating spectral power densities of a ‘white’ wave packet do not couple [57] and the guided field is a linear, weighted sum of its single frequency constituents, the ‘red’ and ‘blue’ components of the field will be spatially separated (Fig. 7.1), similar to the separation of the colours of the visible spectrum and the appearance of a rainbow when sunlight illuminates a transparent prism or falling water droplets. For this reason, we shall henceforth call the stopping and storing of light in such LHHs the ‘trapped rainbow’ effect.

A critical characteristic of the axially varying LHH is that further away from the point where the trapping of the light beam is arranged to occur, i.e. for larger guide thicknesses, it is possible to achieve complete impedance matching with a dielectric waveguide. We recall that for waveguide structures, the characteristic impedance is defined as $Z_0 = (V_0 V_0^*)/(2P_{tot})$, where $V_0$ is a ‘voltage’ defined as the line integral of the electric field along some path, which starts from below the lower interface and ends amply above the upper one [137]. Figure 7.4 illustrates the variation with reduced guide thickness of the analytically calculated (see section 7.5) ordinary and LH waveguide characteristic impedances. At a point sufficiently far from the ‘critical’ thickness of the LHH, its characteristic impedance becomes equal to that of a regular waveguide. At this point ($\approx 12.76 a k_0$ in our case) the two structures also have equal thicknesses. Moreover, the spatial distribution of the guided field at the wide end of the LHH (Fig. 7.4 inset, red line) closely matches the field distribution of a single-mode optical waveguide. Accordingly, a lightwave launched from a dielectric guide to a wide-thickness LHH will experience minimal reflection, mainly owing to minute mode-mismatch, which can be further adjusted and optimised at will.

In summary, we have shown how guided electromagnetic fields can efficiently be brought to a complete standstill whilst travelling inside axially varying LH waveguiding heterostructures. By nature, the scheme invokes solid-state materials and, as such, is not subject to low-temperature or atomic coherence limitations. Moreover, it inherently allows for high in-coupling efficiencies and broadband function, since the deceleration of light does not rely on refractive index resonances.
This ‘trapped rainbow’ method for storing photons opens the way to a multitude of hybrid, optoelectronic devices to be used in ‘quantum information’ processing, com-

Figure 7.4: Honed conditions for waveguide coupling: Simultaneous impedance, thickness and mode matching in adjoining RHH and LHHs. The optical parameters of the LHH are similar to those in Fig. 7.2. For the RHH we have $e_{r_1}^{RHH} = 1.5625$, $e_{r_2}^{RHH} = 1.44$, $e_{r_3}^{RHH} = 1.21$, $\mu_{r_1}^{RHH} = \mu_{r_2}^{RHH} = \mu_{r_3}^{RHH} = 1$. We note that the characteristic impedance of the dielectric waveguide (blue) exhibits a minimum at $ak_0 \approx 4.53$, after which it grows monotonically. A similar trend is found for the characteristic impedance of the LHH (red). Here, the minimum value occurs at a guide thickness $ak_0 \approx 0.89$ and at the ‘critical’ LHH thickness it diverges. The two curves cross at $ak_0 \approx 12.77$. The inset shows the profile of the fundamental (blue) and second-order oscillatory (red) mode of the dielectric waveguide and LHH, respectively, at the cross point. The darker the shaded region in the inset, the higher is the magnitude of the refractive index.
munication networks and signal processors, and conceivably heralds a new realm of combined metamaterials and slow light research.

7.2 Derivation of Spatiotemporal Field-Component Equations Used in the Adiabatic Variation

For the applicability of the adiabatic approximation one needs to ensure that the variation of core half-thickness $a$ with propagation distance $z$ is properly slow. Starting from Maxwell's equations and by deploying coupled-mode theory, we can formally show [57, Chs. 19 & 28] that the requirement for slow core thickness variation is fulfilled when the length of each tapered waveguide segment is large compared with the largest distance over which the guided fields can change appreciably owing to phase differences between the supported local modes. This leads to the following 'axial variation criterion':

$$\frac{da}{dz} \ll \frac{a(\beta_2 - \beta_3)}{2\pi},$$

(7.5)

where $\beta_2$ and $\beta_3$ are the scalar propagation constants of the second- ($m = 1$) and third-order ($m = 2$) backward waveguide modes $TM_{m+1}$ of the LHH [55]. This criterion can also take the form:

$$\frac{da}{dz} \ll \frac{(ak_o)\Delta n_{2,3}^3}{2\pi},$$

(7.6)

which is used in the main text of the letter.

