OPTOELECTRONICS

Handout 1

Dr Graham S McDonald
Newton 165
g.s.mcdonald@salford.ac.uk
http://www.cse.salford.ac.uk/profiles/gsmcdonald/

Course Syllabus

The aim of this module is to introduce basic concepts associated with optoelectronic systems and their design. This will be done by considering the following areas:

- **Optical sources**

Production of light. The excited state. What is an electromagnetic wave? The question of coherence. Laser light. How can we make a material lase? Population inversion. Maintaining the radiation density. The polarisation properties of laser light. The coherence properties of laser light. Coherence length and coherence time.

- **Incoherent sources (the LED)**

What is a semiconductor? Generation of light at a semiconductor junction. Construction of an LED.

- **Coherent sources (the laser)**

The HeNe laser. Three-level versus four-level lasers. The semiconductor laser. The output of a semiconductor laser. The shape of the output beam.

- **Transmission systems**

Mirrors. Lenses. Lens aberrations.

- **Optical fibres**

The structure of an optical fibre. Types of optical fibre. Manufacture of glass optical fibre. Transmission of light down an optical fibre. Modes in optical fibres. Coupling light to optical fibres. Losses in optical fibres. Dispersion in optical fibres.

- **Photodetectors**

Performance parameters of photodetectors. The photodiode. The PIN photodiode. The avalanche photodiode. Photodiode circuits.

- **An optoelectronic system** (if time permits)

Electronics & Optoelectronics Assessment

The assessment of the module is broken down as follows:

Semester 1	
· Electronics Continuous Assessment	20%
Semester 2	
· Optoelectronics Course tests (2 tests, each 10%)	20%
· 'Electronics & Optoelectronics' Examination (in May)	25%
Semesters 1 & 2	
· ***Electronics & Optoelectronics Laboratory***[†]	***35%***

[†] critical module component that you must attempt, otherwise the University will initiate revoking your registration.

Assessed Course Materials

These notes contain a great deal of background theory, explanations and references to current developments in the field of optoelectronics. The assessed materials (in tests and the exam) will be taken from the 'Optical Sources' to the 'Optical Fibres' sections, inclusive. So, all materials flagged up as "Further Information" is simply there to make the course more complete. For example, the topic 'Photodetectors' will neither be tested nor examined.

Tutorials and Course Handouts

Tutorials will be held *during* the two-hour lecture slots. Copies of the handouts and tutorials will be deposited in the appropriate filing cabinet drawers outside my office.

Online Materials

Handouts without gaps will be available in the Course Notes section within the Restricted area of the physics website:

http://www.cse.salford.ac.uk/physics/restricted/index.php#coursenotes

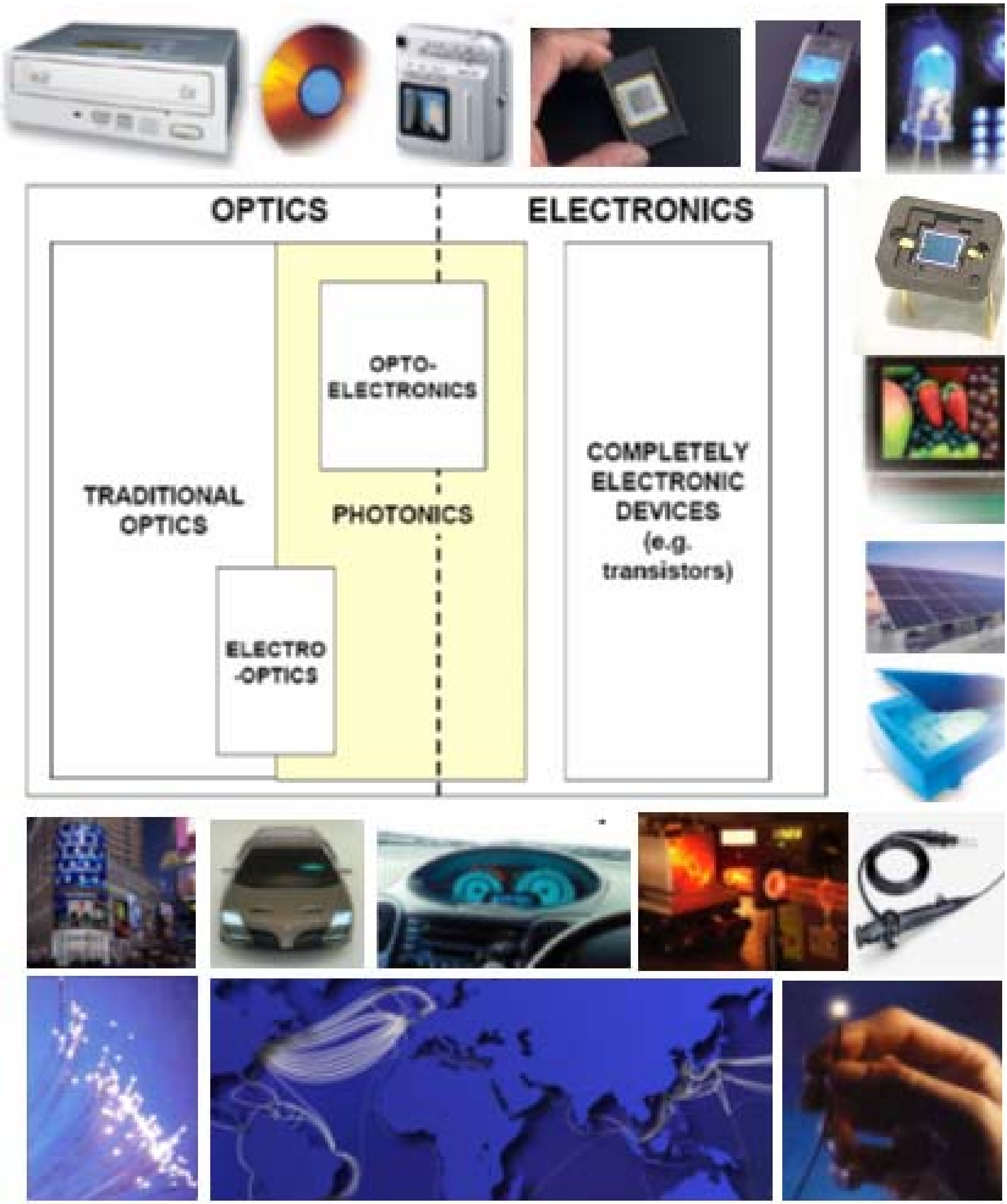
the usual username and password apply (these are different from university network login details). Alongside the course notes, there is also a link to additional materials for further reading and online java simulations.

Recommended Texts

The recommended text for the module is:
The Essence of Optoelectronics, K Booth & S Hill, Prentice Hall International.

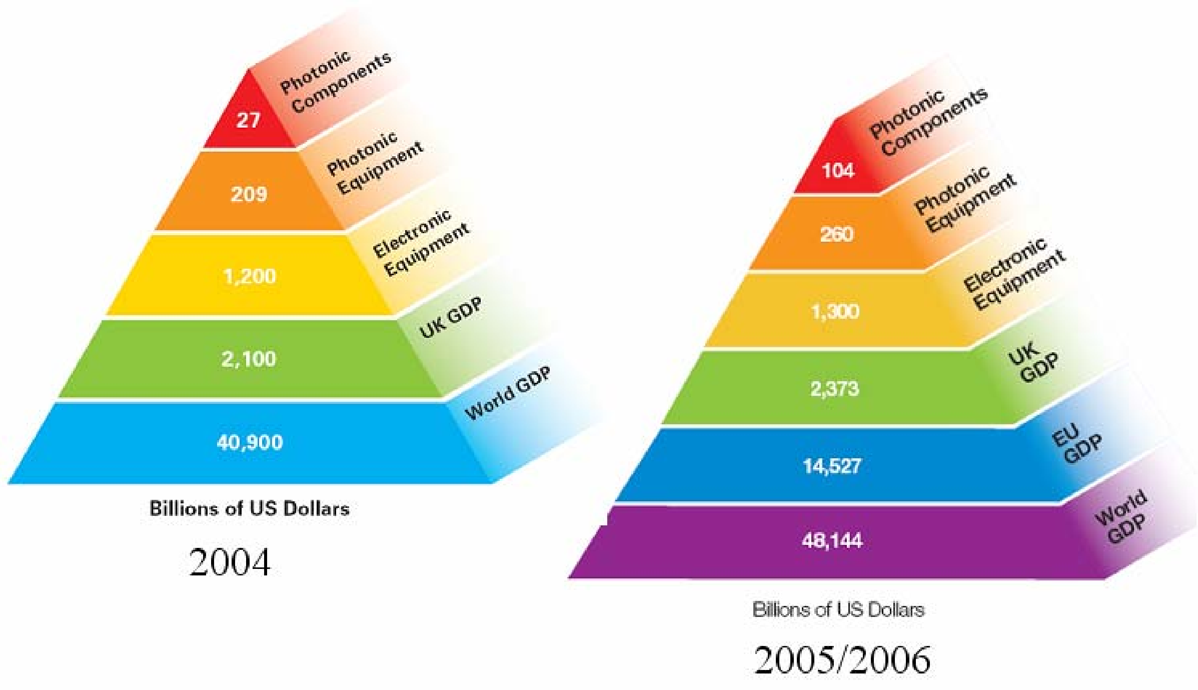
A higher level text which you may also wish to consult is
Optoelectronics: An Introduction, Wilson & Hawkes, Prentice Hall International.

Optoelectronics – What is it?



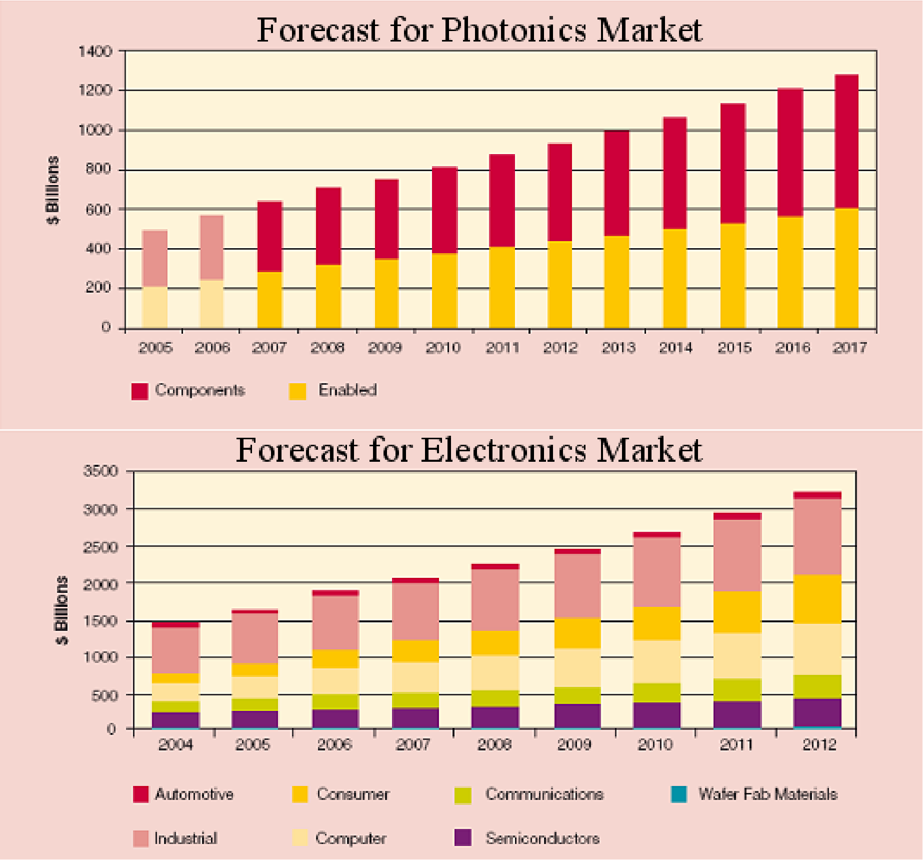
Optoelectronics – Global Market Growth

Sources: North-West Development Agency, DTI, UK Technology Strategy Board



2008 Photonics products: > 600 Billion US Dollars
\+ Electronics products >2,000 Billion US Dollars

2006 – 2017



Photonics Sectors & Products I

1. ICT & Consumer Photonics	**2. Industrial Photonics**
Displays • Microdisplays • Televisions 2D and 3D • Holographic displays • Advertising, indoor and outdoor messaging systems • Low-power displays • Projection and light shutters • Low-resolution and High resolution displays	*Lasers* • Cutting, welding, drilling and marking • Micromachining • Non-contact metrology and reverse engineering • Material processing including surface hardening • Scanning and positioning
Optical Communications • Internet: Voice over IP (VoIP), music, email, Web, eCommerce • Fibre To The premises (FTTx): Video on Demand, VoIP, telemedicine *Optical Data Storage* • CD and DVD media recorders, and now Blu-ray • Workstation and PC network data storage	*Sensing* • Structural health monitoring • Security and intrusion • Chemical agents • Biological agents • Smart industrial and power utility monitoring • Biophotonic/medical applications • Positioning and control • Traffic monitoring
Consumer Products • Mobile phones: displays, cameras and LEDs • Flat panel high definition LCD TVs: displays and backlights • Portable music and video players: displays • Digital camera: displays, image sensor and flash • Entertainment: displays, data storage, communications	*Imaging* • Quality control and inspection • Automated recognition • Thermal imaging • Security certification • Document scanning

Photonics Sectors & Products II

3. Defence and Security

Command, Control, Communications Computers and Intelligence
- Secure communications: Quantum computing and Quantum Cryptography
- Optical fibre: RF signal processing and transmission

Surveillance, Targeting Acquisition and Reconnaissance
- Visual sights, periscopes
- Displays
- Lasers: range finding, laser imaging, systems
- Remote sensing: typically employing multi-spectral techniques or passive IR sensors
- Forward Looking Infra-red Imaging (thermal)
- Perimeter security: IR sensors, image processing, optical sensors
- Checkpoint/Port/Airport security: Terahertz imaging,

Raman spectroscopy and image processing
- Target imaging and identification
- Unmanned Aerial Vehicles (UAV)
- Electro-optical countermeasures
- Sensing: Chemical & biohazard, structural health monitoring, optical strain imaging, embedded NDE

4. Life Sciences and Healthcare Biophotonics

- Diagnostics: non invasive imaging
- Treatment: laser surgery (cutting, drilling), Photodynamic Therapy (PDT)
- Pathogen detection: Raman spectroscopy
- Drug development
- Biometrics

Instrumentation
- Microscopy (confocal, FLM etc)
- Endoscopy
- Genomics/proteomics
- Biosensors: Pulse Oximetry
- Cellular and Clinical imaging: OCT, 4D studies,

Tomography

5. Lighting and Energy

Solid State Lighting
- Architectural, amenity, egress, signage, messaging systems, signal, commercial and residential lighting: Organic Light Emitting Diodes, Inorganic Light Emitting Diodes

Photovoltaics
- Solar cells

Sensing for Process Optimisation
Fuel Analysis
Energy Infrastructure Security

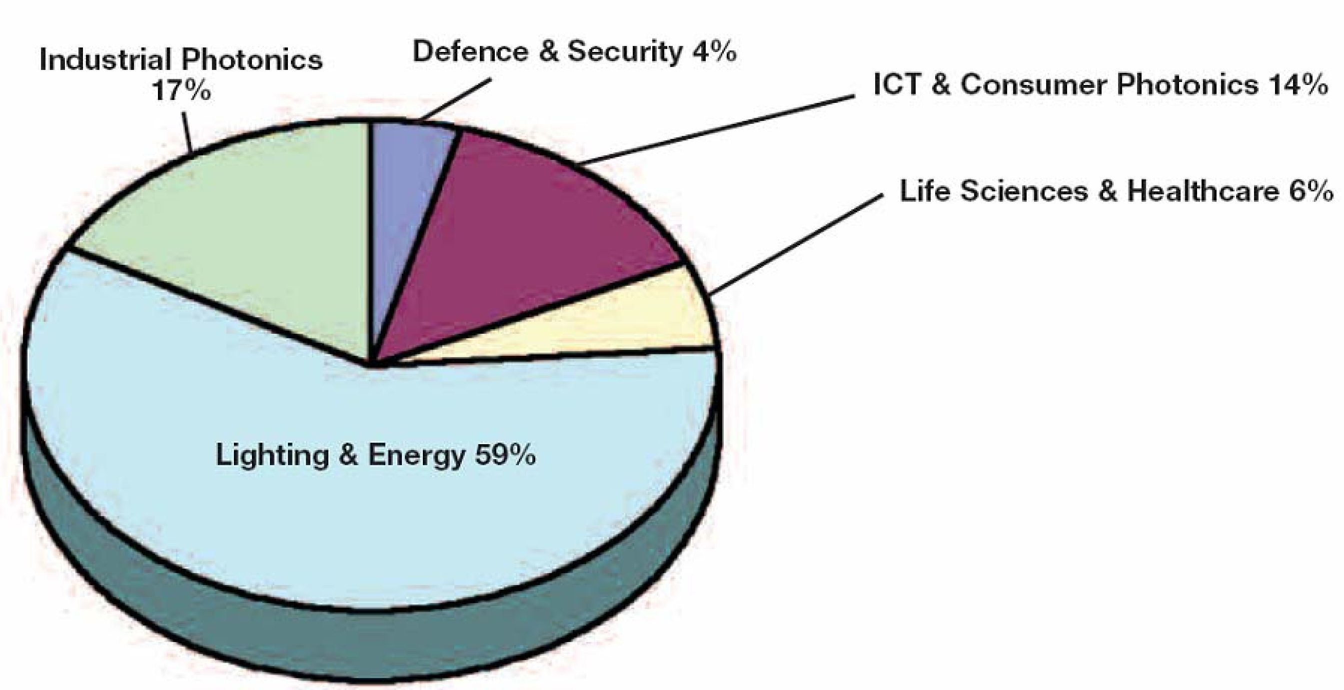
Photonics Employers in UK North West

(2005 data)

NW Region	Photonics	Turnover	Employees
Cheshire	29	£88 million	618
Cumbria	15	£51 million	663
Greater Manchester	88	£331 million	3283
Lancashire	34	£144 million	914
Merseyside	22	£46 million	523
Total	**188**	**£659 million**	**6001**

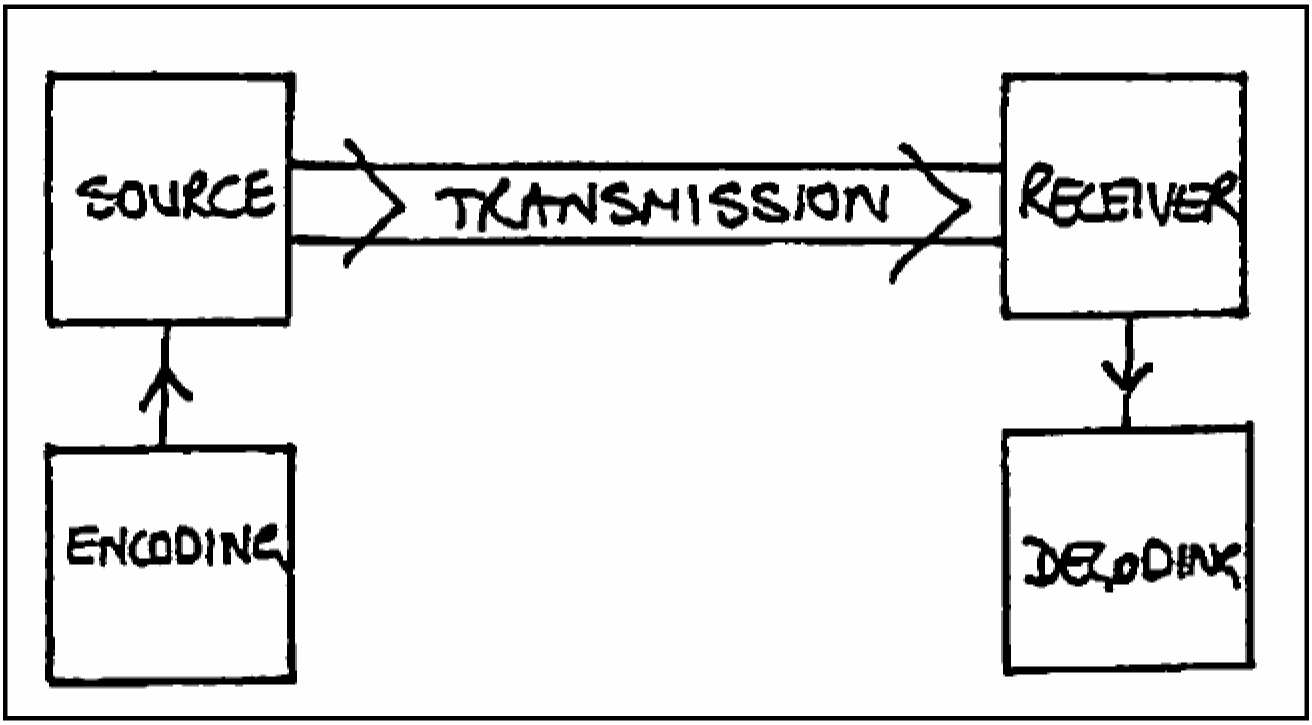
Number of Companies

Photonics Sector Activity in UK North West



A typical optoelectronic system?

The main components of an optoelectronics system are:



Basic components of an optoelectronic system

The choice of which of a range of components to use in any given situation depends on the parameters of the problem.

For instance, we may wish to consider

source wavelength, coherence & power output
size, weight & cost of source/detector
imaging or non-imaging detector?
guided or free space optics?
analogue or digital transmission?
safety
expected/required lifetime & reliability of system?

