

# Materials for Optics and Optoelectronic Devices

Dr Gareth Siôn Jones  
University of Oxford  
Michaelmas 2025

## **Abstract**

These are working lecture notes prepared for Year 3 Options 1 - Materials for Optics and Optoelectronic Devices module at the Department of Materials, University of Oxford, delivered in Michaelmas term, 2025. This course was originally designed by Prof. Jason Smith, but in recent years has been delivered by Dr Lapo Bogani and Prof. Andrew Watt. These notes are strongly derived from, and further built upon, the original lecture notes of Prof. Smith. I have created an updated set of lecture notes based upon the original. here I am keeping with the original syllabus, but I provide my own thoughts in addition. Prof. Smith's lecture notes are also uploaded onto Canvas for reference, and I encourage you to review these, as well as the recorded lectures from Profs. Smith and Watt.

## Recommended Texts

1. Fox, M. *Optical Properties of Solids* (Oxford University Press, 2010)
2. Hecht, E. *Optics* 5th ed. eng (Pearson, Boston, Mass, 2017)
3. Grant, I., & Phillips, W.R. *Electromagnetism*, (John Wiley & Sons, 2013)
4. Wilson, J. & Hawkes, J. F. *Optoelectronics - An Introduction* (2nd edition (1989))
5. Senior, J. M. & Jamro, M. Y. *Optical Fiber Communications: Principles and Practice* (Pearson Education, 2009)
6. Rogers, A. *Essentials of Optoelectronics with Applications* (CRC Press, 1997)
7. Liu, J-M, *Photonic Devices* (Cambridge University Press, 2005)
8. Yariv, A. & Yeh, P. *Optical Waves in Crystals* (Wiley New York, 1984)
9. Nelson, J. A. *The Physics of Solar Cells* (World Scientific Publishing Company, 2003)

# Table of Contents

<b>1</b>	<b>Lecture 1 - The Classical Theory of Light</b>	<b>1</b>
1.1	The Classical Theory of Light . . . . .	1
1.1.1	The Wave Equation . . . . .	2
1.1.2	Solutions to Maxwell's Equations . . . . .	4
1.1.3	Polarisation . . . . .	5
1.1.3.1	Linear Polarisation . . . . .	6
1.1.3.2	Circular Polarisation . . . . .	7
1.1.3.3	Elliptical Polarisation . . . . .	8
1.1.3.4	Unpolarised . . . . .	8
1.1.4	Wave-Particle Duality . . . . .	9
1.2	The Electromagnetic Spectrum . . . . .	10
1.3	Interaction of Light with Bulk Matter . . . . .	12
1.3.1	Dielectrics . . . . .	12
1.3.2	Polarisation in Dielectric Materials . . . . .	14
1.3.3	The Lorentz Oscillator Model . . . . .	15
1.3.3.1	Macroscopic Permittivity and Frequency Dependence . . . . .	16
1.3.3.2	Refractive Index . . . . .	16
1.3.4	Complex Dielectric Function . . . . .	17
1.3.5	Dispersion of Light in a Transparent Medium . . . . .	18
1.3.5.1	The Sellmeier Equation . . . . .	18
1.3.6	Light Propagation: Superposition and Huygens' Principle . . . . .	19
1.4	Reflection/Refraction . . . . .	21
1.4.1	<i>s</i> and <i>p</i> Polarised Light . . . . .	22
1.4.2	Example Application - Polarised Sunglasses . . . . .	24
1.5	Optical Coatings . . . . .	25
1.5.1	Anti-Reflection Coatings . . . . .	25
1.5.2	Distributed Bragg Reflectors (Laser Mirrors) . . . . .	26
1.5.2.1	Photonic StopBand . . . . .	27
1.5.2.2	High and Low Index Materials . . . . .	28
<b>2</b>	<b>Lecture 2 - Optical Waveguides</b>	<b>31</b>
2.1	Planar Waveguides . . . . .	31
2.1.1	Mode Distribution in Planar Waveguides . . . . .	32
2.1.2	Applications of Planar Waveguides . . . . .	34
2.1.2.1	Photonic Integrated Circuits . . . . .	34
2.1.2.2	Semiconductor Lasers . . . . .	36
2.2	Non-Planar Waveguides (Optical Fibres) . . . . .	37
2.2.1	Mode Distribution in Optical Fibres . . . . .	39

2.2.1.1	Manufacture of Optical Fibres . . . . .	40
2.3	Degradation of Optical Signals in Fibres . . . . .	42
2.3.1	Attenuation . . . . .	42
2.3.2	Dispersion . . . . .	43
2.3.2.1	Inter-Modal Dispersion . . . . .	44
2.3.2.2	Intra-Modal Dispersion . . . . .	45
2.4	Wavelength Division Multiplexing . . . . .	47
<b>3</b>	<b>Lecture 3 - Controlling Light Propagation in Optical Materials</b>	<b>50</b>
3.1	Light in Optical Materials . . . . .	50
3.2	Light Propagation in Birefringent Materials . . . . .	50
3.2.1	Polarisation Control Using Wave Plates . . . . .	51
3.3	Electro-Optic Materials . . . . .	54
3.3.1	Derivation of the Electro-Optic Tensor . . . . .	55
3.3.2	Electro-Optic Response of LiNbO <sub>3</sub> . . . . .	58
3.4	Electro-Optic Modulation . . . . .	60
3.4.1	Amplitude Modulation . . . . .	61
3.4.2	Phase Modulation . . . . .	62
3.5	Frequency Conversion . . . . .	63
3.5.1	Sum-Difference Frequency Mixing . . . . .	64
3.5.2	Second Harmonic Generation . . . . .	66
3.6	Magneto-Optics . . . . .	66
<b>4</b>	<b>Metamaterials</b>	<b>68</b>
4.1	Engineering Negative Index Materials . . . . .	68
4.1.1	Negative Index Materials . . . . .	68
4.1.1.1	Engineering Negative $\epsilon_r$ . . . . .	69
4.1.2	Engineering Negative $\mu_r$ . . . . .	70
4.2	Transformation Optics . . . . .	71
4.2.1	Mathematical Formulation . . . . .	72
4.3	Invisibility Cloaks and Negative-Index Metamaterials . . . . .	73
4.3.1	The Cloaking Transformation . . . . .	74
4.3.2	Implementation with Metamaterials . . . . .	74
4.3.3	State-of-the-Art and Challenges . . . . .	76
4.4	The Perfect Lens . . . . .	76
4.4.1	The Conventional Diffraction Limit . . . . .	76
4.4.2	Pendry's Perfect Lens . . . . .	77
4.4.3	State-of-the-Art and Challenges . . . . .	79
4.5	Photonic Crystals . . . . .	80
4.5.1	Electronic Band Gaps . . . . .	80
4.5.2	Photonic Band Gaps . . . . .	81

4.5.2.1	Periodicity and Bloch's Theorem . . . . .	82
4.5.2.2	One-Dimensional Reduction . . . . .	83
4.5.3	Band Structures . . . . .	83
4.5.4	Photonic Crystal Fibres . . . . .	84
4.5.5	On-chip photonic crystal waveguides and cavities . . . . .	84
4.6	Further Modern Optics . . . . .	85
<b>5</b>	<b>Lecture 5 - The Semi-Classical Theory of Light</b>	<b>88</b>
5.1	Radiative Transitions in Atoms . . . . .	88
5.1.1	Frameworks . . . . .	90
5.1.1.1	Classical Theory . . . . .	90
5.1.1.2	Semi-Classical Theory . . . . .	90
5.1.1.3	Quantum Theory . . . . .	91
5.1.2	Einstein Coefficients . . . . .	91
5.1.2.1	Absorption . . . . .	92
5.1.2.2	Spontaneous Emission . . . . .	92
5.1.2.3	Stimulated Emission . . . . .	92
5.2	Radiative Transition Rates . . . . .	92
5.3	Electric Dipole Transitions . . . . .	92

# 1 | Lecture 1 - The Classical Theory of Light

## 1.1 The Classical Theory of Light

Prior to the mid-19th century, the study of electricity and magnetism existed as a collection of separate empirical laws and experimental observations. Scientists such as Charles-Augustin de Coulomb, Hans Christian Ørsted, André-Marie Ampère, and Michael Faraday had each uncovered pieces of the puzzle that described how electric charges and currents interact, but there was no unifying framework connected these discoveries. **Coulomb** had established that electric charges exert forces on one another inversely proportional to the square of their separation, laying the groundwork for electrostatics. **Ørsted's** experiments revealed that an electric current could deflect a magnetic needle, showing for the first time a direct link between electricity and magnetism. **Ampère** formulated mathematical laws describing how currents generate magnetic fields, while **Faraday** discovered that changing magnetic fields could in turn induce electric currents — a phenomenon known as electromagnetic induction. Despite these advances, the laws were viewed as distinct and phenomenological, describing specific interactions without suggesting a deeper unity or the existence of electromagnetic fields as physical entities in space. It was James Clark Maxwell who recognised in the 1860's that these disparate results could be understood as manifestations of a single, continuous electromagnetic field permeating all of space, governed by a set of elegant and universal equations.

Maxwell's equations are among the most profound and unifying achievements in physics, encapsulating the entire behaviour of electric and magnetic fields in a compact mathematical form. Maxwell's insight was to recognise deep symmetries among the prior laws and to introduce a single missing term — the displacement current — that completed the symmetry and ensured the self-consistency of the theory. With this addition, he showed that a time-varying electric field could act as a source of magnetic fields just as currents do. This insight not only unified the known laws of electromagnetism but also predicted the existence of electromagnetic waves that propagate through space at the speed of light, establishing light as an electromagnetic phenomenon.

The four equations can be viewed as the axioms from which all classical electromagnetic behaviour emerges. **Gauss's law** relates the electric field to the charge distribution, asserting that electric field lines begin and end on electric charges. **Gauss's law for magnetism**, on the other hand, states that magnetic field lines are continuous loops with no beginning or end—an expression of the empirical fact that isolated magnetic monopoles have never been observed. **Faraday's law**

**of induction** describes how a changing magnetic field generates an electric field, forming the basis for electric generators and transformers. Finally, the **Ampère-Maxwell law** connects magnetic fields to electric currents and time-varying electric fields, revealing how electromagnetic fields can sustain each other in the absence of charges or currents, leading to self-propagating electromagnetic waves.

In differential form, Maxwell's equations elegantly express local field behaviour through divergence and curl operations, while their integral forms describe global conservation laws over finite regions of space. Together with the Lorentz force law, which dictates how fields act on charged particles, they provide a complete and self-consistent description of classical electrodynamics. From the design of antennae and optical cavities to the understanding of light propagation and radiation, every classical electromagnetic phenomenon can be derived from these four fundamental relations. Maxwell's equations are not just a set of mathematical statements, but a conceptual bridge linking matter, fields, and the very nature of light. The four equations are shown below in differential form:

$\nabla \cdot \vec{D} = \rho_{\text{free}}$	(Gauss's Law for Electrostatics)
$\nabla \cdot \vec{B} = 0$	(Gauss's Law for Magnetism)
$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$	(Faraday's Law of Induction)
$\nabla \times \vec{H} = \vec{J} + \frac{\partial \vec{D}}{\partial t}$	(Ampère's Circuital Law)

### 1.1.1 The Wave Equation

The propagation of electromagnetic waves through a medium is governed by a second-order partial differential equation known as the **wave equation**. To derive this, first assume a linear, homogeneous, isotropic medium (constant  $(\epsilon, \mu)$ ), and recall that  $\vec{D}=\epsilon\vec{E}$  and  $\vec{B}=\mu\vec{H}$ .

Now, we can derive a wave-equation for the electric field  $\vec{E}$  by starting with Faraday's Law:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} = -\mu \frac{\partial \mathbf{H}}{\partial t} \quad (1.1)$$

If we take the curl of both sides, we have:

$$\nabla \times (\nabla \times \mathbf{E}) = -\mu \frac{\partial}{\partial t} (\nabla \times \mathbf{H}) \quad (1.2)$$

To proceed, we can apply a well-known vector identity:

$$\nabla \times (\nabla \times \mathbf{F}) = \nabla(\nabla \cdot \mathbf{F}) - \nabla^2 \mathbf{F} \quad (1.3)$$

If we apply this to the left hand side of equation 1.2 and substitute the Ampère-Maxwell law on the right, we have:

$$\nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\mu \frac{\partial}{\partial t} \left( \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \right) \quad (1.4)$$

Substitute ( $\nabla \cdot \mathbf{E} = \nabla \cdot (\mathbf{D}/\varepsilon) = \rho/\varepsilon$ ) and ( $\mathbf{D} = \varepsilon \mathbf{E}$ ) to get:

$$\nabla \left( \frac{\rho}{\varepsilon} \right) - \nabla^2 \mathbf{E} = -\mu \frac{\partial \mathbf{J}}{\partial t} - \mu \varepsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} \quad (1.5)$$

Finally, we rearrange to find the standard inhomogeneous wave equation for ( $\mathbf{E}$ ):

$$\boxed{\nabla^2 \mathbf{E} - \mu \varepsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} = \mu \frac{\partial \mathbf{J}}{\partial t} + \nabla \left( \frac{\rho}{\varepsilon} \right)} \quad (1.6)$$

In a source-free, homogeneous medium ( $\rho = 0$ ,  $\mathbf{J} = 0$ ), for a wave equation with phase speed ( $v = 1/\sqrt{\mu\varepsilon}$ ), this reduces to

$$\boxed{\nabla^2 \mathbf{E} - \mu \varepsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0} \quad (1.7)$$

### Example 1: Wave Equation

Starting from the Ampère-Maxwell law, repeat this to show that the wave equation for  $\mathbf{H}$  is:

$$\boxed{\nabla^2 \mathbf{H} - \mu \varepsilon \frac{\partial^2 \mathbf{H}}{\partial t^2} = -\nabla \times \mathbf{J}} \quad (1.8)$$

Now, with both forms for the electromagnetic wave-equations, we can define the constant  $c_0$ :

$$c_0 = \frac{1}{\sqrt{\mu_0 \epsilon_0}} = 2.998 \times 10^8 \text{ ms}^{-1} \quad (1.9)$$

.... and then the wave equations can be re-expressed as:

$$\nabla^2 \vec{E} = \frac{1}{c_0} \frac{\partial^2 E}{\partial t^2} \quad (1.10)$$

$$\nabla^2 \vec{H} = \frac{1}{c_0} \frac{\partial^2 H}{\partial t^2} \quad (1.11)$$

### 1.1.2 Solutions to Maxwell's Equations

The solutions to the two wave equations in free space are just freely propagating plane waves with angular frequency  $\omega$  and wave vector  $\kappa$ , related by  $\omega=c\kappa$ :

$$\vec{E} = E_0 e^{i(\omega t \pm \vec{\kappa}, \vec{r})} \quad (1.12)$$

$$\vec{H} = H_0 e^{i(\omega t \pm \vec{\kappa}, \vec{r})} \quad (1.13)$$

The electric field  $\vec{E}$  and magnetic field  $\vec{H}$  strengths are similar quantities, with units of  $\text{Vm}^{-1}$  and  $\text{Am}^{-1}$  respectively. The ratio of the two fields  $E/H$  in light is a fixed constant with units of  $\text{V/A} = \text{Ohms}$ . This quantity is known as the Impedance of free space,  $Z_0$ :

$$\frac{\vec{E}}{\vec{H}} = Z_0 = \sqrt{\frac{\mu_0}{\epsilon_0}} \approx 377\Omega \quad (1.14)$$

Whilst  $\vec{E}$  and  $\vec{H}$  are similar quantities, they propagate through space slightly differently - in fact, perpendicular to one another. To see this, we can substitute each plane wave solution back into Maxwell's curl equation for Faraday's Law and compute each term, giving:

$$\nabla \times \mathbf{E} = i\mathbf{k} \times \mathbf{E}, \quad \frac{\partial \mathbf{H}}{\partial t} = -i\omega \mathbf{H} \quad (1.15)$$

So,

$$i\mathbf{k} \times \mathbf{E} = -\mu_0(-i\omega)\mathbf{H} \Rightarrow \mathbf{k} \times \mathbf{E} = \omega\mu_0\mathbf{H} \quad (1.16)$$

Now, if we take the dot product with  $\mathbf{E}$

$$\mathbf{E} \cdot (\mathbf{k} \times \mathbf{E}) = \omega\mu_0, \mathbf{E} \cdot \mathbf{H} \quad (1.17)$$

But  $\mathbf{E} \cdot (\mathbf{k} \times \mathbf{E}) = 0$  because the scalar triple product of a vector with its own cross product is zero. Therefore:

$$\mathbf{E} \cdot \mathbf{H} = 0 \quad (1.18)$$

That is, **E and H are perpendicular** at every point in space and time. Both the direction of propagation of the light and the transmitted power density are given by the cross product of the two field vectors, called the Poynting vector  $\vec{S}$ :

$$\vec{S} = \vec{E} \times \vec{H} \quad (1.19)$$

For a deeper discussion on Maxwell's equations and the wave equations, review [1].

### 1.1.3 Polarisation

Light polarisation describes the orientation of the electric field vector in an electromagnetic wave. Since light is a transverse wave, its electric field and magnetic field oscillate perpendicular to the direction of propagation. Considering a wave traveling in the z-direction, the electric field lies in the x-y plane, and its specific orientation, amplitude, and phase difference between components determine the polarisation state. Mathematically, the electric field can be written as

$$\mathbf{E}(z, t) = \hat{x}E_x \cos(kz - \omega t + \phi_x) + \hat{y}E_y \cos(kz - \omega t + \phi_y) \quad (1.20)$$

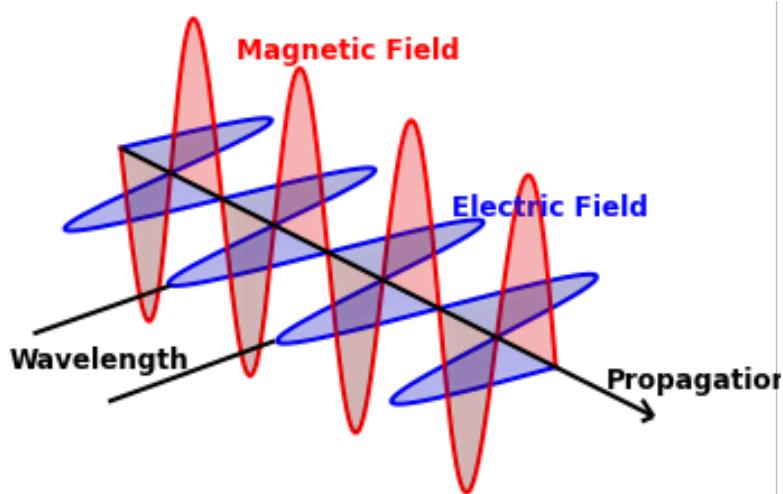
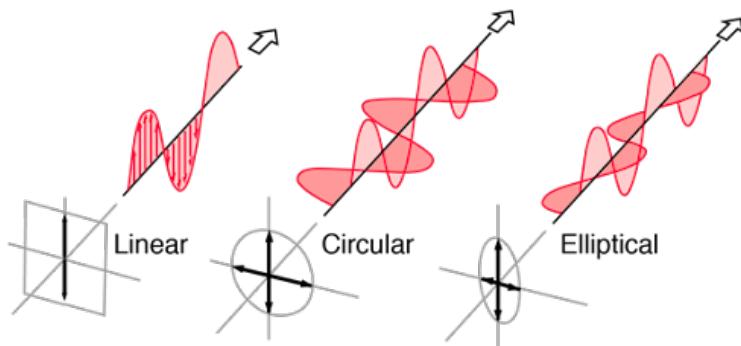


Figure 1.1: A propagating electromagnetic field. The electric and magnetic field components are perpendicular to each other at all points.

.... where  $E_x$  and  $E_y$  are the amplitudes along  $x$  and  $y$ , and  $\phi_x, \phi_y$  are their phases.

Polarisation states can be manipulated using optical elements. Polaroid filters can convert unpolarised light into linearly polarised light, while quarter-wave plates can transform linear polarisation into circular or elliptical polarisation depending on orientation. Half-wave plates rotate the orientation of linear polarisation. These techniques are widely used in optics, communication, and microscopy to control and analyse the behaviour of light according to its polarisation.



### 1.1.3.1 Linear Polarisation

If the electric field vector oscillates along a single, fixed line, the light is said to be linearly or plane-polarized. This occurs when the  $x$  and  $y$  components are in phase or exactly out of phase, so that the tip of the electric field traces a straight line in

the x-y plane. For instance, if the y-component is zero, the polarization is along the x-axis, and if the components are equal and in phase, the polarization lies at 45 degrees. Plane-polarized light is the simplest polarization state and can be easily visualized as the electric field swinging back and forth along one direction.

In general, a linear polarised electric field can be expressed as:

$$\mathbf{E}(z, t) = \hat{x}E_x \cos(kz - \omega t) + \hat{y}E_y \cos(kz - \omega t) \quad (1.21)$$

If the propagation is along the x-axis ( $E_y=0$ ):

$$\mathbf{E}(z, t) = \hat{x}E_0 \cos(kz - \omega t) \quad (1.22)$$

On the other hand, if the propagation is at 45° ( $E_x=E_y=E_0$ ):

$$\mathbf{E}(z, t) = \hat{x}E_0 \cos(kz - \omega t) + \hat{y}E_0 \cos(kz - \omega t) \quad (1.23)$$

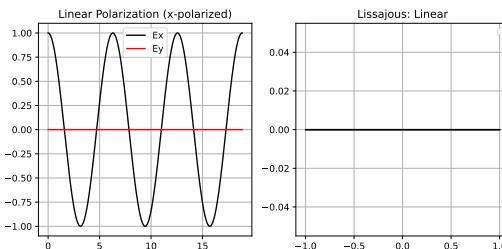


Figure 1.2: Electric field components  $E_x$  and  $E_y$  (R) and lissajous (L) for linearly polarised light.

### 1.1.3.2 Circular Polarisation

When the magnitude of the electric field remains constant but its direction rotates uniformly as the wave propagates, the light is circularly polarised. Circular polarisation occurs when the  $x$  and  $y$  components have equal amplitudes and a phase difference of  $\pm 90$  degrees. The sign of the phase difference determines the handedness: right circular polarisation occurs when the  $y$ -component leads the  $x$ -component by 90 degrees, and left circular polarisation when it lags by 90 degrees. In this case, the tip of the electric field vector traces a perfect circle as the wave moves forward.

If each electric field component  $x$  and  $y$  have equal amplitudes and  $\pm 90^\circ$  phase difference, Right Circular Polarisation (RCP,  $y$  leads  $x$  by  $\pi/2$ ) is:

$$\mathbf{E}(z, t) = E_0 [\hat{x} \cos(kz - \omega t) + \hat{y} \sin(kz - \omega t)] \quad (1.24)$$

On the other hand, Left Circular Polarisation (LCP,  $y$  lags  $x$  by  $\pi/2$ ), so:

$$\mathbf{E}(z, t) = E_0 [\hat{x} \cos(kz - \omega t) - \hat{y} \sin(kz - \omega t)] \quad (1.25)$$

The tip of the electric field traces a circle in the  $x$ - $y$  plane.

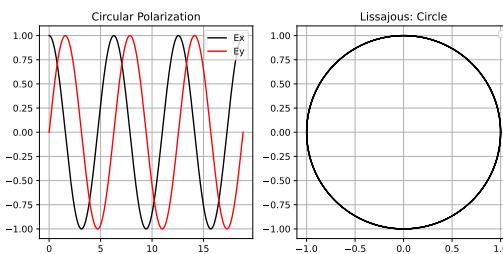


Figure 1.3: Electric field components  $E_x$  and  $E_y$  (R) and lissajous (L) for circularly polarised light.

### 1.1.3.3 Elliptical Polarisation

Elliptical polarisation represents the most general case of light polarisation. When the  $x$  and  $y$  components have unequal amplitudes and are not perfectly in or out of phase, the resulting trajectory of the electric field is an ellipse. Linear and circular polarisations are special cases of elliptical polarisation: linear corresponds to a degenerate ellipse that is a straight line, while circular corresponds to a degenerate ellipse that is a circle. Elliptically polarised light can thus describe nearly any possible orientation of the electric field in the transverse plane.

For unequal amplitudes and arbitrary phase difference  $\delta$ :

$$\mathbf{E}(z, t) = \hat{x}E_x \cos(kz - \omega t) + \hat{y}E_y \cos(kz - \omega t + \delta) \quad (1.26)$$

If  $\delta=0$  or  $\pi$ , this reduces to linear polarisation. If  $E_x=E_y$  and  $\delta=\pm\pi/2$ , this reduces to circular polarisation. The tip of  $\mathbf{E}$  traces an ellipse in the transverse plane.

### 1.1.3.4 Unpolarised

Unpolarized light consists of many waves with random polarization directions that fluctuate rapidly over time. In unpolarized light, there is no fixed phase rela-

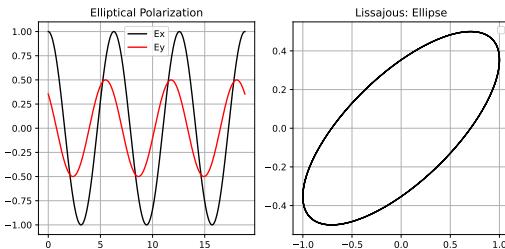


Figure 1.4: Electric field components  $E_x$  and  $E_y$  (R) and lissajous (L) for elliptically polarised light.

tionship between the x and y components, and the electric field has no preferred direction on average. Sunlight and light from incandescent bulbs are typical examples of unpolarized light. In such cases, the time-averaged intensities along x and y are equal, and there is no correlation between the components.