The electromagnetic fields $G = E$ or $H$ at a distance $z = z_t$ inside an axially varying LHH, which satisfies the above criterion, are given by:

$$G(x, y, z_t, t) = F(z_t)g(x, y, \beta(z_t))\exp\left[ i \sum_{z=0}^{z_t} \beta(z)\Delta z - i\omega t \right].$$

(7.7)
Here the parameter $F$, which is the positive constant used in the solution ansatz to the wave equation (see Eq. (S14)), is chosen in such a way that $\partial P_{\text{tot}}^{\text{con}}/\partial z = 0$, where $P_{\text{tot}}^{\text{con}} = \sum_{i=1}^{3} |P_{i}^{*z}|$ is the conserved total time-averaged power flow in the LHH [55], [56] and $P_{i}^{*z}$ the time-averaged power flow in the $i$-layer ($i = 1, 2, 3$), propagating in the increasing $+z$ direction. In order to enforce the conservation of $P_{\text{tot}}^{\text{con}}$ in the analytic computations, we normalize the fields in such a way that, at each tapered waveguide segment, they carry a total (conserved) time-averaged power flow equal to $P_{\text{tot}}^{\text{con}}$. To this end, we start by calculating $P_{i}^{*z} = 1/2 \int_{A_i} \text{Re}(E \times H^*) dA$ in each waveguide layer and, after some algebraic manipulations [55], [56] we arrive at:

$$P_{\text{tot}}^{\text{con}} = \frac{F^2}{4\omega \varepsilon_0 |\varepsilon_{ri}|} \frac{\alpha \beta}{\varepsilon_{ri}} \left( \frac{W_3^2 + \sigma_e^2 U^2}{\sigma_e^2 U^2} \right) \left( 2 + \frac{\rho_e}{W_2} \frac{U^2 - W_2^2}{W_2 W_3^2 + \rho_e^2 U^2} + \frac{\sigma_e}{W_3} \frac{U^2 - W_3^2}{W_3 W_3^2 + \sigma_e^2 U^2} \right),$$

(7.8)

From Eq. (7.8), it follows that by requiring at each segment of the tapered waveguide:

$$F = 2 \sigma U \left[ \frac{\omega \varepsilon_0 |\varepsilon_{ri}| P_{\text{tot}}^{\text{con}}}{\alpha \beta (W_3^2 + \sigma_e^2 U^2) \Theta} \right]^{1/2},$$

(7.9)

with $\Theta$ defined in Eq. (7.3), we ensure that the guided electromagnetic field carries a constant (conserved) total power flow, equal to $P_{\text{tot}}^{\text{con}}$, throughout the LHH. Note that the fields are normalised with respect to $P_{\text{tot}}^{\text{con}}$, not $P_{\text{tot}}^{*z} = 1/2 \int_{-\infty}^{\infty} \text{Re}(E \times H^*) dx$, since the latter one does not remain constant along the LHH, as in regular dielectric guides but, instead, it continuously decreases until it becomes zero at the ‘critical’ guide thickness. Normalising the fields with $P_{\text{tot}}^{*z}$ instead of $P_{\text{tot}}^{\text{con}}$, would have caused their unphysical divergence at the point where they are stopped.
7.3 Derivation of the Expressions for Light-Ray Goos-Hänchen Spatial Displacements