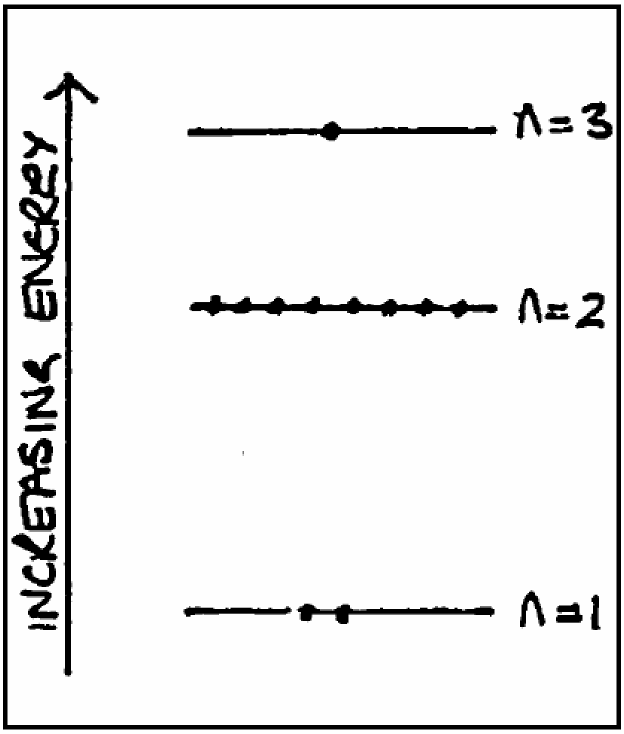
Our choice of components depends on what we want to do.

1. Optical Sources

- Optical sources convert other types of energy (usually electrical energy) into optical energy.
- The conversion may not be direct and is always inefficient.
- Sources may be divided up into two categories: those that are ***coherent*** and those that are ***incoherent***. In order to understand the concept of coherence we need to think about the way in which light is produced.

1.1 Production of Light

Consider the element sodium (symbol Na). The sodium atom has 11 electrons which orbit the nucleus. The energy of these electrons is said to be ***quantised***, that is only certain energies are allowable. Additionally, a principle called the ***Pauli Exclusion Principle*** dictates how many electrons may possess a particular energy. Thus we may draw an ***energy level diagram*** showing the energy levels and the number of electrons in each level. Electrons usually fill up energy levels closest to the nucleus first, as this minimises the energy of the system. The energy level diagram for sodium is shown to the right.



Energy level diagram for sodium (with electrons in their lowest energy states)

It is important to note that the energies of the allowable energy levels are unique to each atom or molecule. The variable *n* is called a *quantum number* and is one of several such numbers used to label individual electrons. Other quantum numbers relate to the spin of an electron and its angular momentum and the orientation of its orbit. One way of stating the Pauli Exclusion Principle is that each electron, associated with this sodium atom, must have its own unique set of quantum numbers (see Further Information below).

<u>Further Information</u> (Bohr Model + Quantum Numbers + Periodic Table)

i) Principle quantum number, n, and the Bohr model

$$\boxed{E = hf}$$ (Photon energy)

where $h = 6.626 \times 10^{-34}$ Js is called Planck's constant.

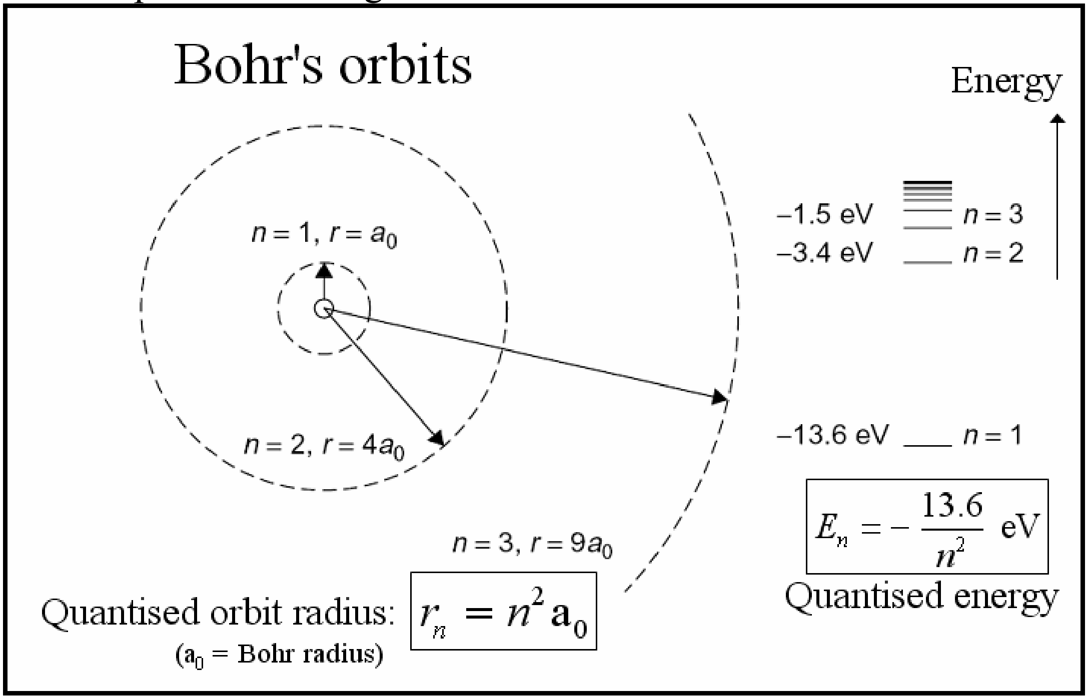
<u>1900</u> Max Planck postulated that emission and absorption of **light** is quantised into packets of energy E proportional to frequency f.

$$\boxed{L = rp = n\frac{h}{2\pi}}$$ (Electron angular momentum)

where r is the orbit radius, p is linear momentum, and $n = 1, 2, 3, \ldots$

<u>1913</u> Niels Bohr proposed quantisation of **electron states** of the hydrogen atom: a discrete set of orbits with quantised angular momenta.

Quantised angular momentum in these well-defined orbits also implies quantisation of the electron speeds and energies…



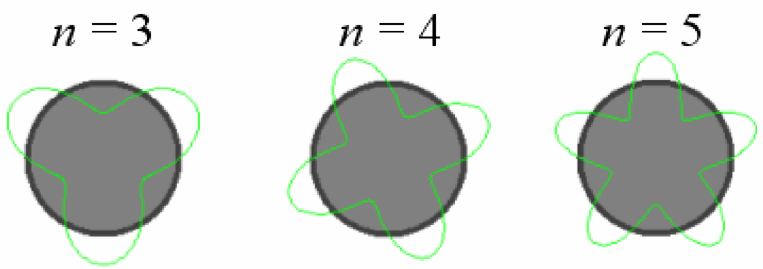
Bohr model electron orbits (shown dashed) around the nucleus

$$\boxed{p = \frac{h}{\lambda}}$$

(Momentum p and de Broglie wavelength λ)

<u>1924</u> Louis de Broglie proposed wave-particle duality in which both **light and matter** have an associated wavelength λ.
Note that this follows from Einstein's $E = pc$ and Planck's $E = hf$.
Moreover, the Bohr hypothesis of $rp = nh/2\pi$ becomes:

$rh/\lambda = nh/2\pi$ implying that $\boxed{2\pi r = n\lambda}$

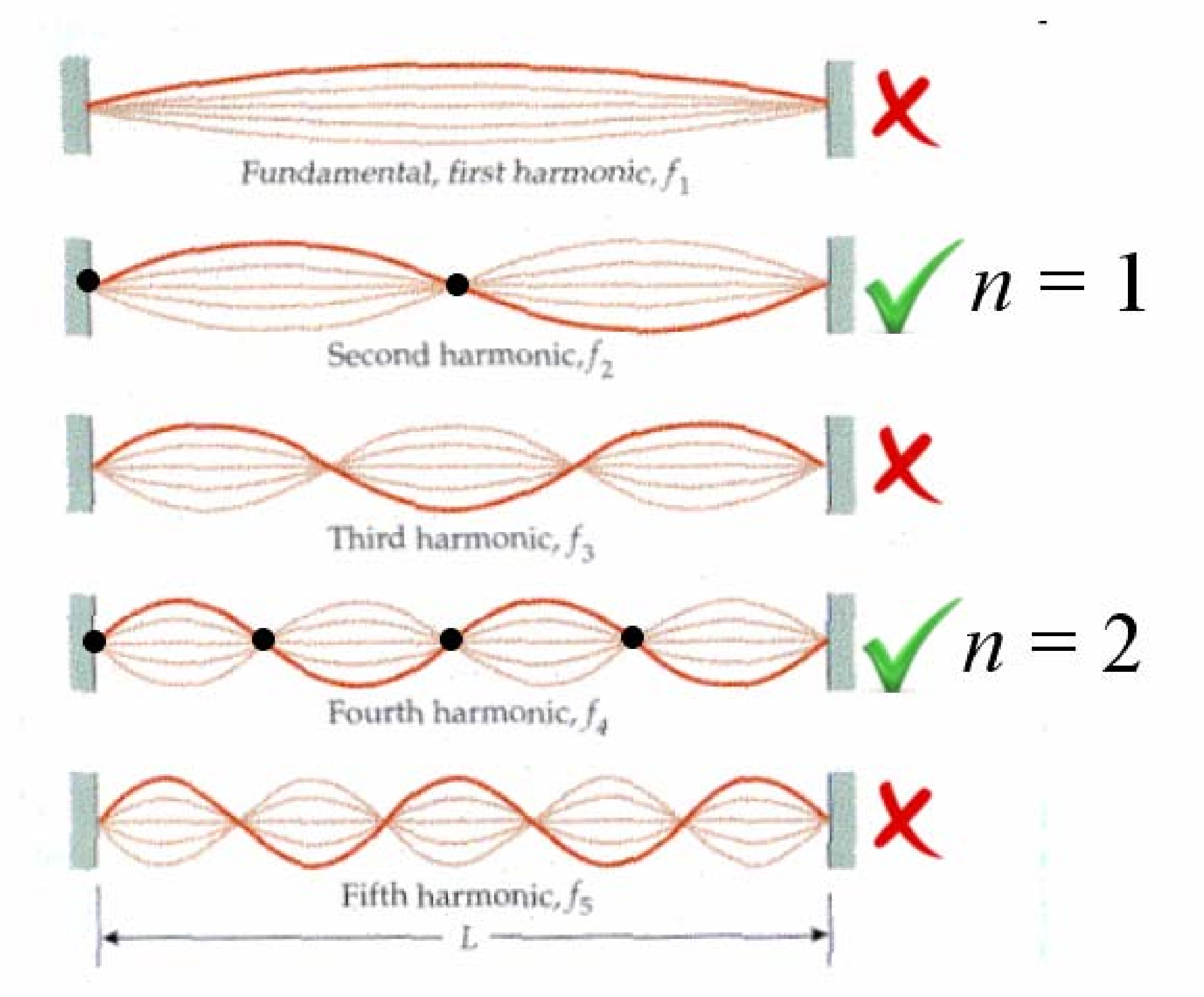


Electron circular standing waves explain quantised Bohr orbits

Recall ***standing waves on a string*** that is fixed at both ends.

Not all of these patterns can be sustained on a circular geometry.

We need a continuous variation of the underlying wave as we look around the circle …



Only multiples of a full wavelength can be sustained in the circular standing wave patterns.

Note that zero amplitude of the wave at the rightmost point will be the ***same point*** in circular geometry.

→ Number of nodes (zeroes) = $2n$

ii) Quantum numbers beyond the Bohr model

In the years after the proposal of the Bohr model, quantum mechanics was fully developed and revealed a much more complex picture. For example, the Pauli Exclusion Principle states that no two electrons may have the same set of quantum numbers. In this context, the full set of quantum numbers are:

the *principle quantum number,* $\boldsymbol{n}$
(major contribution to energy level of the electron, analogous to Bohr model),
the *orbital angular momentum quantum number,* $\boldsymbol{l}$
(the shape of the orbit, determining the *size* of the orbit's angular momentum),
the *azimuthal/magnetic quantum number,* $\boldsymbol{m_l}$
(the orientation of the orbit, and hence *direction* of orbit's angular momentum),
the *electron spin quantum number,* $\boldsymbol{m_s}$
(intrinsic spin of the electron itself, in addition to orbital spin).

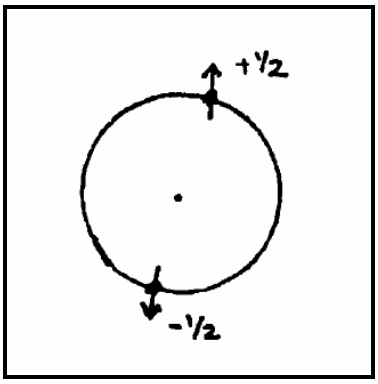
Each of these quantum numbers has only certain allowed values:

$$\begin{aligned} n &= 1, 2, 3, \ldots. \\ l &= 0, 1, 2, \ldots., (n-1) \\ m_l &= -l, -(l-1), \ldots.,+(l-1), +l \\ m_s &= +\tfrac{1}{2}, -\tfrac{1}{2} \end{aligned}$$

Case A: $n = 1$

Then l goes up to $n-1$, so $l = 0$
m_l varies between $-l$ and $+l$, so $m_l = 0$
m_s can be (for each m_l) $m_s = +\frac{1}{2}$ or $-\frac{1}{2}$.

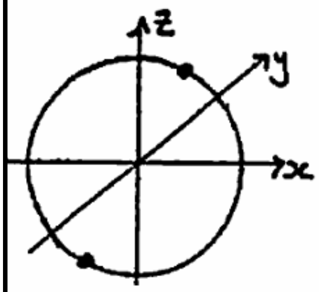
When $l = 0$ the orbital is ***spherical***. It may contain only two electrons since the only unique sets of quantum numbers (n, l, m_l, m_s) are $(1, 0, 0, +\frac{1}{2})$ and $(1, 0, 0, -\frac{1}{2})$.



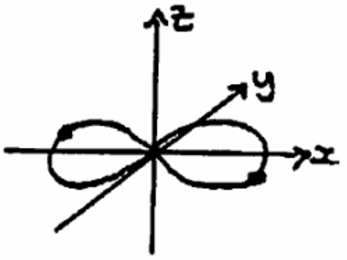
Case B: *n* = 2

Then l goes up to $n-1$, so $l = 0$ or $l = 1$
m_l varies between $-l$ and $+l$, so $m_l = 0$ or $m_l = -1, 0$ or $+1$
m_s can be (for each m_l) $m_s = +½$ or $-½$.

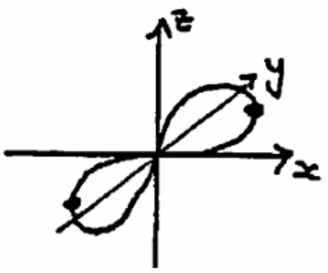
As before, when $l = 0$ the orbital is spherical and may only contain 2 electrons. The $l = 1$ orbital is ***dumbbell*** shaped and there are three of them ($m_l = -1, 0$ or $+1$). The value of m_l decides the orientation of the axis of a given orbital. The orbital axes lie along the *x, y, z* directions of Cartesian co-ordinates. As before, each of these orbitals may only hold a maximum of two electrons ($m_s = +½$ or $-½$). Thus the four orbitals at the $n = 2$ energy level may hold a total of 8 electrons.



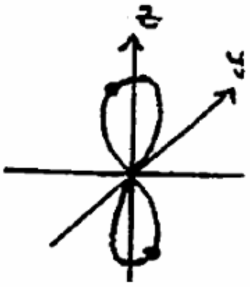
$(2,0,0,\pm1/2)$
$(n,l,m_l,\ m_s)$



$(2,1,1,\pm1/2)$
$(n,l,m_l,\ m_s)$



$(2,1,-1,\pm1/2)$
$(n,l,m_l,\ m_s)$



$(2,1,0,\pm1/2)$
$(n,l,m_l,\ m_s)$

We may carry out a similar analysis for each energy level (*n*) of the atom. It should be noted that if there is no external field applied to the atom then the energies of all the orbitals with the same *n* are said to be ***degenerate***, that is each orbital has the same energy. However, if, for example, a magnetic field is applied to the atom the energies of the orbitals shift slightly, and some have slightly more energy than before and some have slightly less. Thus energy degeneracy is lost.

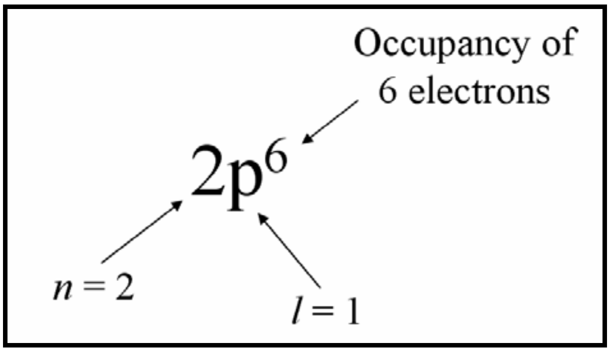
A ***notation*** commonly used to describe electronic states was one devised by spectroscopists long before quantum mechanics. Features of atomic spectra emission lines were defined as **s**harp, **p**rincipal, **d**iffuse, and **f**undamental, based on how they appeared when viewed. This notion maps onto different values of *l*, as shown below:

	orbital angular momentum quantum number, *l*					
	0	1	2	3		4, 5, 6, …
spectroscopic notation	s (sharp)	p (principal)	d (diffuse)	f (fundamental)	and thereafter alphabetical:	g, h, i, …

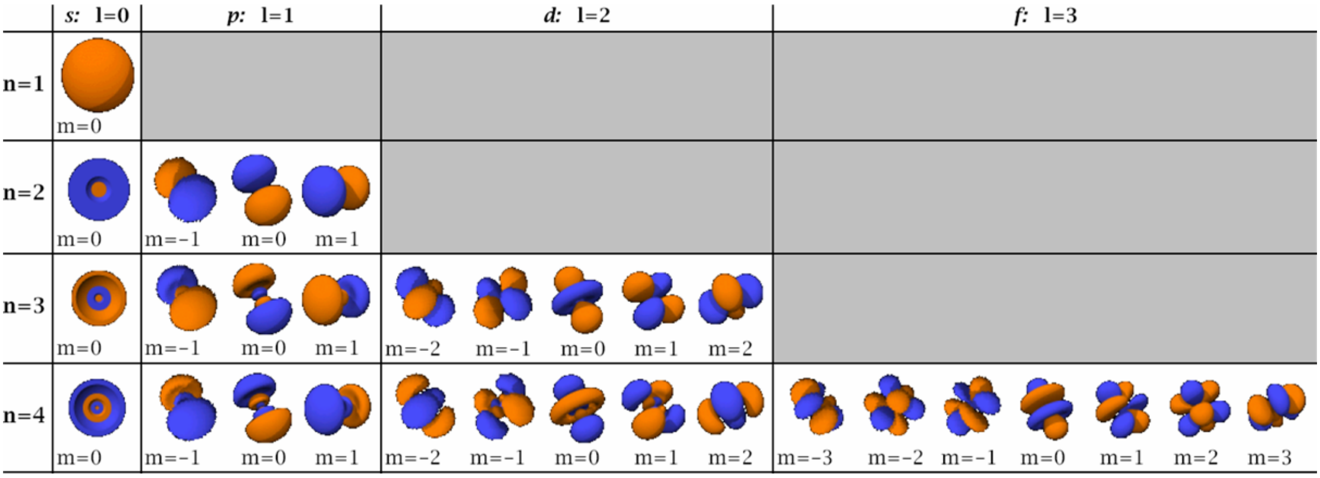
Returning to the example of the *11 electrons* in their lowest possible energy states of the *sodium atom*, there were:

2 electrons at $n = 1$, which have $l = 0$ and $m_l = 0$,
8 electrons at $n = 2$, of which 2 had $l = 0$
and 6 had $l = 1$ (pairs with $m_l = -1$, 0 or +1), and
1 electron at $n = 3$.

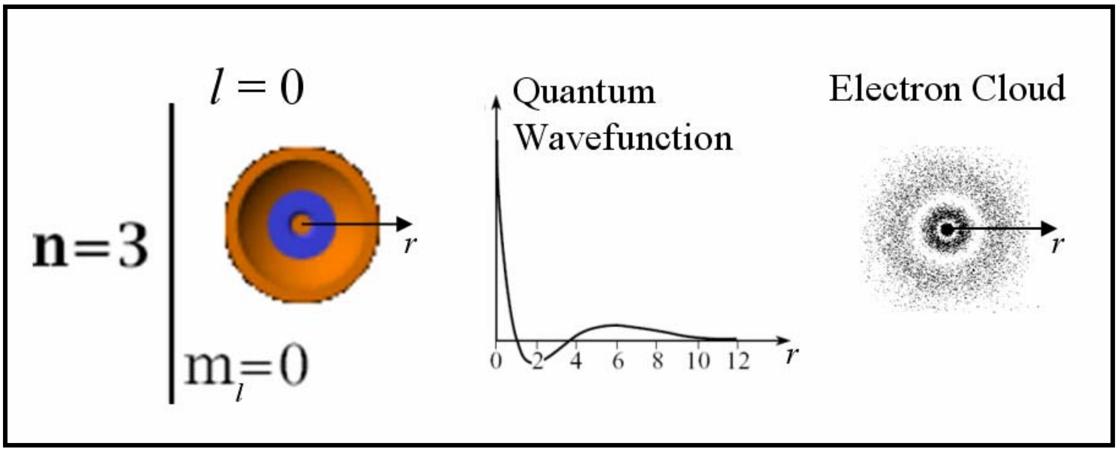
The notation for this configuration of electrons is: $1s^2$, $2s^2$, $2p^6$, $3s^1$
where, for example,



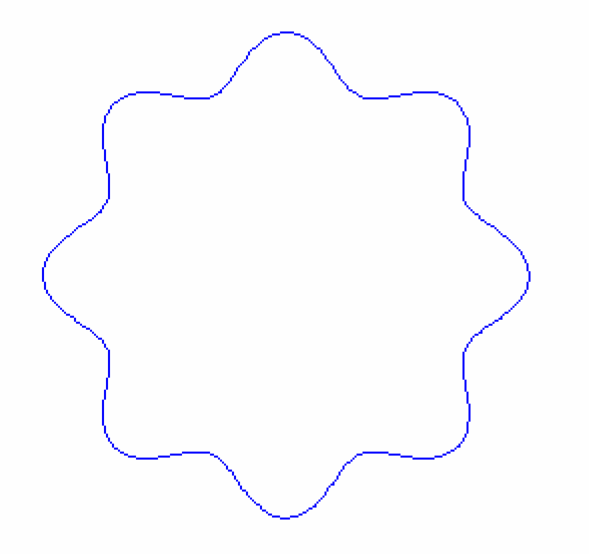
The progressive increase in complexity of the shape of orbitals as the quantum numbers increase is illustrated below:



where colours represent phase of wavefunction and *electron probability distribution*.