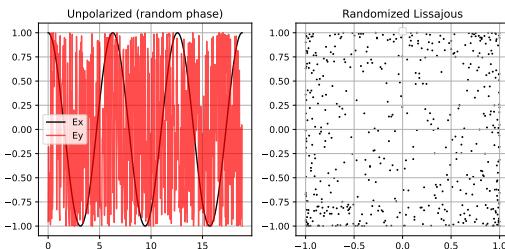


Figure 1.5: Electric field components  $E_x$  and  $E_y$  (R) and lissajous (L) for unpolarised light.

For a deeper discussion on polarisation, review Hecht [2].

### 1.1.4 Wave-Particle Duality

In lecture 5 we will start to consider the quantum mechanical description of light-matter interactions in more detail, but it is instructive at this point to recall the 'Wave-particle duality'. This is a foundational concept in quantum mechanics, expressing the idea that all physical entities, from photons to electrons, exhibit both wave-like and particle-like properties depending on the circumstances of observation. The origin of this idea lies in the early twentieth century with Max Planck's study of blackbody radiation. Planck proposed that electromagnetic energy is not emitted continuously, but rather in discrete packets, or quanta, each carrying energy proportional to the frequency of the radiation. This relationship is expressed by the equation  $E=h\nu$ , where  $E$  is the energy of a single quantum,  $h$  is Planck's constant, and  $\nu$  is the frequency of the wave.

Albert Einstein extended Planck's idea in 1905 while explaining the photoelectric effect. In Einstein's view, light consists of photons, each carrying energy  $h\nu$ . When a photon strikes a metal surface, it can transfer its energy to an electron, liberating it from the material if the energy is sufficient to overcome the binding potential. This was one of the first direct demonstrations of light's particle-like behaviour, complementing its known wave phenomena such as interference and diffraction. The duality of light was thus established conceptually and experimentally.

Louis de Broglie, in 1924, proposed a radical generalisation: if light waves can behave like particles, then perhaps particles of matter can also behave like waves. He introduced the relation  $\lambda=\frac{h}{p}$ , associating a wavelength  $\lambda$  with the momentum  $p$  of any material particle. This means that an electron, for example, possesses a wave nature with a wavelength inversely proportional to its momentum. The idea was soon confirmed in 1927 by the Davisson-Germer experiment, which showed that electrons diffract off crystal lattices just as X-rays do. The concept of the matter wave provided a crucial bridge between classical mechanics and quantum mechanics, showing that all matter has an associated wavelength that becomes significant at atomic scales.

The wave and particle descriptions can be unified using the relativistic relation between energy and momentum, given by  $E^2=(pc)^2 + (m_0c^2)^2$ , where  $m_0$  is the rest mass and  $c$  is the speed of light. For particles with mass, this formula combines rest energy and kinetic energy, reducing at low speeds to the familiar form  $E \approx m_0c^2 + \frac{p^2}{2m_0}$ . For massless particles such as photons, the rest mass term vanishes, leaving  $E=pc$ . Combining this with Planck's relation  $E=h\nu$  gives  $p = \frac{h\nu}{c} = \frac{h}{\lambda}$ , confirming that light waves carry momentum in accordance with de Broglie's formula.

## 1.2 The Electromagnetic Spectrum

The spectrum of electromagnetic waves is a continuum encompassing all frequencies of electromagnetic radiation. The spectrum is categorised into distinct regions depending on the frequency. Frequencies up to about 300 GHz are referred to as radio frequency (RF) - this region encompasses so-called *microwave* and *millimetre wave* frequencies, and sometimes these terms are used interchangeable with RF (although an electrical engineer may take some offence to this!). At the higher energy range with frequencies of the order  $10^{19}$  we have  $\gamma$ -rays.

As an aside, the practical upper energy limit is for frequencies between  $10^{22} - 10^{28}$  Hz, corresponding to photon energies greater than several MeV. These include

the so-called hard gamma rays and the ultra-high-energy (UHE) photon regime (photon energies of GeV to TeV and even PeV). Such radiation is detected, for example, from cosmic sources like supernova remnants, pulsars, active galactic nuclei, and gamma-ray bursts. At these frequencies, electromagnetic photons behave more like particles than waves, undergoing pair production  $\gamma \rightarrow e^+ + e^-$  in the presence of nuclei or background fields. In addition, they interact strongly with matter, and are easily absorbed by even small amounts of material. Due to this, they are not practically usable. As such, energies beyond gamma rays exist only as extremely high-energy photons produced by astrophysical or particle-accelerator processes.

On the other side of the spectrum, whilst Maxwell's equations apply all the way to DC, this spectral region does not *radiate* in the usual sense, since the wavelength at these frequencies is so large as to make the EM wave quasi-stationary rather than propagating.

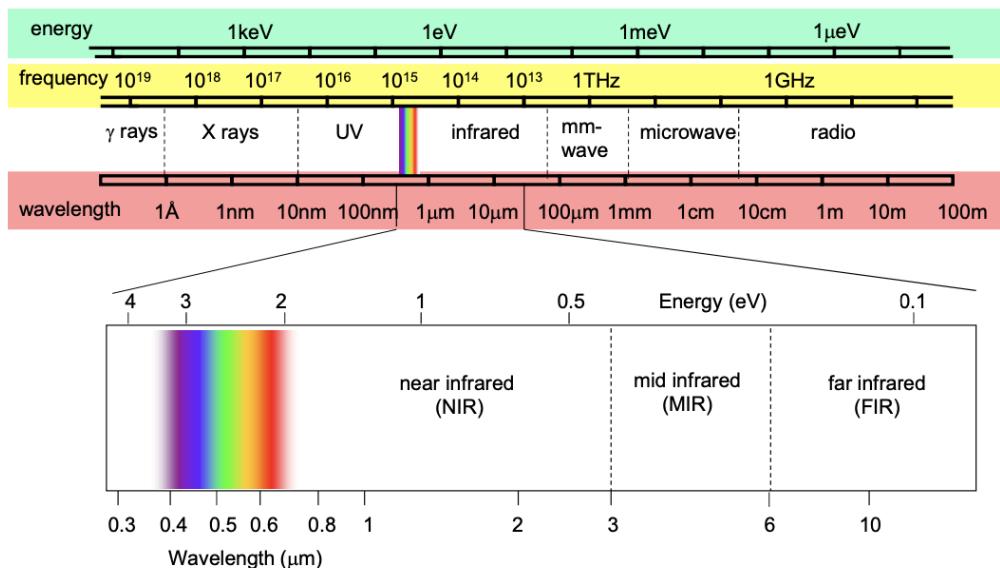


Figure 1.6: The 'usable' electromagnetic spectrum.

In opto-electronics, we exploit the a range of spectra in-between these two extrema, including infrared (IR), visible, and ultra-violet (UV). Infrared radiation is primarily associated with heat, since it arises from the vibrational motion of atoms and molecules. Objects at ordinary temperatures emit infrared radiation as part of their thermal emission spectrum, which forms the basis of infrared imaging and remote sensing technologies.

The visible region constitutes only a tiny fraction of the entire electromagnetic

spectrum, roughly from 400 nm (violet) to 700 nm (red). This is the range detectable by the human eye, corresponding to photon energies of about 1.8–3.1 eV. The colors we perceive are directly related to wavelength, with shorter wavelengths appearing bluish and longer ones reddish. The Sun emits most of its radiation in this band, which has shaped the evolution of visual perception in living organisms.

Just beyond the violet end lies ultraviolet radiation, with wavelengths between about 10 nm and 400 nm. Ultraviolet photons carry higher energies and can induce electronic transitions in atoms and molecules, often leading to chemical reactions such as those responsible for sunburn and photochemical processes in the atmosphere. UV radiation is divided into near, far, and extreme ultraviolet regions, each with distinct interactions with matter and degrees of atmospheric absorption.

Across the spectrum, the photon energy  $E=hf$  increases linearly with frequency. As energy rises, the interaction between electromagnetic waves and matter transitions from collective oscillations of charges (as in radio waves) to molecular vibrations (infrared), electronic transitions (visible and ultraviolet), and finally to ionization and nuclear effects (X-rays and gamma rays).

To convert between wavelength and energy, a useful equation to remember is:

$$E(\text{eV}) = \frac{hc}{e\lambda} = \frac{1240}{\lambda(\text{nm})} \quad (1.27)$$

## 1.3 Interaction of Light with Bulk Matter

### 1.3.1 Dielectrics

When light passes through a material, the electric field 'pushes' bound electrons in the material a little bit causing them to become displaced. In response, this creates a polarisation per unit volume in the material. That is, in a small region of the material proportional to the wavelength of the applied field, individual dipole moments are affected. It is the collective response of these dipoles that determines how light travels within the material. Consider a plane electromagnetic wave entering a dielectric - the oscillating electric field of the light exerts a force on the bound electrons, causing them to oscillate in phase with the driving field (at least approximately, for transparent materials and low absorption). Each oscillating charge distribution behaves like a tiny dipole that, in turn, radiates its own electromagnetic field. The total field in the material is therefore the superposition of the original incident wave and the radiation from all these induced dipoles. The

key is that these secondary radiations interfere in such a way that the net effect is a wave that appears to travel more slowly than it would in vacuum.

The polarisation is the total dipole moment of all the microscopic charges inside that small volume, divided by the volume itself. The relationship between displacement  $\vec{D}$  of the electrons and the resulting polarisation  $\vec{P}$  is described from Maxwell's equations as:

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}, \quad (1.28)$$

The field  $\mathbf{D}$  includes both the vacuum contribution  $\epsilon_0 \mathbf{E}$  and the bound-charge contribution through  $\mathbf{P}$ . It simplifies Gauss's law for electric fields to:

$$\nabla \cdot \mathbf{D} = \rho_{\text{free}}, \quad (1.29)$$

where  $\rho_{\text{free}}$  denotes the free charge density. Thus,  $\mathbf{D}$  is sometimes interpreted as the “source-free” field, in contrast to  $\mathbf{E}$ , which responds to both free and bound charges.

At the microscopic level, each atom or molecule can acquire an induced dipole moment when placed in an electric field. If  $\mathbf{E}_{\text{loc}}$  denotes the local electric field acting on a given atom, then

$$\mathbf{p} = \alpha \mathbf{E}_{\text{loc}}, \quad (1.30)$$

where  $\alpha$  is the **polarizability**, a measure of how easily the electron cloud of the atom can be distorted by the field. For a material containing  $N$  such dipoles per unit volume, the macroscopic polarization becomes

$$\mathbf{P} = N\mathbf{p} = N\alpha \mathbf{E}_{\text{loc}}. \quad (1.31)$$

In the case of a linear, isotropic dielectric, it is often sufficient to express  $\mathbf{P}$  as proportional to the macroscopic field  $\mathbf{E}$ :

$$\mathbf{P} = \epsilon_0 \chi_e \mathbf{E}, \quad (1.32)$$

where  $\chi_e$  is the **electric susceptibility**. Relating the microscopic and macroscopic pictures gives, approximately,

$$\chi_e \approx \frac{N\alpha}{\epsilon_0}. \quad (1.33)$$

This relation is only approximate because the local field  $\mathbf{E}_{\text{loc}}$  generally differs from the macroscopic field  $\mathbf{E}$  due to local-field effects (e.g., the Lorentz correction).

### 1.3.2 Polarisation in Dielectric Materials

You will have come across polarisation in dielectric media in previous courses (namely 'Electrical, Optical, and Magnetic Properties of Materials'). To recap, the different ways in which electric polarisation in materials is manifest is summarised in Figure 1.7, reproduced from Fujiwara [3].

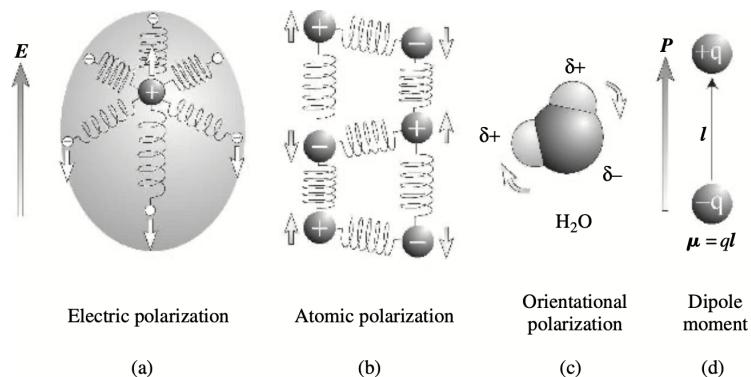


Figure 1.7: Dielectric polarisation in dielectrics. Figure from [3]

When light interacts with a dielectric material, the responsive polarisation can become manifest in several different ways. In semiconductors, the most relevant polarisation response is **electronic polarisation**, where microscopic dipoles are created by perturbing electron wave functions relative to nuclei (Figure ??(a)). This arises when induced electric fields distort the positions of the electrons and the nucleus in opposite directions.

On the other hand, ionic crystals such as NaCl are formed of both electrically positive and negative ions which gives rise to an ionic distortion by an electric field, known as **atomic/ionic polarisation** (Figure 1.7(b)).

When an electric field is applied to a material that contains atoms which are charged due to the difference in electronegativity (such as H<sub>2</sub>), the orientation of the molecule becomes aligned to the direction of the electric field. This is known as **dielectric polarisation**, and is the basis for how food is heated by a microwave oven. This is shown in Figure 1.7(c).

When free-charge carriers move to create internal electric fields, this is known as **Space Charge Polarisation**.

### 1.3.3 The Lorentz Oscillator Model

In a dielectric, electrons are bound to nuclei, so unlike in metals they cannot move freely. If we displace an electron slightly from its equilibrium position, the Coulomb attraction from the nucleus pulls it back toward the centre. For small displacements, this restoring force is approximately linear:

$$F_{restoring} = -kx \quad (1.34)$$

This is a classic harmonic oscillator. That is, each bound electron behaves like a microscopic mass-spring-damper system, where the mass  $m$  is the electron mass, the spring constant  $k$  relates to binding strength, and the natural resonance frequency is  $\omega_0 = \sqrt{k/m}$ . The displacement of the bound charge is governed by a differential equation:

$$m\ddot{x} + m\gamma\dot{x} + m\omega_0^2 x = -eE_0 e^{-i\omega t}, \quad (1.35)$$

where  $\gamma$  is the damping rate due to radiation or collisions. In the steady state, the displacement of the electron is found by solving for  $x$ , and in doing so it is clear that the displacement takes on an frequency dependence. SO:

$$x(\omega) = \frac{e/m}{\omega_0^2 - \omega^2 - i\gamma\omega} E_0. \quad (1.36)$$

The induced dipole moment is

$$p = -ex = \alpha(\omega)E_0, \quad (1.37)$$

from which the frequency-dependent polarisability follows as

$$\alpha(\omega) = \frac{e^2/m}{\omega_0^2 - \omega^2 - i\gamma\omega}. \quad (1.38)$$

Therefore,  $\alpha$  (and consequently  $\mathbf{P}$  and  $\mathbf{D}$ ) are inherently frequency-dependent.

### 1.3.3.1 Macroscopic Permittivity and Frequency Dependence

The macroscopic polarisation is the average of the microscopic polarisation; that is  $\mathbf{P} = N\alpha(\omega)\mathbf{E}$ . From Eq. (1.28), we have

$$\mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P} = \epsilon_0 \left[ 1 + \frac{N\alpha(\omega)}{\epsilon_0} \right] \mathbf{E}. \quad (1.39)$$

Since the polarisation is frequency dependent, so too is the permittivity. That is:

$$\boxed{\epsilon(\omega) = \epsilon_0 \left[ 1 + \frac{Ne^2/m}{\epsilon_0 (\omega_0^2 - \omega^2 - i\gamma\omega)} \right].} \quad (1.40)$$

The physical intuition behind this is that each atom or molecule has natural oscillation frequencies associated with electronic transitions (in the UV-visible range), vibrational transitions (in the infrared), and Lattice vibrations or phonons (in the far-IR or THz range). When the frequency of light  $\omega=2\pi c/\lambda$  approaches one of these natural frequencies, the electrons' motion lags behind the field, producing a large and rapidly varying polarisation. This delayed response modifies the dielectric constant  $\epsilon(\omega)$ , and since  $n^2(\omega)=\epsilon(\omega)$  (for non-magnetic materials), we get a frequency-dependent refractive index.

### 1.3.3.2 Refractive Index

The refractive index is simply the square root of the dielectric function. Clearly then, if the dielectric constant is complex, so too is the refractive index:

$$\tilde{n}(\omega) = n(\omega) + i\kappa(\omega) \quad (1.41)$$

The complex coefficient  $\kappa$  is known as the "extinction coefficient", and it describes how much an incident electric field is attenuated by the material. When a plane wave  $E(z, t)=E_0e^{i(kz-\omega t)}$  travels through the medium, the wavevector is:

$$k = \frac{\omega}{c}\tilde{n} = \frac{\omega}{c}(n + i\kappa) \quad (1.42)$$

So:

$$E(z, t)=E_0e^{-\frac{\omega}{c}kz}e^{i(\frac{\omega}{c}nz-\omega t)} \quad (1.43)$$

The oscillatory component  $n$  describes the phase velocity, and the exponential decay  $\kappa$  describes how fast the wave dies in the material. This has a simple relation to **absorption**. For an electric field intensity  $I_0$ , the decay is through the medium is:

$$I(z) = I_0 e^{-\alpha z} \quad (1.44)$$

....where the absorption coefficient  $\alpha$  is:

$$\alpha = \frac{4\pi\kappa}{\lambda_0} \quad (1.45)$$

### 1.3.4 Complex Dielectric Function

The dielectric function varies strongly with frequency and is generally complex,

$$\epsilon = \epsilon_1 - i\epsilon_2 \quad (1.46)$$

where  $\epsilon_1$  and  $\epsilon_2$  represent the real and imaginary components respectively. These are related to the refractive index  $n$  and extinction coefficient  $k$  by

$$\epsilon_1 = n^2 - k^2, \quad \epsilon_2 = 2nk \quad (1.47)$$

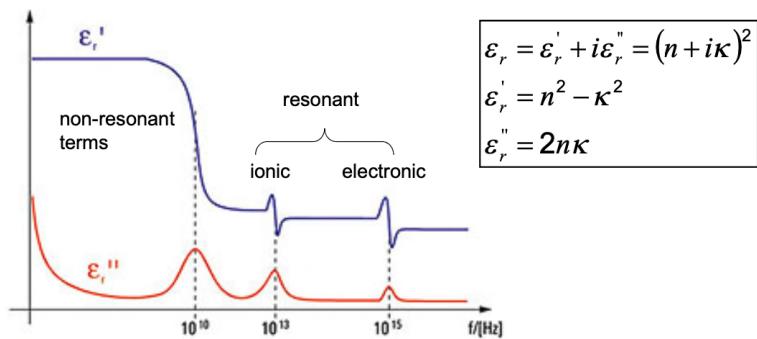


Figure 1.8: Real  $\epsilon_1$  and imaginary  $\epsilon_2$  part of the complex dielectric function  $\epsilon$ .

At low angular frequencies,  $\epsilon_1$  corresponds to the static dielectric constant  $\epsilon_s$ , which includes contributions from atomic and electronic polarisations.

Since  $\epsilon_2$  is proportional to  $k$ , it describes optical absorption. Peaks in  $\epsilon_2$  correspond to resonant absorption frequencies: orientational polarisation resonates in the microwave region, atomic polarisation in the infrared, and electronic polarisation in the visible/UV. For example, microwave heating of water exploits orientational polarisation of H<sub>2</sub>O molecules.

As  $\omega$  increases beyond the infrared, atomic polarisation cannot follow the oscillating field, reducing  $\epsilon_1$  to the high-frequency limit  $\epsilon_\infty$ . At still higher frequencies, even electronic polarisation fails to respond, and  $\epsilon_1 \rightarrow 1$ , the vacuum value. Thus, the dielectric constant depends on the frequency-dependent polarisation response, known as the dielectric function or dielectric dispersion.

Near resonance, the dielectric function follows the Lorentz model, though experimental spectra usually show multiple overlapping resonances. In covalent crystals such as Si and Ge, where ionicity is absent,  $\epsilon_s = \epsilon_\infty$  because atomic polarisation is negligible.

If free carriers (electrons or holes) are present, additional absorption arises from free-carrier motion. In metals and heavily doped semiconductors, this shifts the dielectric response so that  $\epsilon_1$  becomes negative below the plasma frequency  $\omega_p$ .

### 1.3.5 Dispersion of Light in a Transparent Medium

When light propagates through an optical material, as discussed earlier, it interacts with the bound charges of the medium giving rise to a polarisation. In turn, this modifies the phase velocity  $v_p$  of light. The refractive index  $n$  quantifies this slowing down of light. Since  $n$  is frequency dependent, the electrons in the material do not respond equally to all incident frequencies, and  $n$  is **dispersive**.

Dispersion arises from resonance phenomena. Each atom or molecule has natural oscillation frequencies associated with electronic transitions (in the UV-visible range), vibrational transitions (in the infrared), and Lattice vibrations or phonons (in the far-IR or THz range). When the frequency of light  $\omega = 2\pi c/\lambda$  approaches one of these natural frequencies, the electrons' motion lags behind the field, producing a large and rapidly varying polarisation. This delayed response modifies the dielectric constant  $\epsilon(\omega)$ , and since  $n^2(\omega) = \epsilon(\omega)$  (for non-magnetic materials), we get a frequency-dependent refractive index.

#### 1.3.5.1 The Sellmeier Equation

The Sellmeier equation is an empirical form inspired by the Lorentz model, written in terms of wavelength instead of frequency:

$$n^2(\lambda) = 1 + \sum_{i=1}^N \frac{B_i \lambda^2}{\lambda^2 - C_i} \quad (1.48)$$

where  $B_i$  are dimensionless coefficients related to oscillator strength,  $C_i$  (in units of  $\mu\text{m}^2$  or  $\text{nm}^2$ ) correspond roughly to the square of resonance wavelengths  $\lambda_{0,i}^2$ , and  $N$  is the number of significant resonances (typically 2-3 terms suffice for transparent dielectrics). This form is valid in spectral regions where the material is non-absorbing (the “transparency window”).

At long wavelengths ( $\lambda \gg \lambda_{0,i}$ ), each term behaves as  $\frac{B_i \lambda^2}{\lambda^2 - C_i} \approx B_i \left(1 + \frac{C_i}{\lambda^2}\right)$ . So  $n(\lambda)$  slowly decreases with increasing  $\lambda$ . This is **normal dispersion**.

Approaching resonance ( $\lambda \rightarrow \lambda_{0,i} = \sqrt{C_i}$ ), the denominator approaches zero, so  $n$  increases sharply, and just beyond resonance it decreases, leading to **anomalous dispersion**. This region usually coincides with strong absorption.

In the transparency region, between resonances,  $n(\lambda)$  changes smoothly and monotonically, enabling accurate fitting by a Sellmeier form.

#### Example 2: Fused Silica ( $\text{SiO}_2$ )

The components of the Sellmeier model can be extracted by fitting optical data acquired by an ellipsometer to the Sellmeier equation [3]. A common Sellmeier formula with coefficients for fused silica (wavelengths in  $\mu\text{m}$ ) is:

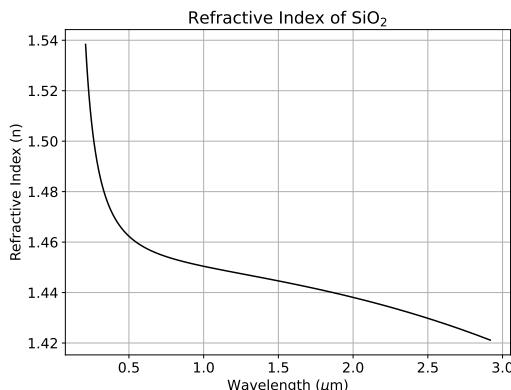
$$n^2(\lambda) = 1 + \frac{0.6961663\lambda^2}{\lambda^2 - 0.0684043^2} + \frac{0.4079426\lambda^2}{\lambda^2 - 0.1162414^2} + \frac{0.8974794\lambda^2}{\lambda^2 - 9.896161^2}$$

Looking at the figure below, a plot of the Sellmeier model for  $\text{SiO}_2$ , what can you tell about the dispersion in different spectral regions?

### 1.3.6 Light Propagation: Superposition and Huygens' Principle

Electric fields follow the principle of superposition, namely that the total electric field vector at a point is the sum of the electric field vectors acting at that point. The propagation of light is then neatly described by Huygens' Principle which states that

*“Every point on a given wavefront acts as a source of secondary spherical wavelets that spread out in all directions with the speed of light. The new wavefront at a*



*later time is the envelope (tangent surface) of all these secondary wavelets.” - Christiaan Huygens (1678)*

Huygen's principle provides a very useful corollary for the propagation of light in a dielectric. Each secondary source can be associated with the radiating dipoles induced by the external time-varying electric field. Their collective emission produces the next wavefront that propagates through the medium, and the reduced speed of propagation simply means that these secondary wavelets advance more slowly, because their phase accumulates at a slower rate due to the delayed response of the induced dipoles.

From this point of view, light propagation in a dielectric can be seen as a continuous process of re-radiation and interference. The apparent wave that we observe traveling through the medium is not merely the same wave that entered it, but rather a coherent reconstruction arising from countless microscopic dipole oscillations that perpetually generate secondary waves. The phase delay per unit length is what gives rise to the refractive index, while the absence of absorption implies that all of these microscopic processes remain perfectly coherent and energy is conserved in the form of the transmitted light field.

Thus, in a dielectric, light does not move as an independent entity unaffected by matter; it is constantly being absorbed and re-emitted on a microscopic scale, but in a way that is perfectly synchronised across the material. The macroscopic consequence of these countless induced dipoles and their mutual interference, viewed through the lens of Huygens' principle, is the smooth, continuous propagation of light with a well-defined refractive index.

A simple mathematical description of this process is to sum the amplitudes of the waves resulting from all possible paths of propagation:

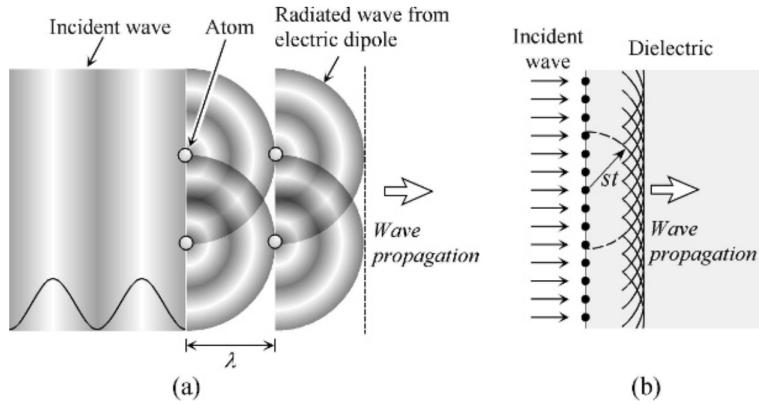


Figure 1.9: (a) Propagation of light in a dielectric. Atoms are shown with the interval of the wavelength  $\lambda$ . (b) Propagation of light according to Huygens's principle.  $st$  shows a distance to which light travels with a speed  $s$  during a time  $t$ . Figure from [3]

$$\vec{E}(\vec{r}) = \sum_{\text{all paths}, l} \vec{E}_l(r) e^{i\phi_l(\vec{r})} \quad (1.49)$$

where  $\phi_l(\vec{r})$  is the phase of the wave at position  $r$ .

## 1.4 Reflection/Refraction

When light is incident upon a material at oblique incidence, its propagation direction is generally affected by refraction. When light advances into a material, light is emitted from the surface atoms, and electric dipole radiation occurs from each atom (Figure 1.10). Light waves emanating from the atoms overlap constructively due to Huygens' principle, and propagate as a transmitted wave.

If the incident light propagates from point B to D, the transmitted wave moves from point A to E and the reflected wave moves to point C. It is clear that:

$$\frac{\sin \theta_i}{\overline{BD}} = \frac{\sin \theta_r}{\overline{AC}} = \frac{\sin \theta_t}{\overline{AE}} = \frac{1}{\overline{AD}} \quad (1.50)$$

Since the speed of light in a medium is given by  $\nu=c/n$ , in time  $t$  light is able to travel by a distance  $dt = ct/n$ . So:

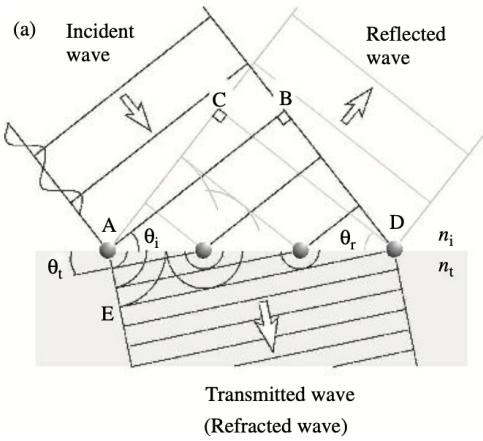


Figure 1.10: Light transmission and reflection at oblique incidence. Figure from [3]

$$\overline{BD} = ct/n_i \quad \overline{AC} = ct/n_r \quad \overline{AE} = ct/n_t \quad (1.51)$$

....where  $n_i$ ,  $n_r$ , and  $n_t$  are the refractive indices of the incident, reflected, and transmitted light respectively. Since  $n_i=n_r$ ,  $BD=AC$ . From this, we can find the **Law of Reflection**:

$$\theta_i = \theta_r \quad (1.52)$$

....and by combining equations 1.50 and 1.51, we can derive **Snell's Law**:

$$n_i \sin \theta_i = n_t \sin \theta_t \quad (1.53)$$

### 1.4.1 *s* and *p* Polarised Light

When light reflects from the boundary between two media, its subsequent behaviour depends on the polarisation of the electric field relative to the plane of incidence. The plane of incidence is defined by the incident wave vector  $\vec{k}_i$  and the normal to the interface, and polarisation components are defined with respect to this plane.

When the reflected electric field vector is perpendicular to the plane of incidence (i.e. tangential to the surface and directed out of the plane), this is known as *s*-

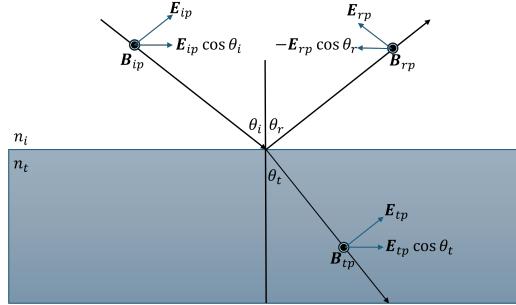


Figure 1.11: Electric field  $E$  and magnetic induction  $B$  for p-polarisation.

polarised light. The terminology  $s$  is from the German word for 'perpendicular', which is *senkrecht*. On the other hand, if the reflected field vector is parallel to the plane of incidence, this is known as  $p$  polarised light (with  $p$  meaning 'parallel')

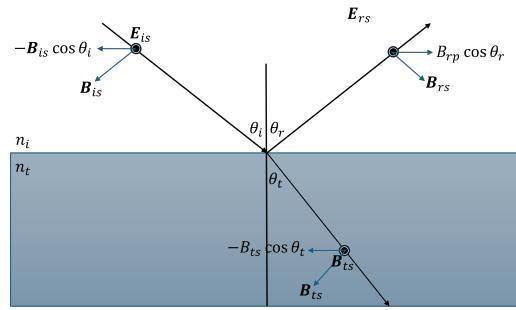


Figure 1.12: Electric field  $E$  and magnetic induction  $B$  for s-polarisation.

At the interface, part of the light wave is reflected, and part is transmitted. The amplitudes of these waves are described by the **fresnel Equations**. For an incident angle  $\theta_i$ , transmitted angle  $\theta_t$ , and refractive indices  $n_1$  and  $n_2$  in the incident and transmitted media respectively, the reflection coefficients are

$$r_s = \frac{n_1 \cos \theta_i - n_2 \cos \theta_t}{n_1 \cos \theta_i + n_2 \cos \theta_t} \quad (1.54)$$

$$r_p = \frac{n_2 \cos \theta_i - n_1 \cos \theta_t}{n_2 \cos \theta_i + n_1 \cos \theta_t} \quad (1.55)$$

and the corresponding transmission coefficients are

$$t_s = \frac{2n_1 \cos \theta_i}{n_1 \cos \theta_i + n_2 \cos \theta_t} \quad (1.56)$$

$$t_p = \frac{2n_1 \cos \theta_i}{n_2 \cos \theta_i + n_1 \cos \theta_t} \quad (1.57)$$

The transmitted angle satisfies Snell's law. Since  $r_s$  and  $r_p$  are generally different in both amplitude and phase, the reflected wave can undergo a change in its polarisation state. A linearly polarised incident wave that contains both  $s$  and  $p$  components will generally be reflected as elliptically polarised light. At normal incidence ( $\theta_i=0$ ), we have  $r_s=r_p$ , so no change in polarisation occurs.

At the **Brewster angle**, defined by

$$\tan \theta_B = \frac{n_2}{n_1}, \quad (1.58)$$

the reflection coefficient for  $p$ -polarized light vanishes ( $r_p=0$ ), and the reflected light is therefore purely  $s$ -polarized. For metallic or absorbing media,  $n_2$  becomes complex, and both  $r_s$  and  $r_p$  acquire complex phase shifts, often resulting in strongly elliptical or even circularly polarised reflected light.

### 1.4.2 Example Application - Polarised Sunglasses

When unpolarised light strikes a reflective surface such as glass, the reflected wave is not equally composed of all polarisation directions. The Fresnel reflection coefficients for  $s$ - and  $p$ -polarised light are generally different in magnitude, which means that one component reflects more efficiently than the other. Near the Brewster angle, the reflection coefficient for the  $p$ -polarised component vanishes ( $r_p=0$ ), and the reflected light is therefore almost purely  $s$ -polarised. Since the electric field of  $s$ -polarised light is perpendicular to the plane of incidence, this corresponds, for a horizontal surface, to an electric field oscillating mainly in the horizontal direction. Consequently, glare from such surfaces is predominantly horizontally polarised.

Polarised spectacles make use of this property to suppress unwanted glare. They depend on selective absorption of one polarisation component. The lenses are made from a thin polymer film (commonly stretched polyvinyl alcohol combined with iodine or another dichroic dye). The stretching aligns long-chain molecules in one direction, giving the material anisotropic absorption. That is, it strongly absorbs light whose electric field oscillates parallel to the molecular alignment and transmits the perpendicular component.

In practice, the polarising axis of the spectacle lens is oriented vertically. Horizontally polarised light, which constitutes most of the reflected glare from flat

surfaces, is absorbed by the filter, while vertically polarised light, carrying most of the desired visual information from the environment, passes through. This selective absorption greatly reduces glare and improves contrast and visibility, especially near water or on roads.

## 1.5 Optical Coatings

For some applications such as solar cells, lasers, optical microcavities, optical filters, and interferometers, a very precise control over the reflectivity of a surface is required. For example, the optical reflection at the surface of a solar cell should be as close to 0 as possible. On the other hand, the reflection of a laser mirror should be as close to 100% as possible. This fine control of transmittivity and reflectivity can be achieved by exploiting the interference properties of light and by coating surface with thin optical films which achieve the desired effect.

When light interacts with a structure made of alternating layers of materials with different refractive indices, the reflected and transmitted light waves at each interface will interfere. This interference can be **constructive** or **destructive**, depending on their relative optical phases. This can be used to make useful optical structures such as *anti-reflective coatings* or *DBR mirrors*.

### 1.5.1 Anti-Reflection Coatings

Anti-Reflection (AR) Coatings, which are thin-film structures designed to minimise the reflection of light from optical surfaces, and thereby maximise transmission through the component. They work by introducing one or more dielectric layers whose thicknesses and refractive indices are engineered to produce destructive interference between reflected waves from each interface. The simplest form is a single quarter-wave layer with refractive index  $n = \sqrt{n_s}$ , where  $n_s$  is the substrate refractive index. Achieving this is tricky, and the most common material used is  $\text{MgF}_2$  with  $n=1.38$ , although some polymers and nanoporous materials are also good options.

The goal is to ensure that the wave reflected from the top surface (air-coating interface) and the one reflected from the coating-substrate interface have equal amplitudes but opposite phases, leading to cancellation. At normal incidence, the reflected field amplitudes are approximately

$$E_{r1} = r_{01}E_0, \quad E_{r2} = t_{01}r_{12}t_{10}E_0e^{i2\phi} \quad (1.59)$$

where  $r_{ij}$  and  $t_{ij}$  are the Fresnel reflection and transmission coefficients, and

$\phi=2\pi n_1 d_1 / \lambda$  is the single-pass phase through the film. Perfect cancellation at the design wavelength occurs when  $2\phi=\pi$  (quarter-wave optical thickness) and when the amplitude condition  $|r_{01}|=|r_{12}|$  is satisfied. These conditions give

$$n_1 = \sqrt{n_0 n_s}, \quad d_1 = \frac{\lambda_0}{4n_1} \quad (1.60)$$

Physically, this means the second reflected wave, having traveled an extra optical path of half a wavelength, returns  $180^\circ$  out of phase with the first reflection and cancels it. As a result, nearly all the incident power transmits into the substrate.

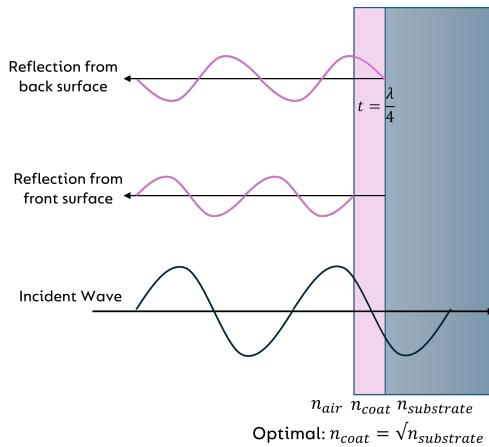


Figure 1.13: Incident and reflecting electromagnetic waves from an anti-reflection coating formed of a layer with  $n=\sqrt{n_s}$ , of thickness  $\lambda/(4n)$ .

However, single-layer coatings are effective only at a specific wavelength and angle of incidence. To achieve broadband or wide-angle performance, modern AR coatings employ multiple layers with alternating high and low refractive indices, creating a gradual impedance match between air and substrate. Such coatings are essential in applications ranging from camera lenses and solar cells to precision laser optics, where even small reflective losses can degrade system efficiency or introduce unwanted interference effects.

### 1.5.2 Distributed Bragg Reflectors (Laser Mirrors)

Rather than introducing a single  $\lambda/4$  layer and relying on destructive interference, if we make multi-layer stack, with a sequence of high-index and low-index layers, and design for constructive interference, we can construct an extremely reflective mirror over a broad bandwidth. Consider a periodic stack of two dielectrics, with refractive indices  $n_H$  and  $n_L$  (high and low), and physical thicknesses  $d_H = \frac{\lambda}{4n_H}$

and  $d_L = \frac{\lambda}{4n_L}$ . Each interface between materials partially reflects and transmits the light. The reflected wave from each interface has a certain phase shift determined by the optical path through the layers.

Each interface between high and low refractive index layers reflects a small portion of the incident light. At certain wavelengths, specifically those satisfying the **Bragg condition**  $2n_{\text{eff}}\Lambda = m\lambda$  for integer  $m$  (you will have come across this in other courses!), the reflections from successive interfaces add **constructively**. This means that in each round-trip through a high-low pair adds a phase of  $2\pi$  to the reflected field at the design wavelength  $\lambda$ . Therefore, the reflections from all interfaces in the stack arrive back in phase and reinforce one another, and the transmitted wave becomes suppressed, leading to **high reflectivity** and **low transmission**.

The total reflected amplitude increases roughly geometrically with the number of periods  $N$ , while the transmitted amplitude decreases exponentially. The total reflection coefficient of the multilayer is

$$R = \left( \frac{(n_H/n_L)^{2N} - (n_s/n_0)}{(n_H/n_L)^{2N} + (n_s/n_0)} \right)^2 \quad (1.61)$$

where  $n_0$  is the refractive index of the incident medium. As  $N$  increases,  $R$  approaches unity, even though each interface by itself may reflect only a few percent.

### 1.5.2.1 Photonic StopBand

In a homogeneous medium, a plane wave has a simple propagation constant  $\beta = nk_0 = \frac{2\pi n}{\lambda}$ , and the field evolves as  $E(z) = E_0 e^{i\beta z}$ . But in a periodic medium, the refractive index  $n(z)$  varies periodically:

$$n(z + \Lambda) = n(z) \quad (1.62)$$

where  $\Lambda = d_H + d_L$  is the period. Because of this periodicity, the wave equation admits **Bloch wave solutions**:

$$E(z) = u(z)e^{iKz} \quad (1.63)$$

where  $u(z)$  has the same periodicity as the structure, and  $K$  is the **Bloch wavevector**.

Furthermore, this effect is broadband, with the spectral region of high reflectivity

known as the **photonic stop band**. To understand this, consider that the periodic alternation of refractive index causes coupling between forward and backward propagating waves. Using **coupled-mode theory**, one can show that:

$$\frac{dA_+}{dz} = i\kappa A_- e^{2i\delta z}, \quad (1.64)$$

$$\frac{dA_-}{dz} = -i\kappa A_+ e^{-2i\delta z} \quad (1.65)$$

Here,  $A_+$  and  $A_-$  are the forward and backward wave amplitudes,  $\kappa$  is the coupling coefficient proportional to the index contrast, and  $\delta=\beta-\pi/\Lambda$  is the detuning from the Bragg condition. Solving this gives propagation constants:

$$K = \sqrt{\kappa^2 + \delta^2} \quad (1.66)$$

Crucially, when  $|\delta| < \kappa$ ,  $K$  becomes imaginary. This means the Bloch wave does not propagate, but rather decays exponentially. The frequency range over which  $K$  is imaginary defines the photonic stop band. If we plot the dispersion relation we find that in allowed bands,  $K$  is real, and therefore there are propagating Bloch waves. In forbidden bands,  $K$  is complex, and we have evanescent fields. This looks analogous to the electronic band gap in a crystal lattice, but here it's for photons instead of electrons. The central gap (first-order stop band) occurs around the Bragg frequency  $\omega_B$  where  $2n_{\text{eff}}\Lambda=\lambda_B$ .

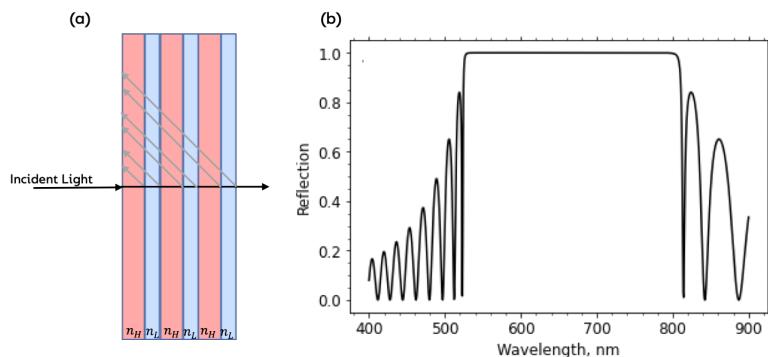
The relative width of the stop band depends on the contrast between high and low refractive index layers:

$$\frac{\Delta\lambda}{\lambda_0} \approx \frac{4}{\pi} \sin^{-1} \left( \frac{n_H - n_L}{n_H + n_L} \right) \quad (1.67)$$

A higher index contrast leads to a wider stop band, whilst lower contrast gives rise to narrower, weaker reflection.

### 1.5.2.2 High and Low Index Materials

In practice, only a small number of materials are suitable candidates for the high and low index layers.  $\text{SiO}_2$  is by far the most common choice of low index media, whilst the higher index media often varies. A summary of commonly chosen materials is given in Table 1.1.



**Figure 1.14: Distributed Bragg Reflector:** (a) A DBR is formed by stacking layers of high-to-low index materials, designed to be a quarter wavelength. As light is incident on the mirror, a portion of the light is reflected, whilst transmitted light is refracted through the mirror. (b) Stacking multiple DBR pairs together forms a high-reflectivity photonic stopband.

Low Index	$n$	High Index	$n$
SiO <sub>2</sub> (silica)	1.4570	TiO <sub>2</sub> (titania)	2.5836
Na <sub>3</sub> AlF <sub>6</sub> (cryolite)	1.3387	ZnS (zinc sulphide)	2.3504
LiF (lithium fluoride)	1.3913	ZrO <sub>2</sub> (zirconia)	2.1517
CaF <sub>2</sub> (calcium fluoride)	1.3770	Ta <sub>2</sub> O <sub>5</sub> (tantalum pentoxide)	2.1203
LaF <sub>3</sub> (lanthanum fluoride)	1.6025	HfO <sub>2</sub> (hafnium dioxide)	1.8943

Table 1.1: Refractive index at  $\lambda=633$  nm

### Example 3: Deposition Methods

What is the current state-of-the-art method for deposition of high index materials? For optical grade materials, what might be the design considerations/inherent trade-offs which would inform the choice of high index material for a particular application? Hint: your answer should consider the specific application - different applications have different requirements / foci.

## References

1. Grant, I. S. & Phillips, W. R. *Electromagnetism* (John Wiley & Sons, 2013).
2. Hecht, E. *Optics* 5th ed. eng (Pearson, Boston, Mass, 2017).

3. Fujiwara, H. *Spectroscopic Ellipsometry: Principles and Applications* (John Wiley & Sons, 2007).

## 2 | Lecture 2 - Optical Waveguides

For many optical/optoelectronic applications, control over the propagation direction is crucial. Often this is highly impractical to achieve with free-space optics (for example, transmitting internet data in free-space would be extremely difficult). This challenge can be mitigated by employing **optical waveguides** - structures which can support **electromagnetic modes**. For microwave frequencies, waveguides are typically made from metals such as copper, gold, or platinum; or superconductors such as niobium. For visible-telecoms frequencies however, these waveguides are typically made from dielectrics.

An optical waveguide consists of a longitudinal high-index dielectric (known as the *core*), transversely surrounded by a lower index dielectric (called the *cladding*). Due to total internal reflection between these layers, a guided optical wave can propagate along the longitudinal direction.

There are two basic types of optical-waveguide - **planar** and **non-planar**. A planar waveguide has optical confinement in only one transverse direction, with a core is sandwiched between cladding layers in only one direction, and an index profile dependent only on one coordinate  $x$  or  $y$ . On the other hand, a non-planar waveguide consists of a core transversally confined in two-dimensions by cladding. In this waveguide structure, the refractive index  $n$  is dependent on both  $x$  and  $y$  coordinates.

### 2.1 Planar Waveguides

The basic planar waveguide is simply a layer of higher refractive index material sandwiched between two layers of lower index material.

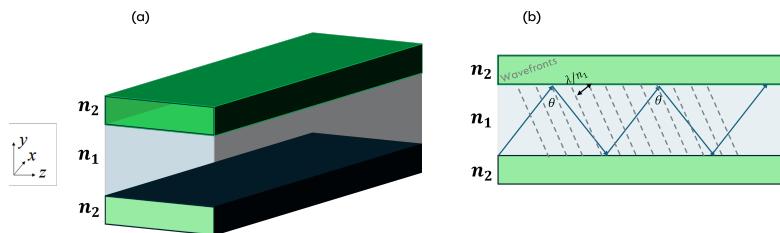


Figure 2.1: (a) Schematic diagram of a planar waveguide structure, formed by a core region surrounded by cladding in one dimension. (b) Guided modes from total internal reflection.

To find the optical modes (ie, states) of this system, we solve the wave equation:

$$\nabla^2 E_x = \frac{d^2 E_x}{dx^2} + \frac{d^2 E_x}{dy^2} + \frac{d^2 E_x}{dz^2} = \frac{n(y)^2}{c^2} \frac{d^2 E_x}{dt^2} \quad (2.1)$$

(only the x component of the E field needs to be considered). Separating variables and recognising that the solution is a wave travelling freely in the x and z directions:

$$E_x = X(x)Y(y)Z(z)T(t) \quad X = e^{ik_x x} \quad Y = e^{ik_y y} \quad Z = e^{ik_z z} \quad T = e^{-i\omega t} \quad (2.2)$$

gives:

$$\frac{d^2 Y}{dy^2} = \left( \frac{n(y)^2 \omega^2}{c^2} - \beta^2 \right) Y \quad (2.3)$$

....where  $\beta = \sqrt{k_x^2 + k_z^2}$  is the component of the wavevector in the plane of the waveguide.

If the RHS of this equation is negative we have solutions of the form

$$Y = A e^{ik_y y} + B e^{-ik_y y} \quad (2.4)$$

And if it is positive then solutions are of the form

$$Y = C e^{\kappa y} + D e^{-\kappa y} \quad (2.5)$$

In the case that  $n_1 > c\beta/\omega > n_2$ , Y is an oscillating function in the core and an exponential function in the cladding. In the cladding we can discard the term that increases with distance from the core, and retain only the decaying term. The result is that no power escapes in the  $\pm y$  direction, but the light is propagating freely in  $x, z$ ; ie a guided mode exists. The solutions in the core and cladding are joined together by continuity conditions, and we find that the core field distribution is quantised.

### 2.1.1 Mode Distribution in Planar Waveguides

In an optical waveguide, electromagnetic waves can only propagate in certain discrete field configurations known as modes. Each mode is a self-consistent so-

lution to Maxwell's equations that satisfies the boundary conditions imposed by the geometry and refractive index profile of the waveguide. These field distributions are stable in time and space and propagate along the longitudinal axis of the waveguide without changing shape, apart from a phase advance determined by the propagation constant  $\beta$ . The electric field of a mode can be expressed as

$$E(x, y, z, t) = \tilde{E}(x, y)e^{i(\beta z - \omega t)} \quad (2.6)$$

....where  $\tilde{E}(x, y)$  describes the spatial distribution of the transverse electric field,  $\beta$  is the propagation constant, and  $\omega$  is the angular frequency of the optical wave.

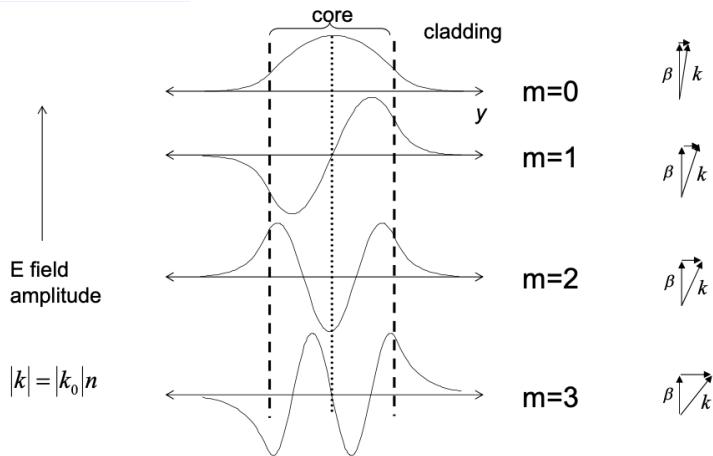


Figure 2.2: One dimensional schematic diagram of mode distributions in planar waveguides

Waveguide modes arise because light confined within a high-refractive-index core experiences total internal reflection at the interfaces with the lower-index cladding. For certain field distributions and propagation constants, the reflections at the boundaries are self-consistent, leading to constructive interference and a standing-wave pattern in the transverse direction. The resulting guided modes maintain their amplitude as they propagate along the z-axis. At each core-cladding interface, the tangential components of both the electric and magnetic fields must be continuous, which leads to quantisation of the allowed propagation constants. For guided propagation to occur, the propagation constant must satisfy

$$n_2 k_0 < \beta < n_1 k_0 \quad (2.7)$$

....where  $n_1$  and  $n_2$  are the refractive indices of the core and cladding respectively,

and  $k_0=2\pi/\lambda_0$  is the free-space wavenumber corresponding to the vacuum wavelength  $\lambda_0$ . The propagation constant  $\beta$  is related to an effective refractive index defined as

$$n_{\text{eff}} = \frac{\beta}{k_0} \quad (2.8)$$

The effective index lies between the indices of the core and cladding: modes with  $n_{\text{eff}} \approx n_1$  are tightly confined to the core, while modes with  $n_{\text{eff}}$  approaching  $n_2$  are weakly guided and extend more into the cladding.

Different families of modes exist depending on the symmetry and dimensionality of the waveguide. In a planar or slab waveguide, where the refractive index varies only in one dimension, the solutions can be separated into transverse electric (TE) and transverse magnetic (TM) polarisations. For TE modes, the longitudinal electric field component vanishes,  $E_z=0$ , while for TM modes, the longitudinal magnetic component vanishes,  $H_z=0$ . The allowed values of  $\beta$  in each case are obtained by solving the dispersion relations derived from the boundary conditions at the core-cladding interfaces.

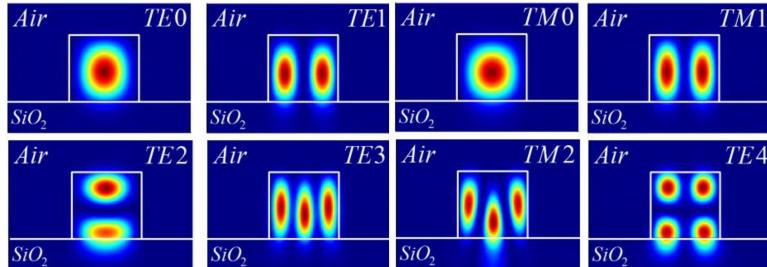


Figure 2.3: Two dimensional schematic diagram of mode distributions in planar waveguides

## 2.1.2 Applications of Planar Waveguides

### 2.1.2.1 Photonic Integrated Circuits

Planar waveguides are the fundamental structures that guide light in integrated photonic circuits, serving as the optical equivalent of electrical wires on a chip. They confine and route optical signals across a planar substrate using total internal reflection. The typical structure consists of a high refractive index core layer sandwiched between lower-index claddings and supported by a substrate. This index contrast confines the optical field in the vertical direction, while lateral confine-

ment is achieved by patterning the waveguide geometry, forming what are called channel waveguides.

Light propagates through these waveguides as guided modes that satisfy boundary conditions at the core-cladding interfaces. Each mode has an effective refractive index  $n_{\text{eff}}$ , which determines the propagation constant  $\beta = n_{\text{eff}} 2\pi/\lambda$ . The waveguide geometry and materials define which modes can exist, how tightly the light is confined, and how dispersion and loss behave. These parameters can be precisely engineered by controlling the waveguide width, thickness, and refractive index contrast.

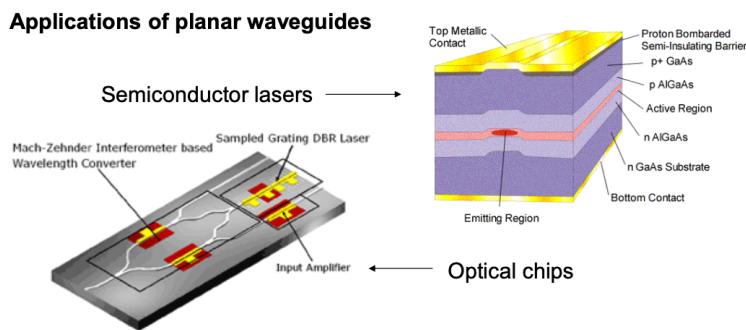


Figure 2.4: Planar waveguides are used routinely in photonic integrated circuits and semiconductor lasers

In integrated optical chips, planar waveguides form the interconnect network that links all active and passive components. Devices such as Y-branch splitters, directional couplers, and Mach-Zehnder interferometers are formed by branching or coupling nearby waveguides to manipulate amplitude and phase. Waveguides can be curved into loops to form ring or disk resonators for filtering, modulation, or sensing. Grating couplers are etched onto the surface of waveguides to diffract light into or out of optical fibres, providing efficient input and output coupling. Other devices, such as modulators and detectors, rely on guiding light through regions where it interacts with electro-optic or photoactive materials.

Several material platforms are used for planar waveguides, depending on the desired performance and wavelength range. Silicon-on-insulator (SOI) is widely used in silicon photonics because of its high index contrast and compatibility with CMOS fabrication. Silicon nitride ( $\text{Si}_3\text{N}_4$ ) waveguides offer low propagation loss and operate across a broad wavelength range, including the visible. III-V semiconductors such as indium phosphide (InP) support optical gain and are used for integrated lasers and amplifiers. Lithium niobate on insulator (LNOI) provides a strong electro-optic response for high-speed modulators. Chalcogenide glasses

are employed for nonlinear and mid-infrared photonic circuits.

Planar waveguides enable a wide range of applications including optical communication, quantum information processing, optical sensing, LiDAR beam steering, and microwave photonic signal processing. Their advantages include compactness, scalability, mechanical stability, and compatibility with established semiconductor manufacturing. By integrating many such waveguides and components on a single chip, complex optical systems can be realised with high performance, low cost, and reduced footprint compared to traditional free-space optics.

### 2.1.2.2 Semiconductor Lasers

In semiconductor lasers, planar waveguides play a central role in confining and guiding both the optical and carrier distributions within the laser structure. The main purpose of the waveguide is to ensure that the light generated in the active region remains confined along the desired direction of propagation, allowing it to undergo stimulated emission and amplification before exiting as a coherent laser beam.

The typical semiconductor laser consists of an active region, usually a quantum well or multiple quantum wells, sandwiched between layers of materials with slightly lower refractive index. This creates an optical waveguide along the plane of the junction. The refractive index contrast between the core (active region) and the surrounding cladding layers ensures total internal reflection, forming a planar optical cavity that confines light in the vertical direction. Lateral confinement is achieved either by etching ridge structures or by index guiding through variations in the refractive index across the wafer.

The optical mode supported by the planar waveguide overlaps strongly with the gain region, maximising optical confinement and gain efficiency. The confinement factor, defined as the fraction of the optical mode that overlaps the active region, is a key parameter that determines threshold current and output power. The thickness and refractive index profile of the waveguide layers are engineered to balance optical confinement with low loss and efficient heat dissipation.

In edge-emitting semiconductor lasers, such as Fabry-Perot, distributed feedback (DFB), and distributed Bragg reflector (DBR) lasers, the planar waveguide defines the propagation path between the cleaved or reflective facets that form the longitudinal cavity. The light is guided along the plane of the chip, and feedback is provided by the cleaved end faces or periodic Bragg gratings integrated within the waveguide. The planar geometry enables precise control of mode confinement and wavelength selectivity, essential for single-mode and tuneable laser designs.

In vertical-cavity surface-emitting lasers (VCSELs), the planar waveguide confines light in the vertical direction between two distributed Bragg reflectors while maintaining lateral confinement through oxide apertures or surface patterning. Although the emission is perpendicular to the wafer plane, the optical guiding principle in the cavity layers remains the same as in planar waveguides: high-index layers confine light where gain occurs, enabling efficient lasing.

Planar waveguides are also used in integrated laser platforms, where the laser cavity is formed monolithically or hybridly with other photonic components on a chip. The guided mode from the laser can be directly coupled into adjacent passive waveguides, enabling on-chip routing of coherent light without the need for bulky alignment. This approach is used in silicon photonics with heterogeneous integration of III-V gain materials bonded onto silicon or silicon nitride waveguides.

## 2.2 Non-Planar Waveguides (Optical Fibres)

Light propagation in an optical fibre can be understood by treating the fibre as a cylindrical dielectric waveguide with refractive index  $n(r)$  that varies only with radius. For a step-index fibre, the refractive index is  $n_1$  in the core ( $r < a$ ) and  $n_2$  in the cladding ( $r > a$ ), with  $n_1 > n_2$ .

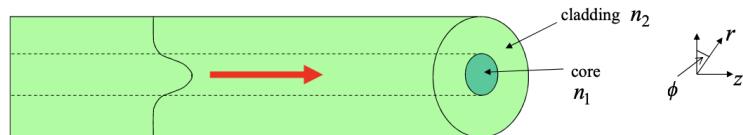


Figure 2.5: Cross-sectional schematic of an optical fibre, showing the core and cladding regions

The vector wave equation for the electric field in an optical fibre is:

$$\nabla^2 \mathbf{E} - \frac{n^2(r)}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0 \quad (2.9)$$

Assuming harmonic time dependence, for a monochromatic field  $\mathbf{E}(r, \phi, z, t) = \mathbf{E}(r, \phi, z)e^{i\omega t}$ , substitution gives the **Helmholtz equation**:

$$\nabla^2 \mathbf{E} + k_0^2 n^2(r) \mathbf{E} = 0 \quad (2.10)$$

Since optical fibre has cylindrical symmetry, we can use cylindrical polar coordinates  $r, \phi, z$ . Assuming the field varies along  $z$  as  $e^{i\beta z}$ , we have:

$$\mathbf{E}(r, \phi, z) = \mathbf{E}(r, \phi)e^{i\beta z} \quad (2.11)$$

Substituting into the Helmholtz equation gives:

$$[\nabla_t^2 + (k_0^2 n^2(r) - \beta^2)] \mathbf{E}_t(r, \phi) = 0 \quad (2.12)$$

where  $\nabla_t^2$  is the **transverse Laplacian**:

$$\nabla_t^2 = \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2}{\partial \phi^2} \quad (2.13)$$

In the **weakly guiding approximation** ( $n_{\text{core}} \approx n_{\text{clad}}$ ), polarisation coupling is negligible, and we can treat one field component (say  $E_z$ ) as satisfying the scalar wave equation:

$$\left[ \frac{1}{r} \frac{d}{dr} \left( r \frac{d}{dr} \right) + \frac{1}{r^2} \frac{d^2}{d\phi^2} + k_0^2 n^2(r) - \beta^2 \right] \psi(r, \phi) = 0 \quad (2.14)$$

where  $\psi$  represents the longitudinal field component (e.g.,  $E_z$  or  $H_z$ ).

Inside the core ( $r < a$ ),  $n(r)=n_1$ :

$$[\nabla_t^2 + (k_1^2 - \beta^2)] \psi = 0 \quad (2.15)$$

In the cladding ( $r > a$ ),  $n(r)=n_2$

$$[\nabla_t^2 + (k_2^2 - \beta^2)] \psi = 0 \quad (2.16)$$

These yield **Bessel function** solutions:

$$\text{Core: } J_l(ur/a) \quad (2.17)$$

$$\text{Cladding: } K_l(wr/a) \quad (2.18)$$

with the transverse propagation parameters  $u=a\sqrt{k_0^2 n_1^2 - \beta^2}$ ,  $w=a\sqrt{\beta^2 - k_0^2 n_2^2}$ , and  $u^2 + w^2 = V^2$ , where

$$V = k_0 a \sqrt{n_1^2 - n_2^2} \quad (2.19)$$

is the **normalised frequency** of the fibre. This sets the number of guided modes: as  $V$  increases, more solutions exist. The fibre is single-mode when  $V < 2.405$ , the first zero of  $J_0$ . The fundamental mode, denoted  $LP_{01}$ , has a nearly Gaussian field distribution described approximately by

$$\Psi_{01}(r) \propto \begin{cases} J_0(ur), & r \leq a \\ \frac{J_0(ua)}{K_0(\mu a)} K_0(\mu r), & r \geq a \end{cases} \quad (2.20)$$

The scalar treatment is accurate for weakly guiding fibres ( $n_1 \approx n_2$ ). In the exact vector formulation, Maxwell's equations yield TE, TM, and hybrid HE and EH modes. The  $LP$  modes are approximate combinations of these; for example,  $LP_{01}$  corresponds closely to the  $HE_{11}$  mode.

The propagation constant  $\beta(\omega)$  determines key dispersive properties. The effective refractive index is  $n_{\text{eff}} = \beta/k_0$ , the group velocity satisfies  $v_g^{-1} = d\beta/d\omega$ , and the group-velocity dispersion is  $\beta_2 = d^2\beta/d\omega^2$ . Dispersion arises from both the material refractive index dependence on frequency and from the waveguide geometry, which causes the modal confinement to vary with wavelength. These combine to give the total dispersion and define the zero-dispersion wavelength of the fibre.

Physically, the fibre confines light through total internal reflection, corresponding mathematically to oscillatory Bessel solutions in the core and exponentially decaying modified Bessel solutions in the cladding. Boundary matching quantises the allowed propagation constants, leading to discrete guided modes. The normalised frequency  $V$  controls how many modes exist and how tightly they are confined, with small  $V$  corresponding to single-mode operation and stronger confinement.

### 2.2.1 Mode Distribution in Optical Fibres

In optical fibers, which are cylindrically symmetric, the fields cannot generally be separated into purely TE or TM polarizations because of the coupling between longitudinal and transverse components. Instead, the modes are hybrid in nature and are denoted as  $HE_{mn}$  and  $EH_{mn}$ , where  $m$  and  $n$  represent the azimuthal and radial mode numbers, respectively. These hybrid modes describe field patterns

that vary both radially and azimuthally while satisfying the cylindrical boundary conditions.

The number of modes that a waveguide can support is determined by its normalised frequency or  $V$ -number, defined as

$$V = \frac{2\pi a}{\lambda_0} \sqrt{n_1^2 - n_2^2} \quad (2.21)$$

where  $a$  is the core radius. The  $V$ -number effectively measures the core size in units of the wavelength, weighted by the index contrast. For a step-index optical fibre, only one guided mode exists when  $V < 2.405$ , corresponding to the fundamental  $\text{HE}_{11}$  mode. When  $V > 2.405$ , higher-order modes become allowed, and the fibre becomes multimode. As  $V$  increases, more transverse field configurations satisfy the boundary conditions, leading to an increasing number of supported modes.

Each guided mode has its own dispersion relation  $\beta(\omega)$ , which describes how the propagation constant varies with optical frequency. The group velocity of a mode is defined by  $v_g = d\omega/d\beta$ , and since different modes have different dispersion relations, their group velocities differ. This leads to modal dispersion, where pulses broaden as they propagate because different modes travel at different speeds. Single-mode fibres avoid this effect by supporting only one spatial mode, which is crucial for high-bandwidth, long-distance communication systems.

In the cladding region, the guided field does not abruptly vanish but decays exponentially. This decay represents the evanescent field associated with total internal reflection. The transverse variation of the electric field in the cladding can be written as

$$E(x) \propto e^{-\kappa x}, \quad \kappa = \sqrt{\beta^2 - n_2^2 k_0^2} \quad (2.22)$$

The parameter  $\kappa$  describes the spatial rate of decay. The presence of this evanescent field allows optical power to couple between adjacent waveguides placed in close proximity, forming the basis of directional couplers and integrated photonic circuits. It is also exploited in optical sensing, where the evanescent field interacts with the surrounding medium.

### 2.2.1.1 Manufacture of Optical Fibres

The classic means of fabricating optical fibres is by the **direct melt method**. In this process, the glass materials are melted and drawn directly into fiber without

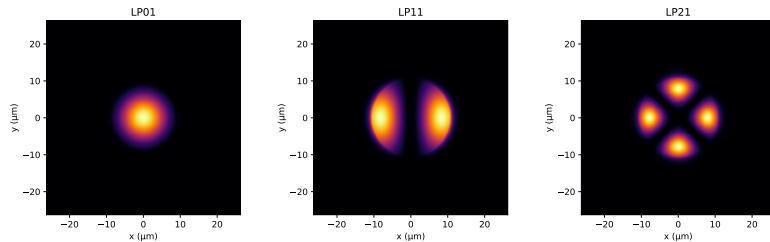


Figure 2.6: A set of LP modes

first producing a solid preform. This technique is most often implemented using a double-crucible system, in which two concentric crucibles contain molten core and cladding glasses of slightly different refractive indices. The inner crucible holds the higher-index core glass, while the outer crucible contains the lower-index cladding glass. Both melts are carefully heated to the appropriate viscosities and allowed to flow simultaneously through aligned orifices at the bottom of the crucibles. As the two molten streams emerge, they fuse together at their interface to form a continuous, concentric core-cladding structure, which is immediately drawn downward into a fine fiber. The fiber solidifies as it cools, after which it is coated with a protective polymer layer and wound onto spools.

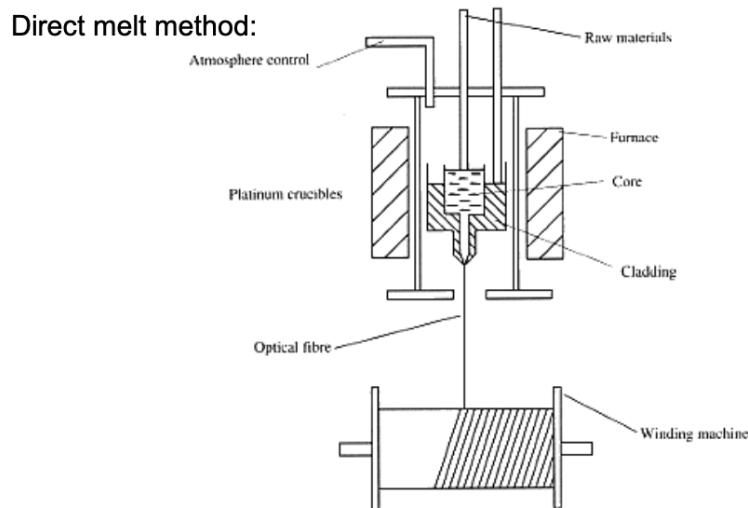


Figure 2.7: Schematic diagram of 'Direct Melt' method for fabricating optical fibres.

The process is simple, fast, and does not require the complex preform fabrication steps used in chemical vapor deposition methods such as MCVD or OVD. However, it offers limited control over the refractive index profile and is more suscep-

tible to contamination from the crucible materials and the atmosphere, resulting in higher optical losses. For these reasons, the direct melt method is not suitable for the production of ultra-low-loss silica fibres used in telecommunications. It remains useful, however, for specialty fibres made from non-silica glasses such as fluoride, tellurite, or chalcogenide compositions, where vapour-phase deposition is impractical. In these cases, the direct melt method provides an efficient means of producing infrared-transmitting fibres and other nonstandard optical materials.

## 2.3 Degradation of Optical Signals in Fibres

Digital communications are achieved by guiding pulses of light down optical fibres. Light intensity above a certain threshold corresponds to a 1 while intensity below the threshold corresponds to a 0. The two main practical considerations in fibre telecomms are:

1. **Attenuation:** How far can I send a 1 before so much is lost that it becomes a 0?

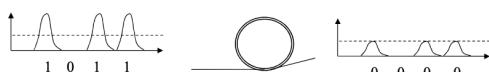


Figure 2.8: Attenuation

2. **Dispersion:** How far and how fast (in bits/s) can I send data before the pulses spread out so much as to become unrecognisable?

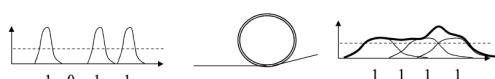


Figure 2.9: Dispersion

### 2.3.1 Attenuation

In practical optical fibres, as light propagates through the length of fibre, its power  $P$  diminishes exponentially  $P(z)=P_0 e^{-\alpha z}$ , where  $\alpha$  is the attenuation constant. Attenuation arises from several mechanisms, including material absorption due to electronic and vibrational resonances in silica, Rayleigh scattering from microscopic index fluctuations, bending and confinement losses, and impurity absorption, such as from  $\text{OH}^-$  ions. In standard single-mode silica fibres at  $1.55 \mu\text{m}$ , the attenuation is extremely low, around  $0.2 \text{ dB/km}$ .

Leakage due to cladding penetration is negligible in typical single mode fibres, as the cladding thickness is 58  $\mu\text{m}$  (9  $\mu\text{m}$  core diameter and 125  $\mu\text{m}$  total diameter). However there are other sources of attenuation to consider:

Attenuation in a silica optical fibre as a function of wavelength:

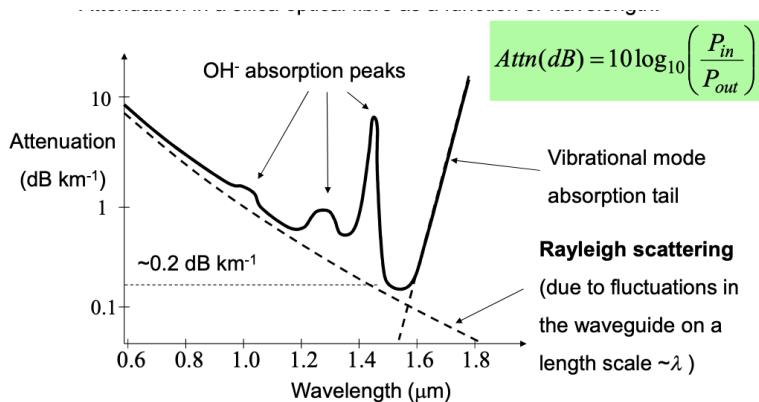


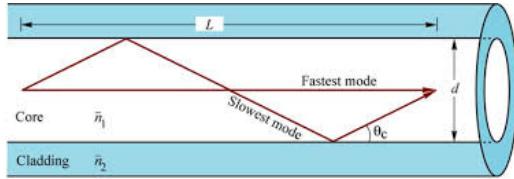
Figure 2.10: <caption>

All long haul telecommunications operate at optical wavelengths of either 1.5–1.6  $\mu\text{m}$  or 1.3  $\mu\text{m}$  to minimise attenuation of the signal.

A possible hi-tech solution? Photonic Crystal fibre with an air core. The guided light has very little interaction with the silica material (invented at Southampton University). More on this in lecture 4!

### 2.3.2 Dispersion

Dispersion in optical fibres refers to the dependence of the propagation constant  $\beta(\omega)$  on frequency. This means that different spectral components of a pulse travel at slightly different velocities, causing temporal broadening of the pulse as it propagates. Dispersion has two primary contributions: material dispersion, which arises from the frequency dependence of the refractive index  $n(\omega)$  of silica, and waveguide dispersion, which originates from the frequency-dependent spatial confinement of the mode. Total dispersion is quantified by the second derivative of the propagation constant,  $\beta_2 = d^2\beta/d\omega^2$ , or equivalently by the dispersion parameter  $D$  used in telecom systems. Silica's material dispersion passes through zero near 1.3  $\mu\text{m}$ , and waveguide dispersion can be engineered to achieve a zero-dispersion wavelength near 1.55  $\mu\text{m}$ , minimising pulse broadening for telecom applications.



### 2.3.2.1 Inter-Modal Dispersion

Different radial modes propagate at different speeds as their wave vectors point in different polar directions. This leads to rapid dispersion of pulses which occupy more than one mode. The time taken to propagate is:

$$t = \frac{Ln_1}{c} \sin \theta_m \quad (2.23)$$

so the time dispersion for this length of fibre is

$$\Delta t = \frac{Ln_1}{c} \left( \frac{1}{\sin \theta_c} - \frac{1}{\sin 90^\circ} \right) = L \frac{(n_1 - n_2)n_1}{n_2 c} \quad (2.24)$$

where  $\theta$  is the angle of propagation relative to the normal to the fibre axis and  $\theta_c$  is the critical angle for total internal reflection at the core/cladding interface.

Eg for a multimode fibre with  $n_1=1.52$ ,  $n_2=1.50$ , we find that  $\Delta t=100$  ns km $^{-1}$ . Such large dispersion would be intolerable for any long distance telecomms, so all long distance telecomms fibres are designed to support only a **single mode**.

A typical single mode telecomms fibre designed to operate at  $\lambda=1.5 - 1.6$   $\mu\text{m}$  has  $n_1=1.52$ ,  $n_2=1.50$ , and a core diameter of about 10  $\mu\text{m}$ .

Different radial modes propagate at different speeds as their wave vectors point in different polar directions. This leads to rapid dispersion of pulses which occupy more than one mode. The time taken to propagate along a length  $L$  of fibre is

$$t = \frac{Ln_1}{c} \sin \theta_m \quad (2.25)$$

so the time dispersion for this length of fibre is

$$\Delta t = \frac{Ln_1}{c} \left( \frac{1}{\sin \theta_c} - \frac{1}{\sin 90^\circ} \right) = L \frac{(n_1 - n_2)n_1}{n_2 c} \quad (2.26)$$

where  $\theta$  is the angle of propagation relative to the normal to the fibre axis and  $\theta_c$  is the critical angle for total internal reflection at the core/cladding interface. The approximate difference in propagation time between the fastest mode (propagating straight down the fibre) and the slowest modes (propagating at the critical angle) is therefore

Eg for a multimode fibre with  $n_1=1.52$ ,  $n_2=1.50$ , we find that  $\Delta t=100 \text{ ns km}^{-1}$ . Such large dispersion would be intolerable for any long distance telecomms, so all long distance telecomms fibres are designed to support only a single mode.

A typical single mode telecomms fibre designed to operate at  $\lambda=1.5 - 1.6 \mu\text{m}$  has  $n_1=1.52$ ,  $n_2=1.50$ , and a core diameter of about  $10 \mu\text{m}$ .

### 2.3.2.2 Intra-Modal Dispersion

This arises from the finite spectral width  $\Delta\lambda_s$  of the light source combined with a wavelength dependence of the propagation speed along the fibre.

The propagation constant  $\beta(\omega)$  depends on frequency. So different spectral components of a pulse travel at slightly different velocities, causing temporal broadening of the pulse as it propagates.

Two primary contributions: **Material dispersion** - arises from the frequency dependence of the refractive index  $n(\omega)$  of silica. **Waveguide dispersion** - originates from the frequency-dependent spatial confinement of the mode. Total dispersion is quantified by the second derivative of the propagation constant,  $\beta_2=d^2\beta/d\omega^2$ , or equivalently by the dispersion parameter  $D$  used in telecom systems.

Intra-modal dispersion arises from the finite spectral width  $\Delta\lambda_s$  of the light source combined with a wavelength dependence of the propagation speed along the fibre.

$$\Delta t = \Delta\lambda \cdot L \cdot |D_T| \quad (2.27)$$

$$D_T = D_M + D_W + D_P \quad (2.28)$$

measured in  $\text{ps nm}^{-1} \text{ km}^{-1}$ .

- $D_M$  is material dispersion due to the dependence of  $n_1$  on wavelength (sign depends on  $\lambda$ ).
- $D_W$  is waveguide dispersion due to dependence of  $\theta_m$  on wavelength (always negative for SMF).

- $D_P$  is profile dispersion due to the dependence of  $n_2$  on wavelength.

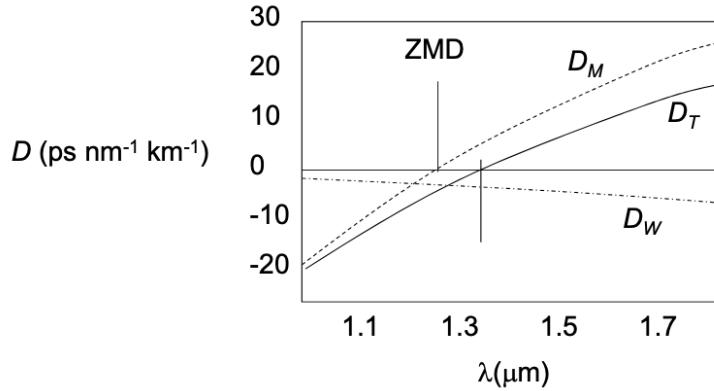


Figure 2.11: <caption>

In a pure silica SMF,  $D_M$  and  $D_W$  dominate. Fortunately we can find a wavelength at which  $D_M + D_W = 0$ , and we can tune this wavelength by doping the fibre.

An electrical pulse in the time domain picture  $E(t)$  is related to a pulse in the frequency domain  $E(\omega)$  by the Fourier transform:

$$\tilde{E}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} E(t) e^{-i\omega t} dt \quad (2.29)$$

$$E(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{E}(\omega) e^{i\omega t} d\omega \quad (2.30)$$

From this Fourier duality, the temporal width of the pulse is directly related to the spectral pulse width:

$$\Delta t \Delta \omega \geq \frac{1}{2} \quad (2.31)$$

So due to Fourier duality, a short optical pulse has a narrow duration  $\Delta t$  but correspondingly a broad spectral width  $\Delta\omega$ , satisfying:

$$\Delta t \Delta \omega \geq 1/2 \quad (2.32)$$

Dispersion in the fibre causes each frequency component to accumulate a different phase,  $E(z, t) = \int A(\omega) e^{i[\beta(\omega)z - \omega t]} d\omega$ , and this frequency-dependent phase

leads to temporal spreading of the pulse. In this sense, the temporal broadening is a classical manifestation of Heisenberg's uncertainty principle: the shorter the pulse, the broader its spectrum, and the more it is affected by dispersion.

In addition, attenuation interacts with dispersion to limit the information-carrying capacity of the fibre. High-power, short pulses with large spectral widths are more susceptible to dispersion-induced broadening, and any differential loss across the spectrum can further distort the pulse. Thus, both attenuation and dispersion enforce practical limits on pulse compression and transmission, providing a time-bandwidth trade-off dictated by wave physics. Narrow pulses require wide frequency content, and the fibre acts as a dispersive filter that converts this spectral width into temporal spreading.

So, there must be an optimal pulse duration which balances pulse length and dispersion. From Equation 2.32

$$c\Delta t \Delta k \geq \frac{1}{2} \quad (2.33)$$

and

$$\frac{2\pi c}{\lambda^2} \Delta t \Delta \lambda \geq \frac{1}{2} \quad (2.34)$$

So, from the uncertainty principle, the spectral width is optimally:

$$\Delta \lambda \geq \frac{\lambda^2}{4\pi c \Delta t} \quad (2.35)$$

A 0.1 ns pulse at  $\lambda=1.55 \mu\text{m}$  has  $\Delta\lambda \sim 0.01 \text{ nm}$  and so  $D > 0.1 \text{ ps/km}$ . After 1000 km the pulse will double in length. This is about the best we can do with a single wavelength.

## 2.4 Wavelength Division Multiplexing

Light travels in silica fibre at roughly  $\nu_g = \frac{c}{n} \approx \frac{3 \times 10^8}{1.45} \approx 2.07 \times 10^8 \text{ m/s}$ . This is the **speed of the photons**, not the data rate. The data rate can be far lower, due to dispersion. The **maximum data rate** depends on the **optical bandwidth**  $B$  of the single channel and how fast the signal can be modulated:

$$R_{\max} \sim 2B \log_2 M \quad (2.36)$$

...where  $M$  is number of discrete signal levels (modulation order, e.g., 2 for binary, 16 for 16-QAM)

The **optical window** for standard single-mode fibres is limited:

Window	$\lambda$	Usable $B$
O	1260-1360 nm	100 nm
E	1360-1460 nm	100 nm
S	1460-1530 nm	70 nm
C	1530-1565 nm	35 nm
L	1565-1625 nm	60 nm

If only one wavelength can be used, the maximum data rate is limited by how fast a laser can be modulated. The practical limit is about 100 – 1000 Gbit/s.

What if we use multiple wavelengths instead of just one? Now, each data stream would be modulated onto a laser of a specific wavelength. A multiplexer (MUX) combines all these wavelengths into a single fibre. At the receiving end, a demultiplexer (DEMUX) separates the wavelengths and sends each to its own detector.

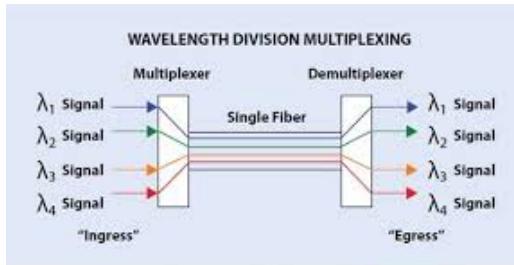


Figure 1: Basic WDM Technology Diagram

Figure 2.12: Figure from Optcore

The 1500 – 1600 nm region of the attenuation window is traditionally split into two: **C band** = 1530 – 1565 nm, and **L band** = 1565 – 1625 nm.

In “dense WDM” (DWDM) each band is divided up into multiple different wavelengths.

Band	Channels	Rate/ $\lambda$	Total per Band
C band	80	400 Gbit/s	32 Tbit/s
L band	60	400 Gbit/s	24 Tbit/s
Total	140	400 Gbit/s	56 Tbit/s

The total data transmission rate is 50-60 Tbit/s! With this capacity there is more than enough fibre already installed to keep the telecommunications industry happy

for many years (decades?) to come. There is currently a huge market for devices that can amplify, modulate, and switch, broadband WDM signals efficiently so that we can make full use of this bandwidth.

# 3 | Lecture 3 - Controlling Light Propagation in Optical Materials

## 3.1 Light in Optical Materials

When light propagates in a crystalline material, it interacts with the atomic structure, and the velocity of the light passing through is reduced by a factor related to the materials permittivity  $\epsilon$ . Permittivity takes the form of a symmetric rank-2 tensor:

$$\epsilon_r = \begin{pmatrix} \epsilon_{11} & \epsilon_{12} & \epsilon_{13} \\ \epsilon_{21} & \epsilon_{22} & \epsilon_{23} \\ \epsilon_{31} & \epsilon_{32} & \epsilon_{33} \end{pmatrix} \quad (3.1)$$

By operating in a rotated frame of reference with axes parallel to the principal axis of the permittivity, this tensor can be diagonalised:

$$\epsilon'_r = \begin{pmatrix} \epsilon_1 & 0 & 0 \\ 0 & \epsilon_2 & 0 \\ 0 & 0 & \epsilon_3 \end{pmatrix} \quad (3.2)$$

So, for each principal axis 1, 2, 3 there is a distinct permittivity  $\epsilon_1, \epsilon_2, \epsilon_3$ . If all these permittivities are equal, the crystal is said to be isotropic, and the refractive indices ( $n_i = \sqrt{\epsilon_i}$ ) are also equal ( $n_1 = n_2 = n_3$ ). However, if the permittivities are not equal, the material is said to be anisotropic. In a biaxial crystal, all axes affect  $n$  differently, so  $n_1 \neq n_2 \neq n_3$ . In a uniaxial crystal one of the crystal axes affects the refractive index differently, so  $n_1 = n_2 \neq n_3$ .

Note that since it does not follow the rules for rotating tensors, the refractive index is not a tensor. However, the refractive index experienced by light passing through the crystal clearly depends on the direction of the electric field, ie the polarisation of the light. In uniaxial materials, this is known as birefringence, and in biaxial materials the effect is known as trirefringence.

## 3.2 Light Propagation in Birefringent Materials

As we just discussed, a birefringent material is an anisotropic medium in which the refractive index depends on the polarisation direction of light. The material possesses two orthogonal principal axes:

Symmetry	Material	$n_1$	$n_2$	$n_3$
Uniaxial	Calcite ( $\text{CaCO}_3$ )	1.66	1.66	1.49
	Lithium Niobate ( $\text{LiNbO}_3$ )	2.27	2.27	2.19
	Barium Titanate $\text{BaTi}_2\text{O}_3$	2.41	2.41	2.32
Biaxial	Calcium Titanate ( $\text{CaTiO}_3$ )	2.30	2.34	2.38
	Potassium Titanyl Phosphate ( $\text{KTiOPO}_4$ )	1.74	1.75	1.83

Table 3.1: &lt;caption&gt;

- the *fast axis*, corresponding to refractive index  $n_f$ , and
- the *slow axis*, corresponding to refractive index  $n_s$ .

If a light wave enters such a crystal, its electric field can be decomposed into components along these two axes. Each component accumulates a different phase during propagation because of the index difference:

$$\phi_i = \frac{2\pi}{\lambda} n_i d, \quad (3.3)$$

where  $d$  is the propagation distance and  $i \in \{f, s\}$ . The phase difference between the two components is

$$\Delta\phi = \frac{2\pi d}{\lambda} (n_s - n_f). \quad (3.4)$$

The birefringent material acts as a phase retarder, introducing a relative phase shift between orthogonal components. Depending on  $\Delta\phi$ , different polarisation transformations occur:

- $\Delta\phi = 0$  or  $2\pi$ : polarisation unchanged.
- $\Delta\phi = \pi/2$ : quarter-wave plate — converts linear polarisation into circular or elliptical.
- $\Delta\phi = \pi$ : half-wave plate — rotates the direction of linear polarisation.

### 3.2.1 Polarisation Control Using Wave Plates

If light passes through a plate of birefringent material such that the direction of propagation of the light is parallel to  $x_1$  or  $x_2$ , then the transverse electric field lies in the plane containing the two principal axes with different  $n$ . This field vector can then be resolved into components along these principal axes, and the two components propagate at different speeds through the crystal, therefore having different wavelengths. When the emerge from the other side, there will be a phase

difference between them

$$\Delta\phi = kL\Delta n = \frac{2\pi L}{\lambda} \Delta n \quad (3.5)$$

where  $L$  is the path length through the crystal.

Suppose the light beam is incident parallel to the  $x$  axis of the crystal and has field amplitude  $E_0$ , and we rotate the plate so that the polarisation direction of the incident beam makes an angle  $\theta$  with the  $y$  axis. The input beam is then given by

$$\vec{E}_{in} = E_0 (\cos \theta e^{i(kx-\omega t)} \hat{y} + \sin \theta e^{i(kx-\omega t)} \hat{z}) \quad (3.6)$$

and the output beam is

$$\vec{E}_{out} = E_0 (\cos \theta e^{i(kx-\omega t)} \hat{y} + \sin \theta e^{i(kx-\omega t+\Delta\phi)} \hat{z}) \quad (3.7)$$

By choosing  $\theta$  and  $\Delta\phi$  we now determine the polarisation state of the output. Suppose we choose the plate thickness so that  $\phi=\pi$ :

$$\vec{E}_{out} = E_0 (\cos \theta e^{i(kx-\omega t)} \hat{y} - \sin \theta e^{i(kx-\omega t)} \hat{z}) \quad (3.8)$$

so effectively we have inverted the  $z$  component and the polarisation has undergone a reflection in the  $x - y$  plane. This is known as a half wave plate. By choosing  $\theta$  we can 'rotate' a linearly polarised beam to any other linear polarisation (eg,  $\theta=45^\circ$  gives a  $90^\circ$  rotation as in the figure below).

Alternatively we could choose a plate of thickness giving  $\phi = \pi/2$ :

$$\vec{E}_{out} = E_0 (\cos \theta e^{i(kx-\omega t)} \hat{y} + i \sin \theta e^{i(kx-\omega t)} \hat{z}) \quad (3.9)$$

Now the output beam is not linearly polarised, and in fact the electric field is non-zero for all  $x,t$ . The beam is elliptically polarised. The special case of  $\theta=45^\circ$  results in:

$$|\vec{E}_{out}|^2 = \frac{E_0^2}{2} \quad (3.10)$$

for all  $x, t$ , and the direction of polarisation rotating with  $x, t$ , once per period. The light is circularly polarised.

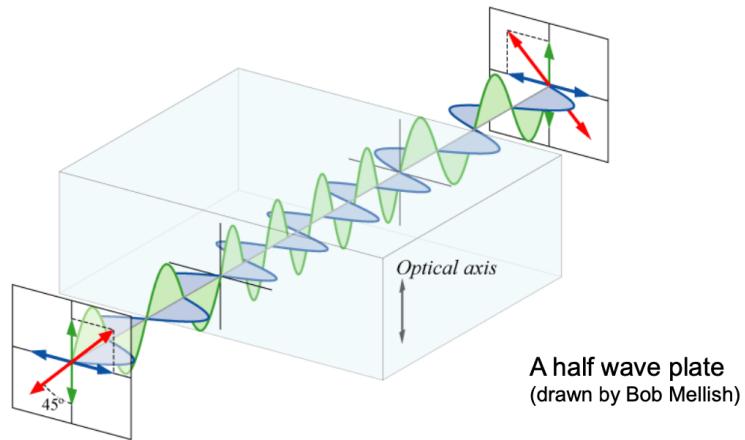


Figure 3.1: Light Propagation in a Half Waveplate

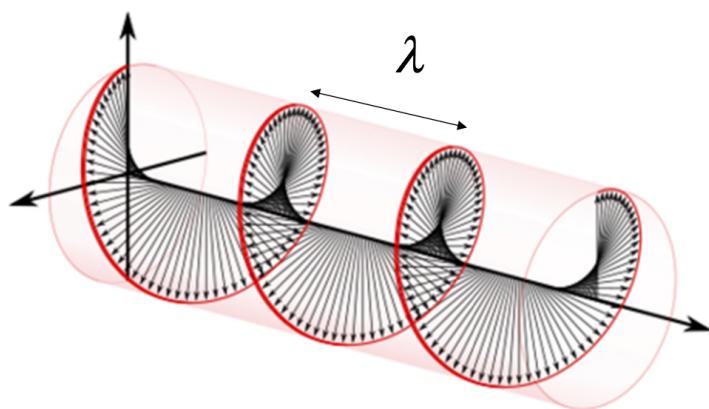


Figure 3.2: Light propagation in a Quarter-Wave Plate

If light enters a birefringent crystal at an angle to the optical axis, in general one polarisation of the electric field will experience only  $n_x, n_y = n_o$ . The other will have a component that experiences  $n_o$  and a component that experiences  $n_z = n_e$ . The wave fronts of this 'extraordinary ray' are elliptical which results in the ray propagating at an angle to the normal (see figure).

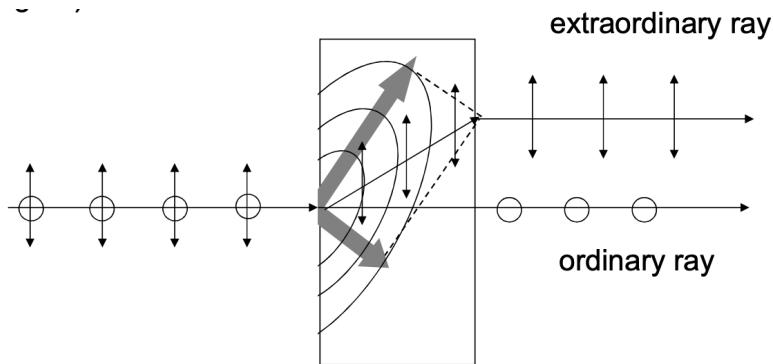
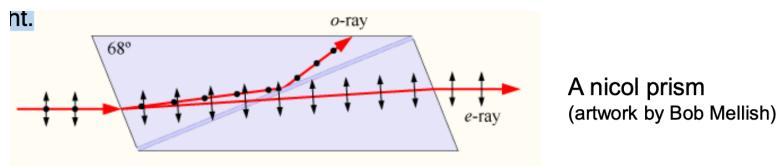


Figure 3.3: Ordinary and Extra-Ordinary Rays

This 'double refraction' is the phenomenon most often associated with birefringence. It is commonly used in a Nicol prism to polarise light.



**Question - Using the Jones Matrix Formalism, derive the relationships for the HWP and QWP waveplates.**

### 3.3 Electro-Optic Materials

Electro-optic materials give a refractive index which can be modified in response to an applied static electric field. There are many known electro-optically active materials, but for most the magnitude of this effect is small.

The figure of merit for linear EO materials is the Pockels Effect, which quantifies the degree to which the cations and anions within the material shift relative to each other in response to an applied electric field, modifying the local electric dipole moments within the crystal, in units of pm/V. As is discussed in the

next section, the Pockels effect is described by a tensor  $r_{ij}$ . For lithium niobate ( $\text{LiNbO}_3$ ), one of the most widely used EO material, the largest Pockels tensor coefficient is  $r_{33}=33 \text{ pm/V}$ . This is considered large for many planar optical modulators, but in practice the change in refractive index is very small - for a bias field of magnitude  $10^7 \text{ MV/m}$ , the relative change in refractive index for  $\text{LiNbO}_3$  would be  $\sim 10^{-4}$ . However, there are many other materials which are known to exhibit an electro-optic response. Ferroelectric materials and perovskites are the most widely studied materials which exhibit an electro-optic response, but many semiconductor alloys such as GaAs also display a modest Pockels effect, as well as certain 2D materials and polymers.

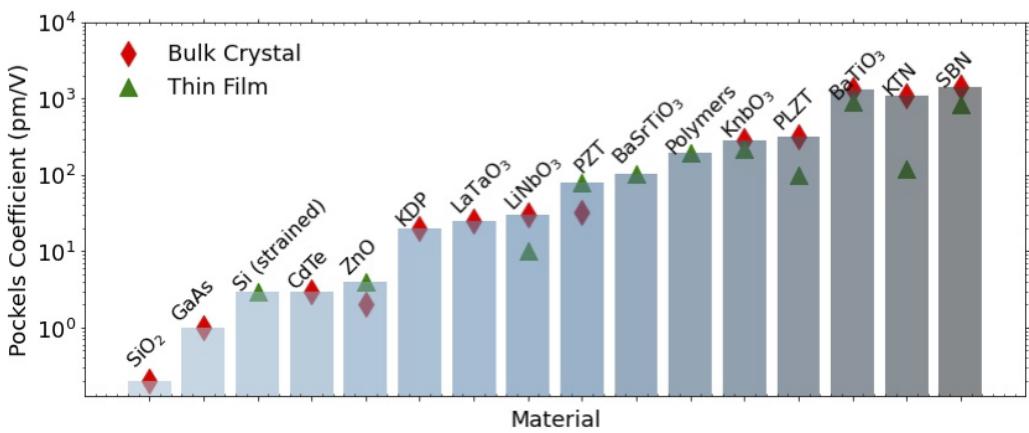


Figure 3.4: **Comparison of literature values for the Pockels effect in thin film form for different known EO materials.** Figure modified and updated from IBM Research [1]. This is not an exhaustive list of all known electro-optic materials, but a summary of the most widely studied ones which have demonstrated a strong Pockels effect and/or have been synthesised in thin-film.

### 3.3.1 Derivation of the Electro-Optic Tensor

This geometric description can be attained by considering the energy density  $U$  of the electromagnetic field  $\vec{E}$  with a flux density  $\vec{D}=\epsilon_0\epsilon\vec{E}$ , given by  $U_E=\frac{\vec{E}\cdot\vec{D}}{2}=\frac{1}{2\epsilon_0\epsilon}\vec{D}\cdot\vec{D}$ . From the definition of the dot-product, this can be decomposed in terms of individual components:

$$U_E = \frac{1}{2\epsilon_0\epsilon} \cdot \left( \frac{D_x^2}{\epsilon_x} + \frac{D_y^2}{\epsilon_y} + \frac{D_z^2}{\epsilon_z} \right) \quad (3.11)$$

From this, the principal cartesian coordinates can be simply recovered as  $x=\frac{D_x}{\sqrt{2\epsilon_0 U_E}}$ ,  $y=\frac{D_y}{\sqrt{2\epsilon_0 U_E}}$ , and  $z=\frac{D_z}{\sqrt{2\epsilon_0 U_E}}$ . The defining equation for the index ellipsoid clearly follows:

$$\frac{x^2}{n_x^2} + \frac{y^2}{n_y^2} + \frac{z^2}{n_z^2} = 1 \quad (3.12)$$

The index ellipsoid is the quadratic representation of the impermeability tensor  $\eta=\epsilon_0\epsilon^{-1}$ , i.e.  $\sum_{ij}\eta_{ij}x_i x_j=1$ , with  $i,j=[1, 2, 3]$ . The expanded impermeability tensor is a  $3 \times 3$  matrix of the form:

$$\eta_{ij} = \begin{pmatrix} \eta_{11} & \eta_{12} & \eta_{13} \\ \eta_{21} & \eta_{22} & \eta_{23} \\ \eta_{31} & \eta_{32} & \eta_{33} \end{pmatrix} \quad (3.13)$$

To simplify notation, the contracted index Voigt notation is used which maps 2-index pairs to a single index. This maps tensor components in the following way:  $\eta_{11} \rightarrow \eta_1$ ;  $\eta_{22} \rightarrow \eta_2$ ;  $\eta_{33} \rightarrow \eta_3$ ;  $\eta_{23}=\eta_{32} \rightarrow \eta_4$ ;  $\eta_{13}=\eta_{31} \rightarrow \eta_5$ ; and  $\eta_{12}=\eta_{21} \rightarrow \eta_6$ . So, the tensor can be expressed as:

$$\eta_{ij} = \begin{pmatrix} \eta_1 & \eta_6 & \eta_5 \\ \eta_6 & \eta_2 & \eta_4 \\ \eta_5 & \eta_4 & \eta_3 \end{pmatrix} \quad (3.14)$$

The ability of an electric field to re-distribute charges within a crystal depends on its capacity to overcome the atomic field binding between charged particles. For most materials, this binding energy is typically of order  $10^8$  V/cm, and therefore an applied field of reasonable intensity is not sufficient to produce an electro-optic effect. However, in certain non-centrosymmetric crystals, this binding may be comparatively weak, and consequently the electro-optic effect may be strong in these materials. In such materials, in the presence of an external electric field, the impermeability tensor is modified by both rotations of the principal axes and stretching and/or squashing of the ellipsoid vertices. Generally, the change of the impermeability tensor with respect to an applied field  $E$  can be understood by finding the impermeability tensor with no applied field  $\eta_{ij}(0)$ , and taking the Taylor expansion for higher order terms:

$$\eta_{ij}(\vec{E}) = \eta_{ij}(0) + \left( \frac{\partial \eta_{ij}}{\partial E_k} \right)_{E=0} E_k + \frac{1}{2} \left( \frac{\partial^2 \eta_{ij}}{\partial E_k \partial E_l} \right)_{E=0} E_k E_l + \dots \quad (3.15)$$

The first order term in this expansion is known as the *Pockels effect*, and is denoted in Einstein summation conventional shorthand as  $r_{ijk}$ , with  $r_{ijk} = \left( \frac{\partial \eta_{ij}}{\partial E_k} \right)_{E=0}$ . The second order term in this expansion is known as the *Kerr effect*, and describes the quadratic electro-optic response of a material in response to an applied field. This is denoted in shorthand as  $s_{ijkl}$ , with  $s_{ijkl} = \left( \frac{\partial^2 \eta_{ij}}{\partial E_k \partial E_l} \right)_{E=0}$ . So, the change in the impermeability tensor in response to an applied field  $\Delta\eta_{ij} = \eta_{ij}(E) - \eta_{ij}(0)$  is:

$$\Delta\eta_{ijk}(\vec{E}) = r_{ijk}E_k + s_{ijkl}E_kE_l \quad (3.16)$$

The Pockels coefficients describe the response of the crystal of an optically active material in response to an applied field. In a non-centrosymmetric crystal, the central cation is moved in a direction relative to the crystal axis, by a magnitude measured in pm/V. The Pockels effect is described by a rank-3 tensor, since it linearly connects a second-rank change in the optical properties of the material dielectric response to a first-rank applied electric field  $E_k$ . The tensor order follows directly from the nature of this physical relationship. connecting the electric field  $E_k$  to the change in the impermeability  $\eta_{ij}$ . The general form of a rank-3 tensor  $T_{ijk}$  for  $i \in \{1, 2, 3\}$  is given in Equation 3.17.

$$T = \left\{ \begin{bmatrix} T_{111} & T_{112} & T_{113} \\ T_{121} & T_{122} & T_{123} \\ T_{131} & T_{132} & T_{133} \end{bmatrix}, \begin{bmatrix} T_{211} & T_{212} & T_{213} \\ T_{221} & T_{222} & T_{223} \\ T_{231} & T_{232} & T_{233} \end{bmatrix}, \begin{bmatrix} T_{311} & T_{312} & T_{313} \\ T_{321} & T_{322} & T_{323} \\ T_{331} & T_{332} & T_{333} \end{bmatrix} \right\} \quad (3.17)$$

A rank-3 tensor therefore contains 27 components, but the impermeability tensor  $\eta_{ij}$  is symmetric, so  $\eta_{ij} = \eta_{ji}$ . Therefore the Pockels tensor  $r_{ijk}$  is also symmetric in  $i, j$ , so  $r_{ijk} = r_{jik}$ , and for each fixed  $k$ , the first two indices  $i, j$  form a symmetric  $3 \times 3$  matrix with 6 independent components. Since  $k \in \{1, 2, 3\}$ , there are 18 independent components of  $r_{ijk}$ . With Voigt notation, the rank-3 Pockels tensor can be represented with two components  $i, j$ , where  $i \in \{1, 2, 3, 4, 5, 6\}, j \in \{1, 2, 3\}$ . So, the Pockels tensor can be reduced to a  $6 \times 3$  matrix which relates the impact of the electric field on each crystal axis:

$$r_{ij} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \\ r_{41} & r_{42} & r_{43} \\ r_{51} & r_{52} & r_{53} \\ r_{61} & r_{62} & r_{63} \end{pmatrix} \quad (3.18)$$

The Voigt notation contraction can be further applied to the Pockels tensor:  $r_{1k}=r_{11k}$ ;  $r_{4k}=r_{23k}=r_{32k}$ ;  $r_{2k}=r_{22k}$ ;  $r_{5k}=r_{11k}=r_{31k}$ ;  $r_{3k}=r_{33k}$ ; and  $r_{6k}=r_{11k}=r_{21k}$ . With this, the index ellipsoid in the presence of an applied electric field can be expanded as:

$$\left(\frac{1}{n_x^2} + r_{1k}E_k\right)x^2 + \left(\frac{1}{n_y^2} + r_{2k}E_k\right)y^2 + \left(\frac{1}{n_z^2} + r_{3k}E_k\right)z^2 + 2r_{4k}yzE_k + 2r_{5k}xzE_k + 2r_{6k}xyE_k = 1 \quad (3.19)$$

The physical interpretation of this is that the electric field linearly maps to an optical response. The three-vector of the electroc field  $E_k$  maps to the change in the optical properties of a material  $\Delta\eta_{ij}$  represented by 6 independent components due to symmetry. The Pockels tensor describes how the electric field in direction  $k$  modifies the optical response  $\eta_{ij}$  in coordinate space  $(i, j)$  as:

$$\Delta\eta_i = \sum_{k=1}^3 r_{ik}E_k \quad (3.20)$$

This describes a linear relationship between the electric field and its optical response. Hence, despite the fact that the Pockels effect is a non-linear change of the polarisation vector  $\vec{P}$  of a material, it is often referred to as a linear electro-optic respoinsne with regards to the dielectric permeability and applied field.

It is evident from eq. 3.19 for certain crystal axes and field orientations, the principal axis of the index ellipsoid will be modified. The principal axes themselves then may require a coordinate transformation to yield the change in each component of  $\eta$  in the presence of an applied field.

### 3.3.2 Electro-Optic Response of LiNbO<sub>3</sub>

LiNbO<sub>3</sub> has a trigonal 3m crystal symmetry group. Therefore, the electro-optic tensor can be reduced to [2]:

$$r_{ij} = \begin{pmatrix} 0 & -r_{22} & r_{13} \\ 0 & r_{22} & r_{13} \\ 0 & 0 & r_{33} \\ 0 & r_{51} & 0 \\ r_{51} & 0 & 0 \\ -r_{22} & 0 & 0 \end{pmatrix} \quad (3.21)$$

Typical measured values from the literature for the different Pockels coefficients are shown below in Table 3.2.

$r_{13}$	10 pm/V
$r_{33}$	33 pm/V
$r_{22}$	7 pm/V
$r_{51}$	30 pm/V

Table 3.2: Typical Pockels coefficients for LiNbO<sub>3</sub>

The equation of the index ellipsoid is found by perturbing this tensor as:

$$\frac{x^2 + y^2}{n_o^2} + \frac{z^2}{n_e^2} - z^2 r_{33} E_z - (x^2 + y^2) r_{13} E_z \quad (3.22)$$

At 633 nm,  $n_o \approx 2.286$  and  $n_e \approx 2.203$

#### **Applying Electric Field along c-axis**

If an electric field  $E_z$  is applied parallel to the *c*-axis of the LiNbO<sub>3</sub> crystal, the index ellipsoid reduces to:

$$\left( \frac{1}{n_o^2} + r_{13} E_z \right) x^2 + \left( \frac{1}{n_o^2} + r_{13} E_z \right) y^2 + \left( \frac{1}{n_e^2} + r_{33} E_z \right) z^2 = 1 \quad (3.23)$$

The impermeability tensor then contains updated values:

$$\boldsymbol{\eta} = \begin{pmatrix} \eta_o + r_{13} E_z & 0 & 0 \\ 0 & \eta_o + r_{13} E_z & 0 \\ 0 & 0 & \eta_e + r_{33} E_z \end{pmatrix} \quad (3.24)$$

Since there are no mixed terms in the index ellipsoid, the principal axis remains unchanged and there is no need to define new principal axes. Applying the differential relation:

$$dn = \frac{1}{2} n^3 d \left( \frac{1}{n^2} \right) \quad (3.25)$$

....the corresponding changes in the refractive indices for each principal component are:

$$n_{x'} = n_o - \frac{1}{2} n_o^3 r_{13} E_z \quad (3.26)$$

$$n_{y'} = n_o - \frac{1}{2} n_o^3 r_{13} E_z \quad (3.27)$$

$$n_{z'} = n_e - \frac{1}{2} n_e^3 r_{33} E_z \quad (3.28)$$

If we had applied the field along a different axis, we would have needed to define new principal axes. This will be left for you as a tutorial task in Tutorial Sheet 1.

#### Example 4: CaTiO<sub>3</sub> Optic Axes

CaTiO<sub>3</sub> is a tri-refringent mineral with  $n_1=2.30$ ,  $n_2=2.34$ , and  $n_3=2.38$ . Using the optical indicatrix, calculate the angle between its two optical axes (hint: you will need to think carefully about the angles of incidence at which a light ray experiences no anisotropy).

## 3.4 Electro-Optic Modulation

The electro-optic effect can be exploited to manipulate the propagation of light within the material. Consider a  $z$ -cut Potassium Dihydrogen Phosphate (KDP) plate with an applied electric field  $E$  parallel to  $z$ . The birefringence of this plate in the  $z$ -propagation direction is:

$$n'_y - n'_x = n_o^3 r_{63} E_z \quad (3.29)$$

If the KDP plate has a finite thickness  $d$ , the **phase retardation** is:

$$\Gamma = \frac{\omega}{c} (n'_y - n'_x) d = \frac{2\pi}{\lambda} n_o^3 r_{63} V \quad (3.30)$$

When the applied voltage yields a phase retardation of  $\Gamma=\pi$ , this is known as the "half-wave voltage," and is given by

$$V_\pi = \frac{\lambda}{2n_o^3 r_{63}} \quad (3.31)$$

where  $\lambda$  is the wavelength of light. This is an import quantity since it denotes the

capacity of an electro-optic material to modulate light. It is therefore a key figure of merit for electro-optic modulators.

### 3.4.1 Amplitude Modulation

Birefringence causes a wave incident at  $z=0$  with polarisation along  $x$  to acquire a polarisation along  $y$  proportionally with the applied voltage. As the component along  $y$  increases, the component along  $x$  decreases until  $V=V_\pi$ , the polarisation becomes parallel to  $y$ . Placing a polariser at the output plane  $z=d$  at  $90^\circ$  to the incident polarisation allows an optical beam to pass freely. On the other hand, when the incident field is not applied, the phase retardation  $\Gamma=0$ , and the output beam is blocked completely by the crossed output polariser. This allows for controllable amplitude modulation of the light.

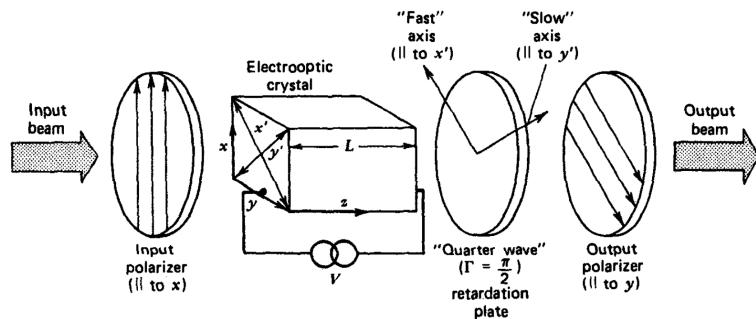


Figure 3.5: Electro-Optic Amplitude Modulation. Figure from [2]

An electro-optic amplitude modulator is typically formed by placing an electro-optic crystal between two crossed polarisers, which are at an angle of  $45^\circ$  to the new principal axes  $x'$  and  $y'$ . An additional birefringent crystal is added to the optical path to create a fixed retardation so that the total retardation  $\Gamma$  is the sum of the retardation due to this crystal and the induced one. The transmission  $T$  through this combination for polarised light is:

$$T = \sin^2 \frac{\Gamma}{2} = \sin^2 \left( \frac{\pi}{2} \frac{V}{V_\pi} \right) \quad (3.32)$$

Typically, the modulator is biased with a fixed retardation  $\Gamma_B=\pi/2$  to 50% transmission by applying a voltage  $V=V\pi/2$ , or by using a naturally birefringent crystal to introduce a  $\pi/2$  phase difference of between the  $x'$  and  $y'$  components. If a small modulating voltage is applied then, the transmitted voltage will modulate sinusoidally also, with a total retardation of:

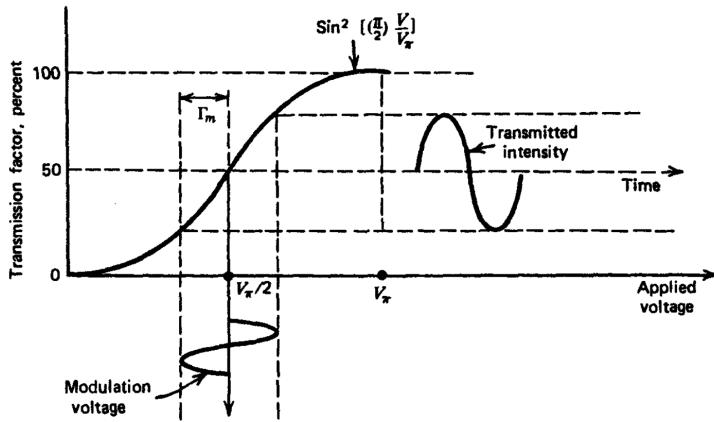


Figure 3.6: Transmission through an amplitude modulator. Figure from [2]

$$\Gamma = \frac{\pi}{2} + \Gamma_m \sin \omega_m t \quad (3.33)$$

....where the modulation retardation  $\Gamma_m$  is related to the amplitude  $V_m$  of the modulation voltage, such that  $\Gamma_m \pi V_m / V_\pi$ . So, the transmission intensity  $I_o$  is:

$$I_o = \frac{I_{in}}{2} (1 + \sin (\Gamma_m \sin \omega_m t)) \quad (3.34)$$

....which for small  $\Gamma_m \ll 1$  is:

$$I_o = \frac{I_{in}}{2} (1 + \Gamma_m \sin \omega_m t) \quad (3.35)$$

### 3.4.2 Phase Modulation

Rather than having equal components along the induced birefringence axes, if the incident beam is polarised parallel to either  $x'$  or  $y'$ , the propagating electric field along  $z$  does not change the state of polarisation, but changes the output phase by

$$\Delta\phi_{x'} = -\frac{\omega d}{c} \Delta n_{x'} \quad (3.36)$$

....where  $d$  is the length  $L$  of the crystal. So, for the case of KDP, this is:

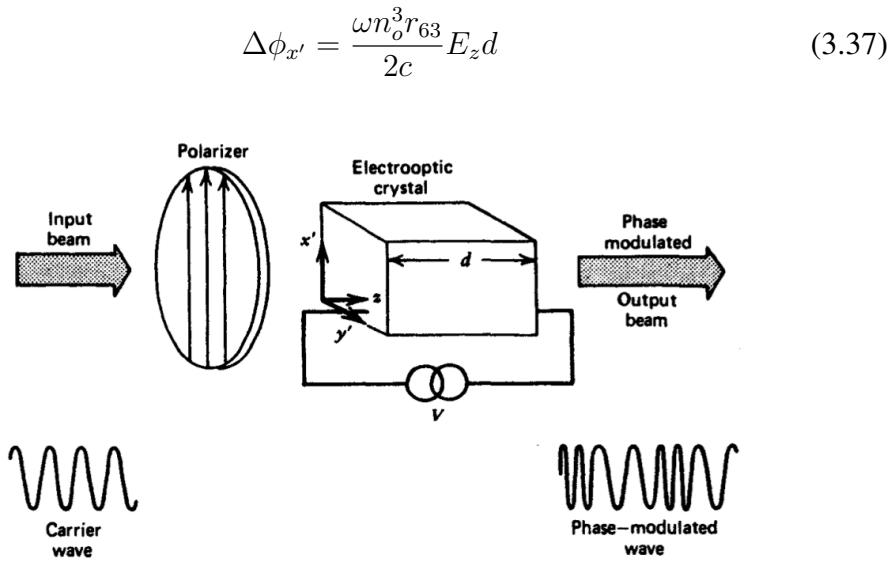


Figure 3.7: Electro-Optic Phase Modulation. Figure from [2]

For a sinusoidal bias field  $E_z = E_m \sin \omega_m t$ , the incident optical field will vary as  $E_{in} = A \cos \omega t$ , producing an output electric field of:

$$E_{out} = A \cos \left[ \omega t - \frac{\omega}{c} \left( n_o - \frac{n_o^3}{2} r_{63} E_m \sin \omega_m t \right) d \right] \quad (3.38)$$

This can be simplified by dropping the constant phase factor such that:

$$E_{out} = A \cos [\omega t + \delta \sin \omega_m t] \quad (3.39)$$

....with the **phase-modulation index**  $\delta$  being:

$$\delta = \frac{\pi n_o^3 r_{63} E_m d}{\lambda} \quad (3.40)$$

### 3.5 Frequency Conversion

Earlier, we described the electro-optic effect in terms of the impermeability tensor. Revisiting here:

$$\eta_{ij}(\vec{E}) = \eta_{ij}(0) + \left( \frac{\partial \eta_{ij}}{\partial E_k} \right)_{E=0} E_k + \frac{1}{2} \left( \frac{\partial^2 \eta_{ij}}{\partial E_k \partial E_l} \right)_{E=0} E_k E_l + \dots \quad (3.41)$$

Recall from lecture 1 that macroscopic polarisation is related to **susceptibility**  $\chi$  as  $P = \epsilon_0 \chi E$ , where  $\chi = \frac{\epsilon}{\epsilon_0} - 1$ . It can be shown (left for you as an exercise!), that in terms of polarisation, the EO effect can be expressed as:

$$\vec{P} = \epsilon_0 \chi_{ij}^{(1)} E_j + \epsilon_0 \chi_{ijk}^{(2)} E_j E_k + \epsilon_0 \chi_{ijkl}^{(3)} E_j E_k E_l + \dots \quad (3.42)$$

The  $\chi^2$  term is the **second-order nonlinearity**, which is the Pockels effect and can be used for **second-order harmonic generation**. The  $\chi^3$  term is the **third-order nonlinearity**, which is the Kerr effect, and can be used for **four-wave mixing**.

This duality often leads to a confusion in the use of the nomenclature. Note the electric field components in Equation 3.42, and how they differ from the impermeability picture. This is because in the impermeability tensor form, the  $E$  field is a static or low-frequency control field that changes the optical properties linearly. In the polarisation expansion though, the material response is expressed in terms of total field, including both optical and static components. It therefore contains more  $E$ -field terms than the native Pockels effect.

The Pockels effect is an intrinsic linear response of the dielectric properties of the medium in response to an electric field, and is therefore called the "Linear Electro-Optic Effect". However, in the overall picture, the Pockels effect is a second order non-linear effect, described by the  $\chi^{(2)}$  coefficient. It is easy to become confused by the nomenclature, and so it is important to clarify what picture is being employed when describing optical non-linearity!

### 3.5.1 Sum-Difference Frequency Mixing

When one field is optical and the other static, the first order terms of Equation 3.42 reduces to the Pockels effect. However, when both fields are optical, we get frequency mixing.

Returning to Maxwell's wave equation for the electric field, we can express it in terms of the nonlinear polarisation  $\mathbf{P}_{NL}$ :

$$\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = \frac{1}{\epsilon_0 c^2} \frac{\partial^2 \mathbf{P}_{NL}}{\partial t^2} \quad (3.43)$$

The nonlinear polarisation acts as a **source** term in the wave equation. Whatever frequency components  $\mathbf{P}_{\text{NL}}$  contains will therefore act as sources that drive new electromagnetic fields at those frequencies.

So if  $\mathbf{P}_{\text{NL}}$  oscillates not only at the original input frequency but also at combinations of input frequencies, then the medium will radiate new waves at those frequencies—this is the physical basis of frequency conversion in nonlinear optics.

To see how this happens, consider the second-order nonlinear response of the material, described by

$$P_i^{(2)}(t) = \varepsilon_0 \sum_{j,k} \chi_{ijk}^{(2)} E_j(t) E_k(t) \quad (3.44)$$

Now suppose that the electric field itself is composed of several monochromatic components:

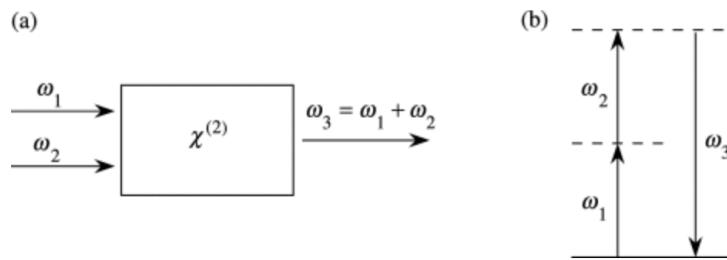
$$\mathbf{E}(t) = \sum_a \mathbf{E}_a e^{-i\omega_a t} + \text{c.c.} \quad (3.45)$$

Substituting into  $P^{(2)}(t)$ , the products of the type  $E_j(t)E_k(t)$  naturally generate **new frequency components**. Specifically, the product of two exponential terms produces oscillations at both the **sum and difference** of their frequencies:

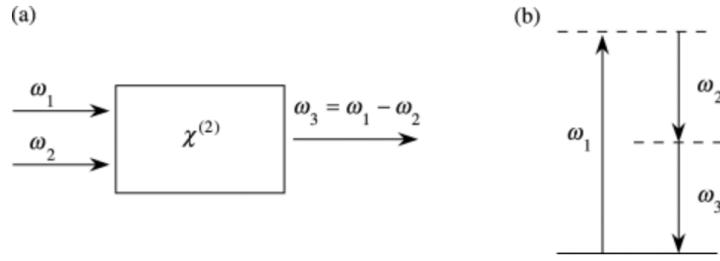
$$E_j(\omega_1)e^{-i\omega_1 t} \times E_k(\omega_2)e^{-i\omega_2 t} = E_j(\omega_1)E_k(\omega_2)e^{-i(\omega_1+\omega_2)t} \quad (3.46)$$

This means that the nonlinear polarisation acquires a component at frequency  $\omega_1 + \omega_2$ , given by

$$P_i^{(2)}(\omega_1 + \omega_2) \propto \chi_{ijk}^{(2)} E_j(\omega_1) E_k(\omega_2) \quad (3.47)$$



In the same way, if we take the complex conjugate combination, we obtain components oscillating at  $\omega_1 - \omega_2$ . These new frequency terms — sums and differences of the original ones — give rise to second-harmonic generation when  $\omega_1 = \omega_2$ , sum-frequency generation  $\omega_3 = \omega_1 + \omega_2$ , and difference-frequency generation  $\omega_3 = \omega_1 - \omega_2$ .



### 3.5.2 Second Harmonic Generation

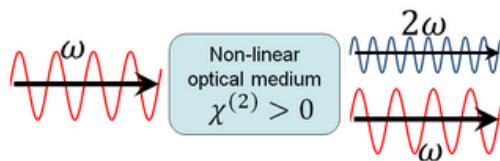
If only one optical field is present and the medium has a second order nonlinearity, the incident field

$$E(t) = E_\omega e^{-i\omega t} + E_\omega^* e^{-i\omega t} \quad (3.48)$$

expands as

$$E(t)E(t) = (E_\omega e^{-i\omega t} + E_\omega^* e^{i\omega t})^2 = E_\omega^2 e^{-2i\omega t} + (E_\omega^*)^2 e^{2i\omega t} + 2|E_\omega|^2 \quad (3.49)$$

The second harmonic term  $E_\omega^2 e^{-2i\omega t}$  and it's c.c. oscillate at twice the original frequency!



## 3.6 Magneto-Optics

There are some things that we need to be able to do with light that electro-optics find difficult, one of which is rectification (passing light in one direction but not

the other). Fortunately there is a solution, in the form of the linear magneto-optic effect or Faraday effect.

When passing through a magneto-optic material parallel or anti-parallel with an applied magnetic field  $\vec{B}$ , the left and right hand polarisations of the light experience slightly different refractive indices. The result is that linearly polarised light is rotated. The sense and magnitude of the rotation is given by:

$$\theta = V \vec{B} \cdot \vec{L} \quad (3.50)$$

Where  $V$  is the Verdet constant (material parameter), and  $L$  is the vector that describes the path of the light through the material (both length and direction). What is unique about the magneto-optic effect is that counter-propagating rays are rotated in opposite senses. By placing a  $45^\circ$  rotator between two polarisers misaligned by  $45^\circ$ , we construct the required rectifier.

## References

1. *Electro-optic Materials on Silicon* <https://research.ibm.com/projects/electro-optic-materials-on-silicon>. Accessed: 2024-12-28.
2. Yariv, A. & Yeh, P. *Optical Waves in Crystals* (Wiley New York, 1984).

# 4 | Metamaterials

## 4.1 Engineering Negative Index Materials

*Metamaterials* are artificially structured media whose electromagnetic properties are governed by sub-wavelength geometric features rather than intrinsic atomic composition. In essence, they allow one to engineer the effective permittivity ( $\epsilon$ ) and permeability ( $\mu$ ) of space itself, thereby controlling how waves propagate through it. Although the concept applies to any kind of wave—acoustic, elastic, or electromagnetic—the term most commonly refers to optical and microwave implementations.

### 4.1.1 Negative Index Materials

What does it mean for the optical response of a material to be 'engineered'? Recall the wave-equation from Lecture 1, and the relationship:

$$\tilde{n} = \sqrt{\epsilon_r \mu_r} = n + i\kappa \quad (4.1)$$

The normal case of polarisation of dielectric materials has that  $n > 1$ , and in the weak absorption limit,  $\kappa \ll 1$  (the weak absorption limit). However, it is not a fixed law of nature that these cases must be met. For example, the Drude model of the polarisation of a free electron gas in a metal, assuming weak damping, gives a dielectric function  $\epsilon_r = 1 - \frac{\omega_p^2}{\omega^2}$ , where  $\omega_p$  is the plasma frequency of the metal. So for any electric field which oscillates at a frequency less than  $\omega_p$ ,  $\epsilon_r$  is negative. This means that the refractive index is purely imaginary, and the metal surface is purely reflective. Further, both  $\mu_r$  and  $\epsilon_r$  can be negative. In this case, their product is again positive, and the refractive index is therefore real again, but **negative**. This can have profound technological implications.

When light interacts with a negative refractive index material, the velocity of propagation is still affected, but not in the same way as for a positive index material. The phase velocity  $\nu_p$  is the velocity of the wave fronts, and is given by:

$$\nu_p = \frac{\omega}{k} \quad (4.2)$$

The group velocity  $\nu_g$  is the velocity of the packet itself (i.e. the photons):

$$\nu_g = \frac{d\omega}{dk} \quad (4.3)$$

When light enters a negative index material at normal incidence the phase velocity is reversed, as  $k = nk_0$  becomes negative. However the group velocity remains positive, so the photons keep going forward. Thus, the wave's phase fronts move in the opposite direction to the net flow of energy.

This may seem somewhat confusing physically. But recall that the phase velocity describes how a surface of constant phase (like a crest) moves through space. It points in the direction of the wavevector  $\mathbf{k}$ . On the other hand, the group velocity describes the velocity at which the envelope of the wave packet (and hence the energy and information) propagates. It points in the direction of energy transport, typically that of the Poynting vector  $\mathbf{S}=\mathbf{E} \times \mathbf{H}$ .

In ordinary materials where  $\epsilon > 0$  and  $\mu > 0$ , the electric field  $\mathbf{E}$ , magnetic field  $\mathbf{H}$ , and wavevector  $\mathbf{k}$  form a *right-handed triad*. The Poynting vector and  $\mathbf{k}$  both point in the same direction, so  $v_p$  and  $v_g$  are both positive (in the same direction). In a negative-index medium where  $\epsilon < 0$  and  $\mu < 0$ , something remarkable happens. Maxwell's equations still allow plane-wave solutions of the form

$$\mathbf{E}, \mathbf{H} \propto e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} \quad (4.4)$$

but the relationship between  $\mathbf{E}$ ,  $\mathbf{H}$ , and  $\mathbf{k}$  changes. The impedance is  $\eta = \sqrt{\mu/\epsilon}$ , which remains positive if both  $\epsilon$  and  $\mu$  are negative. However, the direction of  $\mathbf{H}$  relative to  $\mathbf{E}$  flips because the sign of  $\mu$  changes.

Consequently, the Poynting vector  $\mathbf{S}=\mathbf{E} \times \mathbf{H}$  points **opposite** to  $\mathbf{k}$ . This means:

$$\mathbf{S} \cdot \mathbf{k} < 0 \quad (4.5)$$

Energy flows in one direction, but the phase of the wave (the motion of crests) moves in the opposite direction. Thus, the phase velocity is \*negative\*, while the group velocity, which follows the direction of energy transport, remains positive.

This can be viewed as follows: imagine a wave whose crests appear to move toward the source, even as the energy carried by the wave moves away from it. If we were to track individual crests, they'd seem to run "backwards," but the wave packet as a whole (the pulse of light or energy) still moves forward.

#### 4.1.1.1 Engineering Negative $\epsilon_r$

As we have seen, negative  $\epsilon_r$  is achieved simply with free electrons in metals. Of course a bulk metal on its own wouldn't allow us to engineer  $\mu$ , so we need to create the same 'free electron effect' but with some space for additional components.

The way this is typically achieved is with a 3D grid of wires that allow free flow of electrons along them.

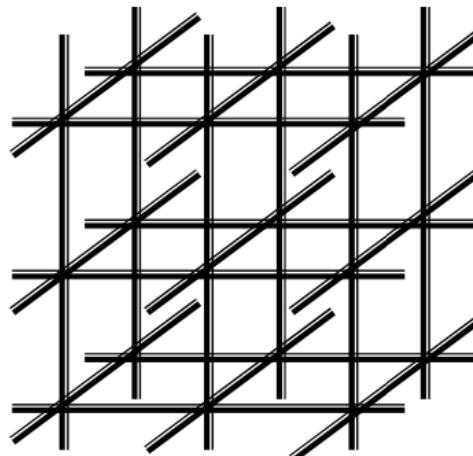


Figure 4.1: 3D wire grid for negative  $\epsilon_r$

#### 4.1.2 Engineering Negative $\mu_r$

Achieving negative  $\mu_r$  is somewhat more challenging, as all magnetic transitions in materials are at frequencies much lower than those of optical waves. The key to achieving negative  $\mu_r$  is to create an artificial resonance slightly lower in frequency than that of the electromagnetic field to be used.

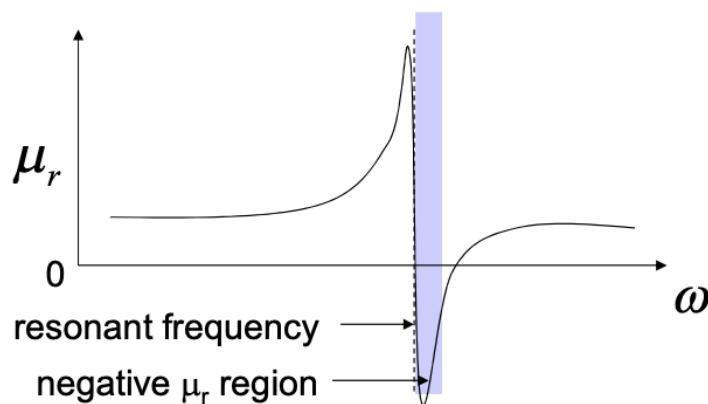


Figure 4.2: Engineering a material resonant frequency response.

To make a resonant electrical circuit in which the impedance drops close to zero

we need a capacitor  $C$  and an inductor  $L$  connected in series whereby (neglecting the resistive component which can be made small)

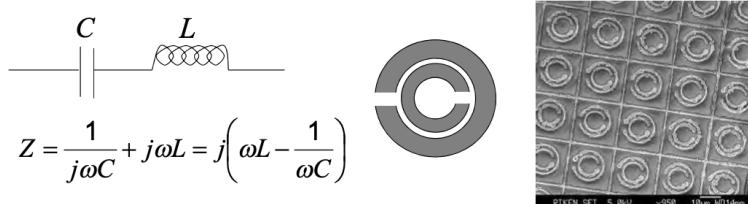


Figure 4.3: Equivalent circuit and SEM image of resonator structures.

and the resonance therefore occurs at  $\omega_{res}=1/\sqrt{LC}$ . A convenient design that achieves this is the split ring resonator, where the ring provides  $L$  and the split provides  $C$ . Making a double ring structure with splits opposed avoids unwanted chirality in the material. An array of these structures combined with the wire grid can thus form the basis for a negative refractive index material (above right).

## 4.2 Transformation Optics

In electromagnetism, the propagation of light depends on how the permittivity and permeability of a medium vary in space. When these quantities are constant, light moves in straight lines, as it does in vacuum. This is the **geodisic** of flat space. However, when they vary, the paths of light bend, slow down, or accelerate in ways that resemble motion through a geometrically distorted space. This observation suggests that it is natural to describe optical propagation not merely in terms of refractive indices and boundary conditions, but in terms of geometry itself. This is a core insight from Einstein's **General Relativity**.

The analytic treatment of a negative index material is not trivial, since such materials possess a reversed spatial orientation. Therefore, to properly describe their properties, a coordinate transformation is required. In 2006, Pendry and Leonhardt independently introduced a powerful mathematical framework—**transformation optics**—for describing electromagnetic fields in media with spatially varying  $\epsilon$  and  $\mu$  [1].

The key insight of this mathematical framework is that:

**Maxwell's equations are form-invariant under coordinate transformations.**

That is, their structure remains identical if space is stretched, compressed, or twisted, provided the material parameters are adjusted accordingly. This invariance means that changing the geometry of space can be perfectly mimicked by

changing the electromagnetic properties of a medium. Instead of actually bending space, one can design a spatially varying, anisotropic material that makes light behave exactly as if the geometry had been deformed. The mathematics of this equivalence is contained in the Jacobian matrix of the transformation, which relates small elements of virtual, flat space to their distorted counterparts in the physical medium. The determinant of the Jacobian captures how volumes are scaled, while its matrix form captures local rotations and stretches. From this information one computes the new permittivity and permeability tensors that ensure Maxwell's equations take the same form in the transformed coordinates.

Treating space as if it were curved is therefore not an aesthetic choice but a practical one. It provides a coordinate-independent and geometrically intuitive way to design how light should move. In ordinary optical design, one might try to infer the refractive index distribution that bends light along a desired path, but this approach becomes intractable for complex trajectories. By contrast, in the geometric view one specifies the desired mapping of space itself—the distortion that would cause light rays to follow the chosen routes—and the transformation equations automatically yield the required material tensors.

Unlike in General Relativity, in transformation optics, no real curvature exists; instead, material anisotropy and inhomogeneity produce an equivalent optical metric that light interprets as curvature. The medium therefore acts as an artificial geometry for electromagnetic fields. This equivalence allows the design of devices that bend, focus, or hide light with precision: e.g. invisibility cloaks correspond to removing a region from space, concentrators to compressing it, and waveguide bends to smoothly rotating it.

### 4.2.1 Mathematical Formulation

Recall Maxwell's curl equations in source-free space:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad \nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t}. \quad (4.6)$$

Imagine electromagnetic fields in some “virtual space” (a simple, empty region of free space). To mimic the distortion of space, a coordinate transformation (for example, stretching, compressing, or bending it) is required. This gives a new ‘physical space’ which possesses a different geometry. Pendry’s deep insight was that you can make light behave as if it were in that distorted geometry by engineering a material with specific permittivity  $\epsilon$  and permeability  $\mu$  tensors. So, rather than actually curving space, you curve the optical metric — effectively making light follow new paths.

If we define physical space coordinates  $x$  and virtual space coordinates  $x'$ , the coordinate transformation mapping the two is:

$$x'^i = f^i(x^1, x^2, x^3), \quad (4.7)$$

The Jacobian is a matrix that describes how a coordinate transformation locally stretches, rotates, or shears space. The Jacobian  $\Lambda$  describing the transformation of coordinates from the virtual/original space  $x'=(x'_1, x'_2, x'_3)$  to physical/transformed space  $x=(x_1, x_2, x_3)$  is:

$$\Lambda^i_j = \frac{\partial x'^i}{\partial x^j}. \quad (4.8)$$

Each entry reveals how one physical coordinate changes with respect to one virtual coordinate. So the Jacobian expresses how small elements of space are deformed by the transformation.

Under this transformation, the constitutive tensors in the new coordinates are related by

$$\varepsilon' = \frac{\Lambda \varepsilon \Lambda^T}{\det(\Lambda)}, \quad \mu' = \frac{\Lambda \mu \Lambda^T}{\det(\Lambda)}. \quad (4.9)$$

In vacuum ( $\varepsilon = \mu = I$ ),

$$\varepsilon' = \mu' = \frac{\Lambda \Lambda^T}{\det(\Lambda)}. \quad (4.10)$$

Thus, the medium with  $\varepsilon'$  and  $\mu'$  reproduces exactly the electromagnetic behavior that light would experience if space itself were distorted according to the mapping  $f$ .

### 4.3 Invisibility Cloaks and Negative-Index Metamaterials

Pendry, Schurig, and Smith demonstrated that transformation optics could be used to design materials that *guide light around an object*, effectively rendering it invisible [1]. If the wave trajectories are smoothly bent around a region and recombine without scattering, the object within that region becomes electromagnetically hidden.

### 4.3.1 The Cloaking Transformation

The canonical example is a spherical cloak that maps a region  $r \in [0, R_2]$  into a shell  $r' \in [R_1, R_2]$ :

$$r' = R_1 + \frac{(R_2 - R_1)}{R_2} r, \quad \theta' = \theta, \quad \phi' = \phi. \quad (4.11)$$

Here, the point  $r = 0$  is mapped to  $r' = R_1$ , so all inner space  $r < R_1$  is compressed into a shell. Light rays then flow smoothly around this inner volume. The physical interpretation of this is that the transformation effectively “bends” space so that light trajectories circumvent the inner region. Maxwell’s equations remain invariant, since the apparent bending of rays is purely due to the engineered material tensors. Since rays never penetrate the core, any object within is electromagnetically isolated.

The resulting material parameters (in Spherical Coordinates) are:

$$\varepsilon_r = \mu_r = \frac{r' - R_1}{r'}, \quad \varepsilon_\theta = \varepsilon_\phi = \mu_\theta = \mu_\phi = \frac{r'}{r' - R_1}. \quad (4.12)$$

Note that  $\varepsilon_\theta, \mu_\theta \rightarrow \infty$  as  $r' \rightarrow R_1$ . Theoretically, this requires infinite phase velocity at the inner boundary. This is impossible in practice, so perfect cloaking is only achievable in an idealised limit.

#### Example:

For a cylindrical cloak with  $R_1 = 1$  cm,  $R_2 = 2$  cm, and  $r' = 1.5$  cm:

$$\varepsilon_r = \frac{1.5 - 1}{1.5} = 0.33, \quad \varepsilon_\theta = \frac{1.5}{0.5} = 3. \quad (4.13)$$

Thus, the cloak must be anisotropic, with spatially varying parameters.

### 4.3.2 Implementation with Metamaterials

In practice, each spatial point in the cloak must realise the required anisotropic  $\varepsilon$  and  $\mu$ . This is achieved through sub-wavelength “unit cells” such as split-ring resonators and metallic wires, whose collective response forms an effective medium. By spatially grading the geometry, one approximates the desired transformation.

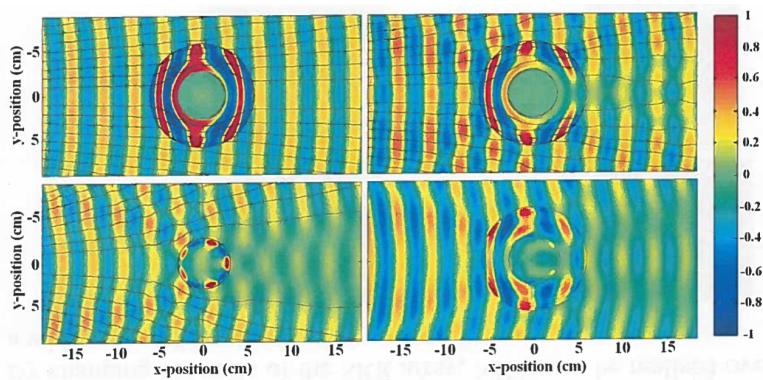


Figure 4.4: Snapshots of time-dependent, steady-state electric field patterns, with stream lines [black lines in (A to C)] indicating the direction of power flow (i.e., the Poynting vector). The cloak lies in the annular region between the black circles and surrounds a conducting Cu cylinder at the inner radius. The fields shown are (A) the simulation of the cloak with the exact material properties, (B) the simulation of the cloak with the reduced material properties, (C) the experimental measurement of the bare conducting cylinder, and (D) the experimental measurement of the cloaked conducting cylinder. Animations of the simulations and the measurements (movies S1 to S5) show details of the field propagation characteristics within the cloak that cannot be inferred from these static frames. The right-hand scale indicates the instantaneous value of the field.

### 4.3.3 State-of-the-Art and Challenges

The first successful demonstrations of the microwave cloak were given in 2006 [2]. Using dielectric metasurfaces [3] and plasmonic [4] approaches, this was pushed to IR/visible. More recently, there has been activity to conceal objects under reflective “bumps”, operational in IR and visible ranges [5].

Although the creation of such an invisibility cloak would be an extraordinary achievement, in practice this is extremely difficult. The main limiting factor is that the negative-index response is inherently dispersive and **narrowband**. Whilst it is indeed possible to cloak a small band of frequencies, consider, for example, the bandwidth of the visible spectrum - clearly the current materials could not achieve true invisibility.

## 4.4 The Perfect Lens

One of the most striking implications of negative-index media is the possibility of creating a **perfect lens**—a flat slab capable of imaging beyond the **diffraction limit** [6].

### 4.4.1 The Conventional Diffraction Limit

In ordinary optics, a lens focuses light by refraction: rays are bent at curved surfaces and converge to form an image. The sharpness of that image, however, is limited by diffraction. Only the spatial frequencies corresponding to propagating waves — those with transverse wavenumber  $|k_x| < k_0 = \omega/c$  — can reach the image plane, because higher spatial frequencies correspond to evanescent waves that decay exponentially with distance. The fine sub-wavelength details of an object reside in these evanescent components, which vanish before a conventional lens can recover them. The result is the familiar diffraction limit, where no lens can resolve features much smaller than about half the wavelength.

To see this more generally, any electric field  $E(x)$  can be written as a superposition of plane waves:

$$E(x) = \int_{-\infty}^{\infty} \tilde{E}(k_x) e^{ik_x x} e^{ik_z z} dk_x \quad (4.14)$$

Here, each  $k_x$  is a **spatial frequency** component, and  $\tilde{E}(k_x)$  is the Fourier amplitude associated with that spatial frequency:

$$\tilde{E}(k_x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} E(x) e^{-ik_x x} dx \quad (4.15)$$

Each component of the superposition satisfies the dispersion relation:

$$k_z = \sqrt{k_0^2 - k_x^2}, \quad k_0 = \frac{\omega}{c} = \frac{2\pi}{\lambda} \quad (4.16)$$

For each  $k_x$  the longitudinal component  $k_z$  determines whether the ray propagates or decays:

$$k_z = \begin{cases} \sqrt{k_0^2 - k_x^2}, & |k_x| \leq k_0 \quad (\text{propagating waves}) \\ i\sqrt{k_0^2 - k_x^2}, & |k_x| > k_0 \quad (\text{evanescent waves}) \end{cases} \quad (4.17)$$

When  $|k_x| \leq k_0$ ,  $k_z$  is real and the component carries energy to the far field. On the other hand, when  $|k_x| > k_0$ ,  $k_z$  becomes imaginary and the corresponding spatial frequency decays as  $e^{-\kappa z}$ , with  $\kappa = \sqrt{k_x^2 - k_y^2}$ . So only components inside the **light cone**  $|k_x| \leq k_0$  can propagate through free space.

This selectivity acts as a spatial bandwidth limit. The highest spatial frequency which can reach an image plane is  $k_x^{max} = k_0$ , and correspondingly the smallest resolvable real-space feature is approximately:

$$\Delta x_{min} = \approx \frac{2\pi}{k_x^{max}} = \frac{\lambda}{2} \quad (4.18)$$

This is the **diffraction limit**. Traditionally this has been seen as the fundamental resolution limit for an optical imaging system. However, Pendry showed how this limit can be surpassed with a metamaterial 'perfect lens'.

#### 4.4.2 Pendry's Perfect Lens

Pendry's concept of the perfect lens arises from the fact that in a negative-index material, rays refract on the same side of the normal, and the medium effectively inverts optical space: phase advances in reverse, as though light were moving backward. This inversion suggests that a flat slab of negative-index material could undo the normal spreading of waves in free space, restoring both propagating and evanescent components of the field.

Suppose a homogeneous slab of thickness  $d$  occupies  $0 \leq z \leq d$ , with an object

located at  $z = -a$  and an image plane at  $z = d + b$ . Consider a single transverse spatial frequency component with amplitude  $A(k_x)$ . In free space, its longitudinal wavenumber is  $k_{z0} = \sqrt{k_0^2 - k_x^2}$  for propagating waves and  $k_{z0} = i\alpha$ ,  $\alpha = \sqrt{k_x^2 - k_0^2}$ , for evanescent waves. Inside the slab, the longitudinal wavenumber is

$$k_{zs} = \sqrt{(nk_0)^2 - k_x^2} \quad (4.19)$$

The multiplicative factor accumulated by this component from the object to the image plane is

$$T(k_x) = e^{ik_{z0}a}, e^{ik_{zs}d}, e^{ik_{z0}b} \quad (4.20)$$

For perfect imaging, every spatial frequency must be transmitted with the same amplitude and phase, so  $T(k_x)$  should be independent of  $k_x$  up to a global constant. This occurs if  $k_{zs} = -k_{z0}$  for all  $k_x$  and if  $a + b = d$ . The first condition ensures that the phase accumulated inside the slab exactly cancels the phase in free space, while the second geometrically aligns the planes. Under these conditions,  $T(k_x) = 1$  for all  $k_x$ , meaning every spatial frequency, including evanescent components, is perfectly reproduced.

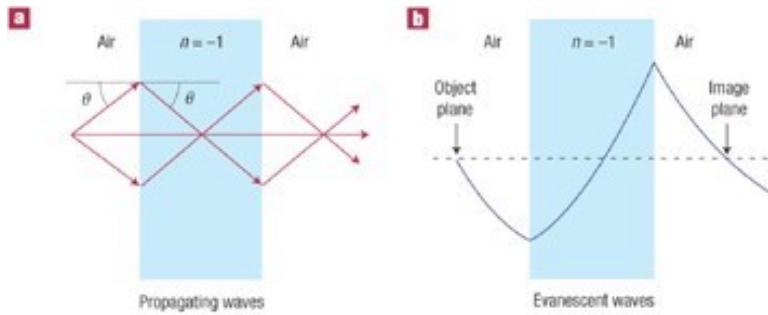


Figure 4.5: Figure from [7]

Choosing  $n = -1$  realises  $k_{zs} = -k_{z0}$  automatically when the correct branch of the square root is taken. For propagating waves, this corresponds to reversed phase velocity inside the slab. For evanescent waves, with  $k_{z0} = i\alpha$ , the correct sign produces exponential growth  $\exp(+\alpha z)$  within the slab, exactly compensating the decay  $\exp(-\alpha a)$  before the slab and  $\exp(-\alpha b)$  after. The total factor for an evanescent component is

$$\exp[-\alpha(a + b - d)] \quad (4.21)$$

which equals unity when  $d=a+b$ . To avoid reflection, the intrinsic impedance  $\eta=\sqrt{\mu/\varepsilon}$  of the slab must match free space. Setting  $\varepsilon_r=\mu_r=-1$  satisfies this automatically, because the ratio remains unity even though each parameter is negative.

Physically, the slab does not act like a conventional lens. Instead, it reproduces the electromagnetic field of the object on the opposite side. Propagating components are refocused via phase reversal, while **evanescent components are amplified, exactly reversing their decay**. The **resulting image is, in principle, an exact replica of the object, containing every spatial frequency**. Importantly, when  $\varepsilon=\mu=-1$ , the lens reproduces the full electromagnetic field, achieving **sub-wavelength resolution**.

The negative-index slab “folds” space back on itself, acting as an optical mirror in which distance and phase are inverted. Geometrically, this can be interpreted through transformation optics. A slab with  $n=-1$  corresponds to a space folding transformation

$$x' = x, \quad y' = y, \quad z' = \begin{cases} z, & z < 0, \\ -z, & 0 < z < d, \\ z - 2d, & z > d. \end{cases} \quad (4.22)$$

....with a negative Jacobian determinant, representing inversion along the propagation axis. In this picture, the perfect lens locally reverses optical space. The condition  $d=a+b$  ensures that the optical path through the inverted region is null for all components, producing perfect imaging. The  $n=-1$  slab thus performs a mirror-like folding of space, mapping a point at  $z=-a$  to an image at  $z=2d-a$ .

#### 4.4.3 State-of-the-Art and Challenges

In reality, achieving  $\varepsilon_r=\mu_r=-1$  without loss is impossible. Real metamaterials exhibit dispersion and absorption, which limit the amplification of evanescent waves and narrow the operational bandwidth. Despite these practical constraints, experiments with plasmonic films and metamaterial slabs have demonstrated sub-diffraction imaging consistent with Pendry’s predictions.

Nevertheless, experimental *superlenses* made from thin silver films ( $\varepsilon \approx -1$ ) have demonstrated near-field sub-diffraction imaging.

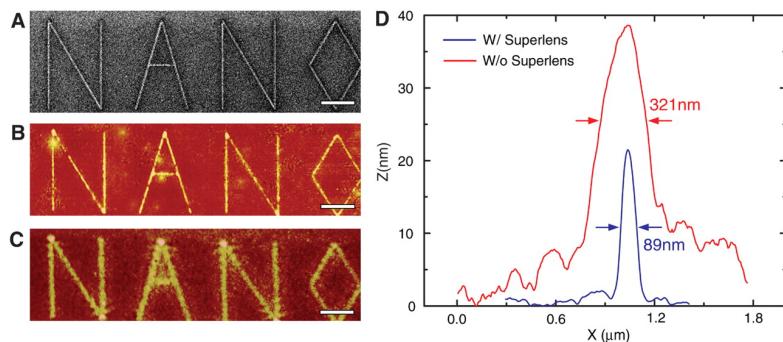


Figure 4.6: An arbitrary object “NANO” was imaged by silver superlens. (A) FIB image of the object. The linewidth of the “NANO” object was 40 nm. Scale bar in (A) to (C), 2  $\mu\text{m}$ . (B) AFM of the developed image on photoresist with a silver superlens. (C) AFM of the developed image on photoresist when the 35-nm-thick layer of silver was replaced by PMMA spacer as a control experiment. (D) The averaged cross section of letter “A” shows an exposed line width of 89 nm (blue line), whereas in the control experiment, we measured a diffraction-limited full width at half-maximum line width of  $321 \pm 10$  nm (red line). Figure from [8]

## 4.5 Photonic Crystals

Photonic crystals can be viewed as a subset of metamaterials, since they modify refractive index via periodic structuring. A photonic crystal is simply a material in which some spatial periodicity exists in the refractive index that is comparable to the wavelength of light, and in which the resulting interference effects modify the optical dispersion (ie, the  $\omega - k$  relationship). In fact we have seen one photonic crystal already - the laser mirror in lecture 1. There, periodic modulation of  $n$  led to a highly reflective mirror, which is effectively a 1- dimensional photonic crystal. The laser mirror can not support the propagation of light normal to the plane within the structure, if the wavelength of that light falls within the stop band of the mirror. Another name for the stop band is a photonic band gap.

### 4.5.1 Electronic Band Gaps

Let’s remind ourselves of electronic band-gaps in solids, since there is a close analogy here to photonic crystals. In solid state physics, Schrödinger equation describes the quantum states of electrons moving in the **periodic potential** created by a crystal lattice of atoms. It’s the mathematical starting point for understanding **electronic band structures** — the energy bands and band gaps in solids that determine whether a material is a metal, semiconductor, or insulator.

For a single electron in a periodic potential  $V(\mathbf{r})$ , the time-independent Schrödinger equation is:

$$\hat{H}\psi(\mathbf{r}) = E\psi(\mathbf{r}) \quad (4.23)$$

where the Hamiltonian operator is

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r}) \quad (4.24)$$

Here,  $\psi(\mathbf{r})$  is the electron's wavefunction,  $E$  is its energy eigenvalue,  $m$  is the electron's mass,  $\hbar$  is Planck's constant divided by  $2\pi$ , and  $V(\mathbf{r})$  is a spatially periodic potential due to the lattice of ions,  $V(\mathbf{r} + \mathbf{R}) = V(\mathbf{r})$  for any lattice vector  $\mathbf{R}$ .

Because the potential is periodic, the solutions take the **Bloch form**:

$$\psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}u_{n\mathbf{k}}(\mathbf{r}) \quad (4.25)$$

where  $u_{n\mathbf{k}}(\mathbf{r})$  has the same periodicity as the lattice. This reduces the problem to solving for  $u_{n\mathbf{k}}$  within a single unit cell, subject to Bloch boundary conditions. The resulting eigenvalues  $E_n(\mathbf{k})$  form **bands** in the  $\mathbf{k}$ -space (momentum space). Each band corresponds to a continuum of allowed energies for an electron whose crystal momentum is  $\hbar\mathbf{k}$ . The **band gaps** (regions of forbidden energy) appear because of **Bragg reflection** of electron waves at the Brillouin zone boundaries.

### 4.5.2 Photonic Band Gaps

We will show here that photonic crystals have very much the same behaviour. To see this, we need to derive the vector eigenvalue problem for electromagnetic fields. In a source-free, lossless, and linear medium with spatially varying permittivity  $\varepsilon(\mathbf{r})$  and permeability  $\mu(\mathbf{r})$ , Maxwell's equations for time-harmonic fields  $\mathbf{E}(\mathbf{r}, t) = \mathbf{E}(\mathbf{r})e^{-i\omega t}$  and  $\mathbf{H}(\mathbf{r}, t) = \mathbf{H}(\mathbf{r})e^{-i\omega t}$  can be written as

$$\nabla \times \mathbf{H} = -i\omega\varepsilon(\mathbf{r})\mathbf{E}, \quad (4.26)$$

$$\nabla \times \mathbf{E} = i\omega\mu(\mathbf{r})\mathbf{H}. \quad (4.27)$$

Eliminating  $\mathbf{H}$  gives the master equation for the electric field:

$$\nabla \times \left( \frac{1}{\mu(\mathbf{r})} \nabla \times \mathbf{E} \right) = \omega^2 \varepsilon(\mathbf{r}) \mathbf{E}. \quad (4.28)$$

Equivalently, by introducing the speed of light  $c = 1/\sqrt{\varepsilon_0 \mu_0}$ ,

$$\nabla \times \left( \frac{1}{\mu(\mathbf{r})} \nabla \times \mathbf{E} \right) = \left( \frac{\omega}{c} \right)^2 \varepsilon(\mathbf{r}) \mathbf{E}. \quad (4.29)$$

Equation (4.29) is a *vector eigenvalue problem* for the operator

$$\mathcal{L} = \nabla \times \frac{1}{\mu(\mathbf{r})} \nabla \times \quad (4.30)$$

with eigenvalue  $(\omega/c)^2$  and eigenfunction  $\mathbf{E}(\mathbf{r})$ .  $\mathcal{L}$  describes how the spatial variations of the electric field interact with the material's permeability distribution to produce the restoring field that determines the allowed electromagnetic modes, and it is the photonic analog of the Schrödinger equation in solid-state physics.

#### 4.5.2.1 Periodicity and Bloch's Theorem

If the dielectric function  $\varepsilon(\mathbf{r})$  and permeability  $\mu(\mathbf{r})$  are periodic in space with lattice vectors  $\mathbf{a}_i$ , i.e.

$$\varepsilon(\mathbf{r} + \mathbf{a}_i) = \varepsilon(\mathbf{r}), \quad \mu(\mathbf{r} + \mathbf{a}_i) = \mu(\mathbf{r}) \quad (4.31)$$

then by Bloch's theorem the eigenmodes of Eq. (4.29) can be written as

$$\mathbf{E}_{n\mathbf{k}}(\mathbf{r}) = \mathbf{u}_{n\mathbf{k}}(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}}, \quad \mathbf{u}_{n\mathbf{k}}(\mathbf{r} + \mathbf{a}_i) = \mathbf{u}_{n\mathbf{k}}(\mathbf{r}), \quad (4.32)$$

where  $\mathbf{u}_{n\mathbf{k}}(\mathbf{r})$  is periodic in the lattice and  $\mathbf{k}$  lies within the first Brillouin zone.

Substituting this form into Eq. (4.29) yields an eigenvalue problem within one unit cell, which can be solved numerically (e.g., using plane-wave expansion or finite-element methods). The resulting discrete eigenfrequencies  $\omega_n(\mathbf{k})$  form the **photonic band structure**. Frequency intervals with no real eigenmodes are called **photonic band gaps**.

#### 4.5.2.2 One-Dimensional Reduction

In a one-dimensional photonic crystal, the material parameters vary only along one axis, say  $\varepsilon(\mathbf{r}) = \varepsilon(x)$ , and the problem can be reduced to a scalar form for a single field component, depending on polarization.

For a transverse electric (TE) mode, with  $\mathbf{E} = E_z(x)\hat{\mathbf{z}}$ , the eigenvalue problem simplifies to

$$\frac{d^2 E_z}{dx^2} + \left(\frac{\omega}{c}\right)^2 \varepsilon(x) E_z = 0, \quad (4.33)$$

assuming  $\mu$  is constant.

Equation (4.33) is a scalar Helmholtz equation with a periodic coefficient  $\varepsilon(x)$ . Because  $\varepsilon(x+a) = \varepsilon(x)$ , Bloch's theorem still applies:

$$E_z(x+a) = E_z(x)e^{iK_a}, \quad (4.34)$$

where  $K$  is the Bloch wavevector.

#### 4.5.3 Band Structures

So, just as we have an electronic band structure, we also have a photonic band structure. In the nearly free electron model, the translational symmetry of the crystal lattice defines Brillouin zones, and when the interaction between the electrons and the lattice is 'switched on', band gaps are formed (see figure):

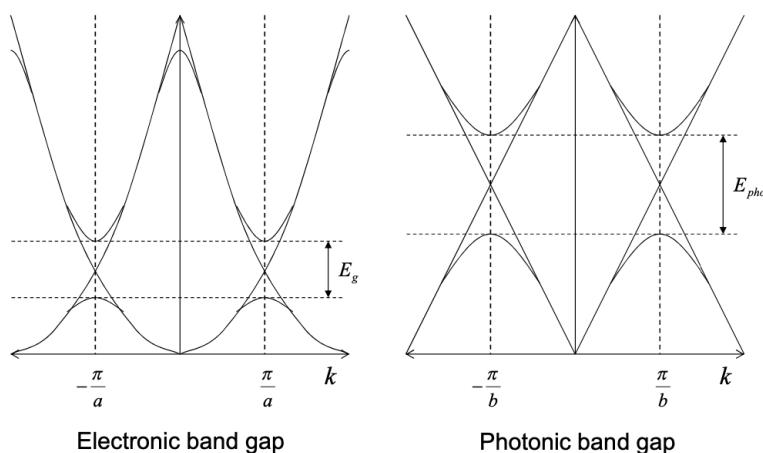


Figure 4.7: Analogy between electronic and photonic band gaps

It is exactly the same with photonic crystals. A periodicity  $b$  is defined, creating Brillouin zones, and the modulation of the refractive index (analogous to the modulation in the electronic potential due to the atomic lattice) causes band gaps to appear. The main difference between the two diagrams above is in the 'free particle' dispersions, where

$$E_{electron} = \frac{\hbar^2 k^2}{2m}; \quad E_{photon=\hbar ck} \quad (4.35)$$

Note also that  $a \sim 0.1$  nm whereas  $b \sim 100$  nm, so the Brillouin zones in the photonic picture are about  $10^3$  times smaller in momentum space than those in the electronic picture.

These ideas also work in 2 dimensions and in 3 dimensions. 3D photonic crystals are generally difficult to fabricate, so most applications are in 2D. The two most commonly encountered are photonic crystal fibres and photonic crystal cavities.

#### 4.5.4 Photonic Crystal Fibres

These were invented by Dr Philip Russell at Southampton University in 1991. Their working principle is that by creating a 2D periodic array of holes around a central core in a silica fibre, a waveguide can be made in which, without total internal reflection, the light can not propagate out of the core and into the periodic cladding. This is a highly engineerable approach that can be used to tailor the fibre properties. Fabrication involves simply drawing bundles of silica rods and tubes stacked in the appropriate pattern.

Higher damage threshold in air-core fibres at high optical powers

Minute control over pulse dispersion properties

Ability to generate highly nonlinear optical response (supercontinuum)

#### 4.5.5 On-chip photonic crystal waveguides and cavities

We know that it is a lot faster and more energy efficient to send information optically rather than electronically, so why not try to make optical chips that can move light around efficiently between devices, and even store it? This is difficult using rectangular waveguides as moving light round sharp corners is not possible with total internal reflection. But with photonic crystals it's no problem. We can therefore conceive highly intricate 2D (possibly even 3D) optical circuits, into which

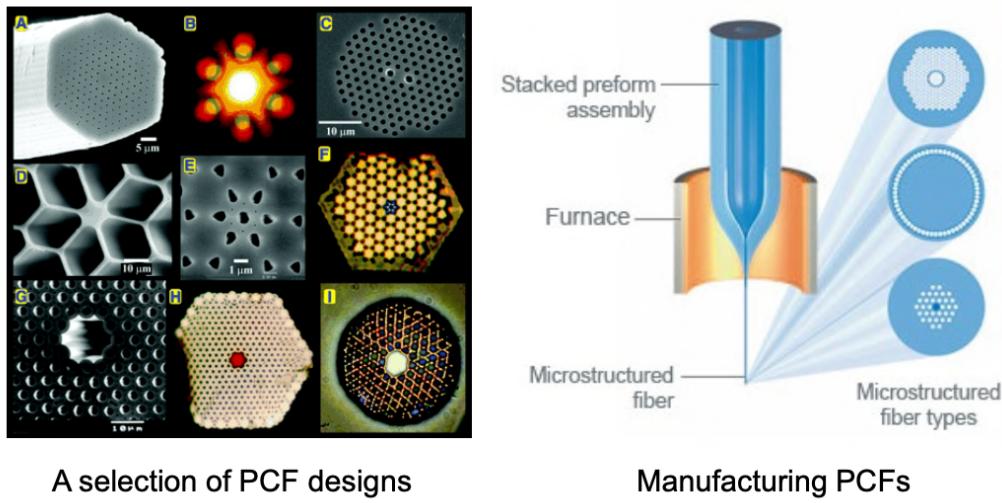


Figure 4.8: Examples of photonic crystal optical fibres, and schematic of fabrication process

components such as beam splitters and cavities can be introduced (more on these later). These can have applications in both classical and quantum computing.

## 4.6 Further Modern Optics

In this lecture we discussed a small subset of modern optical developments. In reality the field is much broader, and there is significant research attention devoted to this. A taxonomy of modern optics is given in Figure 4.10.

## References

1. Pendry, J. B., Schurig, D. & Smith, D. R. Controlling Electromagnetic Fields. *Science* **312**, 1780–1782 (2006).
2. Schurig, D. *et al.* Metamaterial Electromagnetic Cloak at Microwave Frequencies. *Science* **314**, 977–980 (2006).
3. Chen, X. *et al.* Macroscopic invisibility cloaking of visible light. *Nature Communications* **2**, 176 (2011).
4. Renger, J. *et al.* Hidden progress: broadband plasmonic invisibility. *Opt. Express* **18**, 15757–15768 (2010).

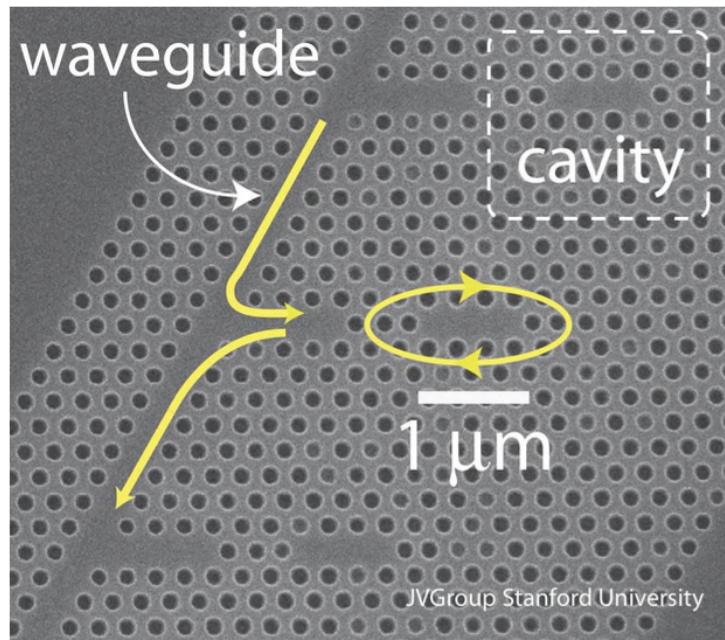


Figure 4.9: A 2D Photonic Crystal Fabricated at Stanford University

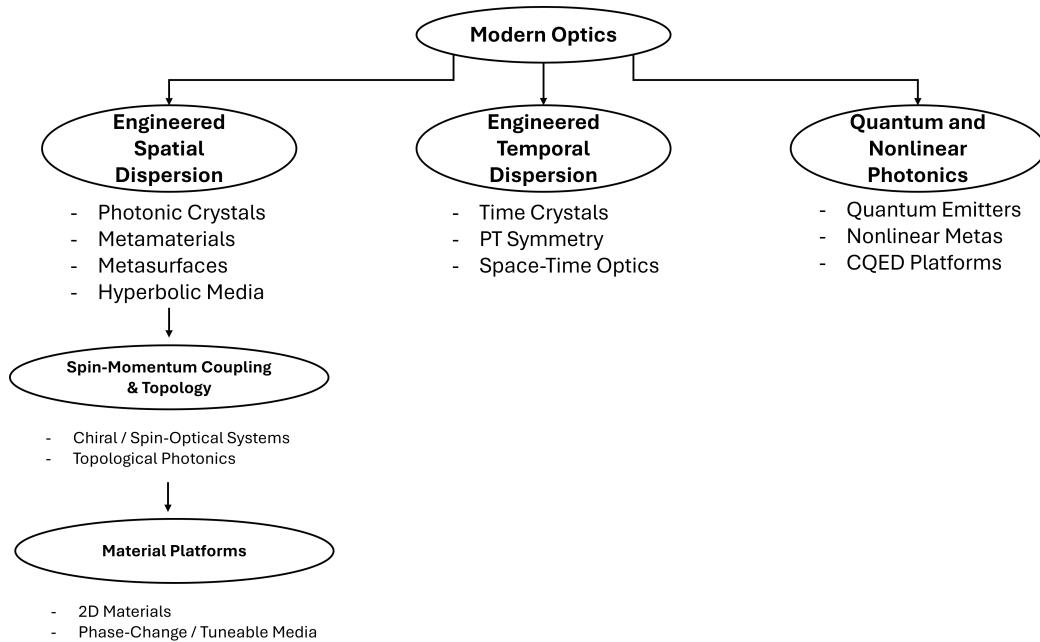


Figure 4.10: Taxonomy of Modern Optics

5. Cheng, J., Jafar-Zanjani, S. & Mosallaei, H. All-dielectric ultrathin conformal metasurfaces: lensing and cloaking applications at 532 nm wavelength. *Scientific Reports* **6**, 38440 (2016).
6. Pendry, J. B. Negative Refraction Makes a Perfect Lens. *Phys. Rev. Lett.* **85**, 3966–3969 (2000).
7. Zhang, X. & Liu, Z. Superlenses to overcome the diffraction limit. *Nature Materials* **7**, 435–441 (2008).
8. Fang, N., Lee, H., Sun, C. & Zhang, X. Sub-Diffraction-Limited Optical Imaging with a Silver Superlens. *Science* **308**, 534–537 (2005).

# 5 | Lecture 5 - The Semi-Classical Theory of Light

The first glimpse of quantum theory was provided by Max Planck in 1900, when he proposed that the Ultraviolet catastrophe of classical physics could be averted by considering light as existing as quanta of energy. The Black Body radiation spectrum was ultimately explained as a competition between two curves - the optical density of states, which increases with frequency as  $\omega^3$ , and the Bose-Einstein distribution which relates the average occupancy of a boson state to the ratio between the state's energy and the available thermal energy  $kBT$ . In the equation  $\rho_\omega$  is the energy density per unit angular frequency and per unit volume of space. It has units  $Jsm^{-3}$ . The graph below shows the equivalent spectral radiance as a function of wavelength (intensity of light emitted by a body per unit area and per unit wavelength =  $c\rho_\omega/4\pi$ ) at various temperatures  $T$ .

$$\rho_\omega = \frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{\exp(\hbar\omega/k_B T) - 1} \quad (5.1)$$

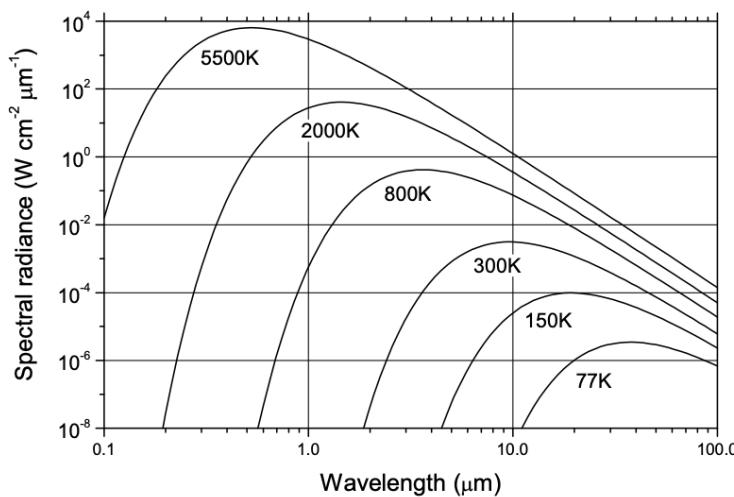


Figure 5.1: <caption>

## 5.1 Radiative Transitions in Atoms

In electromagnetic theory, **radiation** refers to the electromagnetic energy that propagates away from a source. In classical terms, it is the oscillating electric

and magnetic field described by Maxwell's equations, which carries energy and momentum through space. In quantum terms, radiation consists of photons, the quanta of the electromagnetic field. So, radiation is the field itself — the light or electromagnetic wave that exists as a result of emission or absorption processes.

When an atom/particle interacts with an electromagnetic field, a **radiative transition** occurs, whereby the atom's internal energy state is changed through interaction with the electromagnetic field. A radiative transition is the microscopic event that either produces or absorbs radiation. During such a transition, energy is exchanged between the system and the field according to the relation

$$h\nu = E_2 - E_1 \quad (5.2)$$

....where  $E_2$  and  $E_1$  are the energies of the two states involved. Radiative transitions can take several forms: absorption, where a photon is absorbed and the system moves to a higher energy state; spontaneous emission, where the system emits a photon into the vacuum field without external stimulation; and stimulated emission, where the presence of an existing photon induces the emission of another photon of identical frequency, phase, and direction. We will discuss these transitions in the next subsection.

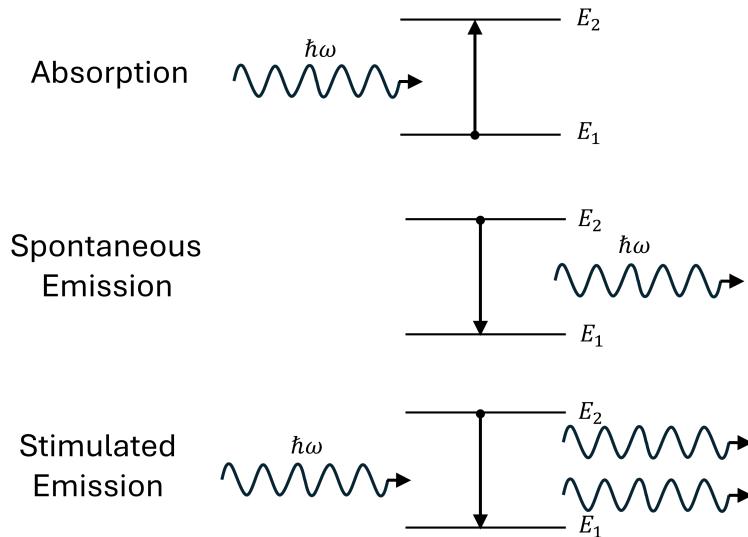


Figure 5.2: Optical transitions between atomic energy states.

Note, not all transitions are radiative. A system can undergo **non-radiative transitions**, in which the energy is transferred to other degrees of freedom rather than to the electromagnetic field. Examples of this include internal conversion, where

electronic energy is converted into vibrational motion, and collisional quenching, where energy is dissipated through particle collisions. These processes change the system's internal state but do not produce or absorb photons.

### 5.1.1 Frameworks

The interaction between light and matter can be described at different levels of abstraction:

- **Classical electrodynamics** - both matter and light are classical.
- **Semi-classical theory** - matter is quantum; the field is classical.
- **Quantum theory of radiation** - both matter and the field are quantised.

Each successive theory resolves the limitations of the previous one.

#### 5.1.1.1 Classical Theory

In the classical picture, light is a **continuous electromagnetic wave** obeying Maxwell's equations. Matter consists of **classical charges** (e.g., electrons) which can oscillate and radiate. Radiation arises from accelerated charges by the Larmor formula:

$$P = \frac{e^2 a^2}{6\pi\varepsilon_0 c^3} \quad (5.3)$$

The fields  $\mathbf{E}(\mathbf{r}, t)$  and  $\mathbf{B}(\mathbf{r}, t)$  are **deterministic**. The energy exchange between atoms and radiation is **continuous**. Atoms are modelled as oscillating dipoles without discrete energy levels.

But, this framework cannot explain **Discrete atomic spectra, Spontaneous emission, or Photon statistics/quantum correlations**. It predicts continuous emission and absorption instead of line spectra.

#### 5.1.1.2 Semi-Classical Theory

In the semi-classical picture, atoms are quantum systems with discrete states. The radiation field is a classical electromagnetic wave. The interaction between the atom and field is described by the Hamiltonian:

$$H = H_{\text{atom}} - \mathbf{d} \cdot \mathbf{E}(t) \quad (5.4)$$

....where  $\mathbf{d}$  is the dipole operator and  $\mathbf{E}(t)$  is a classical field.

This picture successfully describes **coherent light-matter interactions**, and leads to optical Bloch equations and population inversion concepts. However, it fails to explain spontaneous emission, and other quantum processes such as quantum noise and entanglement.

### 5.1.1.3 Quantum Theory

In the quantum picture, both the atoms and the electromagnetic field are treated quantum mechanically. This requires **quantising the electromagnetic field**, which is beyond the scope of this course. In this framework, field modes behave as independent quantum harmonic oscillators, and the interaction between these oscillators and the atomic system is described by the **Jaynes-Cummings Hamiltonian**. This description is outside the scope of this course, so we won't explore this any deeper here. However, the interested reader can consult the Fox book on Quantum Optics [1].

## 5.1.2 Einstein Coefficients

Quantum theory allows us to calculate the rate at which emission or absorption can occur. In 1916, Albert Einstein introduced a mathematical framework to describe the fundamental probabilities of radiative transitions between two discrete energy levels of an atom or molecule [2]. The coefficients of this framework, the **Einstein Coefficients**, provide a bridge between classical radiation theory and the quantum description of light-matter interactions. Ultimately, Einstein's analysis of how atoms interact with radiation fields lead to the theoretical foundation of lasers and the Planck distribution for blackbody radiation.

Einstein considered a simple system with two energy levels,  $E_1$  and  $E_2$ , where  $E_2 > E_1$ . Atoms can move between these levels either by absorbing or emitting electromagnetic radiation of frequency  $\nu = (E_2 - E_1)/h$ . Einstein assumed that three distinct processes can occur in such a two-level system: absorption, spontaneous emission, and stimulated emission. He then introduced coefficients to quantify the rates at which each process takes place.

**5.1.2.1 Absorption**

**5.1.2.2 Spontaneous Emission**

**5.1.2.3 Stimulated Emission**

## **5.2 Radiative Transition Rates**

## **5.3 Electric Dipole Transitions**

## Bibliography

1. Fox, A. M. *Quantum optics: an introduction* (Oxford university press, 2006).
2. Einstein, A. On the quantum theory of radiation. *Physikalische Zeitschrift* **18**, 167–83 (1917).