Let us assume that a ray of $p$-polarised light impinges upon the 1-3 media interface with angle $\theta$ (see Fig. 7.3c). Following a course of analysis similar to that followed for dielectric waveguides, one may show that the associated Goos-Hänchen phase shift will be [57], [135]:

$$\delta_{p13} = -2\tan^{-1}\left(\frac{W_3}{\sigma_3 U}\right). \quad (7.10)$$

For the sake of convenience in the subsequent algebraic manipulations, let us rewrite Eq. (7.10) in the following form:

$$\delta_{p13} = -2\tan^{-1}(f(\theta)) = -2\tan^{-1}\left[\frac{(\sin^2 \theta - \sigma_e \sigma_\mu)^{1/2}}{\sigma_e \cos \theta}\right], \quad (7.11)$$

from whence we obtain:

$$\frac{d f(\theta)}{d \theta} = \frac{(1 - \sigma_e \sigma_\mu) \sin \theta}{\sigma_e \cos^2 \theta (\sin^2 \theta - \sigma_e \sigma_\mu)\sqrt{\sigma_\mu}}. \quad (7.12)$$

$$\frac{d}{d \theta} \{\tan^{-1}[f(\theta)]\} = \frac{\sigma_e (1 - \sigma_e \sigma_\mu) \sin \theta}{(\sin^2 \theta - \sigma_e \sigma_\mu)^{1/2} [(\sigma_e^2 - 1)\cos^2 \theta + (1 - \sigma_e \sigma_\mu)]}. \quad (7.13)$$

The inverted 'penetration' distance, $x_{p13}$ (see Figs. 7.3c-e), can now be calculated by means of the following relationship:

$$x_{p13} = \frac{1}{k_0 |n_1| \sin \theta} \frac{d}{d \theta} \{\tan^{-1}[f(\theta)]\}, \quad (7.14)$$

and we successively have:
where we used the identity $k_0^2 = (\gamma_3^2 + \kappa^2)/(n_1^2 - n_3^2)$. From Eq. (7.15d), it directly follows that:

$$x_{p13} = \frac{\sigma_x}{k_0(n_1^2 \sin^2 \theta - n_3^2)^{1/2}} \left(\frac{\sigma_x^2 - 1}{\cos^2 \theta + (1 - \sigma_x \sigma_\mu)}\right) = \frac{\sigma_x}{\gamma_3 \left(\frac{\sigma_x^2 \kappa^2}{k_0^2} + \frac{\kappa^2}{k_0^2} + n_1^2 - n_3^2\right)} =$$

$$= \frac{\sigma_x}{\gamma_3 \left(\frac{\sigma_x^2 \kappa^2}{\gamma_3^2 + \kappa^2} - \frac{\kappa^2 (n_1^2 - n_3^2)}{\gamma_3^2 + \kappa^2} + n_1^2 - n_3^2\right)} = \frac{\sigma_x}{\gamma_3 \left(\gamma_3^2 + \sigma_x^2 \kappa^2\right)}$$

(7.15a-b)

(7.15c-d)

which is the relation used in section 7.1. In a similar vein, one can prove that:

$$x_{p12} = \frac{\alpha \sigma_x}{W_2} \frac{V_3^2}{W_3^2 + \sigma_x^2 U^2}, \quad \text{which is the relation used in section 7.1.}$$

(7.16)

(7.17)

7.4 Derivation of the Relation Between the Total Time-
Averaged Power Flow and the Effective Thickness of
the Waveguide

The solution ansatz to the wave equation for the $p$-polarised oscillatory waveguide modes supported by the LHH has the following form:

$$H_y = \begin{cases} 
F e^{\kappa x}, & x \leq 0 \\
G \cos(\kappa x) + K \sin(\kappa x), & 0 \leq x \leq 2\alpha, \\
L e^{-(x-2\alpha)n}, & x \geq 2\alpha 
\end{cases}$$

(7.18)
where $G = F$, $K = -\frac{\gamma_3}{\sigma_z \kappa} F$, $L = [\cos(2\alpha \kappa) - \frac{\gamma_3}{\sigma_z \kappa} \sin(2\alpha \kappa)] F$, and $E_z = \frac{j}{\omega \varepsilon} \frac{\partial H_y}{\partial x}$.