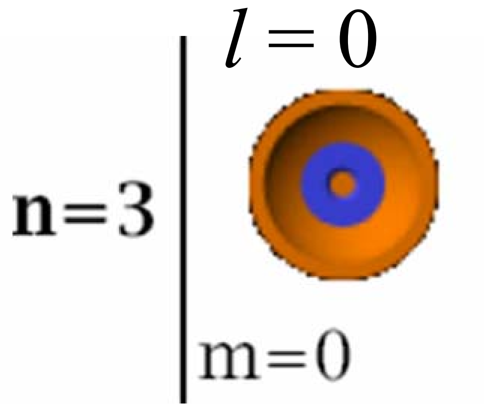
Circular standing waves are essentially one-dimensional (1D).

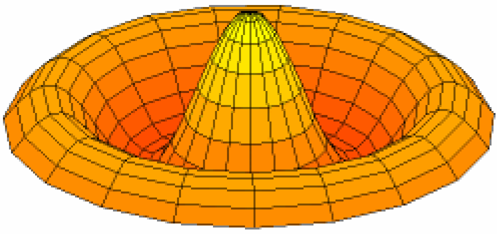


But atoms are three-dimensional (3D).

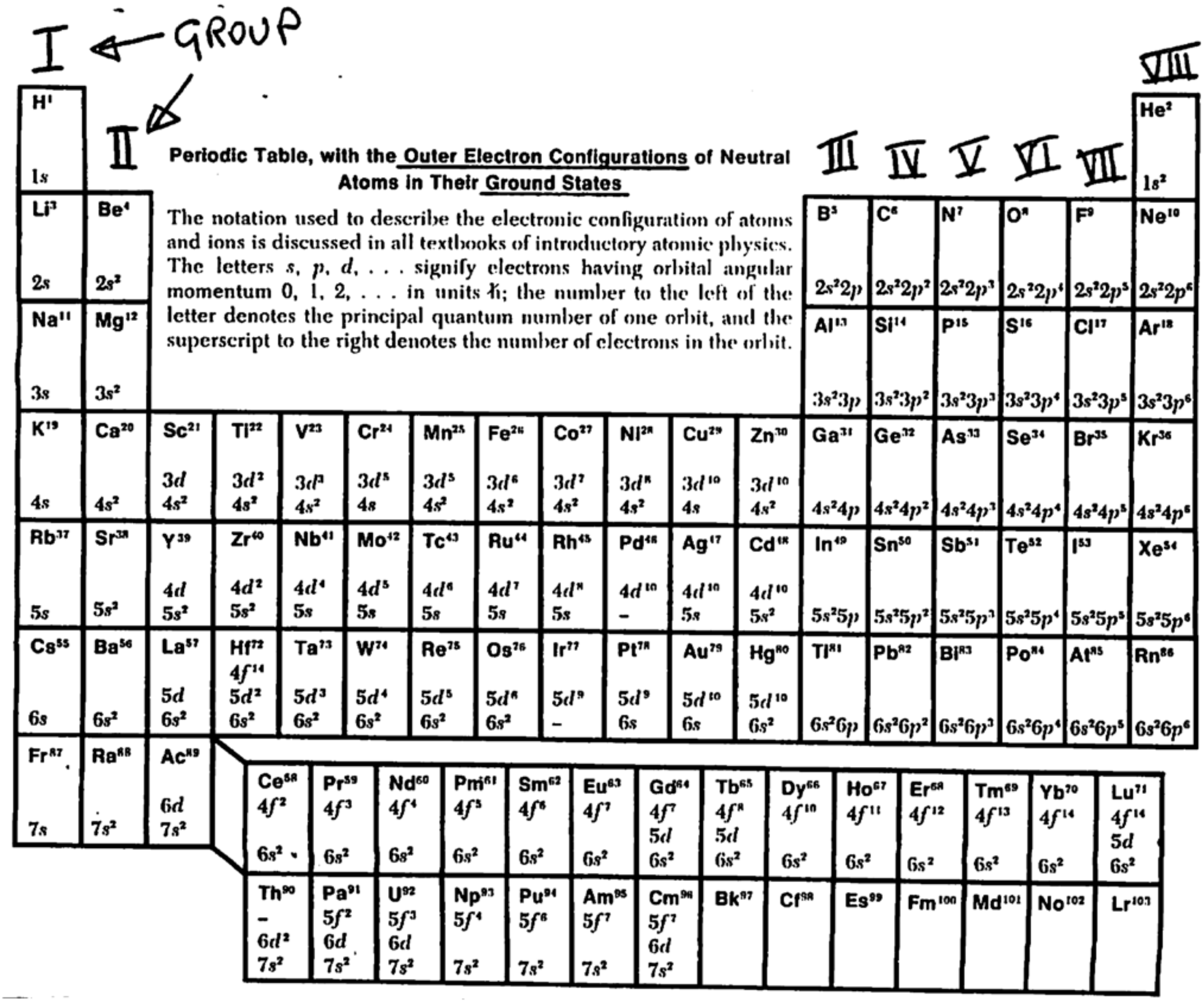
Detailed quantum treatment of the electron states finds that these are ***also standing matter waves***.

Consider waves on a 2D circular membrane, for example:





iii) The Periodic Table

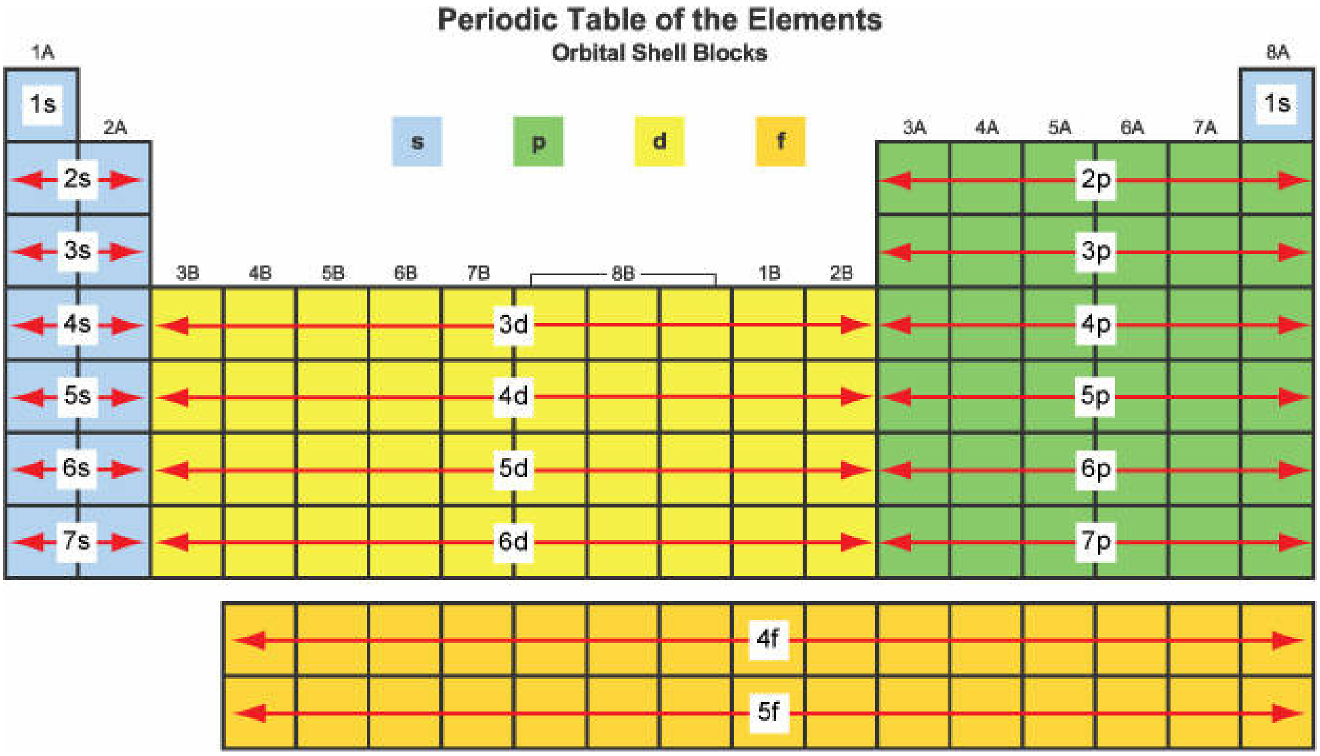


Periodic Table, with the Outer Electron Configurations of Neutral Atoms in Their Ground States

The notation used to describe the electronic configuration of atoms and ions is discussed in all textbooks of introductory atomic physics. The letters s, p, d, . . . signify electrons having orbital angular momentum 0, 1, 2, . . . in units $\hbar$; the number to the left of the letter denotes the principal quantum number of one orbit, and the superscript to the right denotes the number of electrons in the orbit.

I ← GROUP	II											III	IV	V	VI	VII	VIII
H^{1} $1s$																	He^{2} $1s^2$
Li^{3} $2s$	Be^{4} $2s^2$											B^{5} $2s^22p$	C^{6} $2s^22p^2$	N^{7} $2s^22p^3$	O^{8} $2s^22p^4$	F^{9} $2s^22p^5$	Ne^{10} $2s^22p^6$
Na^{11} $3s$	Mg^{12} $3s^2$											Al^{13} $3s^23p$	Si^{14} $3s^23p^2$	P^{15} $3s^23p^3$	S^{16} $3s^23p^4$	Cl^{17} $3s^23p^5$	Ar^{18} $3s^23p^6$
K^{19} $4s$	Ca^{20} $4s^2$	Sc^{21} $3d$ $4s^2$	Ti^{22} $3d^2$ $4s^2$	V^{23} $3d^3$ $4s^2$	Cr^{24} $3d^5$ $4s$	Mn^{25} $3d^5$ $4s^2$	Fe^{26} $3d^6$ $4s^2$	Co^{27} $3d^7$ $4s^2$	Ni^{28} $3d^8$ $4s^2$	Cu^{29} $3d^{10}$ $4s$	Zn^{30} $3d^{10}$ $4s^2$	Ga^{31} $4s^24p$	Ge^{32} $4s^24p^2$	As^{33} $4s^24p^3$	Se^{34} $4s^24p^4$	Br^{35} $4s^24p^5$	Kr^{36} $4s^24p^6$
Rb^{37} $5s$	Sr^{38} $5s^2$	Y^{39} $4d$ $5s^2$	Zr^{40} $4d^2$ $5s^2$	Nb^{41} $4d^4$ $5s$	Mo^{42} $4d^5$ $5s$	Tc^{43} $4d^6$ $5s$	Ru^{44} $4d^7$ $5s$	Rh^{45} $4d^8$ $5s$	Pd^{46} $4d^{10}$ –	Ag^{47} $4d^{10}$ $5s$	Cd^{48} $4d^{10}$ $5s^2$	In^{49} $5s^25p$	Sn^{50} $5s^25p^2$	Sb^{51} $5s^25p^3$	Te^{52} $5s^25p^4$	I^{53} $5s^25p^5$	Xe^{54} $5s^25p^6$
Cs^{55} $6s$	Ba^{56} $6s^2$	La^{57} $5d$ $6s^2$	Hf^{72} $4f^{14}$ $5d^2$ $6s^2$	Ta^{73} $5d^3$ $6s^2$	W^{74} $5d^4$ $6s^2$	Re^{75} $5d^5$ $6s^2$	Os^{76} $5d^6$ $6s^2$	Ir^{77} $5d^9$ –	Pt^{78} $5d^9$ $6s$	Au^{79} $5d^{10}$ $6s$	Hg^{80} $5d^{10}$ $6s^2$	Tl^{81} $6s^26p$	Pb^{82} $6s^26p^2$	Bi^{83} $6s^26p^3$	Po^{84} $6s^26p^4$	At^{85} $6s^26p^5$	Rn^{86} $6s^26p^6$
Fr^{87} $7s$	Ra^{88} $7s^2$	Ac^{89} $6d$ $7s^2$															

Ce^{58} $4f^2$ $6s^2$	Pr^{59} $4f^3$ $6s^2$	Nd^{60} $4f^4$ $6s^2$	Pm^{61} $4f^5$ $6s^2$	Sm^{62} $4f^6$ $6s^2$	Eu^{63} $4f^7$ $6s^2$	Gd^{64} $4f^7$ $5d$ $6s^2$	Tb^{65} $4f^8$ $5d$ $6s^2$	Dy^{66} $4f^{10}$ $6s^2$	Ho^{67} $4f^{11}$ $6s^2$	Er^{68} $4f^{12}$ $6s^2$	Tm^{69} $4f^{13}$ $6s^2$	Yb^{70} $4f^{14}$ $6s^2$	Lu^{71} $4f^{14}$ $5d$ $6s^2$
Th^{90} – $6d^2$ $7s^2$	Pa^{91} $5f^2$ $6d$ $7s^2$	U^{92} $5f^3$ $6d$ $7s^2$	Np^{93} $5f^4$ $7s^2$	Pu^{94} $5f^6$ $7s^2$	Am^{95} $5f^7$ $7s^2$	Cm^{96} $5f^7$ $6d$ $7s^2$	Bk^{97}	Cf^{98}	Es^{99}	Fm^{100}	Md^{101}	No^{102}	Lr^{103}



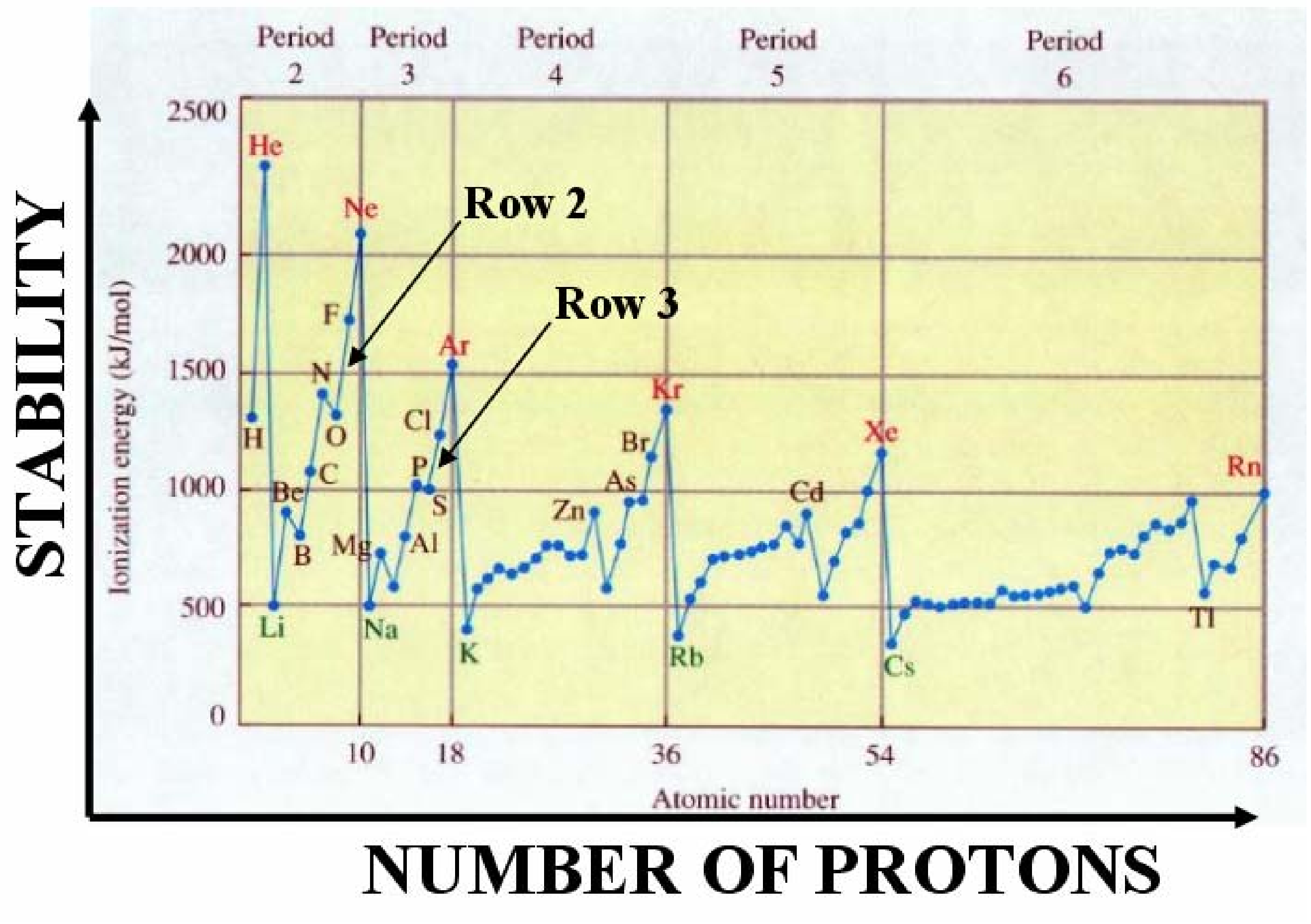
The elements are laid out in the periodic table in terms of their properties and these are largely dictated by:

- atomic number = number of protons
  = total number of electrons in the neutral atom
- number of electrons in the outermost levels
  (degree of being filled)

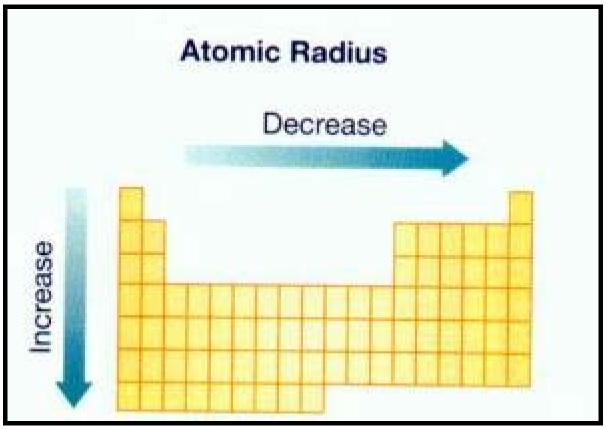
Considering atoms in their ground state and looking right across the third row:

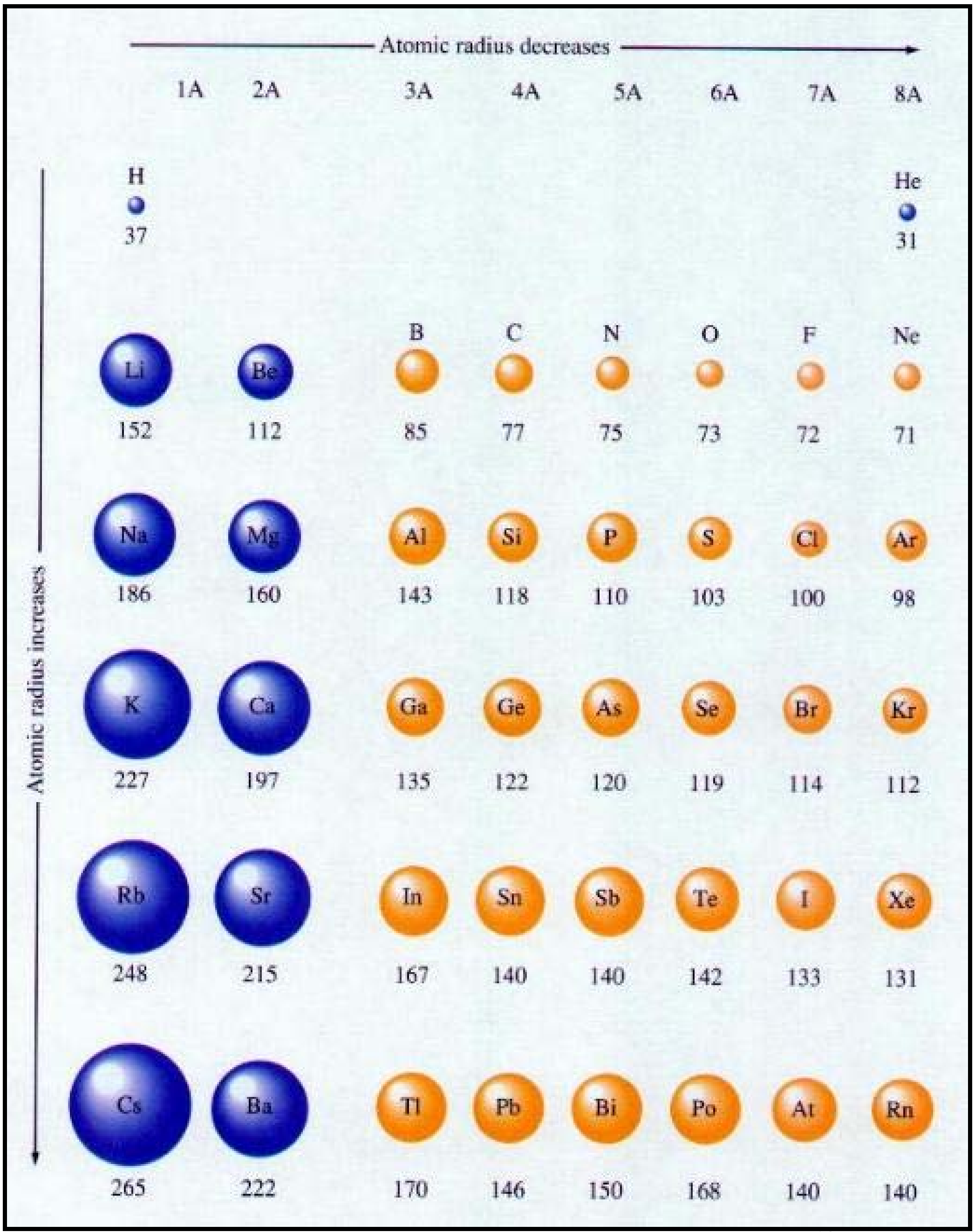
- **Sodium**, Na has 11 protons. The 11 electrons are $1s^2, 2s^2, 2p^6, \underline{3s^1}$.
  There is 1 outer electron ($3s^1$) and Na appear in Column/Group I.
- **Silicon**, Si has 14 protons. The 14 electrons are $1s^2, 2s^2, 2p^6, \underline{3s^2, 3p^2}$.
  There are 4 outer electrons ($3s^2, 3p^2$) and Si appears in Column/Group IV.
- **Argon**, Ar has 18 protons. The 18 electrons are $1s^2, 2s^2, 2p^6, \underline{3s^2, 3p^6}$.
  There are 8 outer electrons ($3s^2, 3p^6$) and Ar appears in Column/Group VIII.

Sodium is highly reactive, Silicon is relatively stable, while Argon is extremely inert (resistant to interactions with itself other elements, it exists as an atomic gas). Similar patterns appear elsewhere as one looks across the rows:



Explaining the trends …





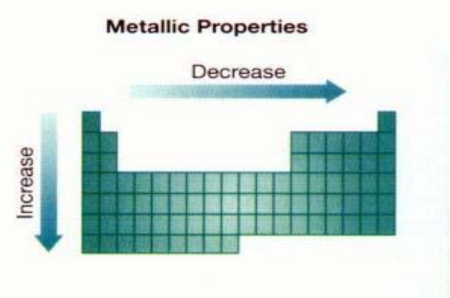
Diameter scale is in units of 1.0 picometre (1x10^{-12}m)

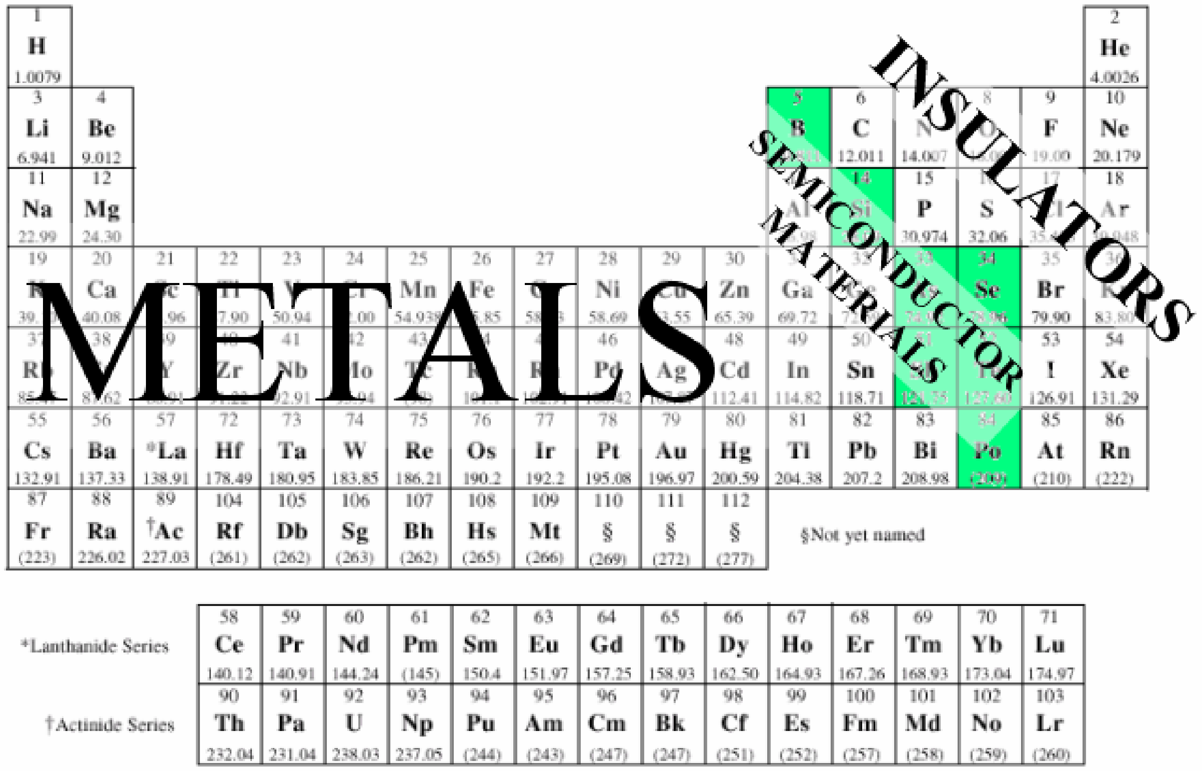
- **Moving Down a Column** in the periodic table, the outermost electrons become less tightly bound to the nucleus. This happens because the number of filled principal energy levels (which shield the outermost electrons from attraction to the nucleus) increases downward within each group. The atoms become larger and the outer electrons more readily interact with other species.

- **Moving Across a Row** (from left to right) in the periodic table, electrons are added one at a time. As this happens, the amount of shielding from inner filled orbitals remains constant but the outermost electrons experience increasingly stronger nuclear attraction (there are more protons), so the electrons become closer to the nucleus and more tightly bound to it. The atoms becomes smaller and elements become less reactive.

- **Bonding of Materials.** Elements interact and bond with each other to lose, or gain share of, outer electrons to leave them with an environment of an effectively filled outer electron shell (maximising stability of the resulting substance).

- **Heat & Electrical Conductivity.** The level of conductivity, and hence metal-like behaviour, is determined by the availability of loosely bound outer electrons to act as the carriers of heat and electricity.





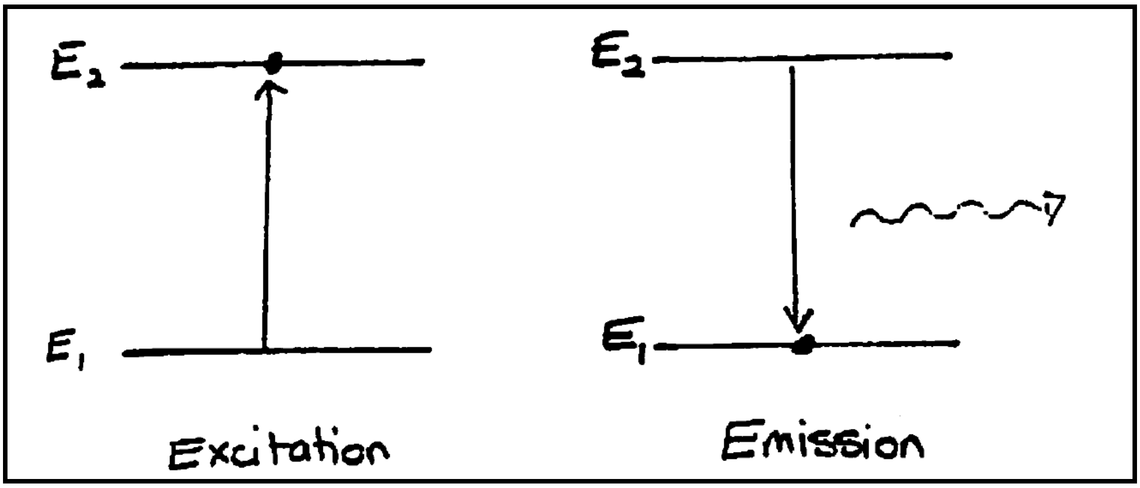
1.2 The Excited State

When the sodium atom has its 11 electrons arranged as shown above then it is said to be in the ***ground state***: it is in its minimum energy configuration, and the electrons are as close as they are allowed to be to the nucleus. If we now give the sodium atom enough energy then we can shift an electron from its normal energy level into one further away from the nucleus, thus giving the atom more energy. The atom is now said to be in an ***excited state***. This is an energetically unfavourable state for the atom to be in and it wishes to lose this energy. The easiest way for this to happen is for the electron to go back to its original energy level and for the atom to give up the excess energy that it has absorbed. The excess energy may be lost in a number of ways, one of which is for the atom to give out the energy as an ***electromagnetic wave***.

For such small energy transitions, the atom emits a single ***photon***. We can determine the wavelength, λ, of the photon using:

$$E = hf = h\frac{c}{\lambda} = E_2 - E_1 = \Delta E \qquad (1)$$

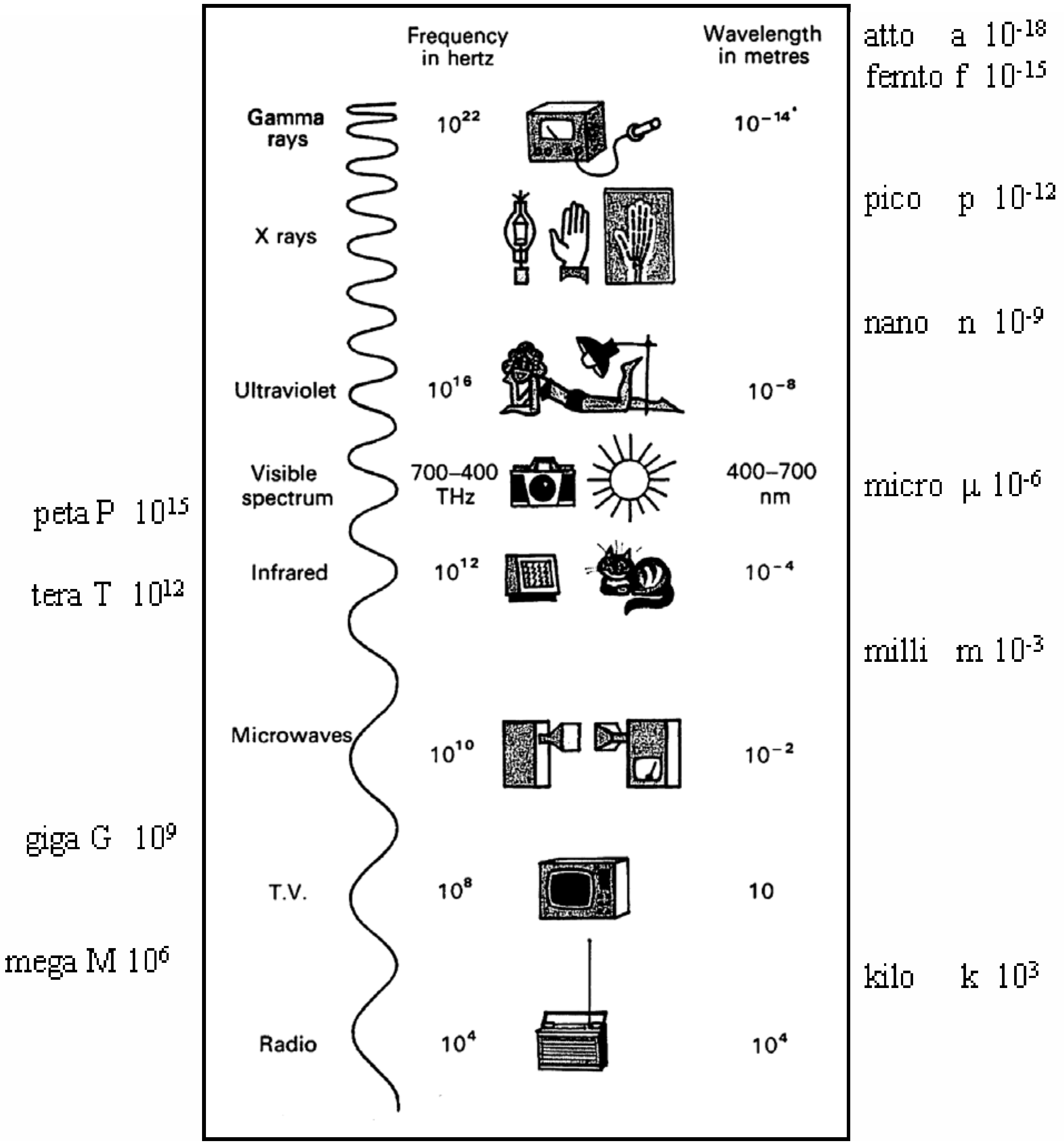
Here, E is the photon energy, f is light frequency, $h = 6.626\times10^{-34}$ Js is Planck's constant, $c = 3\times10^{8}$ ms^{-1} is the speed of light in free space, and λ is wavelength. This is equal to the amount of energy ΔE that the atom has lost.



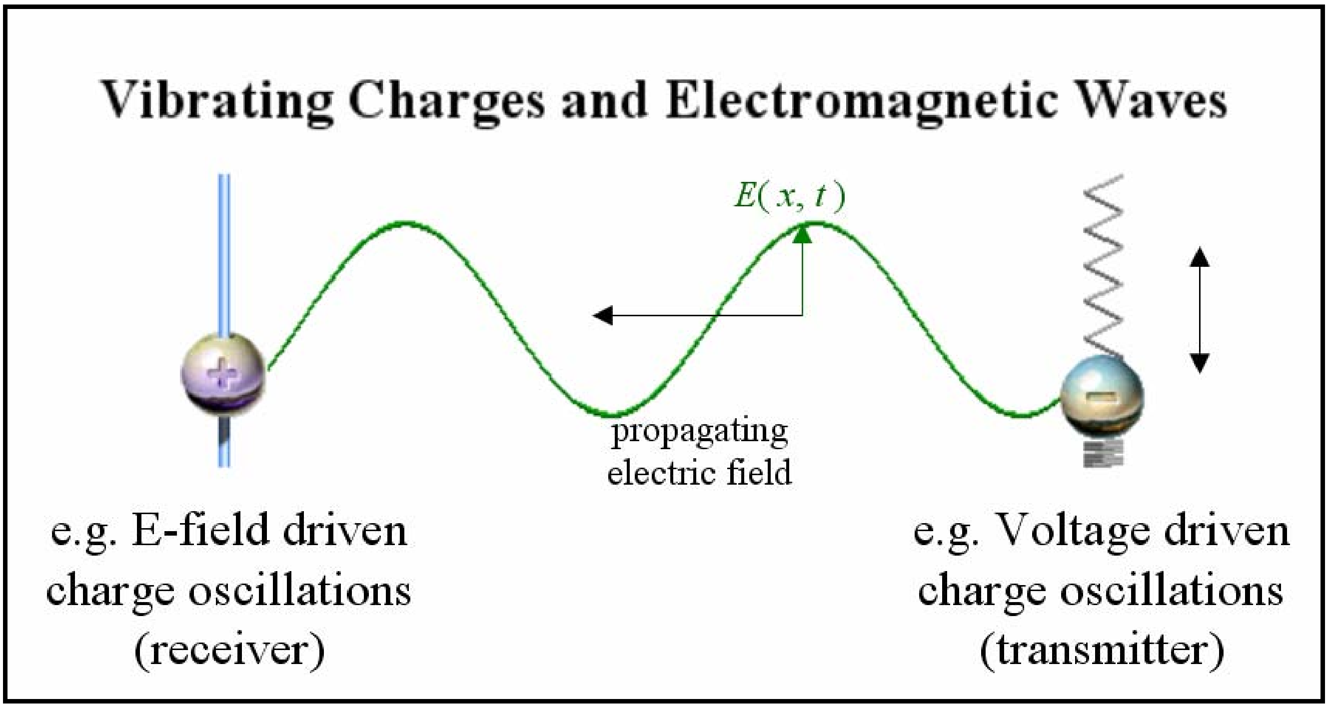
Excitation of an atom and emission of a photon

1.3 What is an Electromagnetic Wave?

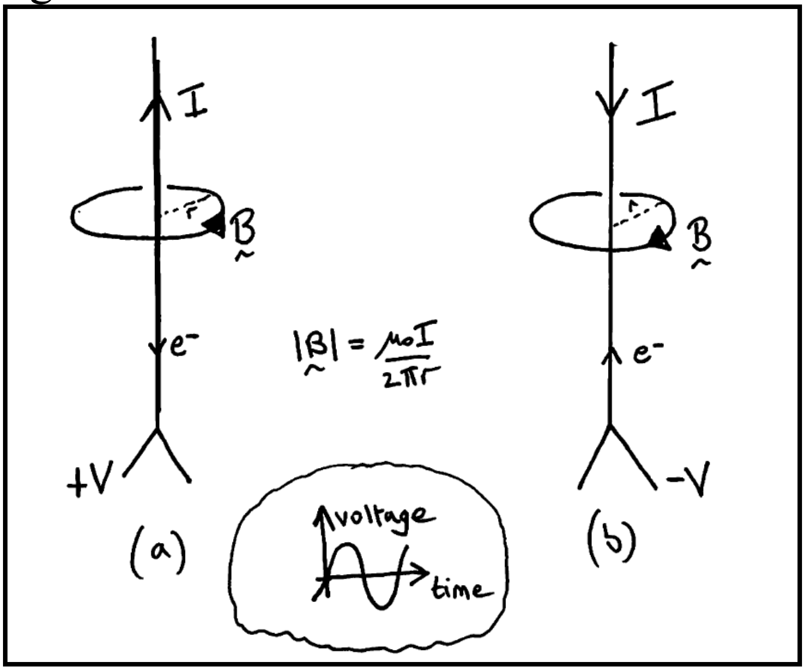
The electromagnetic spectrum includes radio waves (low frequency, f, long wavelength, λ) to gamma rays (high frequency, short wavelength). All electromagnetic radiation travels at a constant speed of $c = f\lambda = 3.0\times10^{8}\ \mathrm{ms}^{-1}$ in a vacuum.



A way to visualise the generation of an electromagnetic wave is to consider a radio antenna connected to +V volts or –V volts via a switch. If the switch is flicked back and forth between the two voltages then we can imagine the electrons in the antenna moving backwards and forwards as they are alternately attracted to and then repulsed from the voltage at the end of the wire.

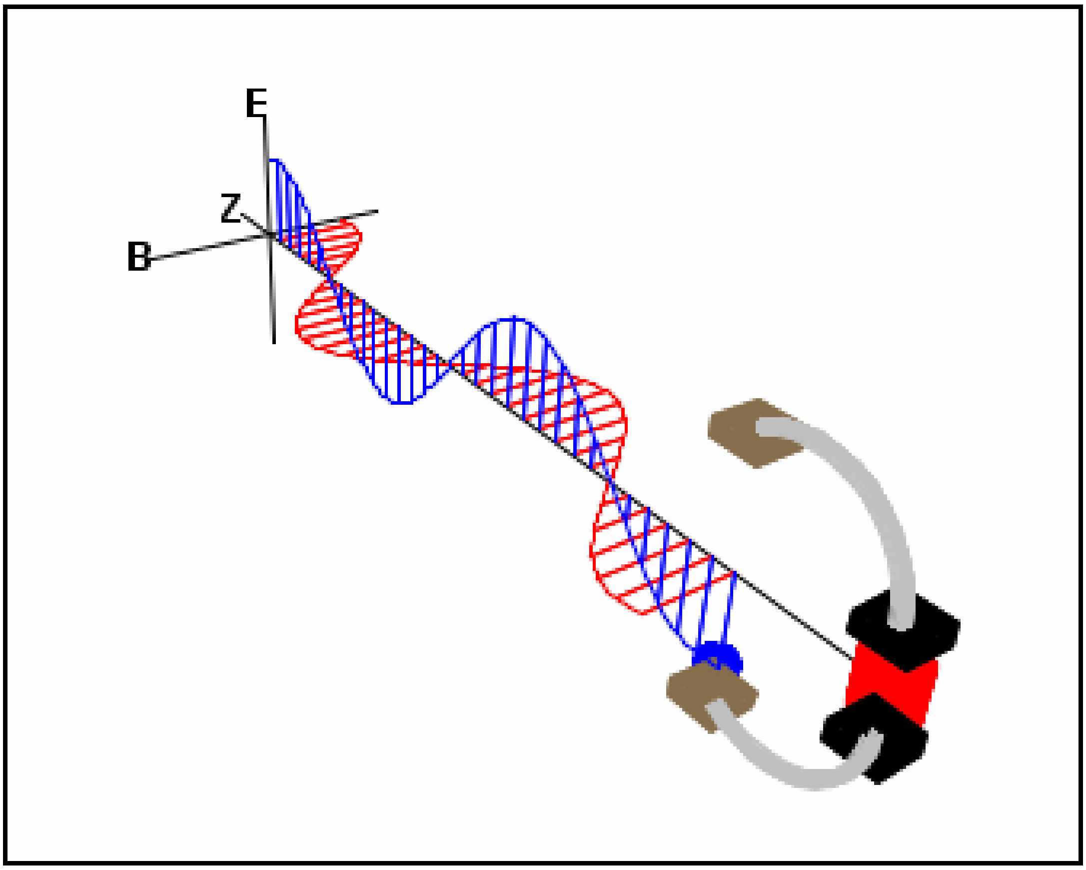


This is the transmission section of a potential optoelectronic system. The signal transmitted can be pre-encoded with information by one, or a combination, of electronic modulations: digital, amplitude (AM), frequency (FM) or phase. But the transmitted electric field pattern is not the whole story. A moving charge (i.e. current *I*) generates a magnetic field *B*.



Oscillating current gives *B* oscillations perpendicular to *E*

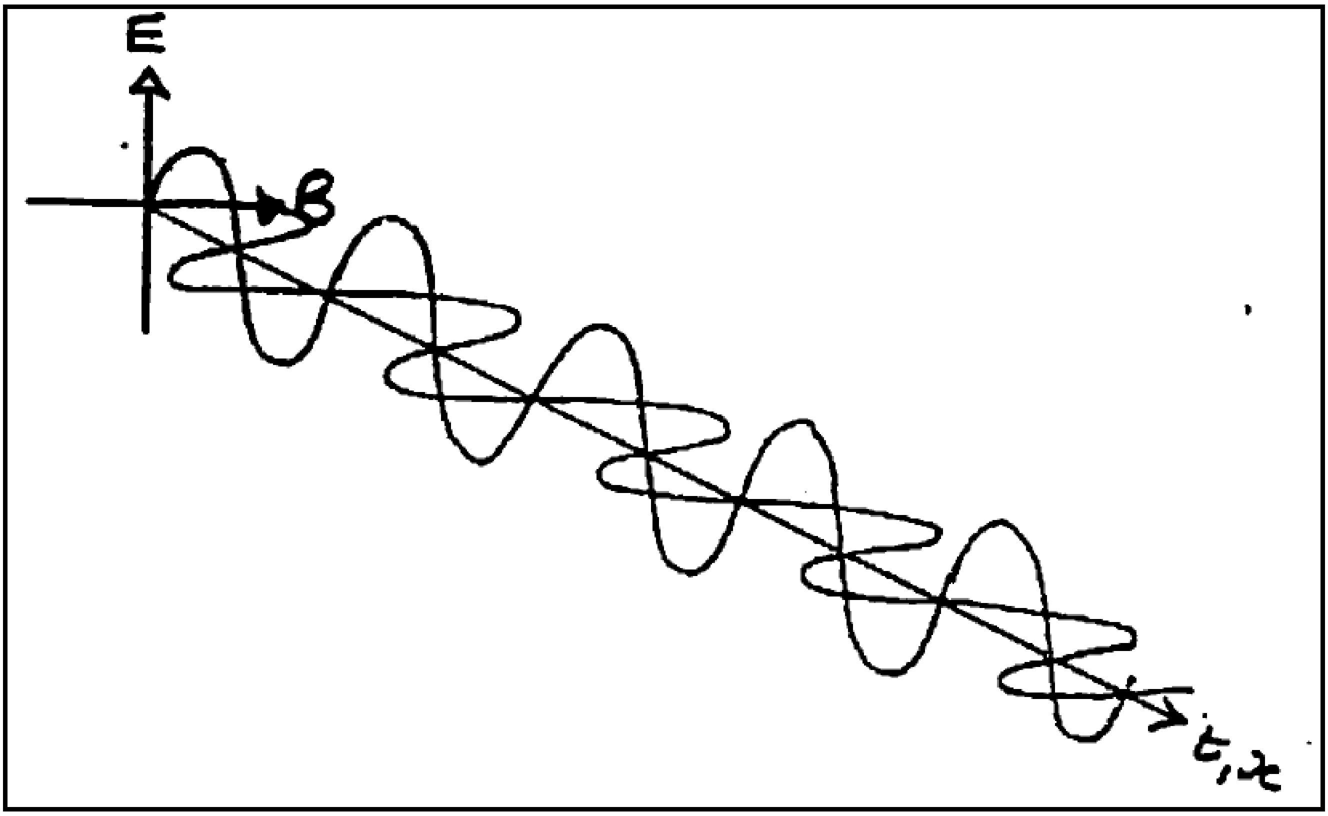
The magnitude and direction of B depends on the speed and direction (velocity) of the electron movement.



Generation of an electromagnetic wave

We can see that the magnitude of the magnetic field is changing with time. According to *Maxwell's equations* a changing magnetic field generates a changing electric field, which generates a changing magnetic field, which generates a changing electric field, and so on. These coupled electric and magnetic fields propagate away from the antenna (recreating each other) as an electromagnetic wave. Electromagnetic waves are unusual in that they do not require a medium in order to propagate.

The maximum electric field is notionally greater than the maximum magnetic field by a factor of $c = 3\times10^8\ \mathrm{ms}^{-1}$. An electromagnetic wave is often represented, and analysed, in a simpler two dimensional form - by considering the electric field alone, although the magnetic field is always implicitly there.



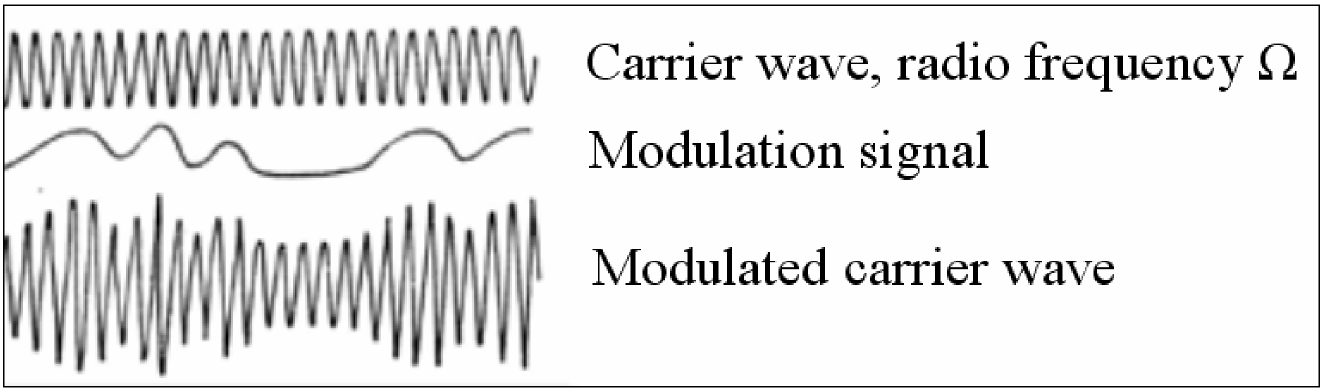
A propagating electromagnetic wave

The key properties to note here are that:

- the magnitudes of the electric and magnetic fields both vary *sinusoidally* with time;
- both fields are associated with particular planes. In free space, the two fields are *in phase* with each other and the magnetic plane is *orthogonal* (at 90°) to the electric plane;
- the *propagation direction* (direction of travel) of the wave is along the line of intersection of the two planes.

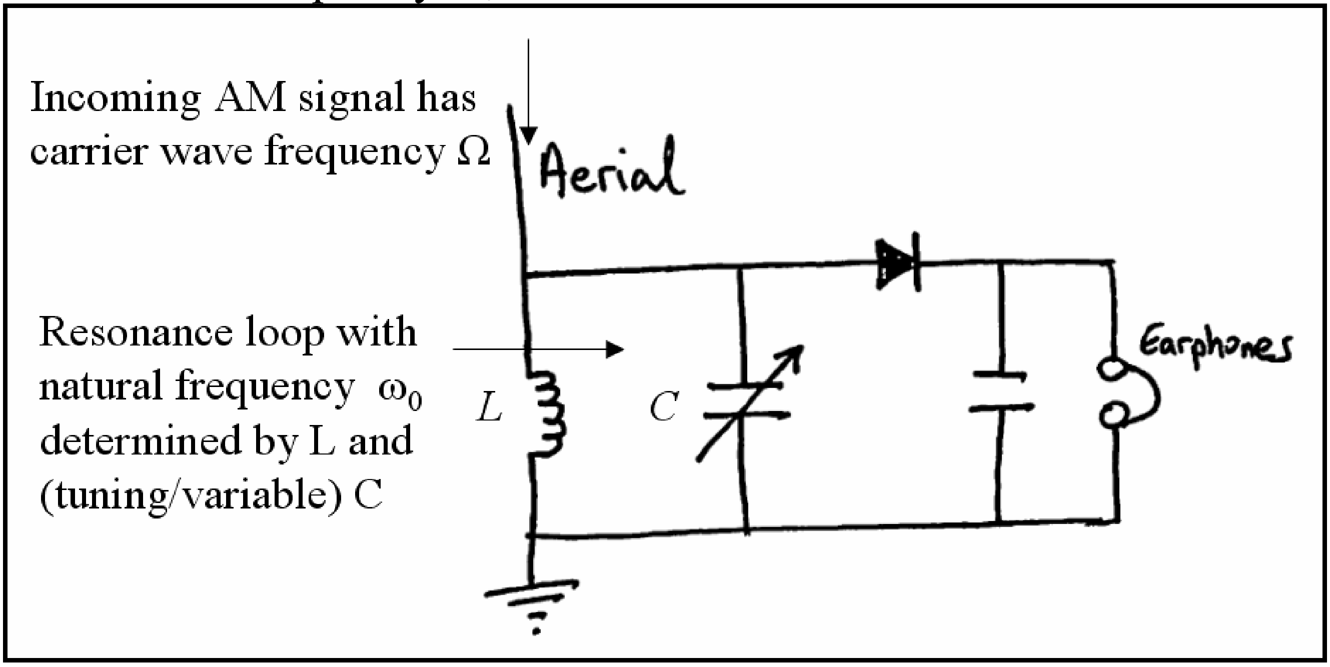
Note that the actual direction of energy travel is given by the vector product ***E*** x ***B*** (which is into the page in the above diagram).

Returning to our optoelectronic system, the de-encoding scheme will depend upon which type of modulation was used to encode the information. A simple case, along the lines of early radio transmissions, is where the amplitude of the carrier (the electromagnetic wave) was modulated, i.e. an AM transmission scheme.



Amplitude Modulation (AM)

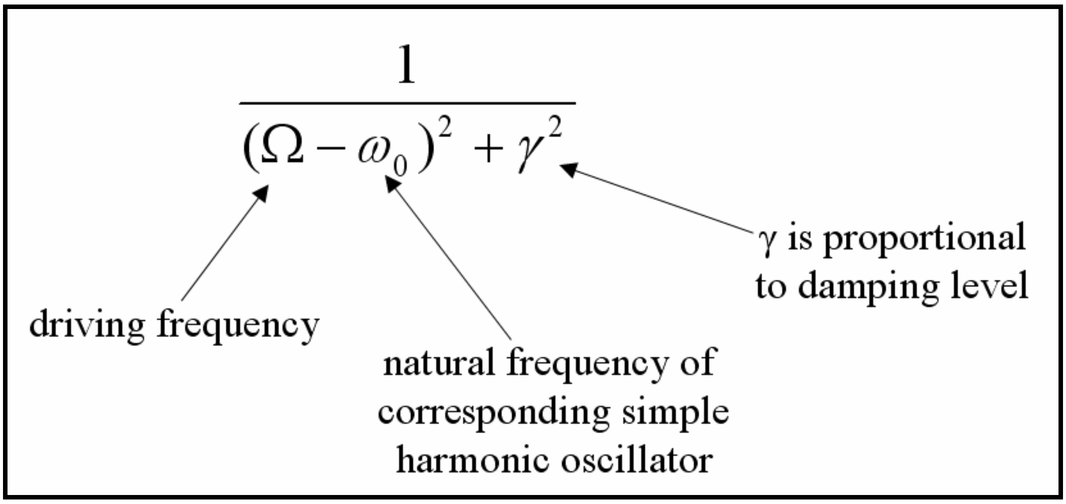
In the circuit below, an inductor (effectively a coil with inductance *L*) and a variable capacitor (with capacitance *C*) create an oscillator with natural frequency ω_0.



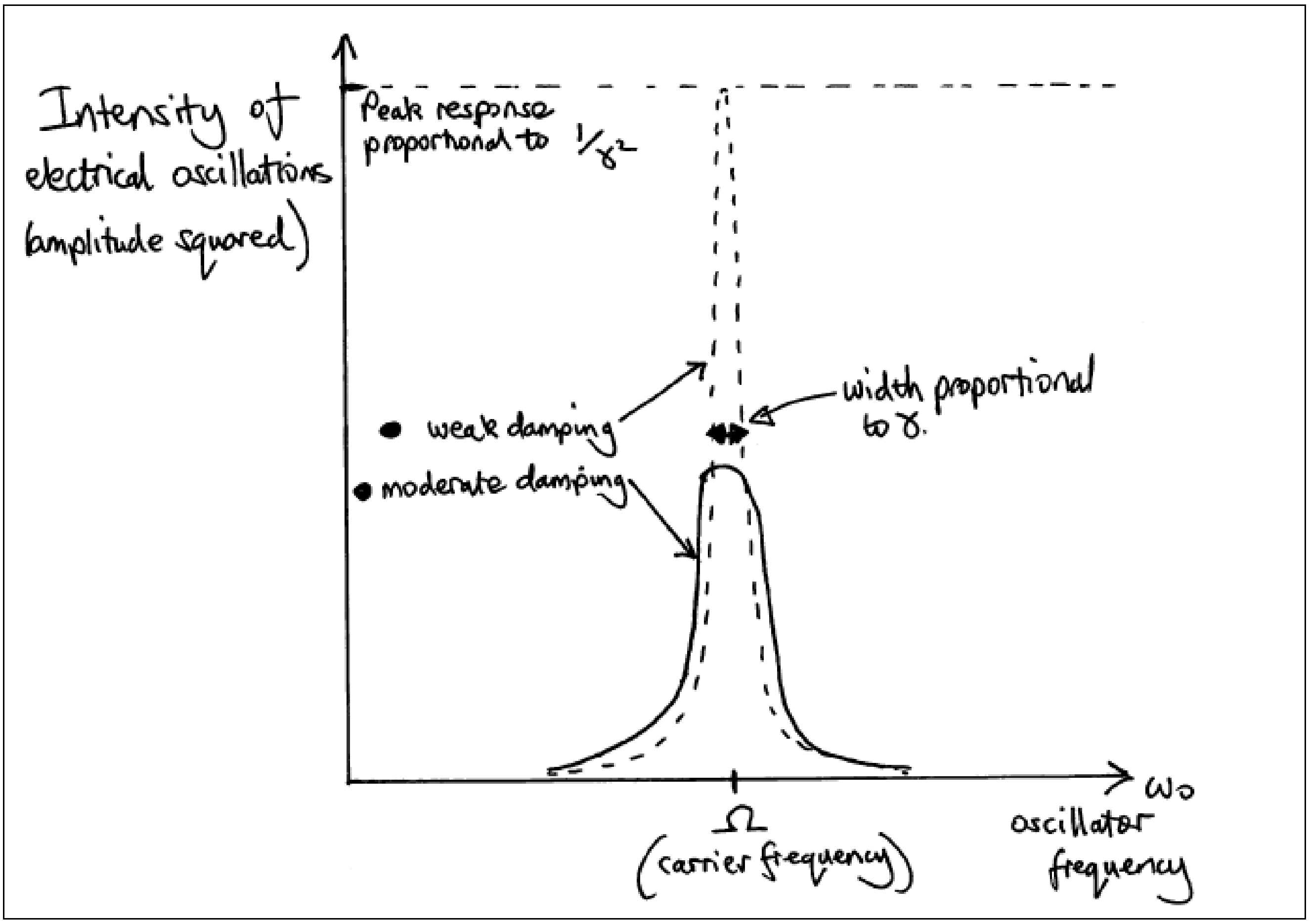
Simple circuit for receiving an amplitude-modulated carrier wave

The amplitude of electrical oscillations in this loop is maximal when the capacitor is tuned to give a resonant matching of the electronic oscillator frequency ω_0 with that of the incoming carrier frequency Ω. The diode allows for a one-way sampling of the resonance loop - the earphones section doesn't feedback ac signals into the aerial oscillator loop. A final capacitor, often in parallel with an additional resistor, is used to filter out the carrier wave – leaving only the modulation signal.

There will also be some damping γ of the electronic oscillations, as there will be some resistance in this circuit. We thus have a classic ***driven-damped oscillator problem***: the *intensity* of oscillations is proportional to:



Intensity of oscillations in driven-damped oscillators



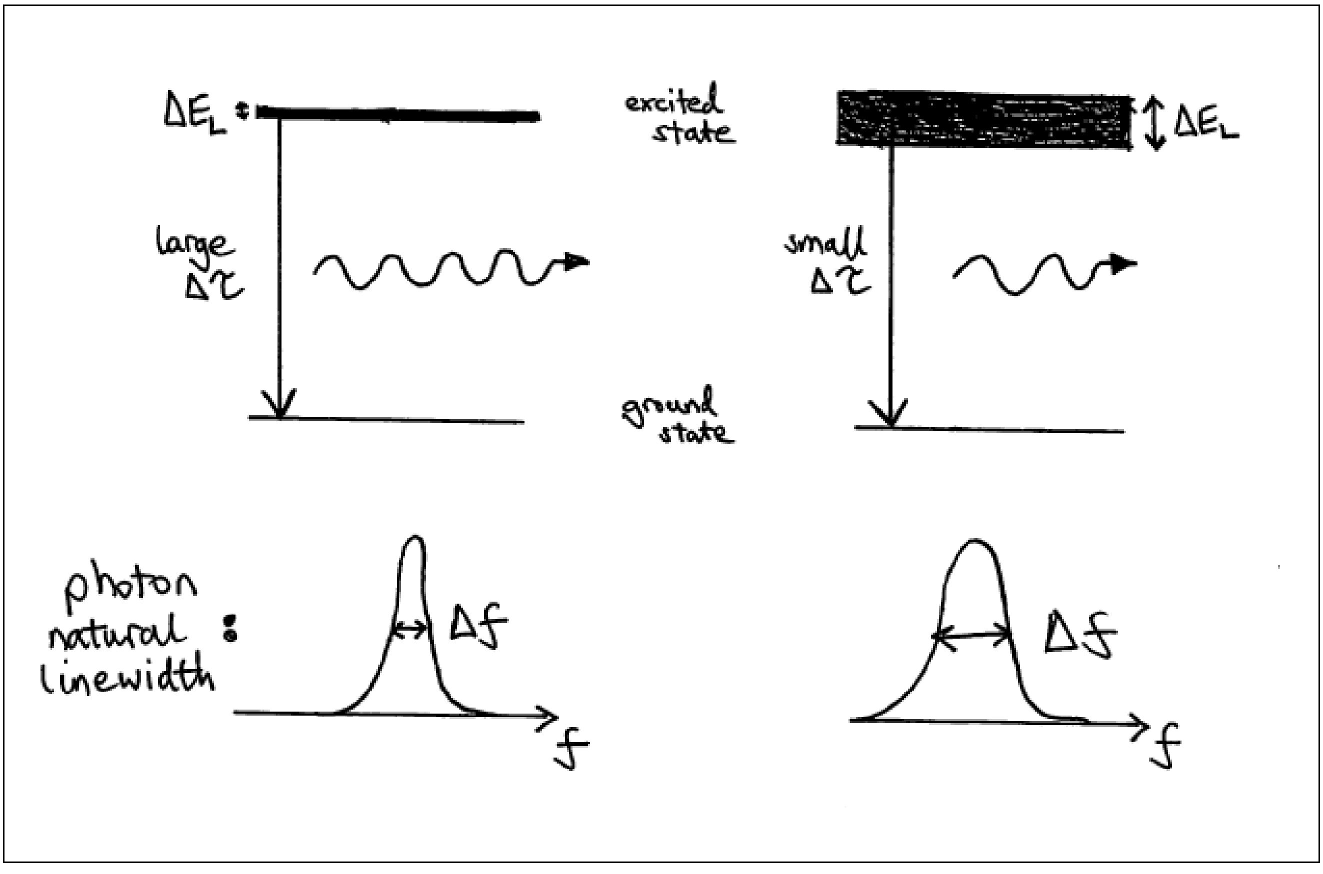
Sketch of resonant response for two levels of damping

1.4 Natural lifetime of excited states

So how do we generate an electromagnetic wave by an electron simply falling from a higher energy level to a lower one? The answer is that the electron does not make a single hop but oscillates backwards and forwards between the two energy levels until it finally comes to rest in the lower one. This raises an interesting concept, that is the *length* of a photon, since the hop takes a finite mount of time, during which the photon is being produced. An important parameter is the ***lifetime, $\Delta\tau$***, of the excited state. This is the amount of time the atom is expected to spend in the excited state, before it loses its excess energy. This (natural) lifetime is governed by the ***uncertainty principle*** of quantum mechanics, and is determined by the energy width ΔE_L *of* the excited state.

$$\boxed{\Delta E_L \Delta\tau \approx \frac{h}{2\pi} \qquad (2)}$$

From this equation, if $\Delta\tau$ is small then ΔE_L must be large.



Faster electron de-excitations result in a broader linewidth

Example

If the excited state lifetime is $\Delta\tau = 10$ ns then what is the natural linewidth of the excited state energy level (and hence also the linewidth of any generated photon) *in frequency units*, Δf ?

Uncertainty Principle: $\Delta E_L = \dfrac{h}{2\pi\Delta\tau}$ (in energy units)

Planck's Hypothesis: $E = hf$ also gives $\Delta E_L = h\Delta f$

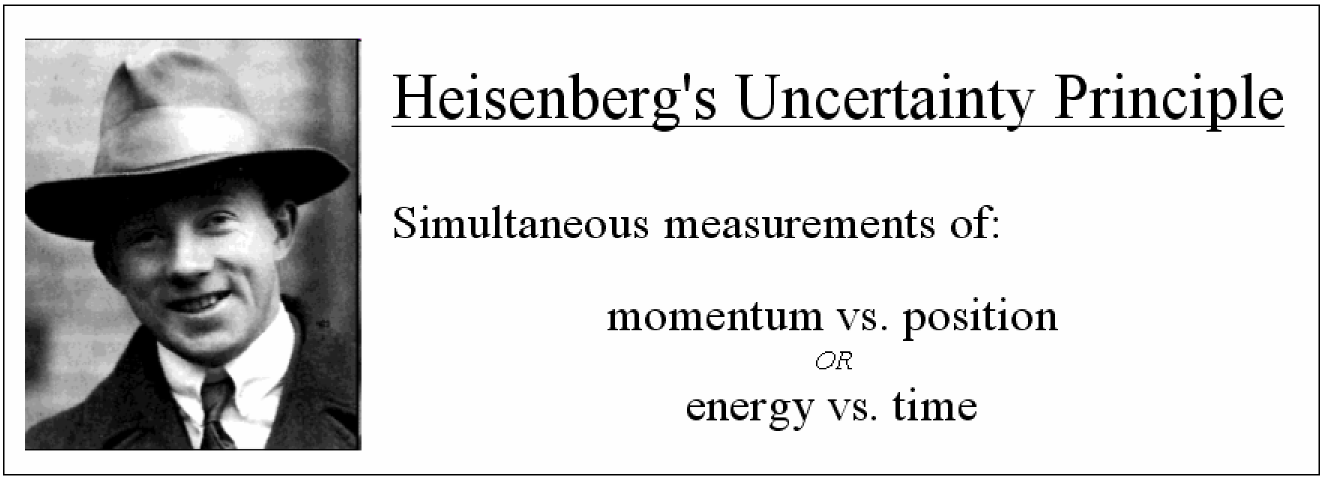
So, $h\Delta f = \dfrac{h}{2\pi\Delta\tau}$ and thus:

$$\boxed{\Delta f = \frac{1}{2\pi\Delta\tau}} \tag{3}$$

Here, $\Delta\tau = 10 \text{ ns} = 10\times10^{-9} \text{ s} = 10^{-8}\text{s}$

giving $\Delta f \approx 0.16\times10^{8} \text{ Hz} = 16\times10^{6} \text{ Hz}$
$= 16 \text{ MHz}$.