$E_z = \frac{B}{\omega e} H_y$. From Eq. (7.18) we deduce that the maximum value for the $H_y$ field component is $H_y^{\text{max}} = F \frac{(W_z^2 + \sigma_z^2 U^2)^{1/2}}{\sigma_z U}$, and occurs within the middle layer-1 at point $x_{\text{max}} = \frac{\alpha}{U} \tan^{-1} \left( -\frac{W_z}{\sigma_z U} \right)$. After some algebraic manipulations, we can analytically calculate the total time-averaged power flow in the increasing +z direction, $P_{\text{tot}}^{+z}$, as [55], [56]:

$$P_{\text{tot}}^{+z} = \left[ \frac{F^2}{4 \omega e_0} \right]^{1/2} \sigma_z U^2 \left| \varepsilon_{\text{r}1} \right| \left( \frac{\rho_z}{W_2} + \frac{V_2^2}{W_2^2 + \rho_z^2 U^2} + \frac{\sigma_z}{W_3} + \frac{V_3^2}{W_3^2 + \sigma_z^2 U^2} - 2 \right) \frac{1}{x_{p12}/a} \frac{1}{x_{p13}/a}, \quad (7.19)$$

from whence we immediately obtain:

$$P_{\text{tot}}^{+z} = \frac{1}{4} \frac{E_x^{\text{max}} H_y^{\text{max}}}{t_{\text{eff}}}, \quad (7.20)$$

where $t_{\text{eff}}$ is defined in the main text. Note that, owing to the negativeness of the permittivity $\varepsilon_{\text{r}1}$ in the core of the LHH, the term $E_x^{\text{max}} H_y^{\text{max}}$ in the right-hand side of Eq. (7.20) is always negative; hence, $P_{\text{tot}}^{+z}$ and $t_{\text{eff}}$ are oppositely signed. It is should be herein noted that a negative $P_{\text{tot}}^{+z}$ corresponds to a negative phase velocity mode ($P_{\text{tot}}^{+z}$ antiparallel to the mode longitudinal propagation constant $\beta$) and a positive $P_{\text{tot}}^{+z}$ to a positive phase velocity mode ($P_{\text{tot}}^{+z}$ and $\beta$ are parallel).
7.5 Characteristic Impedance of Left- and Right-Handed Waveguides

In both cases we begin by calculating the 'voltage' $V_0$ across the waveguide by means of the relation:

$$V_0 = \int_{-\infty}^{\infty} E_x dx = \frac{\beta}{\omega_0} \int_{-\infty}^{\infty} \frac{H_y}{\varepsilon} dx. \quad (7.21)$$

We may then obtain the following expressions for the voltages $V_i (i = 1, 2, 3)$ crosswise each $i$-layer of the LHH:

$$V_{1,2,3}^{LHH} = \left( \frac{F}{\omega_0} \right) \frac{\alpha \beta}{\varepsilon_{ri}} \sqrt{W_3^2 + \sigma_z^2 U^2} \left( \pm \frac{W_2}{\sqrt{W_2^2 + \rho_z^2 U^2}} - \frac{W_3}{\sqrt{W_3^2 + \sigma_z^2 U^2}} \right), \quad (7.22)$$

$$V_2^{LHH} = \left( \frac{F}{\omega_0} \right) \frac{1}{\varepsilon_{ri} \sigma_z W_2} \sqrt{W_3^2 + \sigma_z^2 U^2}, \quad (7.23)$$

$$V_3^{LHH} = \left( \frac{F}{\omega_0} \right) \frac{1}{\varepsilon_{ri} W_3}, \quad (7.24)$$

from whence we find:

$$V_0^{LHH} = \sum_{i=1}^{3} V_i^{LHH} = \left( \frac{F}{\omega_0} \right) \frac{\alpha \beta}{\varepsilon_{ri}} \sqrt{W_3^2 + \sigma_z^2 U^2} \left( \pm \frac{V_2^2}{W_2 \sqrt{W_2^2 + \rho_z^2 U^2}} + \frac{V_3^2}{W_3 \sqrt{W_3^2 + \sigma_z^2 U^2}} \right), \quad (7.25)$$

where the "+" (plus) sign is used for $U \in [(m-1/4)\pi, (m+1/4)\pi]$ and the "-" (minus) sign for $U \in [(m+1/4)\pi, (m+3/4)\pi]$, with $m \in \mathbb{N}, U > 0$.

In a similar vein, using Eq. (7.18) and the parameter definitions that follow it, we obtain the following expressions for the $V_i (i = 1, 2, 3)$ voltages of the RHH:
\[
V_{1RHH} = \left( \frac{F}{\omega \varepsilon_0} \right) \frac{\alpha \beta}{\varepsilon_r} \frac{\sqrt{W'_2 + \sigma_e^2 U^2}}{\sigma_e U^2} \left( \pm \frac{W_2}{\sqrt{W'_2 + \rho_e^2 U^2}} + \frac{W_3}{\sqrt{W'_3 + \sigma_e^2 U^2}} \right),
\]
(7.25)

\[
V_{2RHH} = \pm \left( \frac{F}{\omega \varepsilon_0} \right) \frac{\alpha \beta}{\varepsilon_r} \frac{1}{\sigma_e W'_2} \sqrt{W'_2 + \sigma_e^2 U^2},
\]
(7.26)

\[
V_{3RHH} = \left( \frac{F}{\omega \varepsilon_0} \right) \frac{\alpha \beta}{\varepsilon_r} \frac{1}{\sigma_e W'_3},
\]
(7.27)

from whence we find:

\[
V_{0RHH} = \sum_{i=1}^{3} V_{iRHH} = \left( \frac{F}{\omega \varepsilon_0} \right) \frac{\alpha \beta}{\varepsilon_r} \frac{\sqrt{W'_2 + \sigma_e^2 U^2}}{\sigma_e U^2} \left( \pm \frac{V_{2RHH}^2}{W_2 \sqrt{W'_2 + \rho_e^2 U^2}} + \frac{V_{3RHH}^2}{W_3 \sqrt{W'_3 + \sigma_e^2 U^2}} \right),
\]
(7.28)

where the use of the "+" (plus) and "−" (minus) signs follows the same rule as in the case of the LHH. Moreover, one can show that the time-averaged power flow propagating in the increasing +z direction inside the RHH is given by [55]:

\[
P_{\text{tot}}^{+z} = \left( \frac{F^2}{4\omega \varepsilon_0} \right) \frac{W'_2 + \sigma_e^2 U^2}{\sigma_e^2 U^2} \alpha \beta \left( \frac{2}{\sigma_e W'_2 + \rho_e^2 U^2} + \frac{V_{2RHH}^2}{W_2 \sqrt{W'_2 + \rho_e^2 U^2}} + \frac{V_{3RHH}^2}{W_3 \sqrt{W'_3 + \sigma_e^2 U^2}} \right),
\]
(7.29)

By means of the power-voltage definition of the waveguide characteristic impedance [137]:

\[
Z_0^{pv} = \frac{|V_0|^2}{2P_{\text{tot}}^{+z}},
\]
(7.30)

using Eqs. (7.19) and (7.25) for the LHH, or Eqs. (7.28) and (7.29) for the RHH, one can now directly calculate the impedance for each waveguide. We note from Eq. (7.30) that the characteristic impedance of the LHH diverges at the ‘critical’ guide thickness, as anticipated, since in this case the heterostructure is in the ‘stopped light
regime' and, hence, the corresponding light signal can not penetrate it and propagate inside.
8 Conclusions and Further Work

In this work we have introduced a novel and efficient method for decelerating, over a range of frequencies, and completely ‘stopping’ light (zero group velocity, $v_g = 0$) inside solid-state materials, at room temperature. To this end, we have deployed negative-refractive-index (NRI) metamaterial (MM) waveguides and proved that light propagating inside such structures can be dramatically slowed-down and, even, be brought to a complete halt in a, simultaneously, broadband and efficient manner. In particular, using analytic simulations and ray-tracing analyses, we showed that an adiabatically tapered waveguide having a core of a lossless NRI metamaterial and claddings made of normal dielectrics can ‘trap’ a light pulse in such a way that each individual frequency component of the pulse is stopped at a different point along the waveguide, forming what we have called a ‘trapped rainbow’. It has, further, been shown that light can efficiently be in-coupled inside such a waveguide heterostructure from a normal dielectric waveguide, since with a suitable design one can achieve simultaneous thickness-, mode- and characteristic-impedance-matching between the two waveguides. Upon being injected inside the NRI waveguide, a light pulse propagates smoothly inside it (with negligible back-reflections, owing to the adiabaticity), until it reaches a ‘critical’ region wherein it is stopped and stored. At this point, a pertinent analysis reveals that the optical path length of a ray (associated with a particular frequency component of the pulse), as well as the effective thickness of the NRI waveguide itself become exactly zero, with the ray circulating at the point where it is trapped in such a way that its trajectory forms what we have called (in view of its characteristic hourglass form) an ‘optical clepsydra’.