Further Information



Heisenberg's Uncertainty Principle

Simultaneous measurements of:

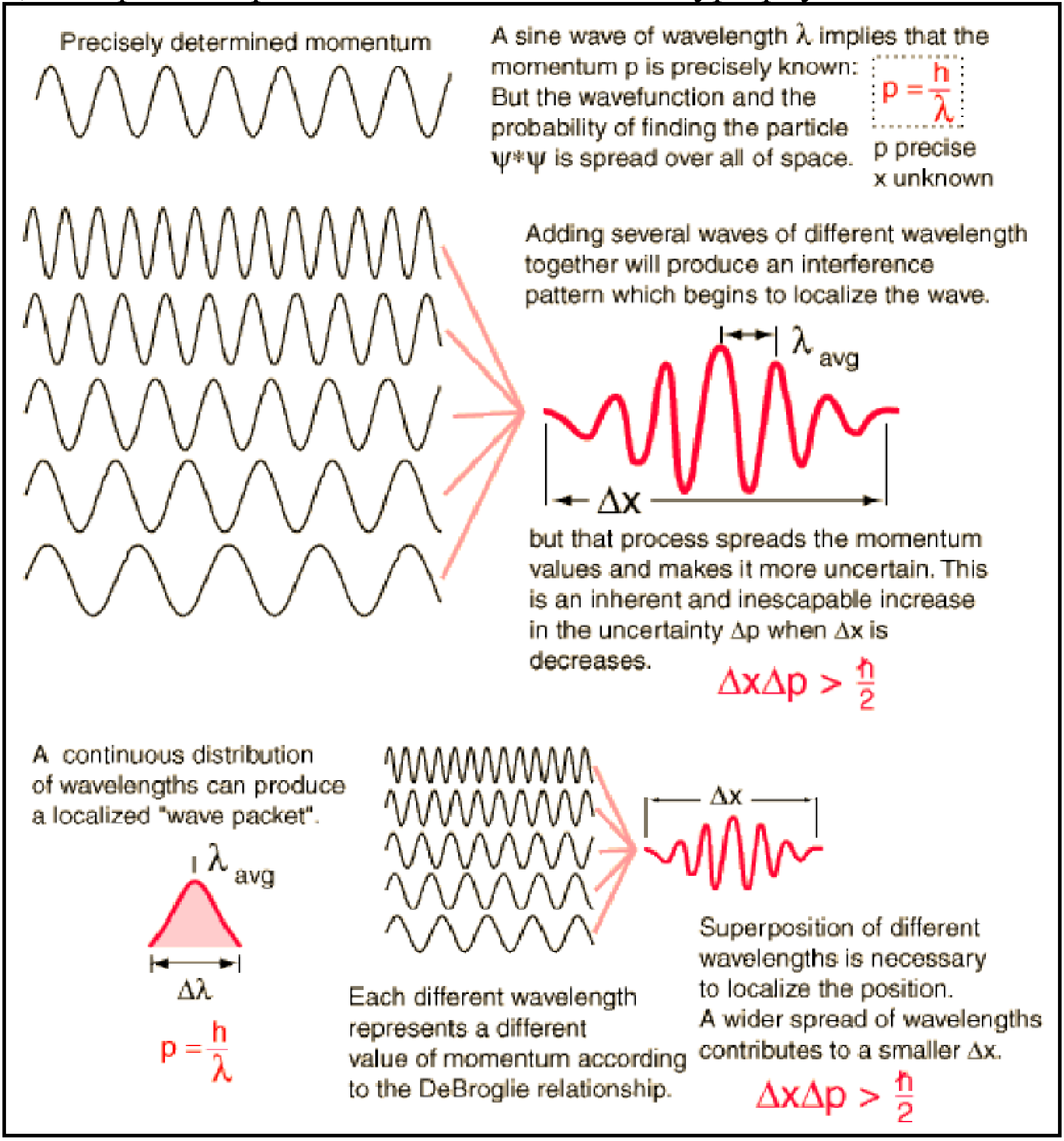
momentum vs. position

OR

energy vs. time

Heisenberg's uncertainty principle takes many forms, involving different pairs of complementary variables, such as energy and time. It can be thought of as arising from the wave character of matter – giving rise to trade-offs in measurement uncertainties.

The following nice account - of the trade-off between the simultaneous determination of the momentum, *p*, and the position, *x*, of a quantum particle – is taken from the hyperphysics website:



A sine wave of wavelength λ implies that the momentum p is precisely known: But the wavefunction and the probability of finding the particle $\psi^*\psi$ is spread over all of space.

$$p = \frac{h}{\lambda}$$

p precise
x unknown

Adding several waves of different wavelength together will produce an interference pattern which begins to localize the wave.

but that process spreads the momentum values and makes it more uncertain. This is an inherent and inescapable increase in the uncertainty Δp when Δx is decreases.

$$\Delta x \Delta p > \frac{\hbar}{2}$$

A continuous distribution of wavelengths can produce a localized "wave packet".

$$p = \frac{h}{\lambda}$$

Each different wavelength represents a different value of momentum according to the DeBroglie relationship.

Superposition of different wavelengths is necessary to localize the position. A wider spread of wavelengths contributes to a smaller Δx.

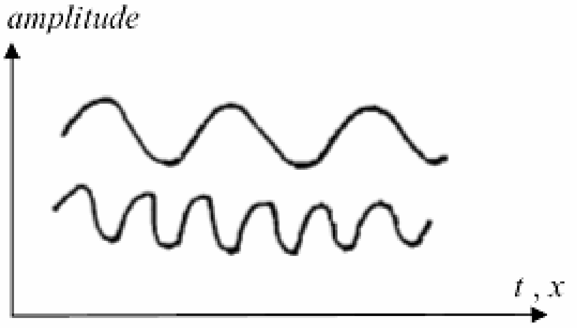
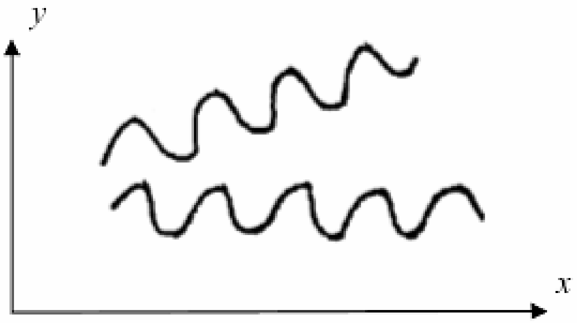
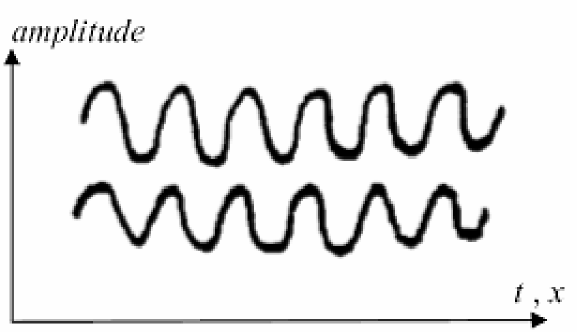
$$\Delta x \Delta p > \frac{\hbar}{2}$$

Notes

- Different statements of the uncertainty principle can vary by factors of 2.
- In the above, the 'h bar' symbol is used: $\hbar = h/2\pi$. In terms of this, $E = hf$ becomes $E = \hbar\omega$ and $p = h/\lambda$ becomes $p = \hbar k$.

1.5 The Question of Coherence

We are now familiar with the idea that a photon is a packet of electromagnetic energy. Each photon has associated with it a given wavelength and it is of a certain length. When we are considering the question of coherence then we have to look in more detail about the relationship between the many photons that are present in a beam of light. Coherence is, therefore, not a property of an individual photon but of a collection of photons, and it is a statement about the ***phase relationship*** between a photon and its neighbours. Consider the following diagrams:

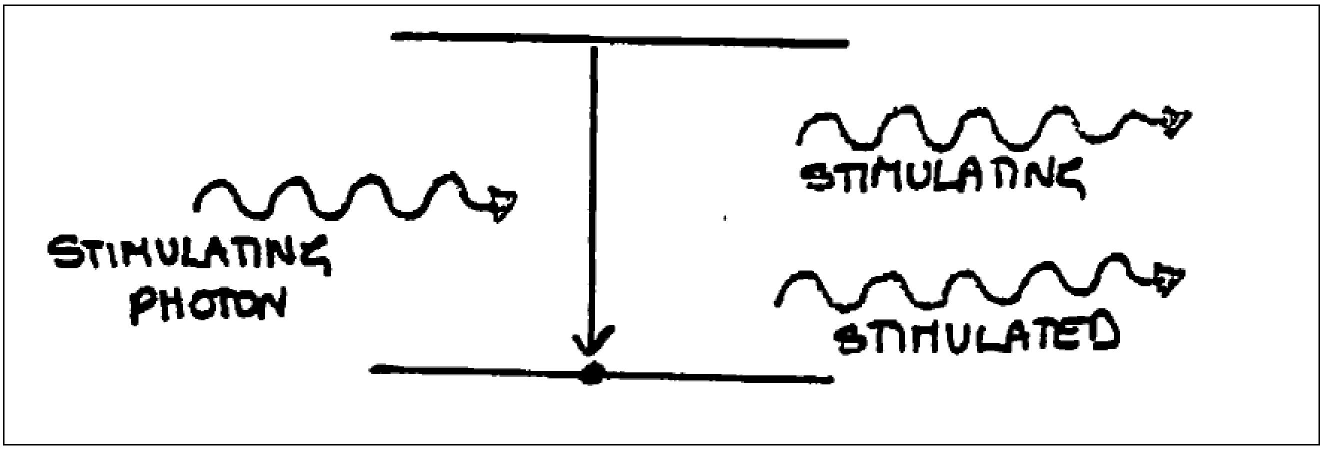
	Both photons start in phase but they quickly become out of phase because they have different wavelengths.
	Both photons start in phase and have the same wavelength. However, one photon is travelling at an angle to the other and thus they move out of phase.
	Both photons start in phase, have the same wavelength and are travelling in the same direction. Thus they remain in phase at all times.

The upper two diagrams represent beams of light which are ***incoherent***. Only the bottom diagram represents a beam of light which is ***coherent***. Light produced by lasers has the potential to be coherent. We shall come back to this characteristic in more detail when we have learnt something about how laser light is produced.

1.6 Laser Light

The word 'LASER' is an acronym for **L**ight **A**mplification by **S**timulated **E**mission of **R**adiation. Two key words in this acronym are 'stimulated emission', for it is this process that makes laser light different from other light.

So far, we have only considered the production of light only by the process of ***spontaneous emission***. In other words, our excited atom has given up its excess energy in its own good time. In the process of ***stimulated emission*** an excited atom is persuaded to produce a photon by another photon of the correct energy passing by.



Stimulated emission of a photon.

When light is produced in this manner there is a special relationship between the properties of the stimulating and stimulated photons.

They have:

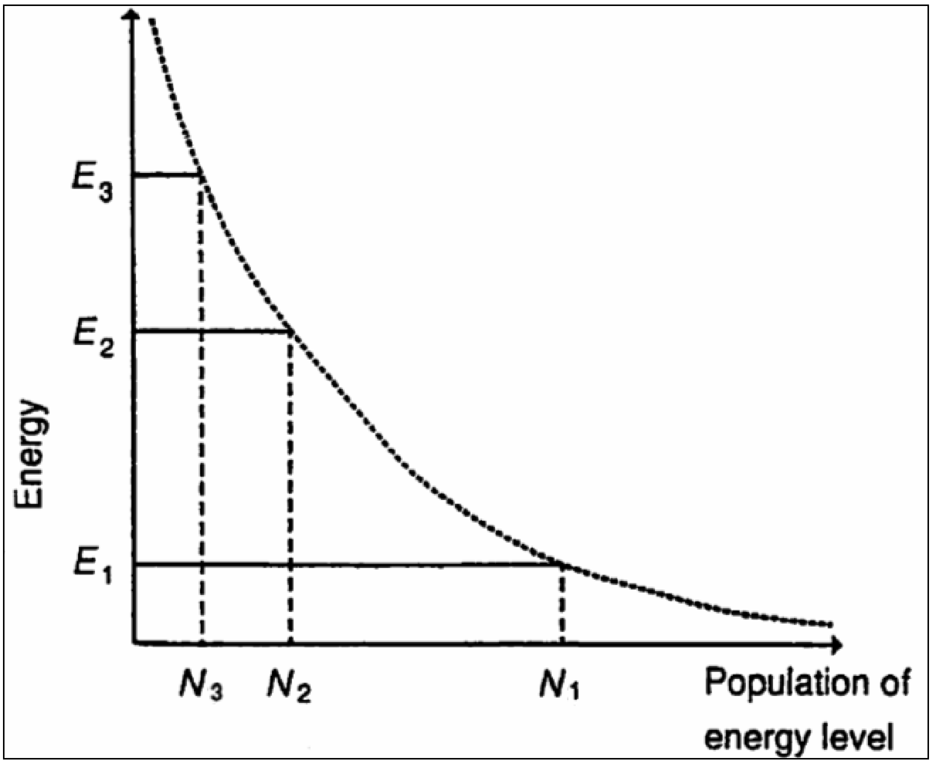
- the same energy (and hence the same wavelength)
- the same phase
- the same direction
- the same polarisation (the electric fields of both photons are oscillating in the same plane).

1.7 Energy at thermal equilibrium: the Boltzmann distribution

Consider a collection of atoms at an absolute temperature of T (kelvins) > 0. At any one time, some of the atoms will have their outer electrons in ground state E_1 while others will have outer electrons in higher energy states E_2 , E_3 , E_4, … etc. due to having been excited by ambient thermal energy. In ***thermal equilibrium***, the overall amount of thermal excitation of atoms reaches a balance with atoms de-exciting and releasing energy. In such equilibrium, a plot of the number of atoms per unit volume versus energy gives a ***Boltzmann distribution***:

$$N_i = N_0 \exp\left(\frac{-E_i}{k_B T}\right) \quad , \quad i = 1, 2, 3, \ldots \qquad (4)$$

where N_i = population density of the i^{th} energy level (energy E_i)
k_B = Boltzmann's constant = 1.381×10^{-23} JK^{-1}
N_0 = normalisation factor (to give the correct total number of atoms per unit volume $N = N_1 + N_2 + N_3 + \ldots$)



Thermal equilibrium distribution of electrons $(T > 0)$

As we may not know the total volume density of atoms, N, a useful alternative form of the Boltzmann distribution expresses the ***ratio of populations*** of two energy levels:

$$\frac{N_2}{N_1}=\frac{N_0\exp\left(-E_2/k_BT\right)}{N_0\exp\left(-E_1/k_BT\right)}=\frac{\exp\left(-E_2/k_BT\right)}{\exp\left(-E_1/k_BT\right)}=\exp\left[\frac{-E_2}{k_BT}-\left(\frac{-E_1}{k_BT}\right)\right]$$

$$\text{i.e.}\quad \frac{N_2}{N_1}=\exp\left[-\frac{\left(E_2-E_1\right)}{k_BT}\right] \qquad (5)$$

Example

The de-excitation of an atom, by an electron falling from a level with energy E_2 to one with energy E_1, may result in the emission of a photon. If this happens when the system is in thermal equilibrium at temperature $T = 300$ K then what relative population of the levels is required for a visible photon?

Wavelengths in the visible spectrum extend from 400 nm to 700 nm, so an average wavelength is $\lambda = 550\text{ nm}$. The energy of this photon and the ratio of populations are:

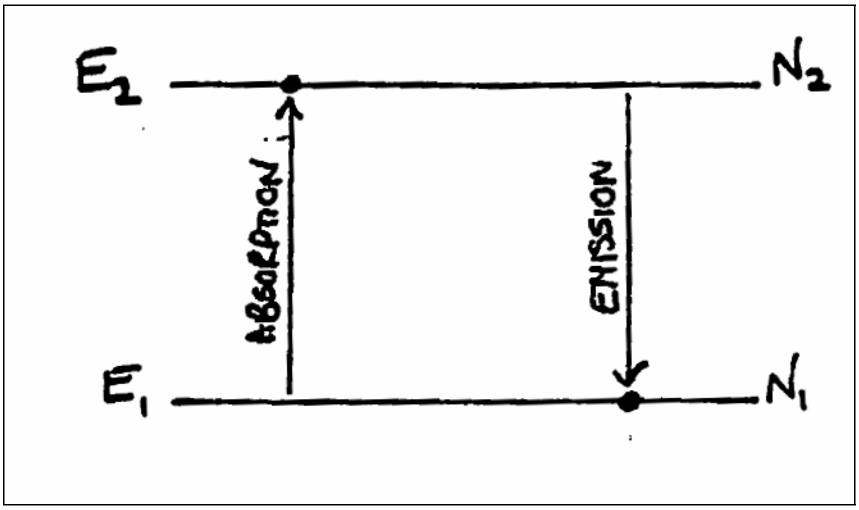
$$E_2-E_1=hf=h\frac{c}{\lambda}\approx 6.626\times10^{-34}\text{ Js }\frac{3\times10^{8}\text{ ms}^{-1}}{550\times10^{-9}\text{m}}\approx 3.6\times10^{-19}\text{ J},$$

$$\frac{N_2}{N_1}=\exp\left[\frac{-\left(E_2-E_1\right)}{k_BT}\right]\approx\exp\left[\frac{-3.6\times10^{-19}}{1.381\times10^{-23}\times300}\right]\approx\exp\left(-87\right)\approx10^{-37}$$

→ Thermal equilibrium contexts will have extremely few atoms with sufficient energy excitation for creating visible photons during de-excitation!

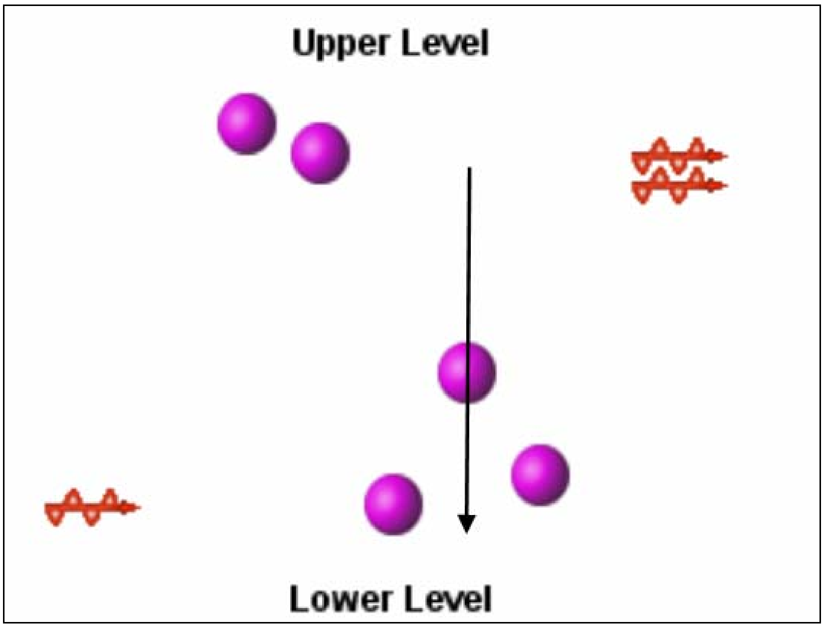
1.8 How can we make a material lase?

In any system where atoms are being excited and then losing their energy in the production of photons there are three energy processes going on: absorption, spontaneous emission and stimulated emission. Consider a simple two energy level system:



Absorption and emission in a two energy level system

The relative populations N_2 and N_1 of the upper and lower energy states, respectively, will be determined by a balance of the rates of the three energy processes occurring:



An electron falls just after a stimulated emission, and a further photon approaches that will be absorbed.

If we have N_1 atoms per unit volume with energy E_1 then the rate at which these can be excited to energy level E_2 depends on both N_1 and the density of the exciting source. In a system in thermal equilibrium, this exciting source may be photons of the correct energy produced by earlier emissions. These photons will have a radiation density that we can express as a number density ρ_f (no. photons per unit volume). Then:

$$\boxed{\textit{Rate of excitation} = N_1\rho_f B_{12} \qquad (6)}$$

Similarly, the rate of stimulated emission depends on N_2 and ρ_f via:

$$\boxed{\textit{Rate of stimulated emission} = N_2\rho_f B_{21} \qquad (7)}$$

Finally, the spontaneous emission rate is independent of photon density and is given by:

$$\boxed{\textit{Rate of spontaneous emission} = N_2 A_{21} \qquad (8)}$$

where $A_{21}=1/\Delta\tau$, B_{12}, and B_{21} are ***Einstein Coefficients.***
The Einstein coefficients tell us how likely a process is to occur.
In order for the system to be in equilibrium we require that:
the rate of absorption = total rate of emission, i.e.

$$\boxed{\begin{array}{ll} & N_1\rho_f B_{12} = N_2\rho_f B_{21} + N_2 A_{21} \\ \Rightarrow & -N_2 A_{21} = (N_2 - N_1)\rho_f B, \quad \text{when } B_{12} = B_{21} = B \\ \Rightarrow & (N_2 - N_1) < 0 \qquad\qquad\qquad (9) \end{array}}$$

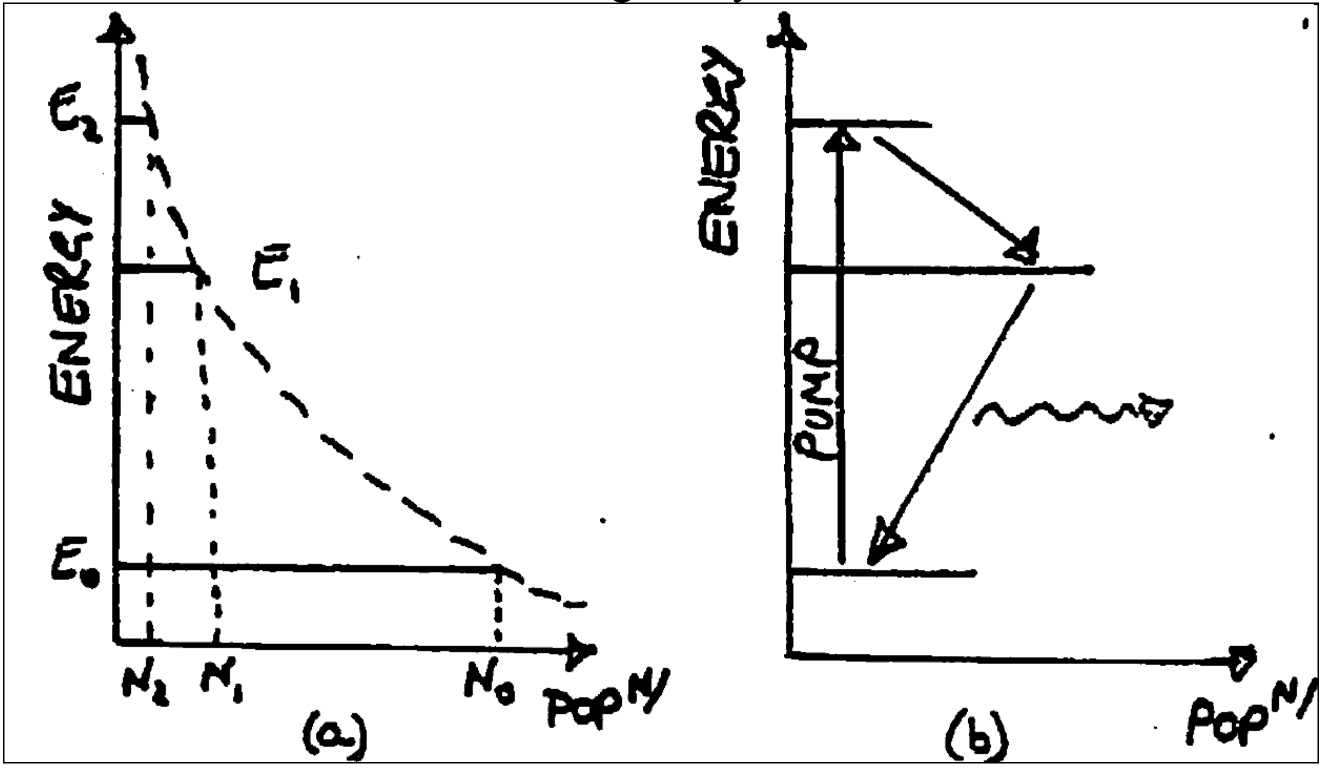
In a system in thermal equilibrium spontaneous emission dominates. In order to increase the proportion of photons produced via stimulated emission we must increase both the number of atoms in the excited state and the radiation density. In fact, in order to achieve a large enough proportion, and get the system to lase, we must make $N_2>N_1$. In other words we must achieve a ***population inversion***.

1.9 Population inversion

If we are to achieve a population inversion then we must supply a large amount of energy to the system, i.e. we must ***pump it***.

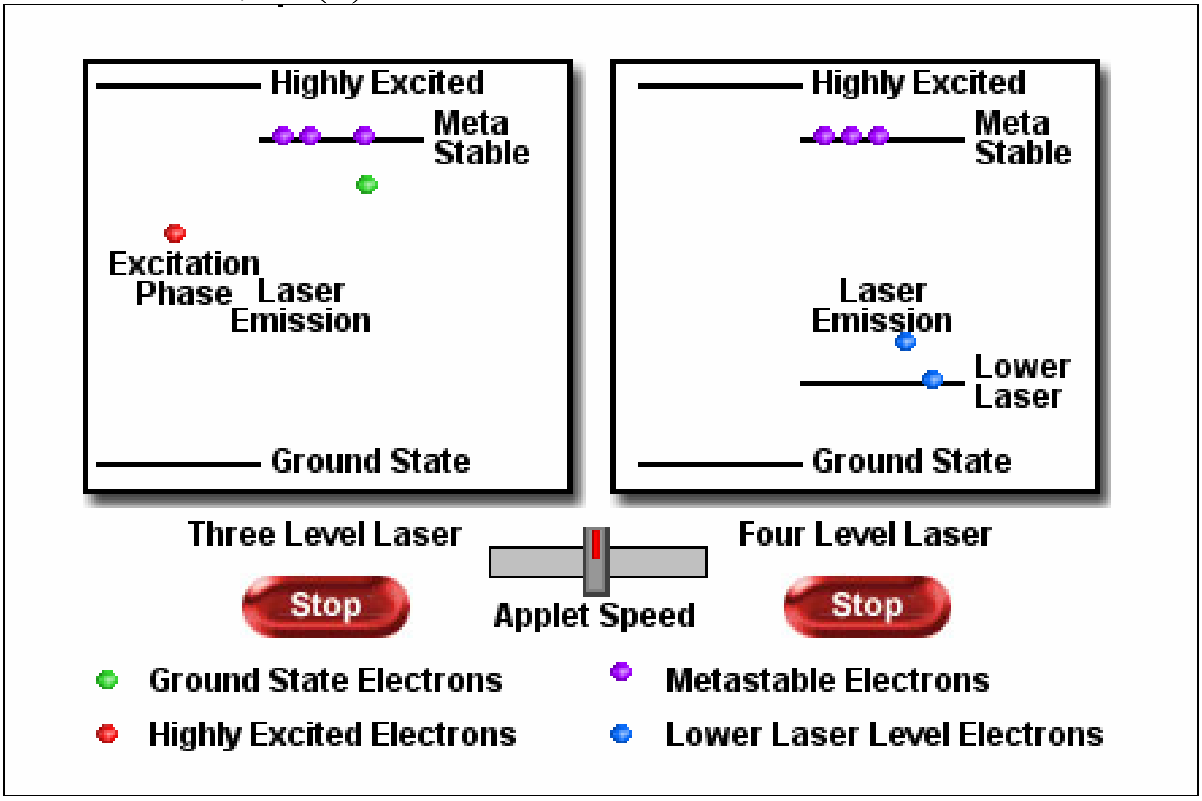
We can see this from equation (9): if we pump hard enough then we can make the spontaneous contribution negligible, through increasing ρ_f. However, since $B_{12} = B_{21}$, with our two level system the best we can do is to make $N_2 = N_1$. This is why virtually all lasers use transitions between at least three energy levels of the system.

Consider the three level system shown below. The equilibrium populations of the energy levels are determined by the Boltzmann distribution, shown in the left figure by the dashed line.



Populations of levels: (a) in equilibrium, (b) under pumping.

The pump source provides the lasing material with enough energy to move electrons up *two* energy levels. Electrons in E_2 then decay down to E_1. This is a ***non-radiative decay*** i.e. the energy is lost, for example, in heating up the laser material. The lasing transition is from E_1 to E_0. Ideally, the transition from E_2 to E_1 should be very quick and the transition from E_1 to E_0 very slow. This means that the electrons will arrive at E_1 faster than they can go down to E_0. E_1 must, therefore, usually be a long-lived state called a ***metastable state***. Population inversion has clearly been achieved between levels E_1 and E_0 in (b) above.

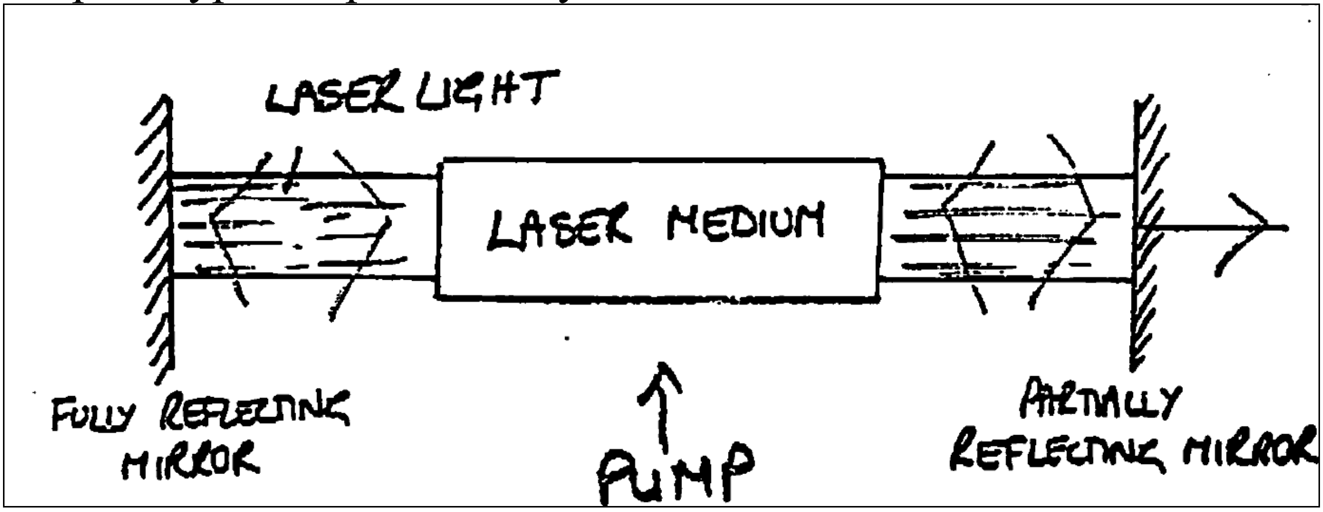


Simulations of 3-level and 4-level systems (shown in the lecture)

It is worth noting that it is better if E_2 is a *band* of closely-spaced energy levels. This makes better use of the pump by absorbing a greater range of energies, and therefore improves the efficiency of the system. Metastable states occur because, under the rules of quantum mechanics, some transitions are ***forbidden***. In reality, they can occur but it takes a longer time that expected from equation (2) for the transition happen.

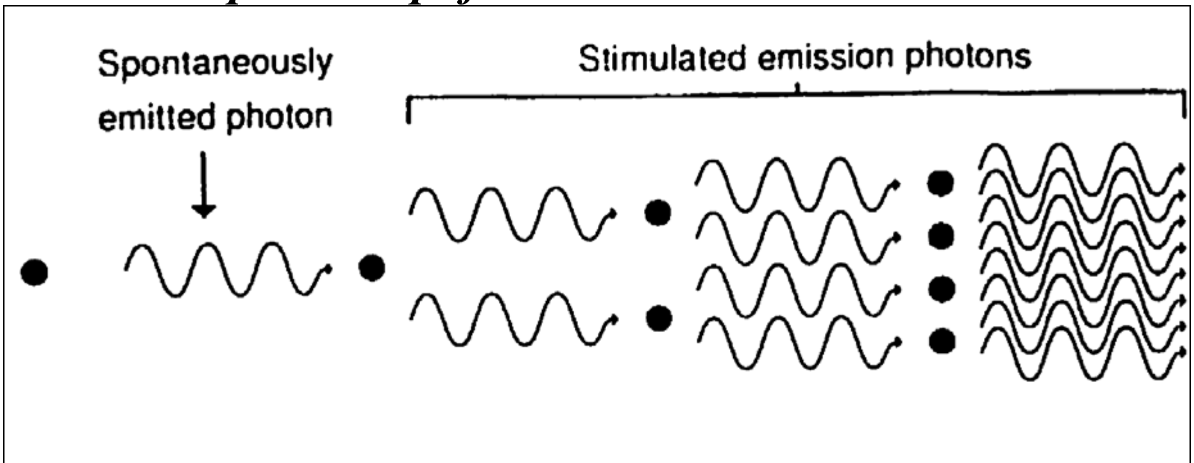
1.10 Maintaining the radiation density

You may recall that there are two criteria for obtaining dominant stimulated emission: the first is achieving population inversion and the second is getting a high radiation density, ρ_f. We can do this partly by pouring lots of energy into the system to generate lots of photons, but once the photons have been generated the density must be maintained at a high level: the photons must not be allowed mostly to escape, but have to be kept where we want them. This is achieved by containing the lasing material in an ***optical cavity***. The simplest type of optical cavity is shown below.



Basic optical cavity

Let us consider a single photon generated spontaneously in the centre of the lasing medium. As this photon travels it will pass by other excited atoms and stimulate them to produce photons with the same energy, direction, phase and polarisation as itself. As these photons reach one of the end mirrors they are reflected back into the cavity to stimulate the production of more photons. Upon reaching the other end mirror they will be reflected back again, producing more photons. Each photon produced by the first photon can stimulate the production of more photons, which can then stimulate the production of even more photons. This is known as a ***cascade*** and results in ***optical amplification***.



The cascade process.

This amplification process is very important, as this is what gives laser light the properties with which it is usually associated, in other words light that is:

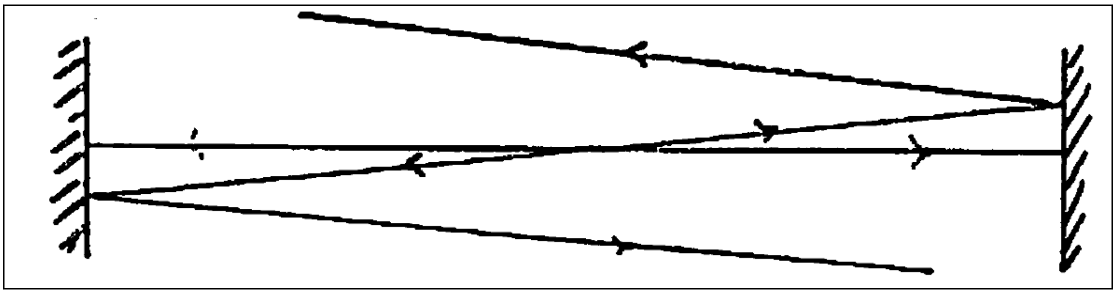
- travelling in one direction (a pencil beam)
- of a single frequency
- polarised
- coherent

1.11 Making laser light be what we expect

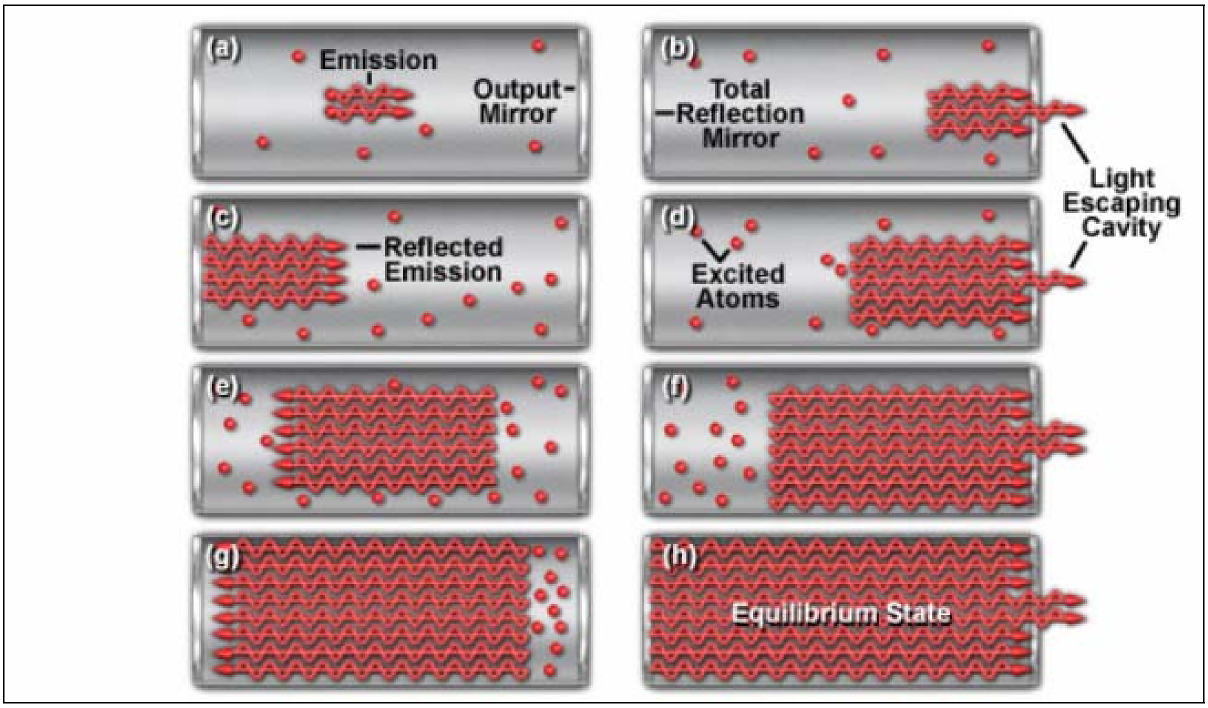
In general, lasers do not produce light with the properties that we expect unless we set them up to do so. It is important to remember that it is highly unlikely that only a single spontaneous emission will occur in the time that the laser takes to 'get going'. Instead, many spontaneously produced photons may be present, all with different: (1) *directions*, (2) *frequencies*, (3) *polarisations*, and (4) *phases* … and all with the potential to reproduce. The trick is in reproducing only those we want at the expense of those we don't. To do this we use the amplification process as a form of natural selection: the more photons of a particular 'species' there are, the more they can produce.

1.11.1 Light all travelling in the same *direction*

If we use mirrors to maintain a high radiation density, then the photons that will remain longest in the cavity are those that are travelling *along its axis*. Photons that are travelling at an angle to the cavity axis will eventually be reflected outside of the cavity region by one of the mirrors. The natural selection process of amplification therefore ensures that, eventually, all photons travel in the same direction.



The effect of mirrors on the path of photons in the cavity

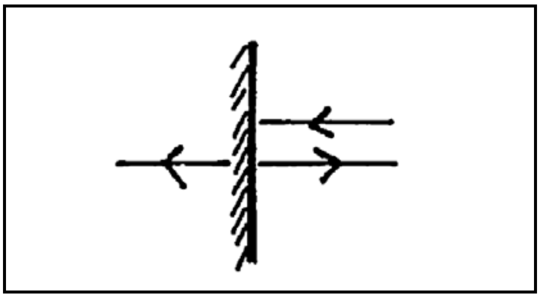


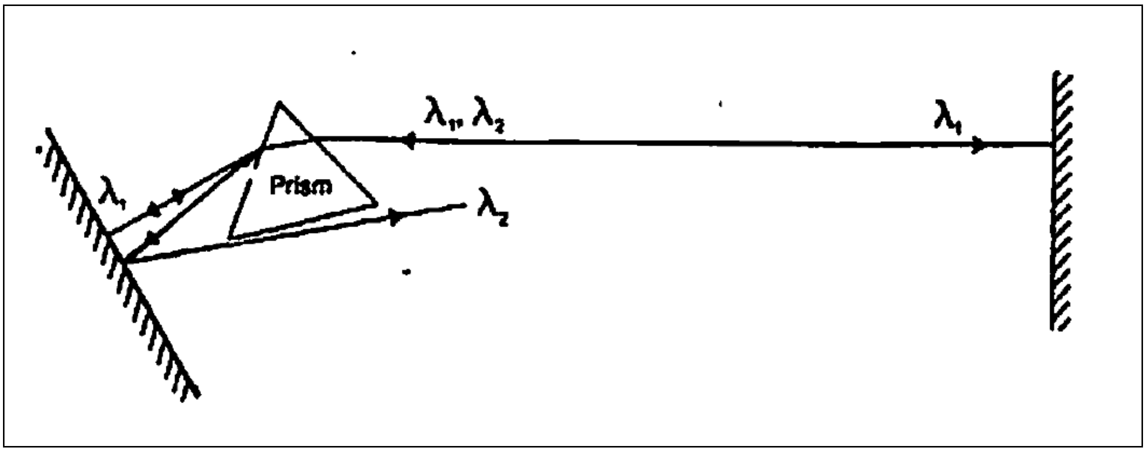
Time sequence, where mirrors select stimulated emission direction

1.11.2 Light all of a single *frequency*

Many lasers will quite happily produce laser light at a number of wavelengths unless we encourage them not to do so.

If a laser is only to lase at one wavelength, then we can coat the end mirrors so that only light of that wavelength is reflected back into the cavity.



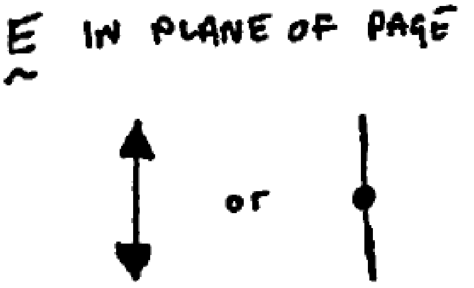
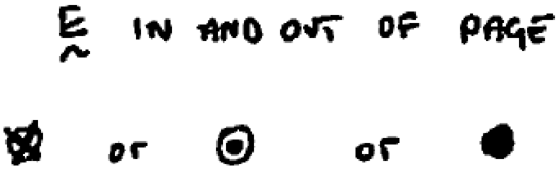
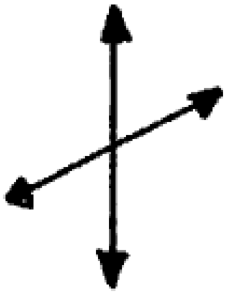


If we want flexibility in choice of wavelength, then we can insert a rotatable prism. In both of these cases, the natural selection of the amplification process does the rest.

1.11.3 Light all of a single *polarisation*

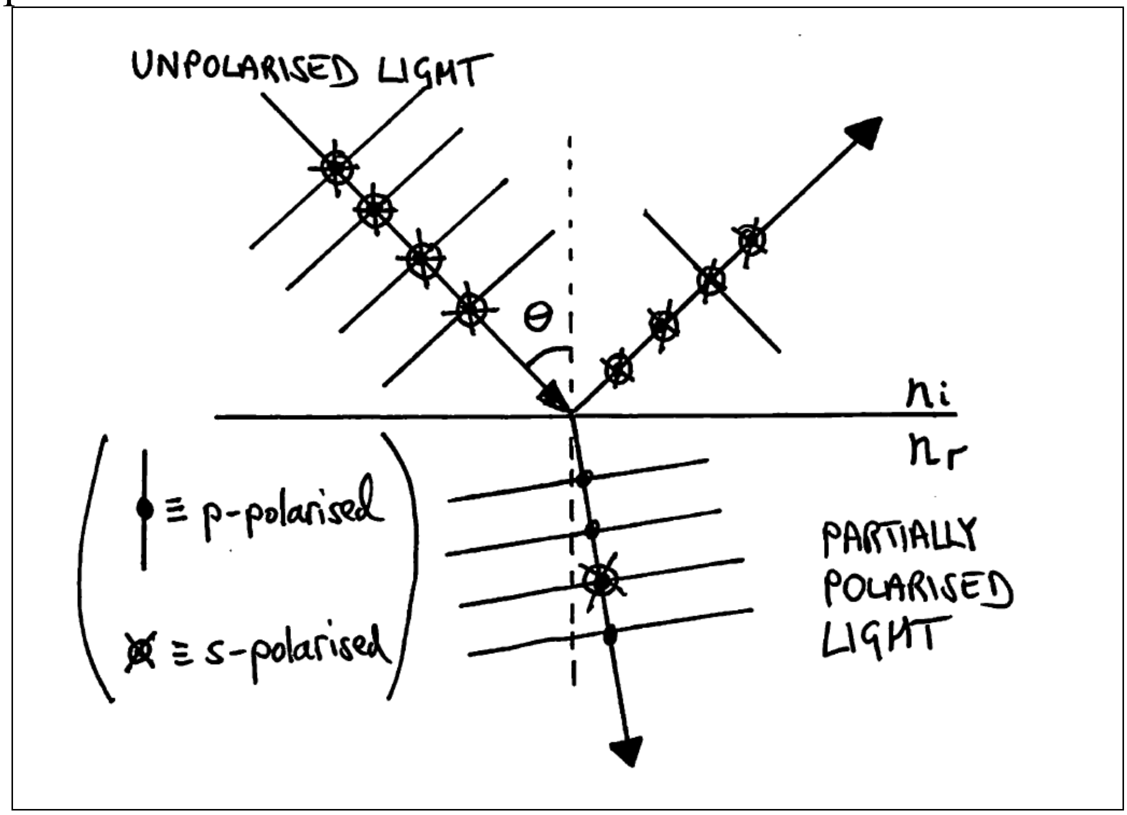
As with the quest of single-wavelength lasing, to produce light of a single polarisation means that we must add, or modify, optical components in the cavity. We cannot just insert a piece of Polaroid, as this absorbs too much light (offsetting optical amplification). We can polarise light, however, by reflection or refraction at a smooth surface.

Consider a beam of light whose direction of travel is along the *z*-axis. If the beam is *unpolarised*, then the plane of oscillation of the *electric field of any photon may be at any angle* between 0° and 360° to the *x*-axis. The only constraint is that the direction of travel of each photon is along the *z*-axis.

The orientation of the polarisation plane of electric field oscillation of a single photon is usually represented by a double arrowed line.	
A circle containing a dot or a cross can be used to represent polarised light with electric field oscillating into or out of the page.	
Unpolarised light is often represented by two double arrowed lines at 90° to each other.	 **OR** 

Representations of light polarisation

Now consider a beam of unpolarised light reflecting from a transparent surface.

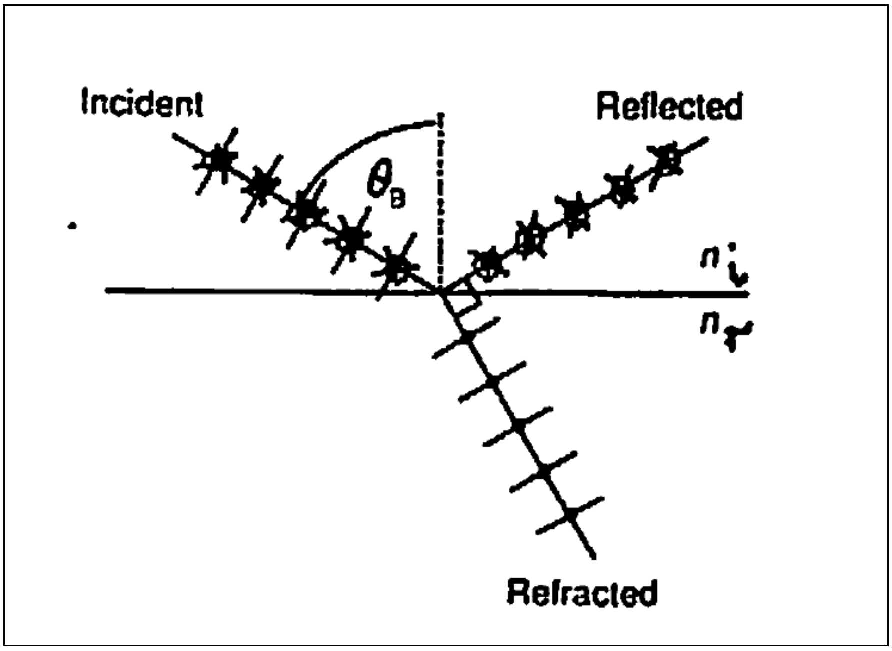


Refraction & reflection of unpolarised light at a transparent surface

If the surface is transparent, then some of the light will be reflected and some will be refracted. If we analysed the polarisation of these two beams, then we would see that:

- the refracted light contained more photons polarised in the plane containing the incident, reflected and refracted beams, the so-called ***p-plane*** (or 'plane of incidence');
- the reflected beam would contain more photons polarised in the direction perpendicular to this plane, the so-called ***s-plane*** (coming 'out of the page').

This effect is a function of angle, and at a particular incidence angle, called ***Brewster's angle*** (or *the polarising angle*), nearly all of the light in the reflected beam is polarised in the s-plane. At this special angle, the reflected and refracted beams are at 90° to one another.

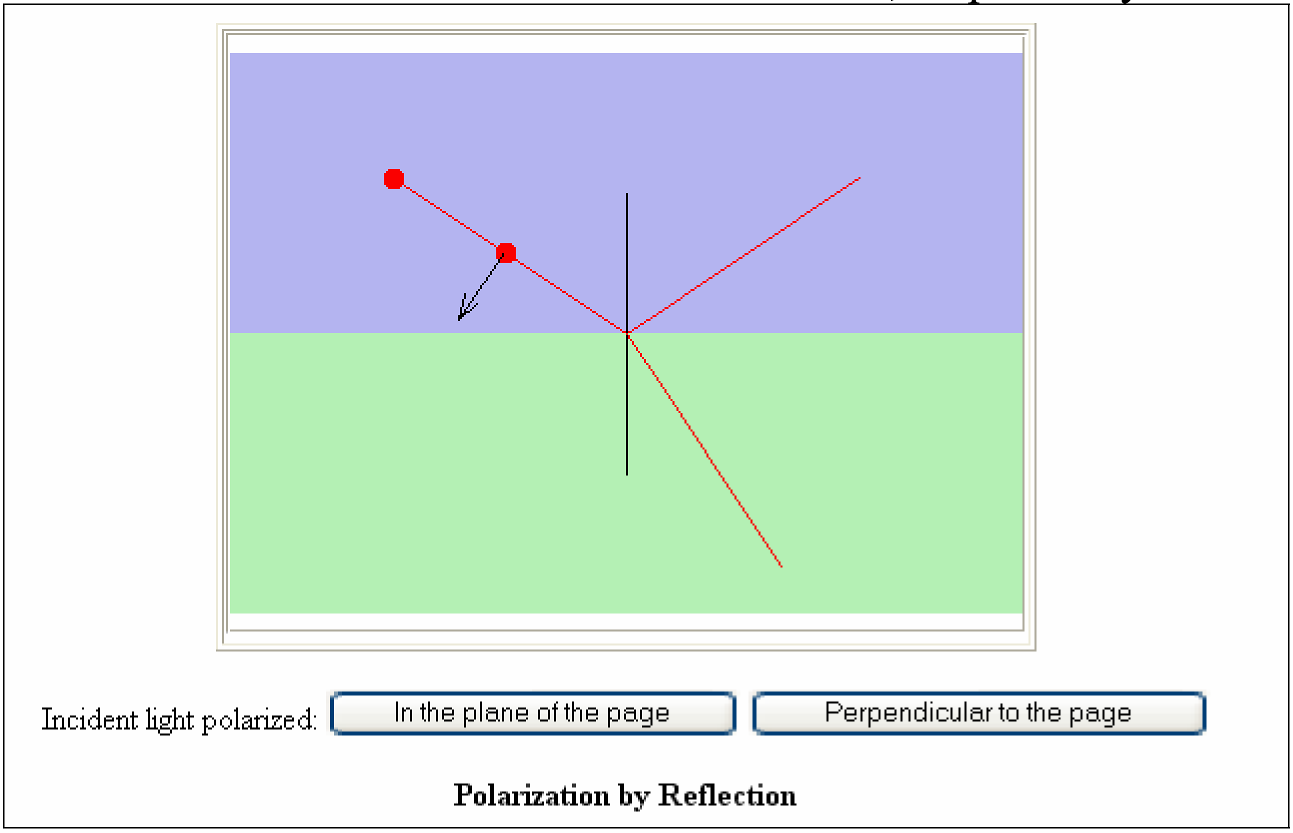


Polarising by refraction at Brewster's angle

Brewster's angle, θ_B , is given by the equation:

$$\tan\theta_B = \frac{n_r}{n_i} \qquad (10)$$

where n_i and n_r are the refractive indices of the materials on the incident and refracted sides of the interface, respectively.



Simulation of polarisation in the lecture

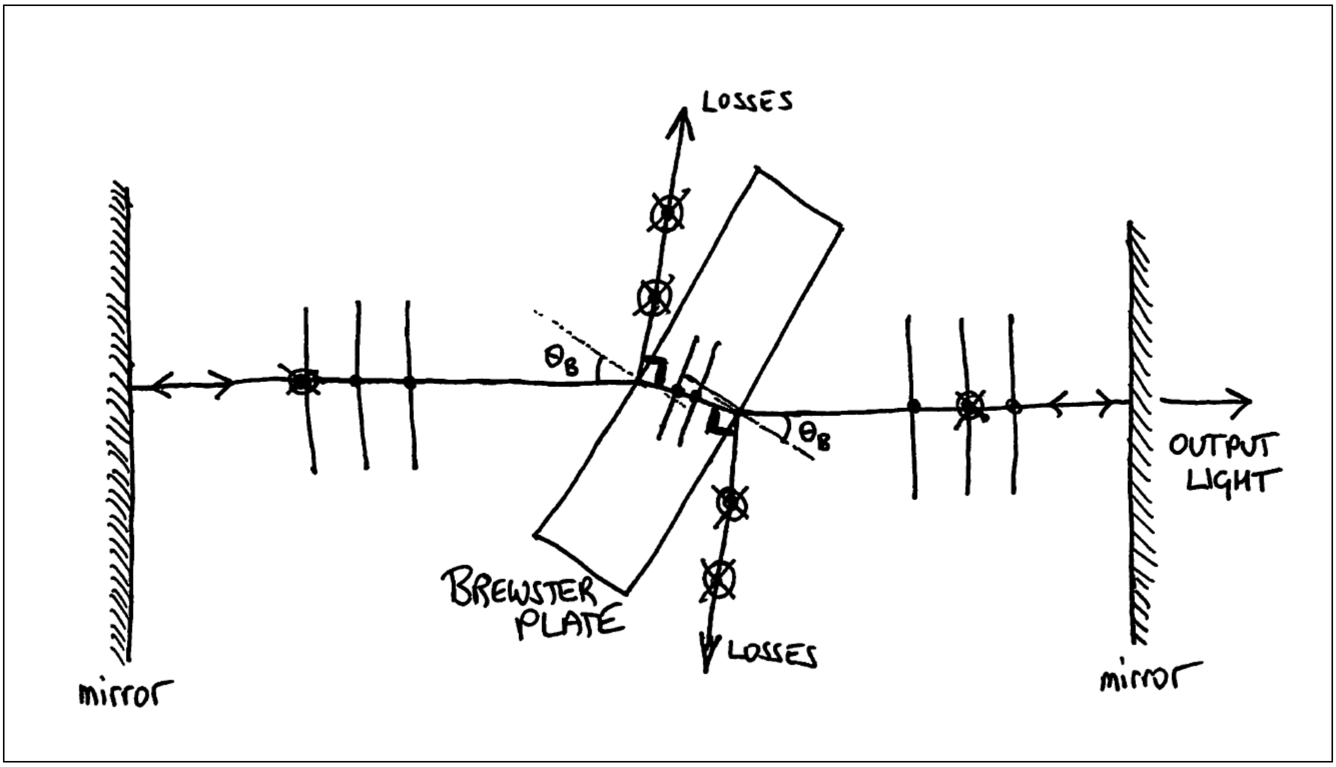
Example

Calculate the value of Brewster's angle for light incident from air on water ($n_{\text{water}} \approx 1.33$).

$$\tan\theta_B = \frac{n_r}{n_i} = \frac{n_{water}}{n_{air}} \approx \frac{1.33}{1}$$

$$\therefore \quad \theta_B \approx \tan^{-1}(1.33) \approx 53^{\circ}$$

Now consider the optical cavity below, in which a piece of glass has been placed at Brewster's angle to the optical axis:



Polarising effect of an intra-cavity Brewster plate

Some of the s-polarised photons will be reflected out of the cavity while all of the p-polarised photons will remain in it. Thus, through the stimulated emission process more p-polarised photons will be produced at the expense of the s-polarised photons, which will continue to be reflected out of the cavity. Hence, we have introduced a natural selection process for polarisation.

In the above diagram, our laser would produce light polarised *vertically*. Even when a so-called ‘Brewster window’ (or ‘Brewster plate’) is introduced into a laser, the degree of polarisation can be limited by other effects such as thermal effects in the laser medium. Note also, from the further information below, that the reflection coefficient for s-polarisation does not become zero when $\theta_i = \theta_B$. A degree of polarisation of 1000:1 (light polarised in one direction relative to light polarised at 90° to this direction) can be typical.

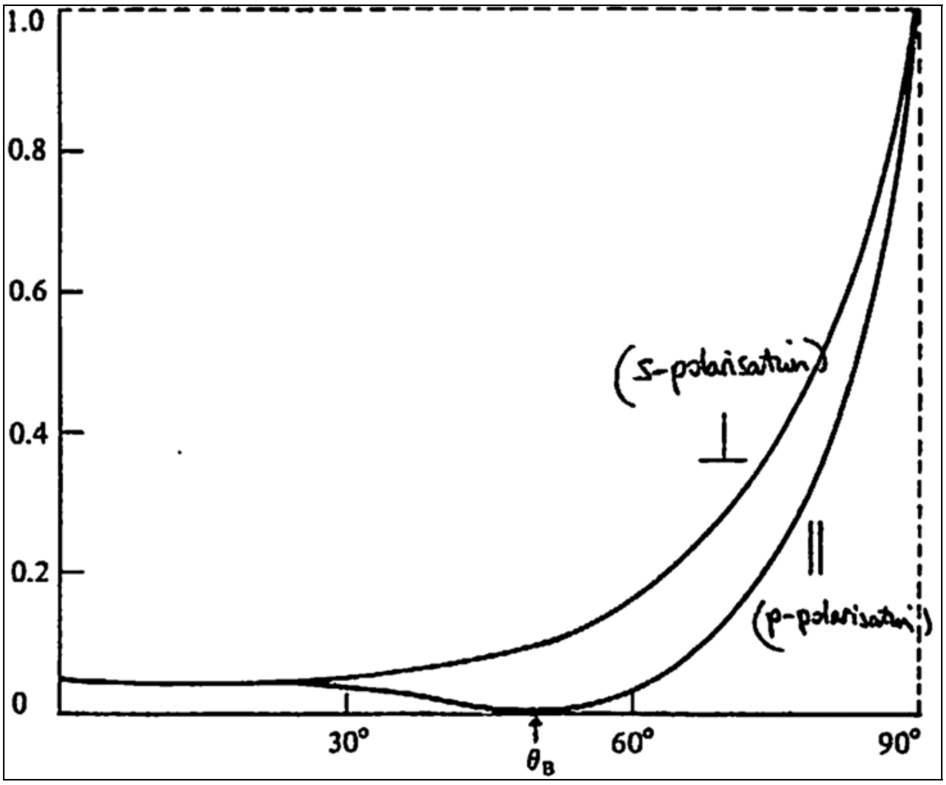
Further Information

- The percentage of light reflected at a boundary between two dielectrics can be calculated using the *Fresnel coefficients*. For the *s-polarised* light (perpendicular to the plane of incidence, ⊥), the percentage of light reflected, R_S , is given by:

$$R_S = \left| \frac{n_i \cos\theta_i - n_r \cos\theta_r}{n_i \cos\theta_i + n_r \cos\theta_r} \right|^2 \times 100\% \ .$$

For the *p-polarised* light (parallel to the plane of incidence, ||), one finds:

$$R_P = \left| \frac{n_i \cos\theta_r - n_r \cos\theta_i}{n_i \cos\theta_r + n_r \cos\theta_i} \right|^2 \times 100\% .$$ At Brewster’s angle, $R_p = 0$.



Fresnel coefficients as a function of incidence angle, θ_i

Further Information

- Proof of equation (10), giving Brewster's angle:

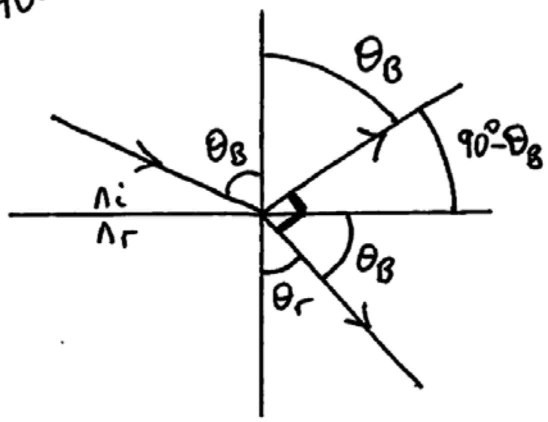
From diagram opposite, $\theta_B + \theta_r = 90^\circ$

Snell's Law : $n_i \sin\theta_i = n_r \sin\theta_r$

$\Rightarrow n_i \sin\theta_B = n_r \sin(90^\circ - \theta_B)$

$\Rightarrow n_i \sin\theta_B = n_r \cos\theta_B$

$\Rightarrow \boxed{\tan\theta_B = n_r/n_i}$



1.11.4 Light that is all *in phase* (coherent)

We saw earlier that coherence is a property relating to the *phase* relationships of photons making up a beam of light. There are two types of coherence: ***temporal coherence*** is a measure of the phase relationship between photons as a function of time, and ***spatial coherence*** is a measure of the phase relationship as a function of position.

- If we looked at the variation of the electric field of a *completely temporally coherent* beam as a function of time, then it would be a pure sine wave.

- If we 'froze' a *completely spatially coherent* laser beam and measured the phase of every photon in the beam at a particular point in the beam's length, then every photon would have the same phase at that point.

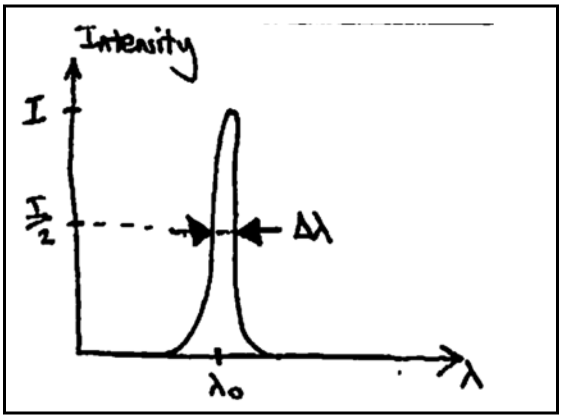
For various reasons, it is not possible to produce completely coherent beams, even with lasers. The best we can do is to produce partly coherent beams and we measure how successful we have been by the ***coherence length*** and ***coherence time*** parameters.

(a) Coherence length, L_c

The coherence length is, effectively, the distance over which the beam looks like a pure sine wave, i.e. before the photons of which it is made up move too far out of phase with each other. The coherence length is given by:

$$\boxed{L_C = \frac{c}{\Delta f} = \frac{\lambda_0^{\,2}}{\Delta\lambda} \qquad (11)}$$

where λ_0 is the central wavelength of the source and $\Delta\lambda$ and Δf are the ***linewidths*** (spread of wavelengths and frequencies) of the source.



Here we have a clue to the origin of incoherence: even laser beams do not produce light that is truly single frequency, although linewidths can be very narrow. In fact, lasers can have coherence lengths varying from millimetres up to around a kilometre.

Note. To get the wavelength form in equation (11), $f = c/\lambda$ gives:

$$\frac{df}{d\lambda} = -\frac{c}{\lambda^2} \text{ then } \Delta f \approx |df| \text{ when } \lambda = \lambda_0. \text{ So, } \Delta f \approx c\frac{\Delta\lambda}{\lambda_0^{\,2}}.$$

(b) Coherence time, t_c

This is the time over which the beam looks like a pure sine wave, and is related to the coherence length by c, the velocity of light:

$$\boxed{t_c = \frac{L_C}{c} \qquad (12)}$$

Example
Calculate the coherence lengths of: (a) a tungsten light, and (b) a laser source with $\Delta f = 165\text{MHz}$.

(a) Tungsten light (for white light, take $\lambda_0 = 550$ nm and $\Delta\lambda = 300$ nm).

$$L_C = \frac{\lambda_0^{\;2}}{\Delta\lambda} = \frac{\left(550\times10^{-9}\right)^2}{300\times10^{-9}}\,\text{m} \approx 1\;\mu\text{m}$$

(b) This laser has $L_C = \frac{c}{\Delta f} = \frac{3\times10^8}{165\times10^6}$ m ≈ 1.82 m

Fortunately, very few laser applications require the source to be highly coherent, and in some it is a positive problem if it is. Holography is one of the few applications that really requires a laser with a reasonable coherence length.