Since maintaining ultra-low or zero material losses over a range of frequencies is a critical prerequisite for any slow-light configuration, we have also developed a novel methodology that allows for obtaining ultra-low- or zero-loss metamaterials over a continuous range of frequencies. In particular, we introduced a blueprint for designing magnetic metamaterials based on equivalent electrical circuits with more than one mesh. It has been analytically proved that such a higher-degrees-of-freedom design methodology leads to metamaterial magnetism with either ultra-high figures-of-merit or with perfectly lossless performance over a broad range of frequencies. Importantly, this methodology does not rely on cancellation of losses by means of gain media, but
only on a judicious design of meta-molecules. Therefore, the so-obtained lossless metamaterial magnetism – which is a crucial component in the design of NRI metamaterials – has a truly intrinsic character, and as such is scalable and can be implemented at any frequency regime, from the radio up to the optical domain. These design recipes are, thus, expected to lead to a new generation of ultra-low- or zero-loss metamaterials that can be exploited in the creation of, amongst other, ‘perfect’ lenses, ‘invisibility’ cloaks, optical nanocircuits, magnetic resonance imaging systems and slow-light waveguides.

Concerning the future directions of this work, one of our first priorities will be the development of accurate and numerically efficient modelling tools based on the recently introduced concepts of ‘exact’ and ‘nonstandard’ finite-differences that we described in chapter 4, which will facilitate even more expeditious and reliable evaluation and optimisation of NRI-MM designs. We, thus, plan to develop a fully-nonorthogonal (FNOG), nonstandard, alternating-direction implicit finite-difference time-domain (NS-ADI-FDTD) algorithm that can run on a desktop PC and that can accurately (when compared with reported experimental results) and efficiently (i.e. without noticeably raised overall computational overhead) describe light propagation in NRI-MM composite structures. A major achievement will also be the development of a parallel version of the aforementioned code, written in Fortran 90, which will lead to significant speed-up of the computations.

Building on the novel MM design recipes that we herein introduced (chapter 6), a further future goal will be to create a general theory for the description of wideband, low- or zero-loss, optical metamaterials (OMMs), followed by an identification of optimal-design proposals and rules. A first success here will be the numerical demonstration of planar, wideband NRI-MM structures in the infrared and visible regimes. A second, and much larger success, will be to numerically demonstrate, either by means of reflection/ transmission numerical spectra or by full-wave modelling of optical pulse propagation inside NRI-MM architectures, minimised or eliminated light energy losses. Inclusion of material optical gain into the aforementioned designs and algorithms will be a priority.

A further important direction that this work will follow is to comprehensively address the issue of controllability of NRI-MMs, which is of critical importance for the construction of controllable, variable optical memories. Success will be to demonstrate controllable negative permittivity and permeability values, either by
means of active elements, such as *pin* diodes and varactors, or by thermo-optic illumination.

Finally, an obvious (but necessary) extension of the results presented in chapter 7 will be to attest by means of *full-wave* solution of Maxwell’s equations without invoking standard paraxial or heuristic approximations that a light pulse propagating along a NRI waveguide can be brought to a complete halt, and then released by switching the optogeometric parameters. Here, success will be to demonstrate high in/out-(ridge/fibre) coupling efficiency, e.g. more than 90% transmission, for both quasi-TE and quasi-TM polarisations, and to present evidence that the so-developed low-loss waveguiding structure allows for *high normalised time delays*, i.e. that the storage time is several *hundreds or thousands* pulse lengths.
References

136. T. Tamir (ed.), *Integrated Optics* (Springer-Verlag, New York, 1979), Ch. 2.
Publications

Book Chapters:


Peer-Reviewed Scientific Journals and Magazines:


International Conferences & Topical Meetings:


Selected Citations:


