

Blackbody Radiation and Planck's Hypothesis

An object at any temperature emits electromagnetic waves in the form of **thermal radiation** from its surface. The characteristics of this radiation depend on the temperature and properties of the object's surface. Careful study shows that the radiation consists of a continuous distribution of wavelengths from all portions of the electromagnetic spectrum. If the object is at room temperature, the wavelengths of thermal radiation are mainly in the infrared region and hence the radiation is not detected by the human eye. As the surface temperature of the object increases, the object eventually begins to glow visibly red. At sufficiently high temperatures, the glowing object appears white, as in the hot tungsten filament of a light bulb.

From a classical viewpoint, thermal radiation originates from accelerated charged particles in the atoms near the surface of the object; those charged particles emit radiation much as small antennas do. The thermally agitated particles can have a distribution of energies, which accounts for the continuous spectrum of radiation emitted by the object. By the end of the 19th century, however, it became apparent that the classical theory of thermal radiation was inadequate. The basic problem was in understanding the observed distribution of wavelengths in the radiation emitted by a black body. A **black body** is an ideal system that absorbs all radiation incidents on it. The electromagnetic radiation emitted by the black body is called **blackbody radiation**.

A good approximation of a black body is a small hole leading to the inside of a hollow object as shown in Figure 1. Any radiation incident on the hole from outside the cavity enters the hole and is reflected a number of times on the interior walls of the cavity; hence, the hole acts as a perfect absorber. The nature of the radiation leaving the cavity through the hole depends only on the temperature of the cavity walls and not on the material of which the walls are made.

The radiation emitted by oscillators in the cavity walls experiences boundary conditions. As the radiation reflects from the cavity's walls, standing electromagnetic waves are established within the three-dimensional interior of the cavity. Many standing-wave modes are possible, and the distribution of the energy in the cavity among these modes determines the wavelength distribution of the radiation leaving the cavity through the hole. The wavelength distribution of radiation from cavities was studied experimentally in the late 19th century. Figure 2 shows how the intensity of blackbody radiation varies with temperature and wavelength.

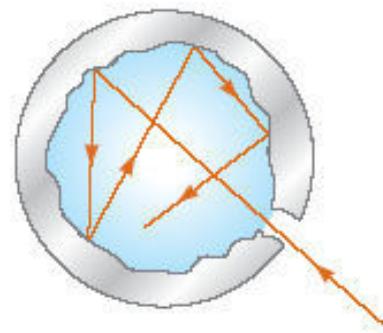


Figure 1. The opening to the cavity inside a hollow object is a good approximation of a black body. Light entering the small opening strikes the interior walls, where some is absorbed and some is reflected at a random angle. The cavity walls reradiate at wavelengths corresponding to their temperature, producing standing waves in the cavity. Some of the energy from these standing waves can leave through the opening.

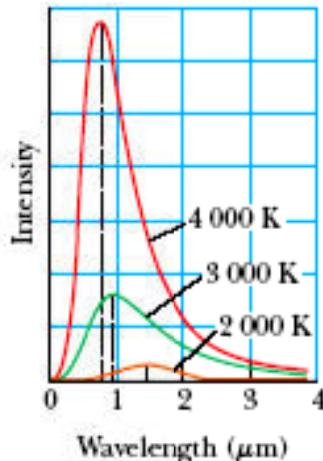


Figure 2. Intensity of blackbody radiation versus wavelength at three temperatures. The amount of radiation emitted (the area under a curve) increases with increasing temperature.

A successful theory for blackbody radiation must predict the shape of the curves in Figure 2. Early attempts to use classical ideas to explain the shapes of the curves in Figure 2 failed.

Let's consider one of these early attempts. To describe the distribution of energy from a black body, we define $I(\lambda, T) d\lambda$ to be the intensity, or power per unit area, emitted in the wavelength interval $d\lambda$. The result of a calculation based on a classical theory of blackbody radiation known as the **Rayleigh–Jeans law** is

$$I(\lambda, T) = \frac{2\pi c k_B T}{\lambda^4} \quad (1)$$

where k_B is Boltzmann's constant. The black body is modelled as the hole leading into a cavity supporting many modes of oscillation of the electromagnetic field caused by accelerated charges in the cavity walls, resulting in the emission of electromagnetic waves at all wavelengths. In the classical theory used to derive the above equation, the average energy for each wavelength of the standing-wave modes is assumed to be proportional to $k_B T$, based on the theorem of equipartition of energy.

An experimental plot of the blackbody radiation spectrum, together with the theoretical prediction of the Rayleigh–Jeans law, is shown in Figure 3. At long wavelengths, the Rayleigh–Jeans law is in reasonable agreement with experimental data, but at short wavelengths, major disagreement is apparent.

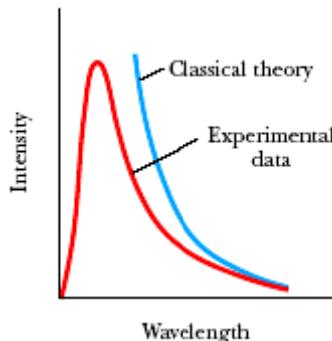


Figure 3. Comparison of experimental results and the curve predicted by the Rayleigh–Jeans law for the distribution of blackbody radiation.

As λ approaches zero, the function $I(\lambda, T)$ given by equation 1 approaches infinity. Hence, according to classical theory, not only should short wavelengths predominate in a blackbody spectrum, but also the energy emitted by any black body should become infinite in

the limit of zero wavelength. In contrast to this prediction, the experimental data plotted in Figure 3 show that as λ approaches zero, $I(\lambda, T)$ also approaches zero. This mismatch of theory and experiment was so disconcerting that scientists called it the *ultraviolet catastrophe*. (This “catastrophe” – infinite energy – occurs as the wavelength approaches zero; the word *ultraviolet* was applied because ultraviolet wavelengths are short.)

In 1900, Max Planck developed a theory of blackbody radiation that leads to an equation for $I(\lambda, T)$ that is in complete agreement with experimental results at all wavelengths. Planck assumed the cavity radiation came from atomic oscillators in the cavity walls in Figure 1. Planck made two bold and controversial assumptions concerning the nature of the oscillators in the cavity walls:

- The energy of an oscillator can have only certain *discrete* values E_n :

$$E_n = nh\nu \quad (2)$$

where n is a positive integer called a **quantum number**, ν is the oscillator’s frequency, and h is a parameter Planck introduced that is now called **Planck’s constant**. Because the energy of each oscillator can have only discrete values given by equation 2, we say the energy is **quantized**. Each discrete energy value corresponds to a different **quantum state**, represented by the quantum number n . When the oscillator is in the $n=1$ quantum state, its energy is $h\nu$; when it is in the $n=2$ quantum state, its energy is $2h\nu$; and so on.

- The oscillators emit or absorb energy when making a transition from one quantum state to another. The entire energy difference between the initial and final states in the transition is emitted or absorbed as a single quantum of radiation. If the transition is from one state to a lower adjacent state – say, from the $n=3$ state to the $n=2$ state – equation 2 shows that the amount of energy emitted by the oscillator and carried by the quantum of radiation is

$$E = h\nu \quad (3)$$

An oscillator emits or absorbs energy only when it changes quantum states. If it remains in one quantum state, no energy is absorbed or emitted. Figure 4 is an **energy-level diagram** showing the quantized energy levels and allowed transitions proposed by Planck. The vertical axis is linear in energy, and the allowed energy levels are represented as horizontal lines. The quantized system can have only the energies represented by the horizontal lines.

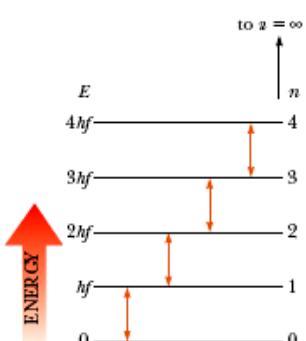


Figure 4. Allowed energy levels for an oscillator with frequency f ($= \nu$). Allowed transitions are indicated by the double-headed arrows.

The key point in Planck’s theory is the radical assumption of quantized energy states. This development – a clear deviation from classical physics – marked the birth of the quantum theory.

In the Rayleigh–Jeans model, the average energy associated with a particular wavelength of standing waves in the cavity is the same for all wavelengths and is equal to $k_B T$.

Planck used the same classical ideas as in the Rayleigh–Jeans model to arrive at the energy density as a product of constants and the average energy for a given energy for a given wavelength, but the average energy is not given by the equipartition theorem. A wave's average energy is the average energy difference between levels of the oscillator, *weighted according to the probability of the wave being emitted*. This weighting is based on the occupation of higher-energy states as described by the Boltzmann distribution law. According to this law, the probability of a state being occupied is proportional to the factor $e^{-E/k_B T}$, where E is the energy of the state.

At low frequencies, the energy levels are close together as on the right in Figure 5, and many of the energy states are excited because the Boltzmann factor $e^{-E/k_B T}$ is relatively large for these states. Therefore, there are many contributions to the outgoing radiation, although each contribution has very low energy. Now, consider high-frequency radiation, that is, radiation with short wavelength. To obtain this radiation, the allowed energies are very far apart as on the left in Figure 5. The probability of thermal agitation exciting these high energy levels is small because of the small value of the Boltzmann factor for large values of E . At high frequencies, the low probability of excitation results in very little contribution to the total energy, even though each quantum is of large energy. This low probability “turns the curve over” and brings it down to zero again at short wavelengths.

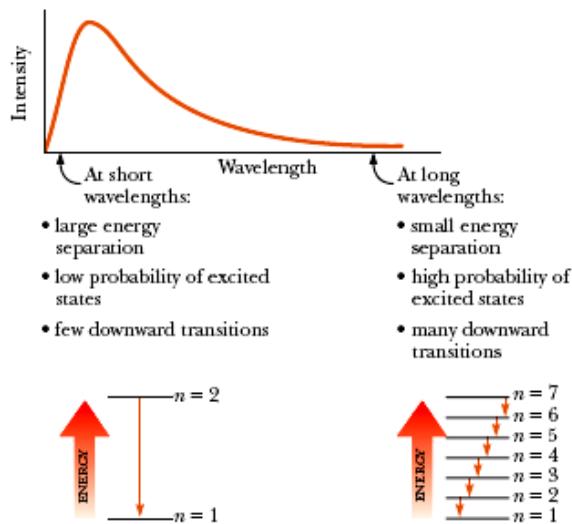


Figure 5. In Planck's model, the average energy associated with a given wavelength is the product of the energy of a transition and a factor related to the probability of the transition occurring. As the energy levels move farther apart at shorter wavelengths (higher energy), the probability of excitation decreases, as does the probability of a transition from the excited state.

Using this approach, Planck generated a theoretical expression for the wavelength distribution that agreed remarkably well with the experimental curves in Figure 2.

$$I(\lambda, T) = \frac{2\pi k c^2}{\lambda^5} \left(e^{hc/\lambda k_B T} - 1 \right) \quad (4)$$

This function includes the parameter h , which Planck adjusted so that his curve matched the experimental data at all wavelengths. The value of this parameter is found to be independent of the material of which the black body is made and independent of the temperature; it is a fundamental constant of nature. The value of Planck's constant, $h = 6.626 \times 10^{-34}$ J·s. At long wavelengths, equation 4 reduces to the Rayleigh–Jeans expression, equation 1, and at short wavelengths, it predicts an exponential decrease in $I(\lambda, T)$ with decreasing wavelength, in agreement with experimental results.

Compton Shift

The incident monochromatic X-rays with frequency ν_0 are regarded as a stream of particles – the photons, with energy

$$E_0 = h\nu_0$$

The linear momentum p_0 of the photon can be obtained from the relativistic relation between energy and momentum

$$E = \sqrt{p^2 c^2 + m_0^2 c^4} \quad (\text{Relativistic relation between energy and momentum for any particle})$$

where E = Energy, p = linear momentum, m_0 = rest mass of the particle and c = speed of light in vacuum.

Since the photon has zero rest mass ($m_0 = 0$), $E = pc$ (or) $p = \frac{E}{c}$. The linear momentum of the photon is

$$p_0 = \frac{E_0}{c} = \frac{h\nu_0}{c}$$

The scattering of X-rays by electrons is treated as elastic collision between photons and electrons which are particles. The electrons which are loosely bound can be treated as almost free particles at rest.

So, the initial linear momentum of the electrons is zero and the initial energy is equal to $m_e c^2$, where m_e is the rest mass of the electron. The Compton scattering of a photon by free electron at rest is shown in Figure 3

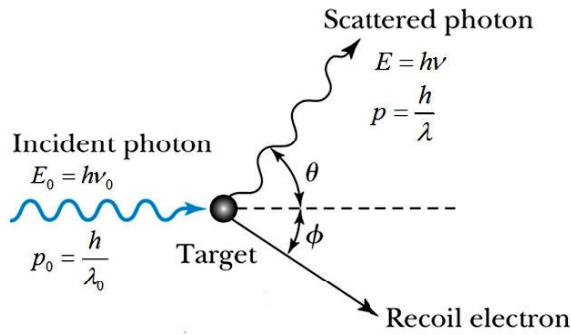


Figure 3. The quantum model for X-ray scattering from an electron. The collision of the photon with the electron displays the particle-like nature of the photon

The photon of energy $E_0 = h\nu_0$ and momentum $p_0 = \frac{h\nu_0}{c}$ hits free electron at rest. The scattered photon moves at an angle θ with direction of incident photon. If ν is the frequency of the scattered X-rays, it has energy $E = h\nu$ and linear momentum $p = \frac{h\nu}{c}$

The electron recoils at an angle ϕ , with the linear momentum p_e , and energy E_e .

The principle of conservation of energy and linear momentum can be applied to the collision process. The sum of the energy of the photon and electron before collision must be the same as the total energy after collision,

$$E_0 + m_e c^2 = E + E_e = E + \sqrt{p_e^2 c^2 + m_e^2 c^4}$$

The conservation of linear momentum gives

$$\vec{p}_0 + 0 = \vec{p} + \vec{p}_e \quad \text{or} \quad \vec{p}_e = \vec{p}_0 - \vec{p}$$

or $p_e^2 = (\vec{p}_0 - \vec{p}) \bullet (\vec{p}_0 - \vec{p}) = p_0^2 + p^2 - 2p_0 \bullet p$

$$= p_0^2 + p^2 - 2p_0 p \cos \theta$$

$$[(E_0 - E) + m_e c^2] = \sqrt{p_e^2 c^2 + m_e^2 c^4}$$

$$[(E_0 - E) + m_e c^2]^2 = p_e^2 c^2 + m_e^2 c^4$$

$$(E_0 - E)^2 + m_e^2 c^4 + 2(E_0 - E)m_e c^2 = p_e^2 c^2 + m_e^2 c^4$$

$$(E_0 - E)^2 + 2(E_0 - E)m_e c^2 = p_e^2 c^2$$

$$= p_0^2 c^2 + p^2 c^2 - 2p_0 p c^2 \cos \theta$$

$$= E_0^2 + E^2 - 2E_0 E \cos \theta$$

$$= E_0^2 + E^2 - 2E_0 E + 2E_0 E - 2E_0 E \cos \theta$$

$$= (E_0 - E)^2 + 2E_0 E (1 - \cos \theta)$$

$$(E_0 - E)m_e c^2 = E_0 E (1 - \cos \theta)$$

$$\frac{1}{E} - \frac{1}{E_0} = \frac{(1 - \cos \theta)}{m_e c^2} \quad \text{or} \quad \frac{1}{\nu} - \frac{1}{\nu_0} = \frac{h(1 - \cos \theta)}{m_e c^2}$$

$$\lambda - \lambda_0 = \frac{h}{m_e c} (1 - \cos \theta) = \lambda_c (1 - \cos \theta) \quad (\text{where Compton wavelength, } \lambda_c = 0.00243 \text{ nm})$$

$$\Delta \lambda = \lambda_c (1 - \cos \theta)$$

The Compton Effect

Arthur H. Compton discovered that when a beam of X-rays is scattered from a target, the wavelengths of the scattered X-rays are slightly greater than the wavelength of incident beam. Figure 1 shows the experimental arrangement Compton used to observe this Compton effect. A beam of X-rays of wavelength λ_0 falls on a carbon (graphite) target. Compton observed the X-rays that were scattered at various angles θ to the incident beam and measured the intensity and the wavelength of the scattered rays at several of these angles.

Figure 2 shows his experimental results for four different values of θ . From the figure, we see that although the incident beam consists essentially of a single wavelength λ_0 , the scattered X-rays have intensity peaks at two wavelengths. One of them (λ_0) is the same as the incident wavelength, but the other (λ) is larger by an amount $\Delta\lambda$. This Compton shift $\Delta\lambda$ varies with the angle θ at which the scattered X-rays are observed.

$$\text{Compton shift } \Delta\lambda = \lambda - \lambda_0 = \frac{h}{m_e c} (1 - \cos \theta)$$

where m_e is the mass of the electron. This expression is known as the **Compton shift equation**.

The unshifted peak at λ_0 in Figure 2 is caused by X-rays scattered from electrons tightly bound to the target atoms. This unshifted peak also is predicted from the above Compton shift equation, as if the bound electron mass is replaced with the mass of a carbon atom, which is approximately 23 000 times the mass of the electron. Therefore, there is a wavelength shift for scattering from an electron bound to an atom, but it is so small that it was undetectable in Compton's experiment.

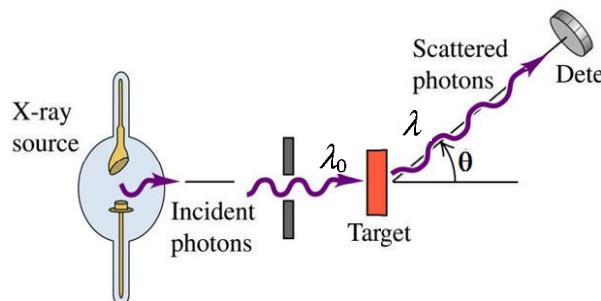


Figure 1. The experimental setup used for observing the Compton effect. The detector, which can be set to any desired angle θ , measure the intensity and the wavelength of the scattered X-rays.

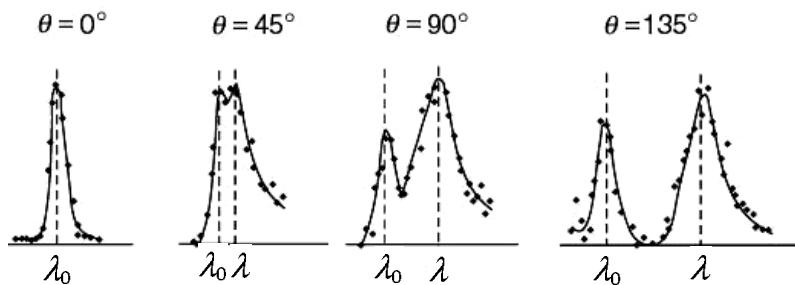


Figure 2. Compton experimental results for four different values of the scattering angle θ . The peak at λ_0 is due to the scattering of the incident radiation from the tightly bound inner electrons of the atom. The second peak represents the radiation scattered from the loosely bound, nearly free outer electrons.

3.5 PARTICLE DIFFRACTION

An experiment that confirms the existence of de Broglie waves

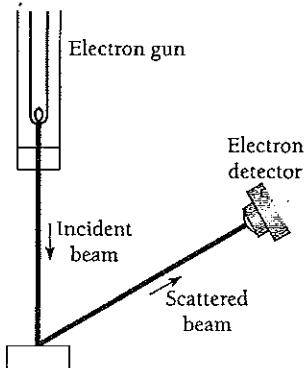


Figure 3.6 The Davisson-Germer experiment.

A wave effect with no analog in the behavior of Newtonian particles is diffraction. In 1927 Clinton Davisson and Lester Germer in the United States and G. P. Thomson in England independently confirmed de Broglie's hypothesis by demonstrating that electron beams are diffracted when they are scattered by the regular atomic arrays of crystals. (All three received Nobel Prizes for their work. J. J. Thomson, G. P.'s father, had earlier won a Nobel Prize for verifying the particle nature of the electron: the wave-particle duality seems to have been the family business.) We shall look at the experiment of Davisson and Germer because its interpretation is more direct.

Davisson and Germer were studying the scattering of electrons from a solid using an apparatus like that sketched in Fig. 3.6. The energy of the electrons in the primary beam, the angle at which they reach the target, and the position of the detector could all be varied. Classical physics predicts that the scattered electrons will emerge in all directions with only a moderate dependence of their intensity on scattering angle and even less on the energy of the primary electrons. Using a block of nickel as the target, Davisson and Germer verified these predictions.

In the midst of their work an accident occurred that allowed air to enter their apparatus and oxidize the metal surface. To reduce the oxide to pure nickel, the target was baked in a hot oven. After this treatment, the target was returned to the apparatus and the measurements resumed.

Now the results were very different. Instead of a continuous variation of scattered electron intensity with angle, distinct maxima and minima were observed whose positions depended upon the electron energy! Typical polar graphs of electron intensity after the accident are shown in Fig. 3.7. The method of plotting is such that the intensity at any angle is proportional to the distance of the curve at that angle from the point of scattering. If the intensity were the same at all scattering angles, the curves would be circles centered on the point of scattering.

Two questions come to mind immediately: What is the reason for this new effect? Why did it not appear until after the nickel target was baked?

De Broglie's hypothesis suggested that electron waves were being diffracted by the target, much as x-rays are diffracted by planes of atoms in a crystal. This idea received

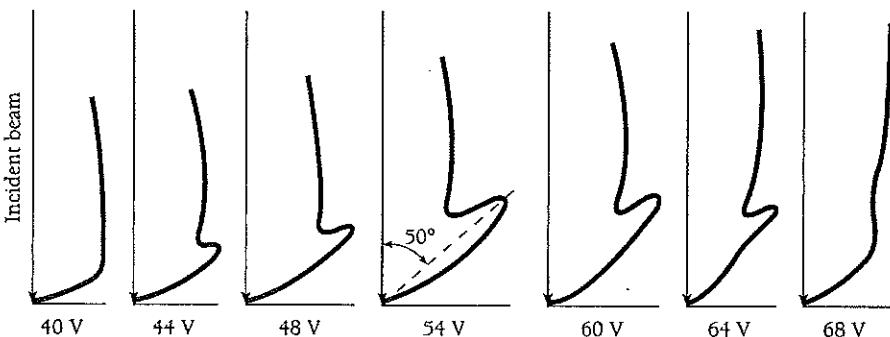


Figure 3.7 Results of the Davisson-Germer experiment, showing how the number of scattered electrons varied with the angle between the incoming beam and the crystal surface. The Bragg planes of atoms in the crystal were not parallel to the crystal surface, so the angles of incidence and scattering relative to one family of these planes were both 65° (see Fig. 3.8).

support when it was realized that heating a block of nickel at high temperature causes the many small individual crystals of which it is normally composed to form into a single large crystal, all of whose atoms are arranged in a regular lattice.

Let us see whether we can verify that de Broglie waves are responsible for the findings of Davisson and Germer. In a particular case, a beam of 54-eV electrons was directed perpendicularly at the nickel target and a sharp maximum in the electron distribution occurred at an angle of 50° with the original beam. The angles of incidence and scattering relative to the family of Bragg planes shown in Fig. 3.8 are both 65° . The spacing of the planes in this family, which can be measured by x-ray diffraction, is 0.091 nm. The Bragg equation for maxima in the diffraction pattern is

$$n\lambda = 2d \sin \theta \quad (2.13)$$

Here $d = 0.091$ nm and $\theta = 65^\circ$. For $n = 1$ the de Broglie wavelength λ of the diffracted electrons is

$$\lambda = 2d \sin \theta = (2)(0.091 \text{ nm})(\sin 65^\circ) = 0.165 \text{ nm}$$

Now we use de Broglie's formula $\lambda = h/\gamma mv$ to find the expected wavelength of the electrons. The electron kinetic energy of 54 eV is small compared with its rest energy mc^2 of 0.51 MeV, so we can let $\gamma = 1$. Since

$$\text{KE} = \frac{1}{2}mv^2$$

the electron momentum mv is

$$\begin{aligned} mv &= \sqrt{2m\text{KE}} \\ &= \sqrt{(2)(9.1 \times 10^{-31} \text{ kg})(54 \text{ eV})(1.6 \times 10^{-19} \text{ J/eV})} \\ &= 4.0 \times 10^{-24} \text{ kg} \cdot \text{m/s} \end{aligned}$$

The electron wavelength is therefore

$$\lambda = \frac{h}{mv} = \frac{6.63 \times 10^{-34} \text{ J} \cdot \text{s}}{4.0 \times 10^{-24} \text{ kg} \cdot \text{m/s}} = 1.66 \times 10^{-10} \text{ m} = 0.166 \text{ nm}$$

which agrees well with the observed wavelength of 0.165 nm. The Davisson-Germer experiment thus directly verifies de Broglie's hypothesis of the wave nature of moving bodies.

Analyzing the Davisson-Germer experiment is actually less straightforward than indicated above because the energy of an electron increases when it enters a crystal by an amount equal to the work function of the surface. Hence the electron speeds in the experiment were greater inside the crystal and the de Broglie wavelengths there shorter than the values outside. Another complication arises from interference between waves diffracted by different families of Bragg planes, which restricts the occurrence of maxima to certain combinations of electron energy and angle of incidence rather than merely to any combination that obeys the Bragg equation.

Electrons are not the only bodies whose wave behavior can be demonstrated. The diffraction of neutrons and of whole atoms when scattered by suitable crystals has been observed, and in fact neutron diffraction, like x-ray and electron diffraction, has been used for investigating crystal structures.

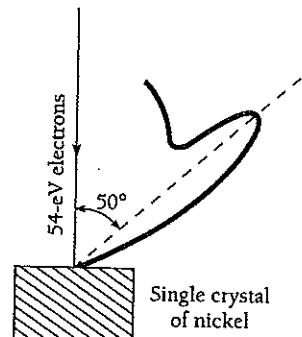


Figure 3.8 The diffraction of the de Broglie waves by the target is responsible for the results of Davisson and Germer.

$$\left(\frac{1}{E'} - \frac{1}{E} \right) = \frac{(1 - \cos \theta)}{m_e c^2}$$

In terms of wavelength

$$\left(\frac{\lambda'}{hc} - \frac{\lambda}{hc} \right) = \frac{(1 - \cos \theta)}{m_e c^2}$$

$$\lambda' - \lambda = \frac{h}{m_e c} (1 - \cos \theta)$$

$$compton\ shift, \Delta \lambda = \frac{h}{m_e c} (1 - \cos \theta)$$

$$E = h\nu = \frac{hc}{\lambda}$$

$$\text{compton shift, } \Delta\lambda = \frac{h}{m_e c} (1 - \cos \theta)$$

Eq. gives the changes in wavelength expected for a photon that is scattered Through the angle θ by the particle of rest mass m_e

This change is independent of wavelength λ of the incident photon

Compton wavelength $\lambda_c = h/mc = 2.426 \times 10^{-12} \text{ m}$

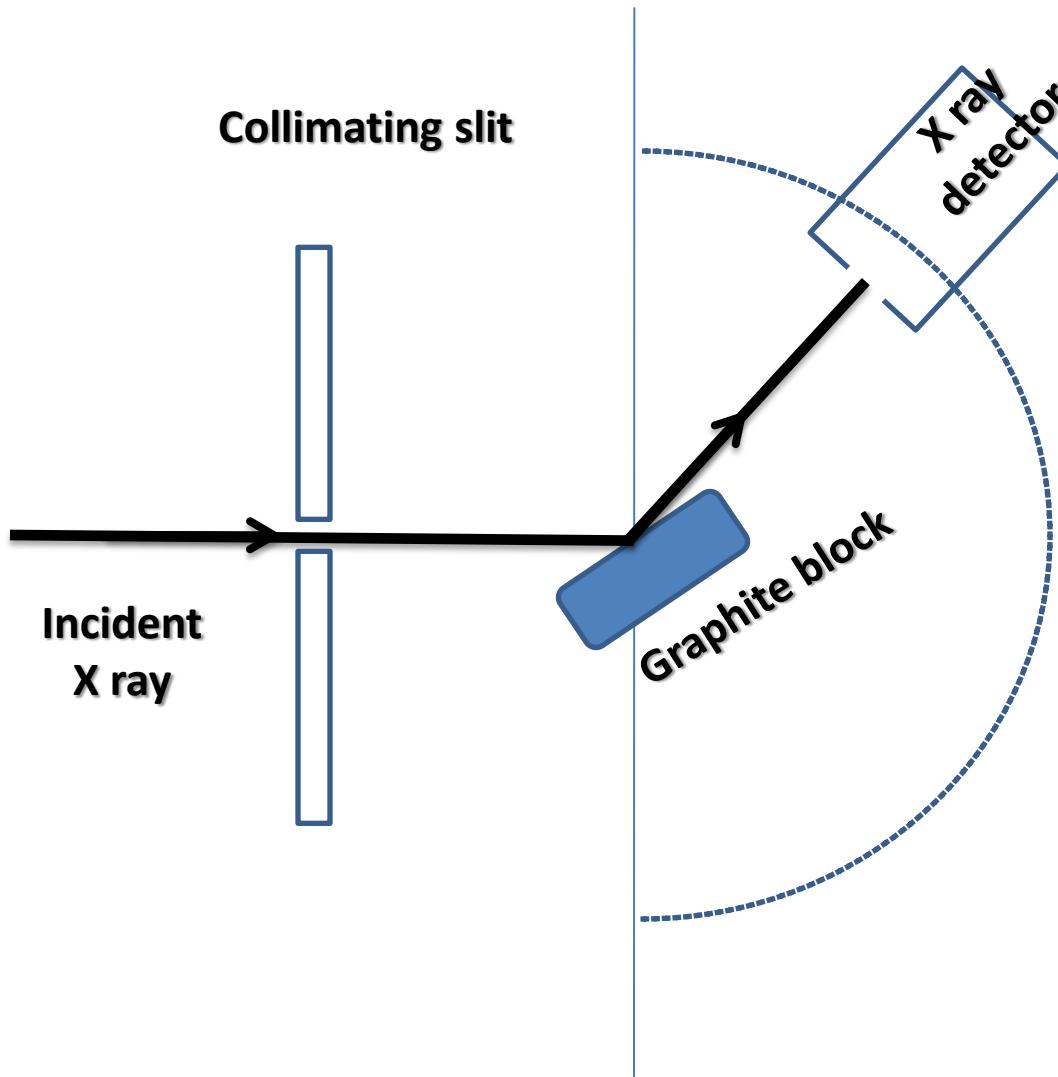
The greatest wavelength change is AT $\theta=180^\circ$

$$\Delta\lambda = 4.852 \times 10^{-12} \text{ m}$$

This changes are observable in x-ray

Shift in wavelength for visible light is less than 0.01 percent of the initial wavelength

Compton Effect : Experimental setup



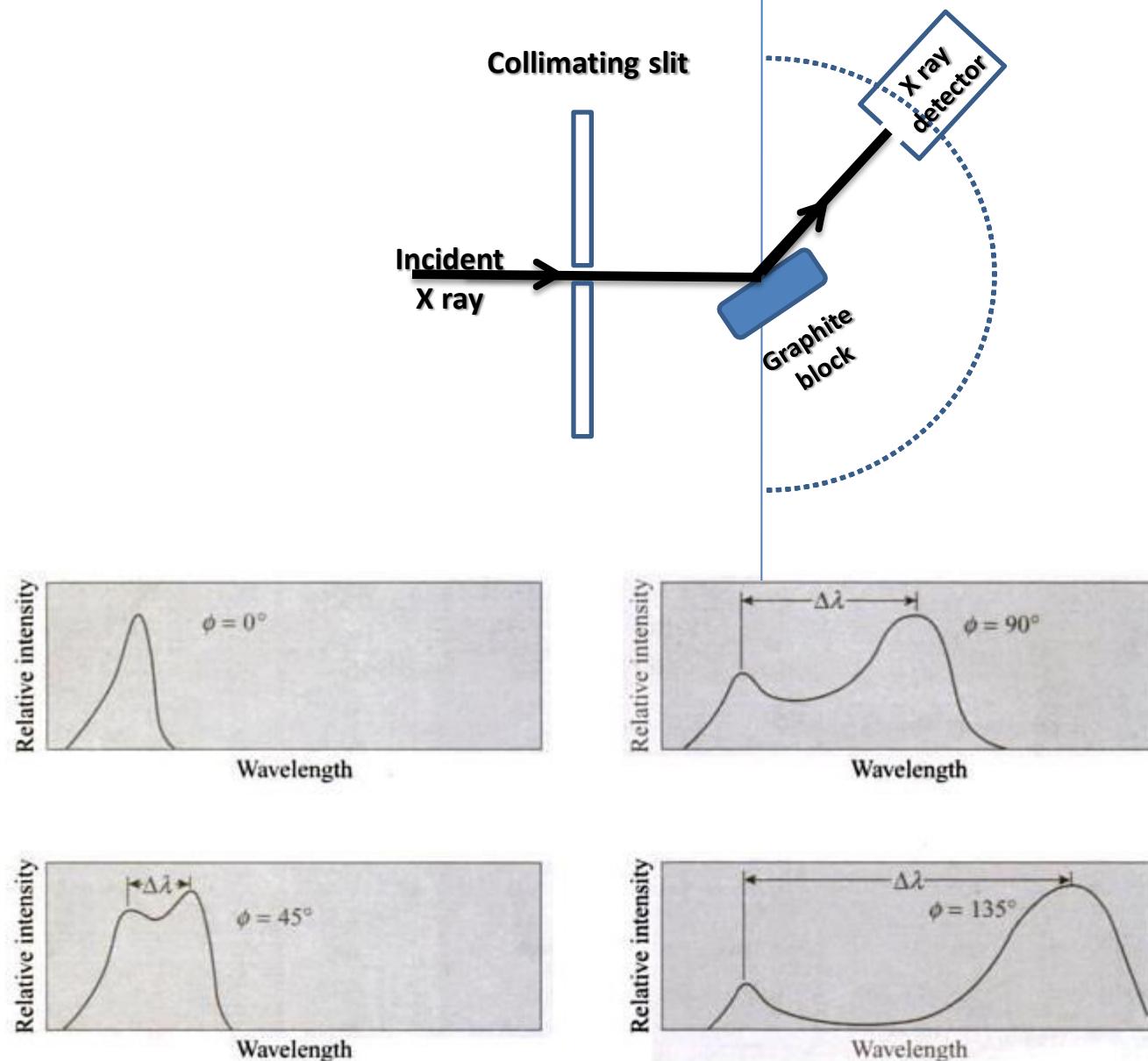


Fig. 2.24 | Experimental confirmation of Compton scattering. The greater the scattering angle, the greater the wavelength change, in accord with Eq. (2.21).

The nature of light.....

- The birth of quantum mechanics is intimately linked with the theories and discoveries relating to the nature of light
- Is the nature of light that of a wave or a particle???

The story of light.....

- Corpuscular theory (**Newton**)
- Wave nature (**Huygens**)
- Double-slit interference experiment (**Young**)

The story of light (contd.).....

- Light is an electromagnetic wave (**Maxwell**)
- Photoelectric effect – existence of light quanta – photons (**Einstein**)
- Photons have momentum (**Compton**).....

Light has a dual nature

- **Wave** (electromagnetic) - Interference
- Diffraction
- **Particle** (photons) - Photoelectric effect
- Compton effect

Wave - Particle Duality for light

What about Matter?

If light, which was traditionally understood as a wave also turns out to have a particle nature, might matter, which is traditionally understood as particles, also have a wave nature?

Louis de Broglie's hypothesis

The **dual** nature of matter

A particle with momentum p has a **matter wave** associated with it, whose wavelength is given by

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

The connecting link – Planck's constant

Dual Nature

Radiation

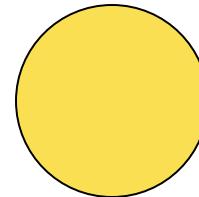
$$E = h\nu$$

Matter

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

Particle

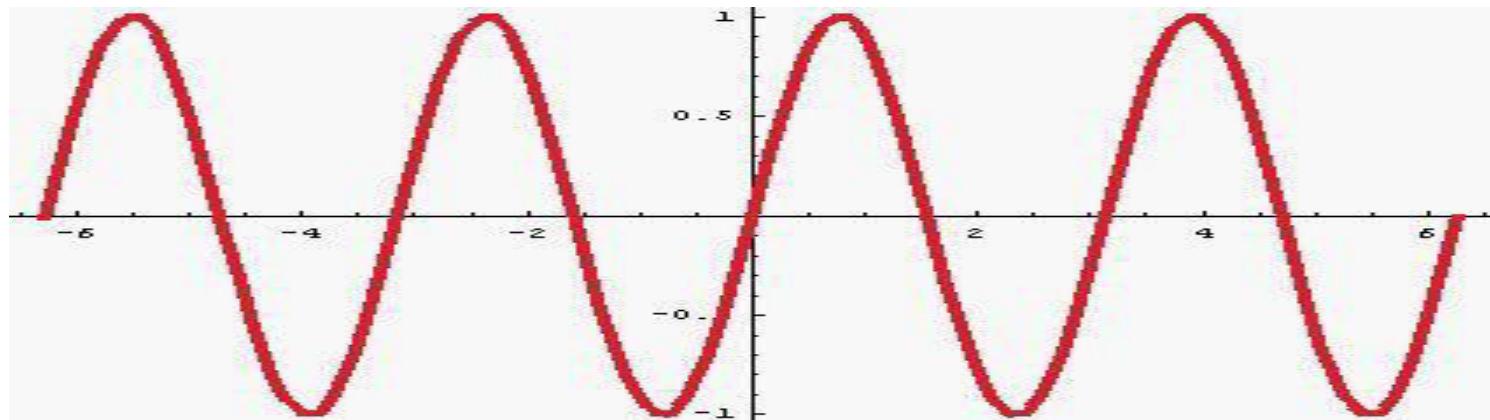
Our traditional understanding of a particle...



**“Localized” - definite position, momentum,
confined in space**

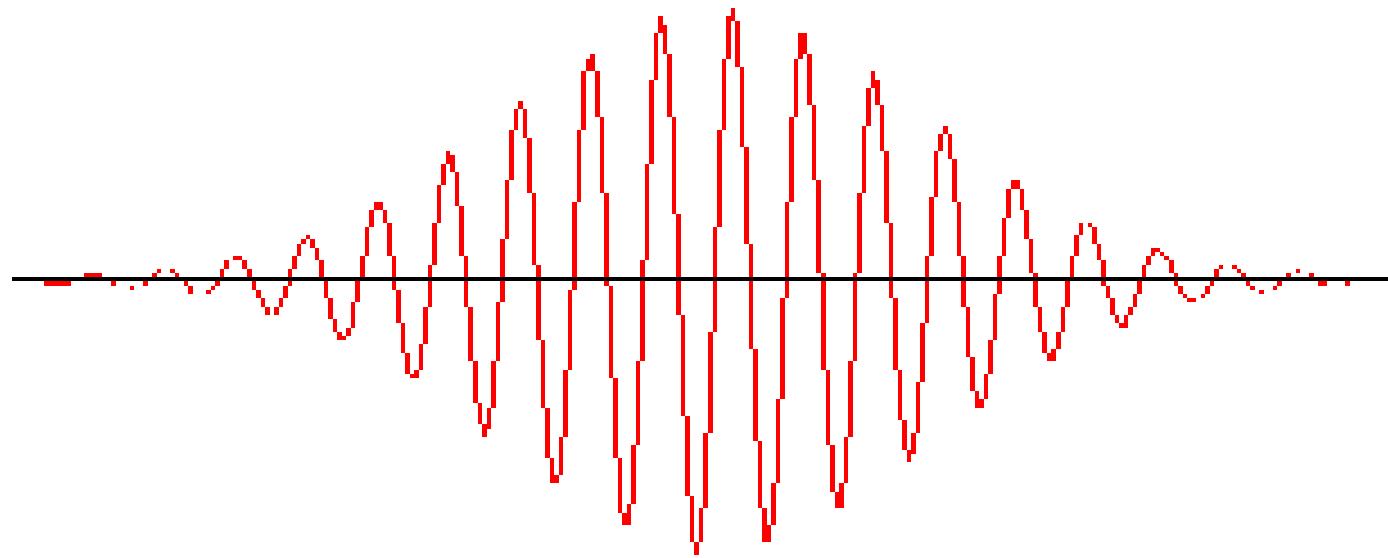
Wave

Our traditional understanding of a wave....



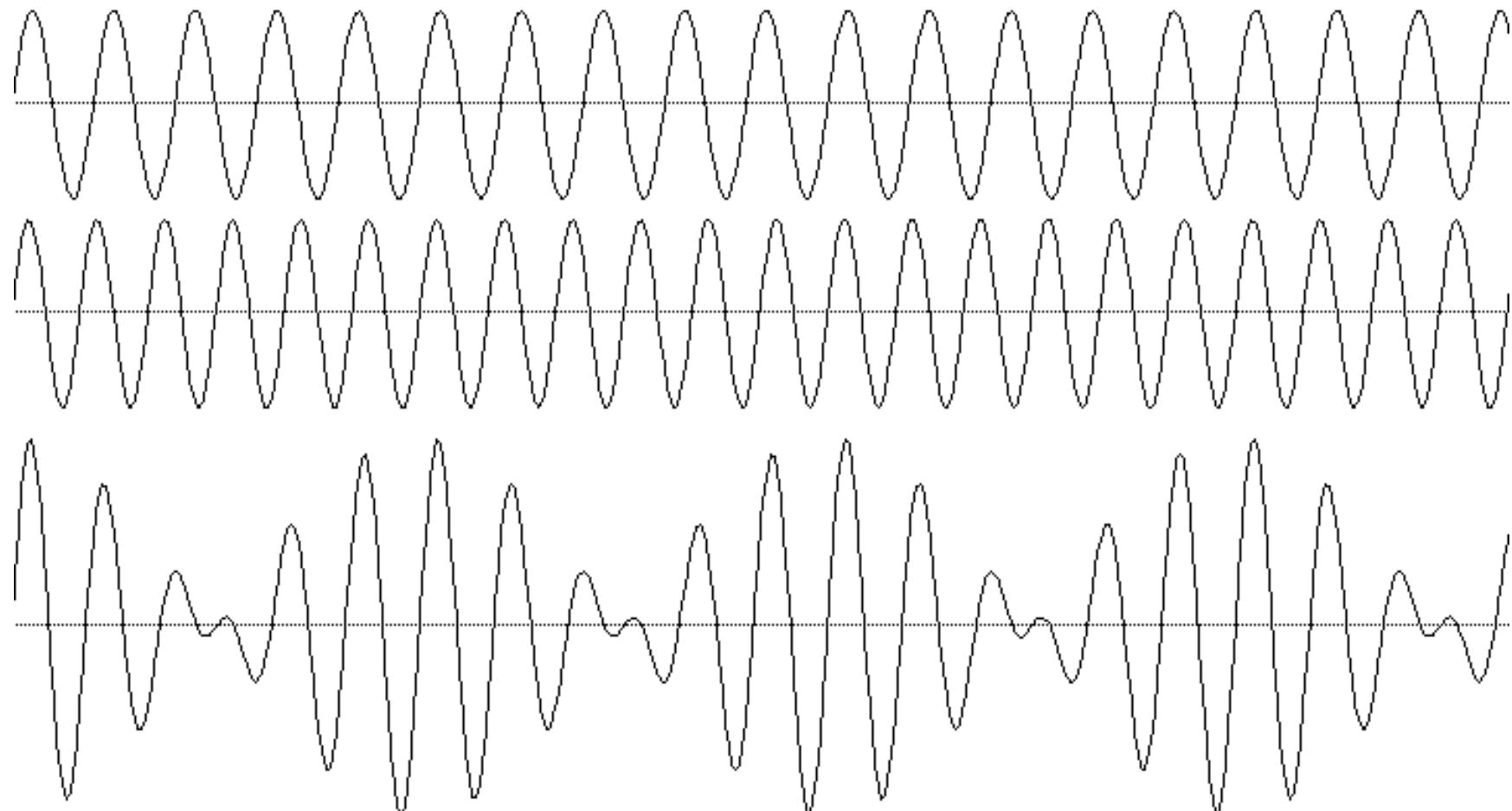
“de-localized” – spread out in space and time

A “Wave Packet”



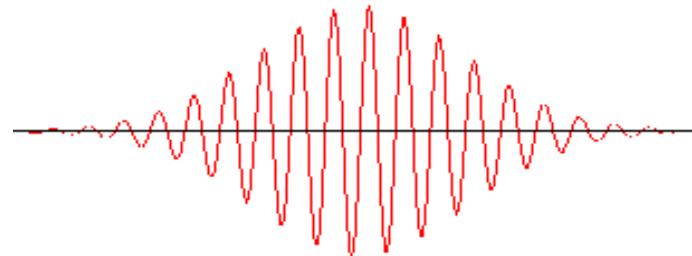
How do you construct a wave packet?

Adding up waves of different frequencies.....



Constructing a wave packet by adding up several waves

If several waves of different wavelengths (frequencies) and phases are superposed together, one would get a resultant which is a **localized wave packet**



A wave packet describes a particle

- A **wave packet** is a group of waves with slightly different wavelengths interfering with one another in a way that the amplitude of the group (envelope) is non-zero only in the neighbourhood of the particle
- A wave packet is **localized** – a good representation for a particle!

Characteristics of Matter Waves

- Only moving material particles exhibit matter waves.
- Smaller is the speed of the particle longer is the wavelength associated with it.
- Matter wave travels faster than the speed of light in vacuum as $v_p > c$
- Matter waves are not real waves and therefore cannot be represented by wave displacement. Variation of Ψ constitute matter waves. They are neither longitudinal nor transversion in nature.

Why Quantum Physics?²⁹

- “Classical Physics”:
 - developed in 15th to 20th century;
 - provides very successful description of “every day, ordinary objects”
 - motion of trains, cars, bullets,....
 - orbit of moon, planets
 - how an engine works,..
 - subfields: mechanics, thermodynamics, electrodynamics,
 - Quantum Physics:
 - developed early 20th century, in response to shortcomings of classical physics in describing certain phenomena (blackbody radiation, photoelectric effect, emission and absorption spectra...)
 - describes “small” objects (e.g. atoms and their constituents)

Some key events/observations that led to the development of quantum mechanics...

- Black body radiation spectrum (Planck, 1901)
- Photoelectric effect (Einstein, 1905)
- Model of the atom (Rutherford, 1911)
- Quantum Theory of Spectra (Bohr, 1913)
- Compton effect (Compton, 1922)
- Exclusion Principle (Pauli, 1922)
- Matter Waves (de Broglie 1925)
- Experimental test of matter waves (Davisson and Germer, 1927)

Thermal radiation



© 2004 Thomson - Brooks/Cole

- Hot filament glows.
- Classical physics can't explain the observed wavelength distribution of EM radiation from such a hot object.
- This problem is historically the problem that leads to the rise of quantum physics during the turn of 20th century

Blackbody Radiation³²

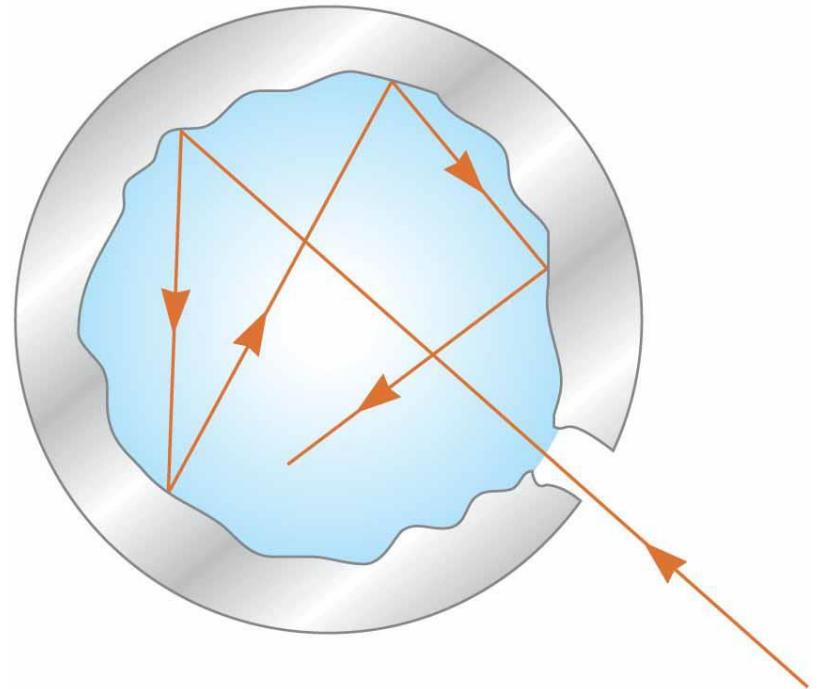
- An object at any temperature is known to emit thermal radiation
 - Characteristics depend on the temperature and surface properties
 - The thermal radiation consists of a continuous distribution of wavelengths from all portions of the em spectrum
- At room temperature, the wavelengths of the thermal radiation are mainly in the infrared region

Blackbody Radiation

- As the surface temperature increases, the wavelength changes
 - It will glow red and eventually white
- The basic problem was in understanding the observed distribution in the radiation emitted by a black body
 - Classical physics didn't adequately describe the observed distribution
- A black body is an ideal system that absorbs all radiation incident on it
- The electromagnetic radiation emitted by a black body is called blackbody radiation

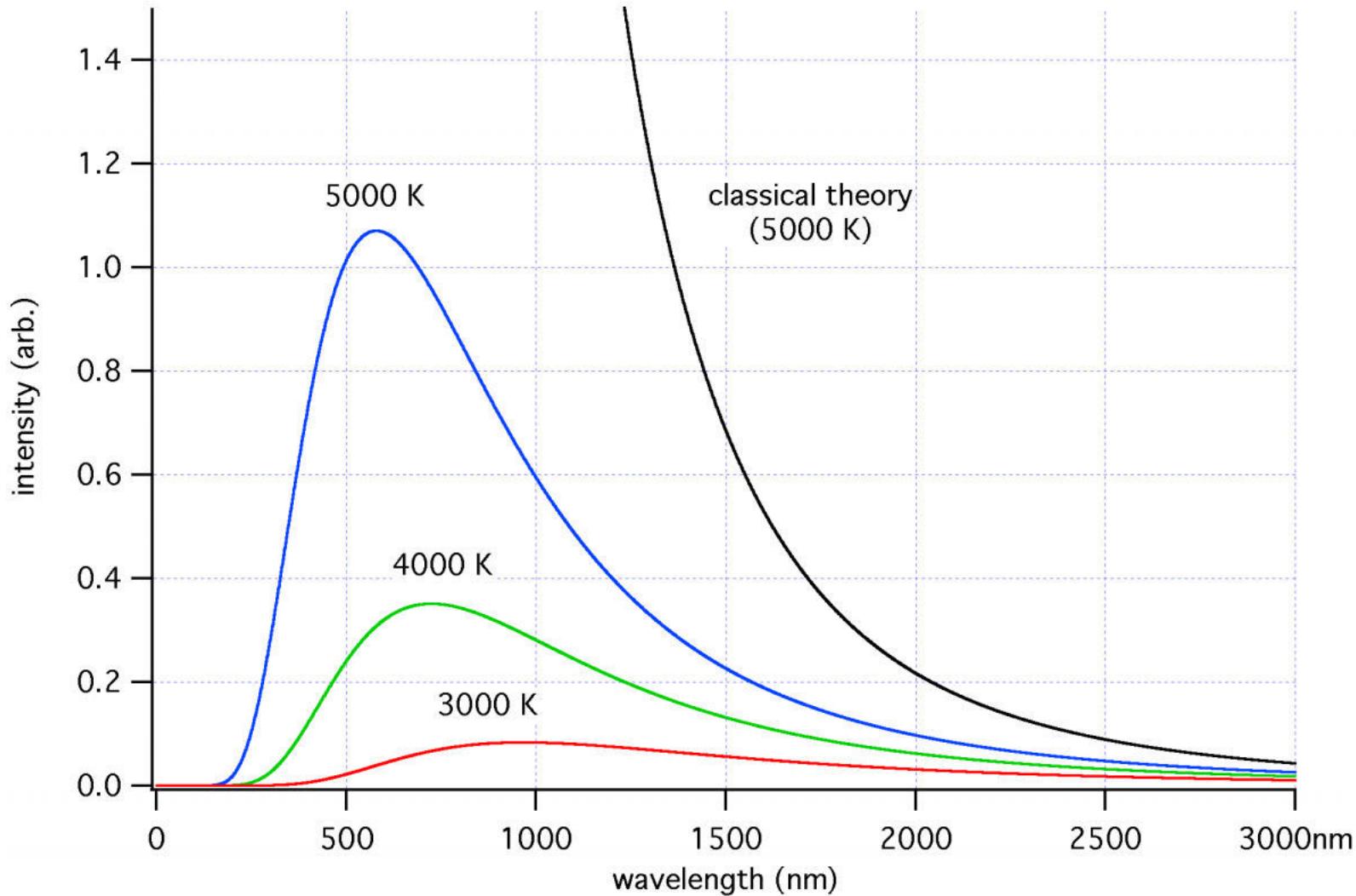
Blackbody Approximation

- A good approximation of a black body is a small hole leading to the inside of a hollow object
- The hole acts as a perfect absorber
- The nature of the radiation leaving the cavity through the hole depends only on the temperature of the cavity



©2004 Thomson - Brooks/Cole

Blackbody Radiation Spectrum



Wien's displacement law (1893):

the peak of the curve shifts towards longer wavelength as the temperature falls

$$\lambda_{peak}T = b$$

where b is called the **Wein's constant**.

This law is quite useful for measuring the temperature of a blackbody with a very high temperature.

The above laws describes the blackbody radiation very well.

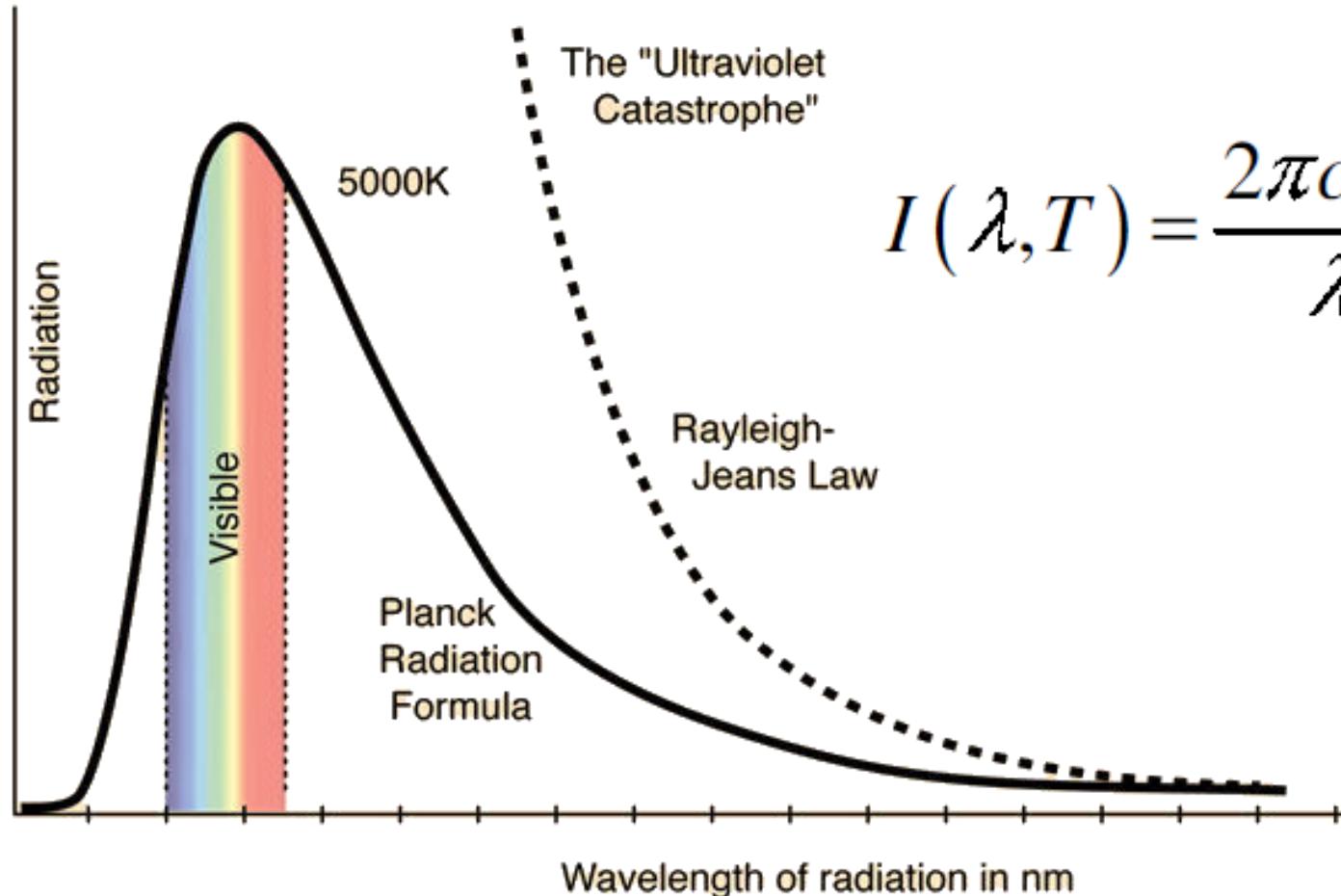
• Rayleigh and Jeans

In 1890, Rayleigh and Jeans obtained a formula using the classical electromagnetic (Maxwell) theory and the classical equipartition theorem of energy in thermodynamics .The formula is given by

$$I(\lambda, T) = \frac{2\pi c k_B T}{\lambda^4}$$

Rayleigh-Jeans formula was correct for **very long wavelength** in the far infrared but hopelessly wrong in the visible light and ultraviolet region.

Ultraviolet catastrophe³⁸



- Wein's formula:

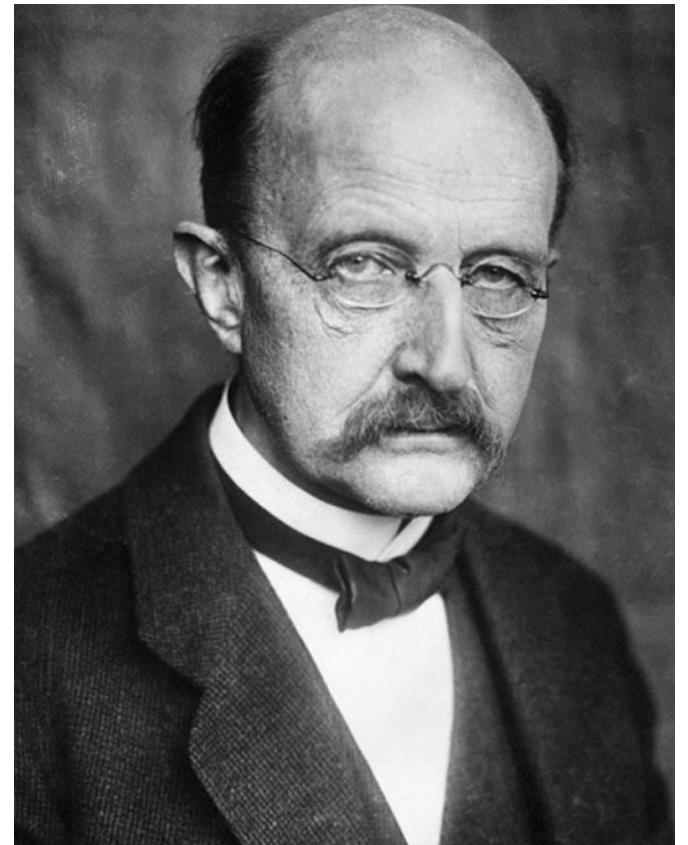
Later on in 1896, Wein derived another important formula using thermodynamics.

$$E_{\lambda}(T) = C_2 \lambda^{-5} e^{-\frac{C_3}{\lambda T}}$$

Unfortunately, this formula is only valid in the region of **short wavelengths**.

Max Planck

- Introduced the concept of **“quantum of action”**
- In 1918 he was awarded the Nobel Prize for the discovery of the quantized nature of energy



© 2004 Thomson - Brooks/Cole

Planck's radiation law

- Planck assumed that the radiation in the cavity was emitted (and absorbed) by some sort of “oscillators” contained in the walls. He used Boltzman’s statistical methods to arrive at the following formula:

$$I(\lambda, T) = \frac{2\pi c^2 h}{\lambda^5} \frac{1}{e^{hc/\lambda kT} - 1}$$

Planck's radiation law

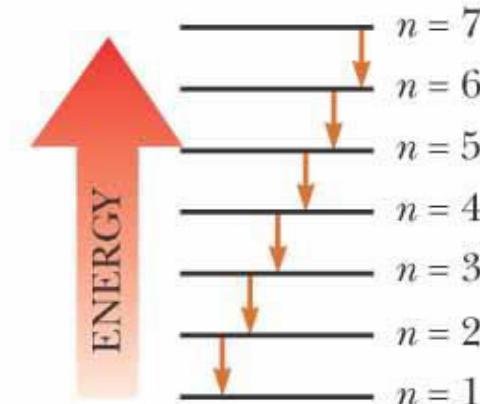
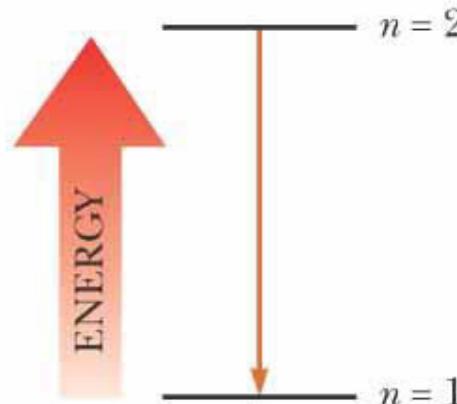
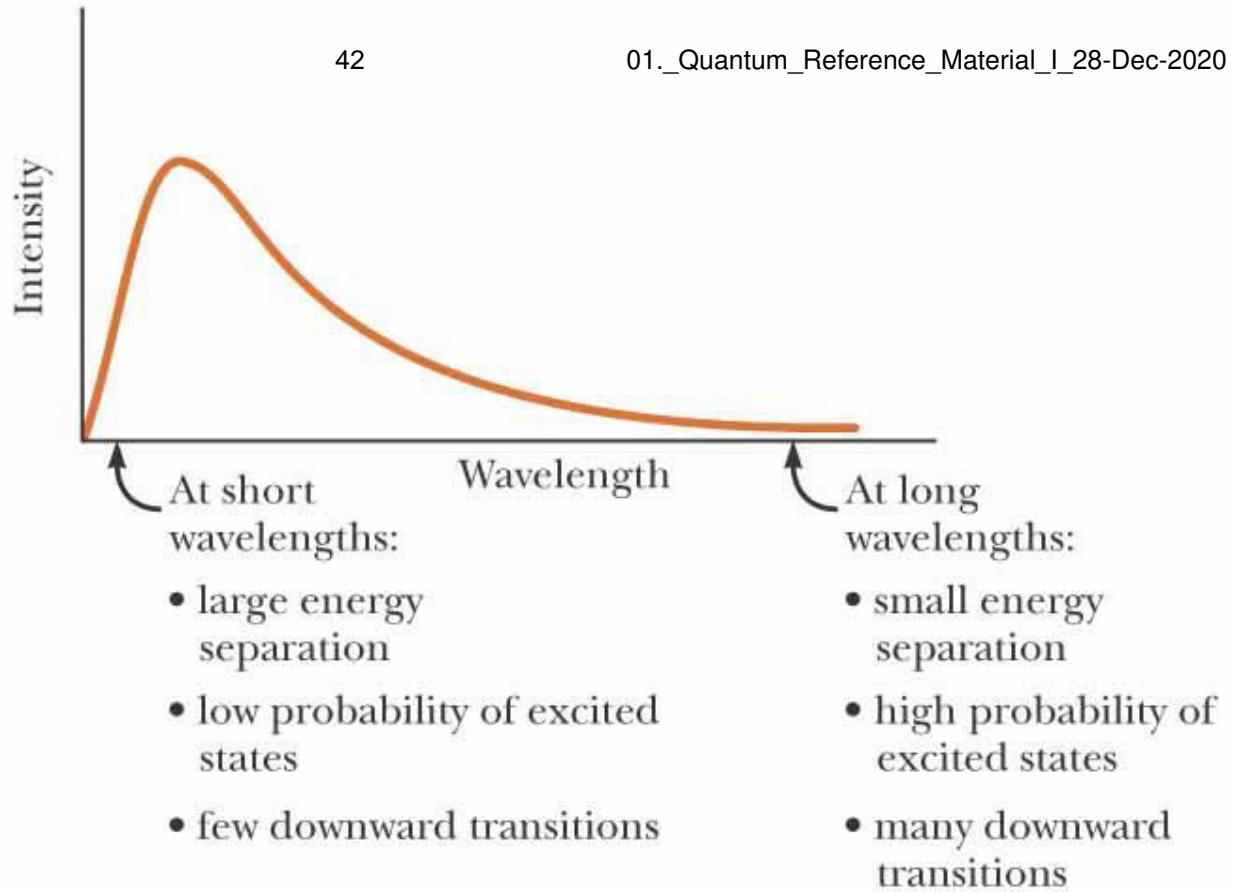
- Planck made two modifications to the classical theory:
 - 1) The oscillators (of electromagnetic origin) can only have certain discrete energies determined by $E_n = nh\nu$, where n is an integer, ν is the frequency, and h is called Planck's constant.

$$h = 6.6261 \times 10^{-34} \text{ J}\cdot\text{s.}$$

- 2) The oscillators can absorb or emit energy in discrete multiples of the fundamental quantum of energy given by

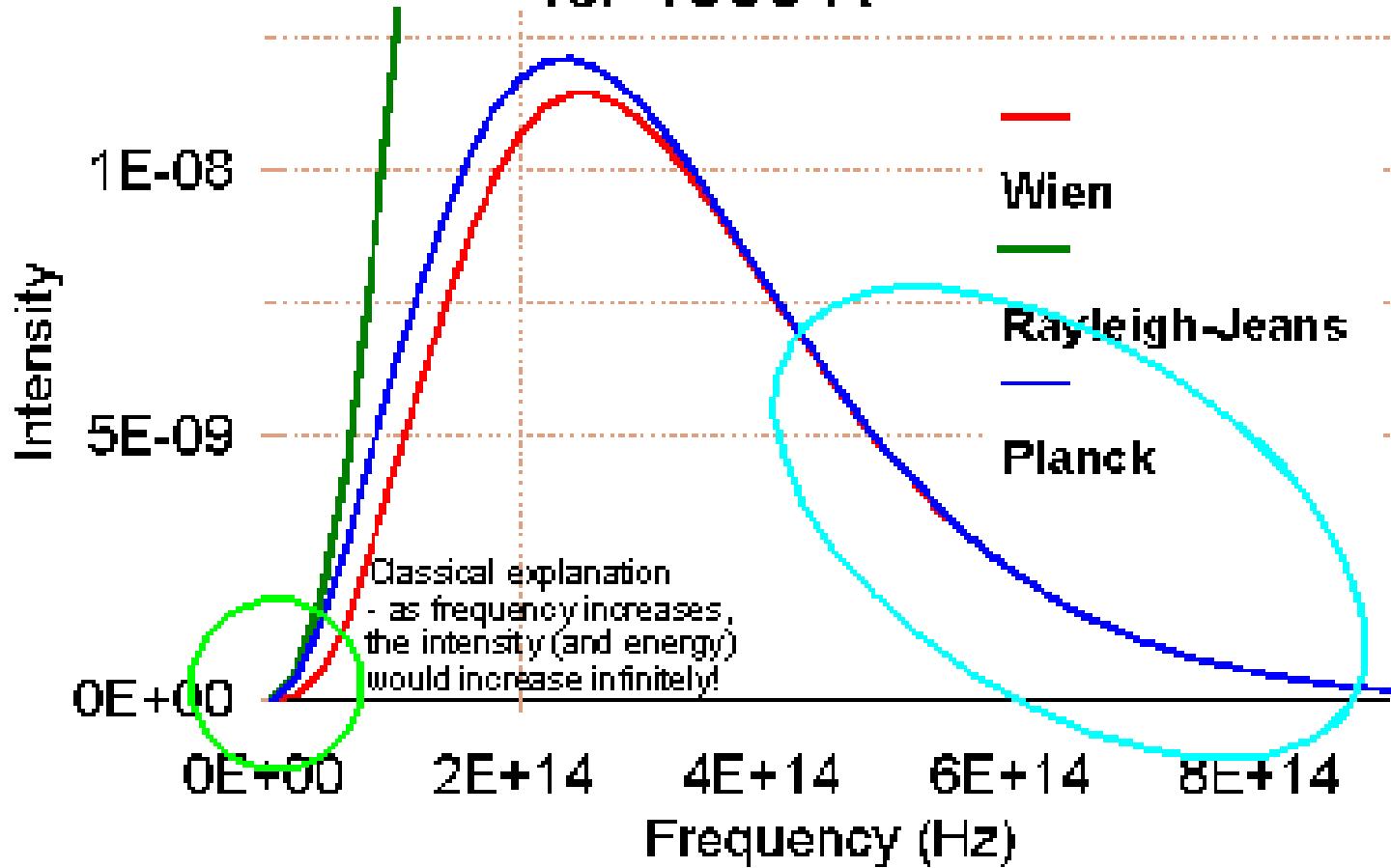
$$\Delta E = h\nu$$

Planck's Model, Graphs



Blackbody Radiation Laws

for 4000 K





"QUANTUM MECHANICS"

Wave Function

- Water waves – the quantity to vary periodically – HEIGHT
- Sound waves – Pressure
- Light waves – electric and magnetic fields
- What is it that varies in the case of matter waves?

Wave Function ψ

- The quantity whose variations make up matter waves is called the wave function

Wave Function

- Wave function ψ cannot be measured directly
- Wave function ψ can be a complex number
- Wave function ψ obeys the principle of superposition

Wave Function cannot be measured directly

- The ψ itself, has no direct physical significance
- Eg. The probability that something be in a certain place at a given time must lie between 0 and 1.
- An intermediate probability, say 0.2, means 20% chance
- But the amplitude of a wave can be –ve and +ve and a –ve probability is meaningless
- This objection does not apply to $|\psi|^2$, Probability density

Probability density

- The probability of experimentally finding the body described by the wave function ψ at the point x,y,z at the time t is proportional to the value of $|\psi|^2$ there at t.
- A large value of $|\psi|^2$ means the strong possibility of the body's presence
- A small value of $|\psi|^2$ means the slight possibility of its presence

Normalization and Probability

- The probability of the particle being between x_1 and x_2 is given by

$$P = \int_{x_1}^{x_2} \Psi^* \Psi \, dx$$

- The wave function must also be normalized so that the probability of the particle being somewhere on the x axis is 1.

$$\int_{-\infty}^{\infty} \Psi^*(x, t) \Psi(x, t) \, dx = 1$$

Wave function can be a complex number

- Wave functions are usually complex with real and imaginary parts
- Probability density $|\psi|^2 = \psi^* \psi$

Wave function

$$\Psi = A + iB$$

where A and B are real functions. The complex conjugate Ψ^* of Ψ is

Complex conjugate

$$\Psi^* = A - iB$$

and so

$$|\Psi|^2 = \Psi^* \Psi = A^2 - i^2 B^2 = A^2 + B^2$$

since $i^2 = -1$. Hence $|\Psi|^2 = \Psi^* \Psi$ is always a positive real quantity, as required.

Wave function obeys the principle of superposition

- Wave

$$\Psi = a_1 \Psi_1 + a_2 \Psi_2$$

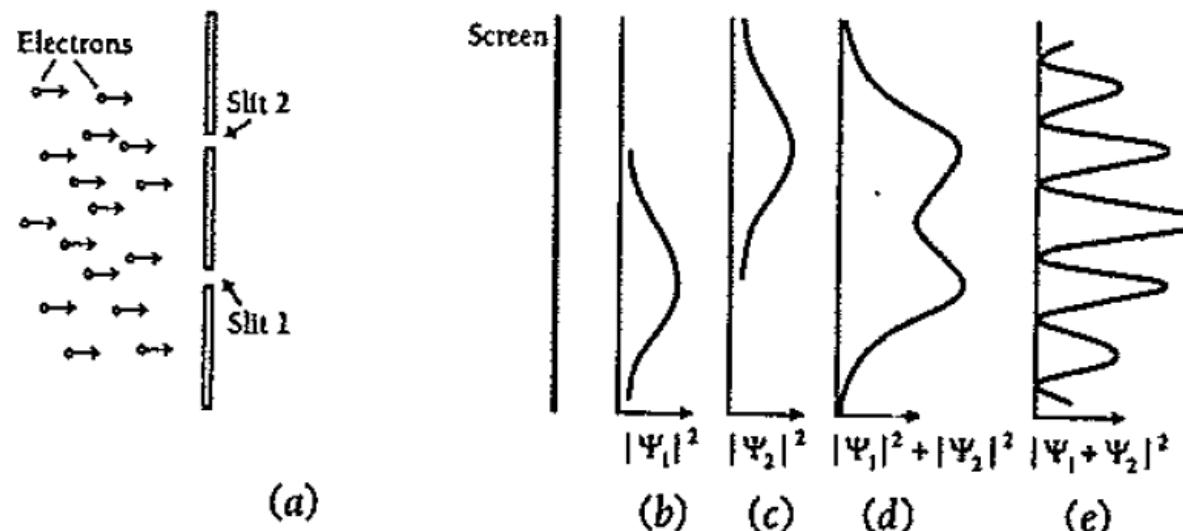


Figure 5.2 (a) Arrangement of double-slit experiment. (b) The electron intensity at the screen with only slit 1 open. (c) The electron intensity at the screen with only slit 2 open. (d) The sum of the intensities of (b) and (c). (e) The actual intensity at the screen with slits 1 and 2 both open. The wave functions Ψ_1 and Ψ_2 add to produce the intensity at the screen, not the probability densities $|\Psi_1|^2$ and $|\Psi_2|^2$.

Properties of Valid Wave Functions

Boundary conditions

- 1) In order to avoid infinite probabilities, the wave function must be finite everywhere.
 - 2) In order to avoid multiple values of the probability, the wave function must be single valued.
 - 3) For finite potentials, the wave function and its derivative must be continuous. This is required because the second-order derivative term in the wave equation must be single valued. (There are exceptions to this rule when V is infinite.)
 - 4) In order to normalize the wave functions, they must approach zero as x approaches infinity.
- Solutions that do not satisfy these properties do not generally correspond to physically realizable circumstances.

W. Heisenberg

and Uncertainty



879

Über quantentheoretische Umdeutung
kinematischer und mechanischer Beziehungen.

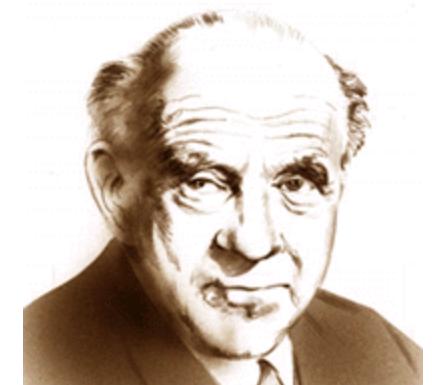
Von W. Heisenberg in Göttingen,
(Eingegangen am 29. Juli 1925.)

In der Arbeit soll versucht werden, Grundlagen zu gewinnen für eine quanten-theoretische Mechanik, die ausschließlich auf Beziehungen zwischen prinzipiell beobachtbaren Größen basiert ist.

Bekanntlich läßt sich gegen die formalen Regeln, die allgemein in der Quantentheorie zur Berechnung beobachtbarer Größen (z. B. der Energie im Wasserstoffatom) benutzt werden, der schwerwiegende Einwand erheben, daß jene Rechenregeln als wesentlichen Bestandteil Beziehungen enthalten zwischen Größen, die scheinbar prinzipiell nicht beobachtet werden können (wie z. B. Ort, Umlaufszeit, d. h. Zeit). Es muß also jenen Regeln offenbar jedes ansehen, daß es ein Mangel daran ist, wenn man nicht will, daß jene binären Beziehungen zwischen den beobachtbaren Größen bestehen.

Heisenberg realised that ...

- In the world of very small particles, one cannot measure any property of a particle without interacting with it in some way
- This introduces an unavoidable uncertainty into the result
- One can never measure all the properties exactly



Werner Heisenberg (1901-1976)

Heisenberg's Uncertainty Principle

$$\Delta x \Delta p \geq \frac{h}{4\pi}$$

↑
uncertainty
in position

↓
uncertainty
in momentum

The more accurately you know the position (i.e., the smaller Δx is) , the less accurately you know the momentum (i.e., the larger Δp is); and vice versa

Implications

- It is impossible to know *both* the position and momentum exactly, i.e., $\Delta x=0$ and $\Delta p=0$
- These uncertainties are inherent in the physical world and have nothing to do with the skill of the observer
- Because h is so small, these uncertainties are not observable in normal everyday situations

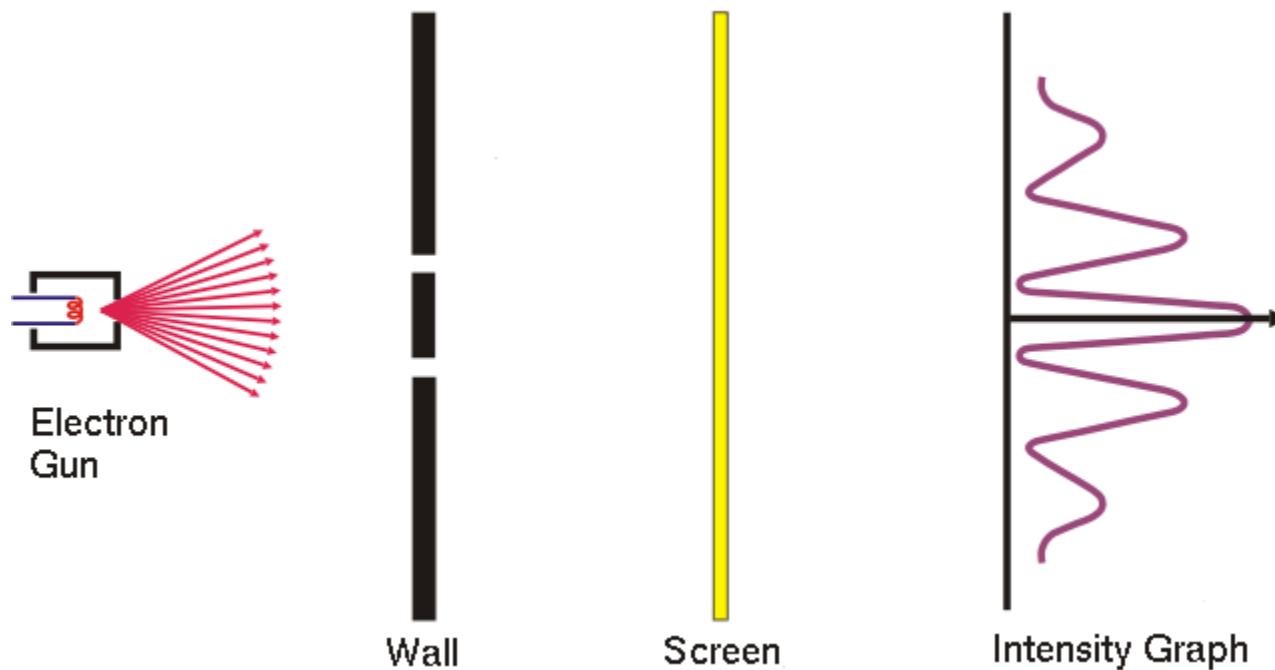
Heisenberg's Uncertainty Principle involving energy and time

$$\Delta E \Delta t \geq \frac{h}{4\pi}$$

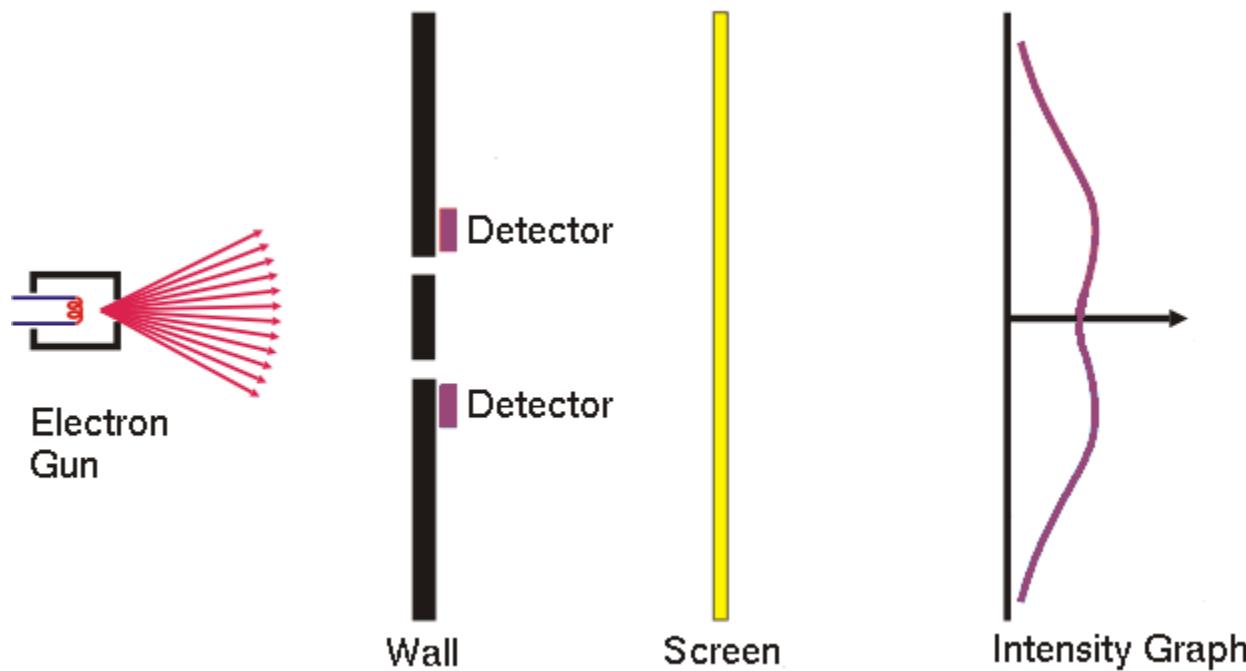
- The more accurately we know the energy of a body, the less accurately we know how long it possessed that energy
- The energy can be known with perfect precision ($\Delta E = 0$), only if the measurement is made over an infinite period of time ($\Delta t = \infty$)

Double-Slit Experiment:

cannot predict where electron would land

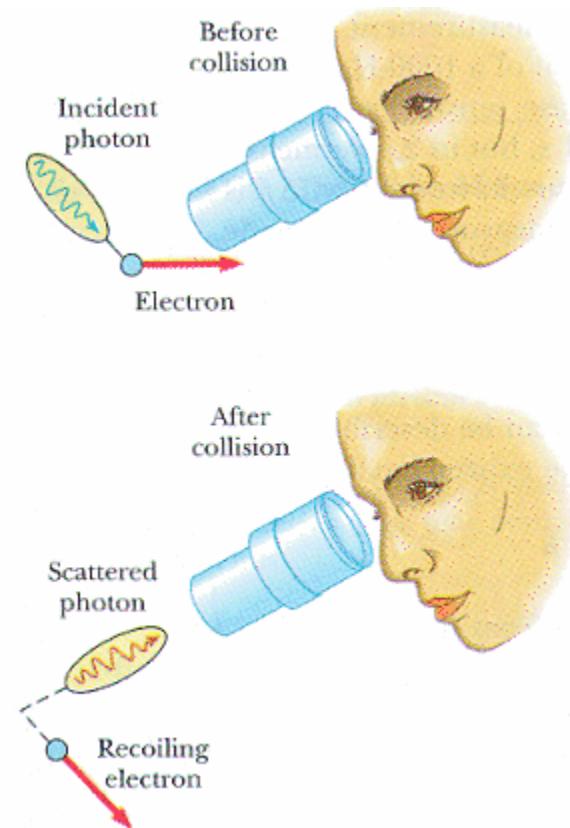


Double-Slit Experiment: act of observation affects behaviour of electron



Measuring the position and momentum of an electron

- Shine light on electron and detect reflected light using a microscope
- Minimum uncertainty in position is given by the wavelength of the light
- So to determine the position accurately, it is necessary to use light with a short wavelength



Measuring the position and momentum of an electron (cont'd)

- By Planck's law $E = hc/\lambda$, a photon with a short wavelength has a large energy
- Thus, it would impart a large 'kick' to the electron
- But to determine its momentum accurately, electron must only be given a small kick
- This means using light of long wavelength!

Fundamental Trade Off ...

- Use light with short wavelength:
 - accurate measurement of position but not momentum
- Use light with long wavelength:
 - accurate measurement of momentum but not position

Example of Baseball

- A pitcher throws a 0.1-kg baseball at 40 m/s
- So momentum is $0.1 \times 40 = 4 \text{ kg m/s}$
- Suppose the momentum is measured to an accuracy of 1 percent , i.e.,

$$\Delta p = 0.01 p = 4 \times 10^{-2} \text{ kg m/s}$$

Example of Baseball (cont'd)

- The uncertainty in position is then

$$\Delta x \geq \frac{h}{4\pi\Delta p} = 1.3 \times 10^{-33} \text{ m}$$

- No wonder one does not observe the effects of the uncertainty principle in everyday life!

Example of Electron

- Same situation, but baseball replaced by an electron which has mass 9.11×10^{-31} kg
- So momentum = 3.6×10^{-29} kg m/s and its uncertainty = 3.6×10^{-31} kg m/s
- The uncertainty in position is then

$$\Delta x \geq \frac{h}{4\pi\Delta p} = 1.4 \times 10^{-4} \text{ m}$$

Particle in a 1-Dimensional Box

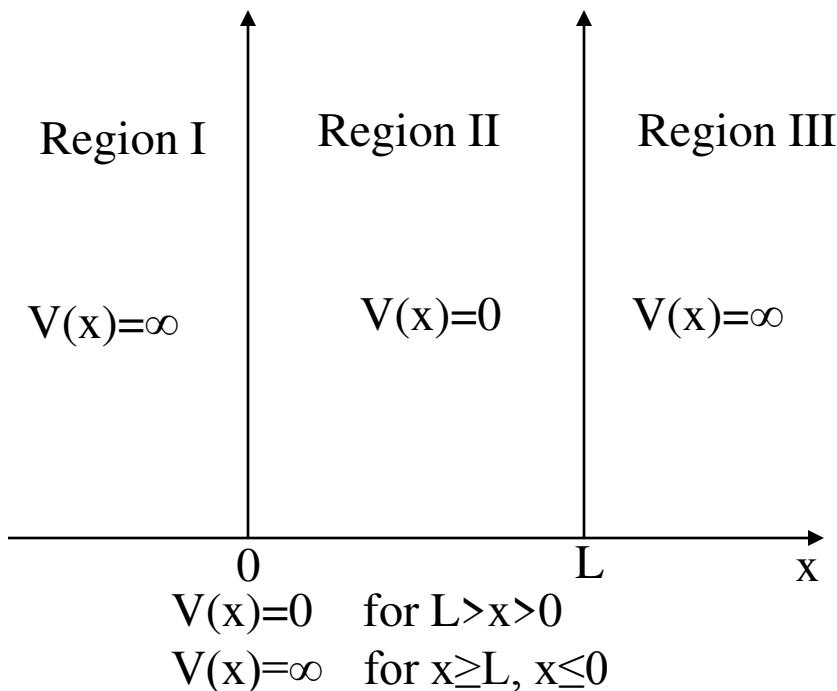
Time Dependent Schrödinger Equation

$$\frac{-\hbar^2}{2m} \frac{d^2\Psi}{dx^2} + V(x)\Psi = E\Psi$$

KE

PE

TE



Classical Physics: The particle can exist anywhere in the box and follow a path in accordance to Newton's Laws.

Quantum Physics: The particle is expressed by a wave function and there are certain areas more likely to contain the particle within the box.

Wave function is dependent on time and position function:

$$\Psi(x, t) = \overbrace{f(t)}^1 \psi(x)$$

Time Independent Schrödinger Equation

$$\frac{-\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} + V(x)\psi = E\psi$$

Applying boundary conditions:
 Region I and III:

$$\frac{-\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} + \infty^* \psi = E\psi \longrightarrow |\psi|^2 = 0$$

Region II:

$$\frac{-\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} = E\psi$$

Finding the Wave Function

$$\frac{-\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} = E\psi \longrightarrow -\frac{d^2\psi(x)}{dx^2} = \frac{2m}{\hbar^2} E\psi$$

This is similar to the general differential equation:

$$-\frac{d^2\psi(x)}{dx^2} = k^2\psi \rightarrow \psi = A \sin kx + B \cos kx$$

So we can start applying boundary conditions:

$$\begin{aligned} x=0 \quad \psi=0 \\ 0=A \sin 0k + B \cos 0k \rightarrow 0=0+B*1 \therefore B=0 \end{aligned}$$

$$\begin{aligned} x=L \quad \psi=0 \\ 0=A \sin kL \quad A \neq 0 \rightarrow kL=n\pi \quad \text{where } n=\mathbb{N}^* \end{aligned}$$

Calculating Energy Levels:

$$\begin{aligned} k^2 &= \frac{2mE}{\hbar^2} \rightarrow E = \frac{k^2\hbar^2}{2m} \rightarrow E = \frac{k^2\hbar^2}{2m4\pi^2} \\ \hbar &= \frac{h}{2\pi} \end{aligned}$$

$$E = \frac{n^2\pi^2}{L^2} \frac{h^2}{2m4\pi^2} \rightarrow E = \frac{n^2h^2}{8mL^2}$$

Our new wave function:

$$\psi_{II} = A \sin \frac{n\pi x}{L} \quad \text{But what is 'A'?}$$

Normalizing wave function:

$$\int_0^L (A \sin kx)^2 dx = 1$$

$$|A|^2 \left[\frac{x}{2} - \frac{\sin 2kx}{4k} \right]_0^L = 1$$

$$|A|^2 \left[\frac{L}{2} - \frac{\sin 2\frac{n\pi}{L}}{4\frac{n\pi}{L}} \right] = 1$$

Since $n=\mathbb{N}^*$

$$|A|^2 \left(\frac{L}{2} \right) = 1 \rightarrow |A| = \sqrt{\frac{2}{L}}$$

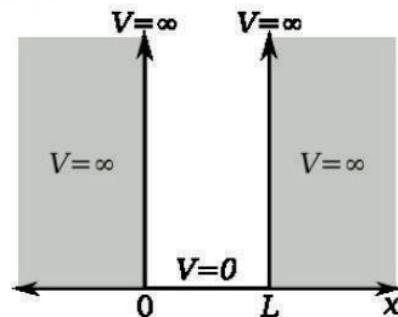
Our normalized wave function is:

$$\psi_{II} = \sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L}$$

Particle in a box

A simple system for which the time-independent Schrödinger equation can be solved exactly is the particle in a box. The particle in a box is represented by a small region in space where inside the box the potential is zero and at its wall and outside is infinite. Thus, the particle is free to move inside the box since it experiences no forces, but remain in the box since the walls are infinitely high. While this system might seem very abstract it can to some degree be used to gain insight into the optical properties of e.g. quantum dots and conjugated molecules.

Particle in a 1D-box



The particle in a 1D box of length L is represented by a potential given by

$$V(x) = \begin{cases} \infty & \text{for } x \leq 0 \\ 0 & \text{for } 0 < x < L \\ \infty & \text{for } x \geq L \end{cases}$$

From this we can see that outside the box the wavefunction must be zero since

$$\infty\psi_I(x) = \frac{\hbar^2}{2m} \frac{d^2\psi_I(x)}{dx^2}$$

$$\psi_I(x) = \frac{1}{\infty} \frac{\hbar^2}{2m} \frac{d^2\psi_I(x)}{dx^2} = 0$$

Therefore, we only need to solve the Schrödinger equation for the particle inside the box

$$-\frac{\hbar^2}{2m} \frac{d^2\psi_{II}(x)}{dx^2} + V(x) = E\psi_{II}(x)$$

$$\begin{aligned}\frac{\hbar^2}{2m} \frac{d^2\psi_{II}(x)}{dx^2} + E\psi_{II}(x) &= 0 \\ \frac{d^2\psi_{II}(x)}{dx^2} + \frac{2m}{\hbar^2} E\psi_{II}(x) &= 0 \\ y'' + qy &= 0\end{aligned}$$

Which we recognize as a linear homogenous second-order differential equation with constant coefficients. We know that the solution is of the form $y = e^{sx}$ and we can write the auxiliary equation as

$$\begin{aligned}s^2 + 2mE\hbar^{-2} &= 0 \\ s^2 + q &= 0 \\ s &= \pm\sqrt{-q} \\ s &= \pm i\sqrt{q} \\ s &= \pm i\sqrt{2mE}/\hbar\end{aligned}$$

Therefore, the solution is given by

$$\psi_{II}(x) = c_1 \exp^{i\sqrt{2mE}x/\hbar} + c_2 \exp^{-i\sqrt{2mE}x/\hbar}$$

We can now manipulate the solution in the following way

$$\psi_{II}(x) = c_1 \exp(i\theta) + c_2 \exp(-i\theta)$$

where $\theta = \sqrt{2mE}x/\hbar$. Using that $\exp(i\theta) = \cos(\theta) + i\sin(\theta)$ we get

$$\begin{aligned}\psi_{II}(x) &= c_1 \exp(i\theta) + c_2 \exp(-i\theta) \\ &= c_1 \cos(\theta) + ic_1 \sin(\theta) + c_2 \cos(-\theta) + ic_2 \sin(-\theta) \\ &= c_1 \cos(\theta) + ic_1 \sin(\theta) + c_2 \cos(\theta) - ic_2 \sin(\theta) \\ &= (c_1 + c_2) \cos(\theta) + (ic_1 - ic_2) \sin(\theta) \\ &= A \cos(\theta) + B \sin(\theta)\end{aligned}$$

where A and B are two new arbitrary constants. We can determine these constants by applying the boundary conditions. We will postulate, reasonable, that the wavefunction needs to be continuous, therefore, we require that the wavefunction is also continuous at the boundary

$$\begin{aligned}\lim_{x \rightarrow 0} \psi_I &= \lim_{x \rightarrow 0} \psi_{II} \\ 0 &= \lim_{x \rightarrow 0} \psi_{II} \{A \cos(\theta) + B \sin(\theta)\} \\ 0 &= A \cos(0) + B \sin(0) = A\end{aligned}$$

where n is the quantum number. Draw the first few solution of the wavefunction on the board, and the square of the solutions as well. Compare this with what is expected from classical physics ? Where is the particle ? What happens when n goes toward ∞ ? In this limit we reach the classical results of uniform probability density. This is known as the Bohr correspondence principle.

We know that the wavefunction is normalized so

$$\int_{-\infty}^{\infty} \psi_i^* \psi_i dx = 1$$

However, what is we use two wavefunctions with different quantum numbers

$$\int_{-\infty}^{\infty} \psi_i^* \psi_j dx = 2/L \int_0^L \sin(n_i \pi x/L) \sin(n_j \pi x/L) dx = 0$$

where we have used the integral table in Appendix. There, two wavefunctions with different quantum numbers are orthogonal

$$\int_{-\infty}^{\infty} \psi_i^* \psi_j dx = \delta_{ij}$$

where δ_{ij} is Kronecker delta, which is one if $i = j$ otherwise zero.

Therefore,

$$\psi_{II}(x) = B \sin(\theta)$$

Applying the boundary condition at $x = L$ we get

$$\psi_{II}(x) = B \sin(\sqrt{2mE}L/\hbar) = 0$$

This is satisfied for

$$\sqrt{2mE}L/\hbar = \pm n\pi \quad n = 1, 2, 3, \dots$$

What about $n = 0$? This corresponds to $\psi_{II}(x) = 0$ and thus no particle. This gives us the energy

$$E = \frac{n^2\pi^2\hbar^2}{2mL^2} = \frac{n^2h^2}{8mL^2}$$

Therefore, we see that quantization of the energy for the particle in a box comes directly as a result of the Schrödinger equation. The wavefunction is

$$\psi_{II}(x) = B \sin\left(\frac{n\pi x}{L}\right), \quad n = 1, 2, 3, \dots$$

How do we then determine B? We use the normalization requirement

$$\begin{aligned} \int_{-\infty}^{\infty} |\psi(x)|^2 dx &= 1 \\ \int_0^L |\psi(x)|^2 dx &= 1 \\ |B|^2 \int_0^L \sin^2\left(\frac{n\pi x}{L}\right) dx &= 1 \\ |B|^2 \frac{L}{2} &= 1 \\ |B| &= \sqrt{2/L} \end{aligned}$$

where we have used that

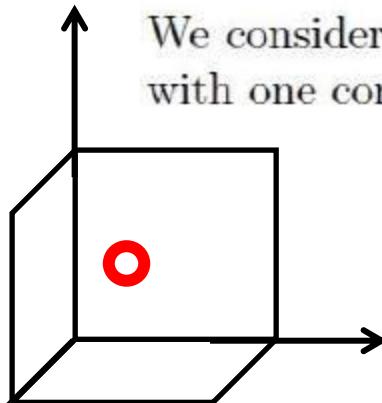
$$\int \sin^2 bx dx = \frac{x}{2} - \frac{1}{4b} \sin(2bx)$$

from Table A.5 on page 721 on Levine. Therefore, the final solution for the wavefunction is

$$\psi_{II}(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right) \quad n = 1, 2, 3, \dots$$

Particle in a 3D-box

We consider a particle in a box with lengths a , b , and c which is centered with one corner at the origin. The potential is then given by



$$V(x, y, z) = \begin{cases} 0 & \text{if } 0 < x < a \\ 0 & \text{if } 0 < y < b \\ 0 & \text{if } 0 < z < c \\ \infty & \text{elsewhere} \end{cases}$$

Now lets consider a cube, i.e. $a = b = c$. The energy of a particle in cubic box becomes

$$E = \frac{\hbar^2}{8m} \left(\frac{n_x^2}{a^2} + \frac{n_y^2}{a^2} + \frac{n_z^2}{a^2} \right) = \frac{\hbar^2}{8ma^2} (n_x^2 + n_y^2 + n_z^2)$$

and a wavefunction

$$\psi(x, y, z) = \left(\frac{8}{abc} \right)^{1/2} \sin \left(\frac{n_x \pi x}{a} \right) \sin \left(\frac{n_y \pi y}{b} \right) \sin \left(\frac{n_z \pi z}{c} \right)$$

1.1 Scanning Tunneling Microscopy

The scanning tunneling microscope (STM) is the ancestor of all scanning probe microscopes. It was invented in 1981 by Gerd Binnig and Heinrich Rohrer at IBM Zurich. Five years later they were awarded the Nobel prize in physics for its invention. The STM was the first instrument to generate real-space images of surfaces with atomic resolution.

STMs use a sharpened, conducting tip with a bias voltage applied between the tip and the sample. When the tip is brought within about 10Å of the sample, electrons from the sample begin to “tunnel” through the 10Å gap into the tip or vice versa, depending upon the sign of the bias voltage. (See Figure 1-2.) The resulting tunneling current varies with tip-to-sample spacing, and it is the signal used to create an STM image. For tunneling to take place, both the sample and the tip must be conductors or semiconductors. Unlike AFMs, which are discussed in the next section, STMs cannot image insulating materials.

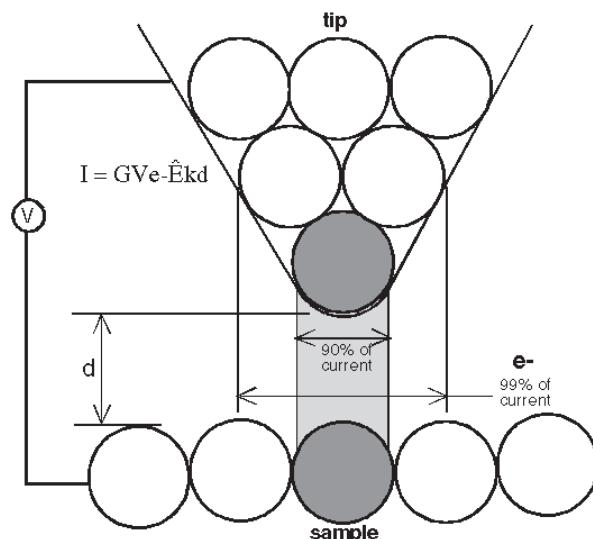


Figure 1-2. Schematic of tip and sample interaction for STM.

The tunneling current is an exponential function of distance; if the separation between the tip and the sample changes by 10% (on the order of 1Å), the

SPM Techniques 3

tunneling current changes by an order of magnitude. This exponential dependence gives STMs their remarkable sensitivity. STMs can image the surface of the sample with sub-angstrom precision vertically, and atomic resolution laterally.

STMs can be designed to scan a sample in either of two modes: *constant-height* or *constant-current* mode, as shown in Figure 1-3.

In constant-height mode, the tip travels in a horizontal plane above the sample and the tunneling current varies depending on topography and the local surface electronic properties of the sample. The tunneling current measured at each location on the sample surface constitute the data set, the topographic image.

In constant-current mode, STMs use feedback to keep the tunneling current constant by adjusting the height of the scanner at each measurement point. For example, when the system detects an increase in tunneling current, it adjusts the voltage applied to the piezoelectric scanner to increase the distance between the tip and the sample.

In constant-current mode, the motion of the scanner constitutes the data set. If the system keeps the tunneling current constant to within a few percent, the tip-to-sample distance will be constant to within a few hundredths of an angstrom.

Each mode has advantages and disadvantages. Constant-height mode is faster because the system doesn't have to move the scanner up and down, but it provides useful information only for relatively smooth surfaces. Constant-current mode can measure irregular surfaces with high precision, but the measurement takes more time.

4 Chapter 1

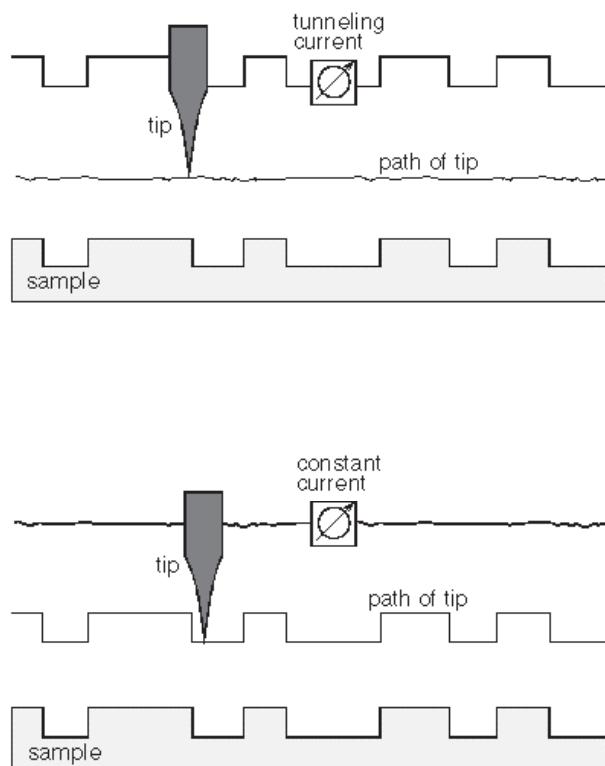


Figure 1-3. Comparison of constant-height and constant-current mode for STM.

As a first approximation, an image of the tunneling current maps the topography of the sample. More accurately, the tunneling current corresponds to the electronic density of states at the surface. STMs actually sense the number of filled or unfilled electron states near the Fermi surface, within an energy range determined by the bias voltage. Rather than measuring physical topography, it measures a surface of constant tunneling probability.

Infinite Square-Well Potential

- The simplest such system is that of a particle trapped in a box with infinitely hard walls that the particle cannot penetrate. This potential is called an infinite square well and is given by

$$V(x) = \begin{cases} \infty & x \leq 0, x \geq L \\ 0 & 0 < x < L \end{cases}$$

- Clearly the wave function must be zero where the potential is infinite.
- Where the potential is zero inside the box, the Schrödinger wave

equation becomes $\frac{d^2\psi}{dx^2} = -\frac{2mE}{\hbar^2}\psi = -k^2\psi$ where $k = \sqrt{2mE/\hbar^2}$.

- The general solution is $\psi(x) = A \sin kx + B \cos kx$.

Quantization

- Boundary conditions of the potential dictate that the wave function must be zero at $x = 0$ and $x = L$. This yields valid solutions for integer values of n such that $kL = n\pi$.
- The wave function is now $\psi_n(x) = A \sin\left(\frac{n\pi x}{L}\right)$
- We normalize the wave function

$$\int_{-\infty}^{\infty} \psi_n^*(x) \psi_n(x) dx = 1 \quad A^2 \int_0^L \sin^2\left(\frac{n\pi x}{L}\right) dx = 1$$

- The normalized wave function becomes

$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right)$$

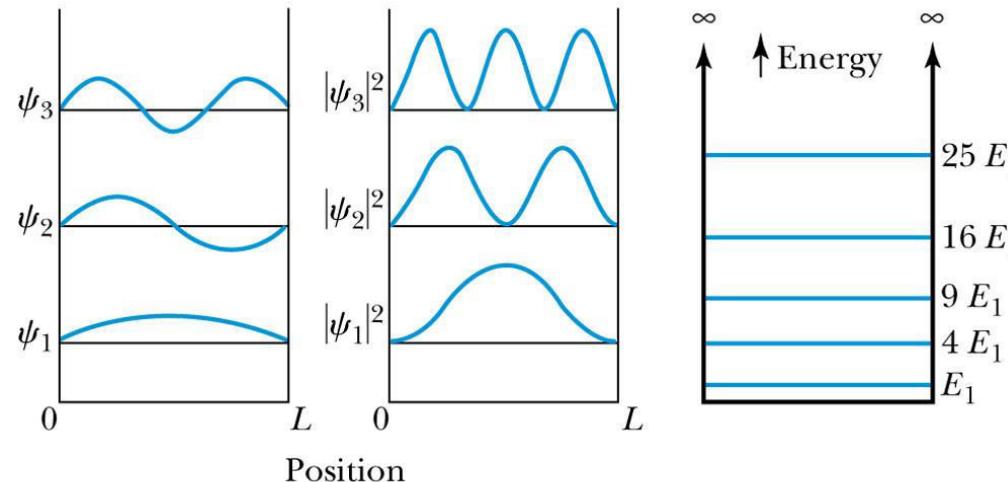
- These functions are identical to those obtained for a vibrating string with fixed ends.

Quantized Energy

- The quantized wave number now becomes $k_n = \frac{n\pi}{L} = \sqrt{\frac{2mE_n}{\hbar^2}}$
- Solving for the energy yields

$$E_n = n^2 \frac{\pi^2 \hbar^2}{2mL^2} \quad (n = 1, 2, 3, \dots)$$

- Note that the energy depends on the integer values of n . Hence the energy is quantized and nonzero.
- The special case of $n = 0$ is called the ground state energy. $E_1 = \frac{\pi^2 \hbar^2}{2mL^2}$



Classical PictureC
1 electron → electric field

in classical physics, the electron is repelled by an electric field as long as energy of electron is below energy level of the field

Quantum Picture

electron wave →



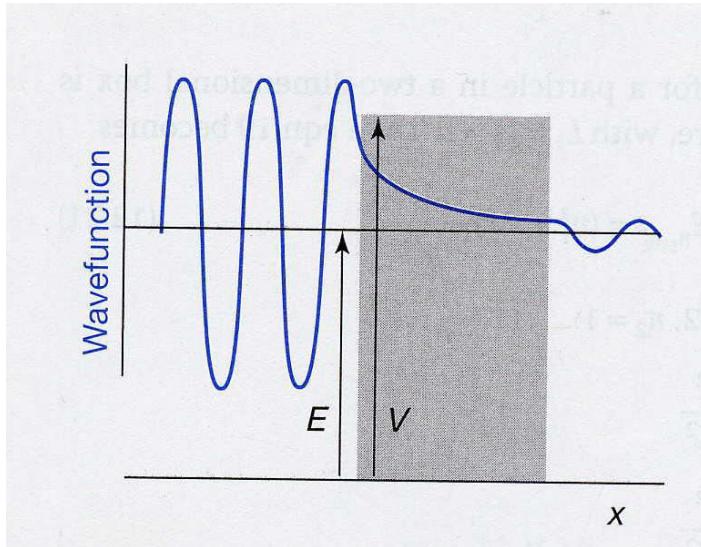
in quantum physics, the wave function of the electron encounters the electric field, but has some finite probability of tunneling through



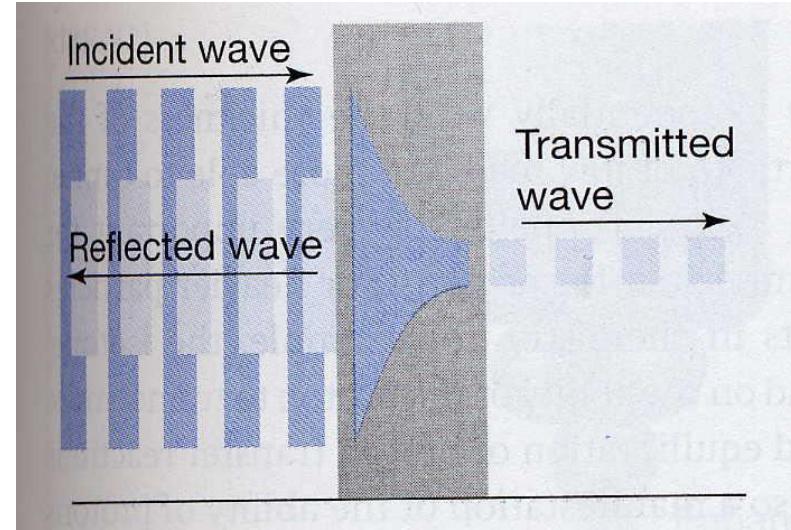
this is the basis for transistors



Quantum Mechanical Tunneling



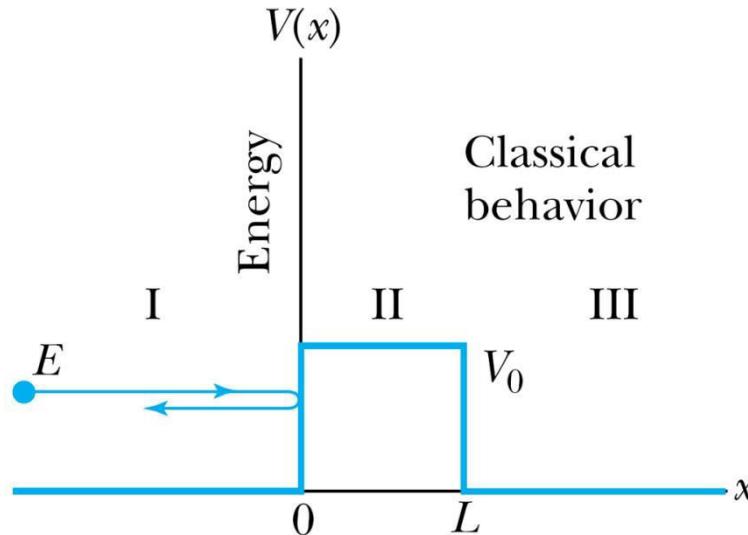
| L |



| L |

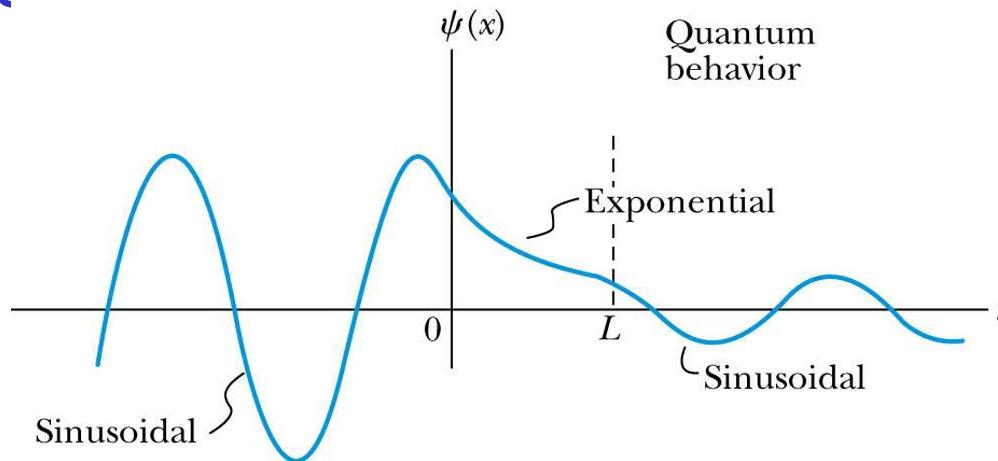
- Quantum mechanics allows a small particle, such as an electron, to overcome a potential barrier larger than its kinetic energy.
- Tunneling is possible because of the wave-like properties of matter.

Tunneling



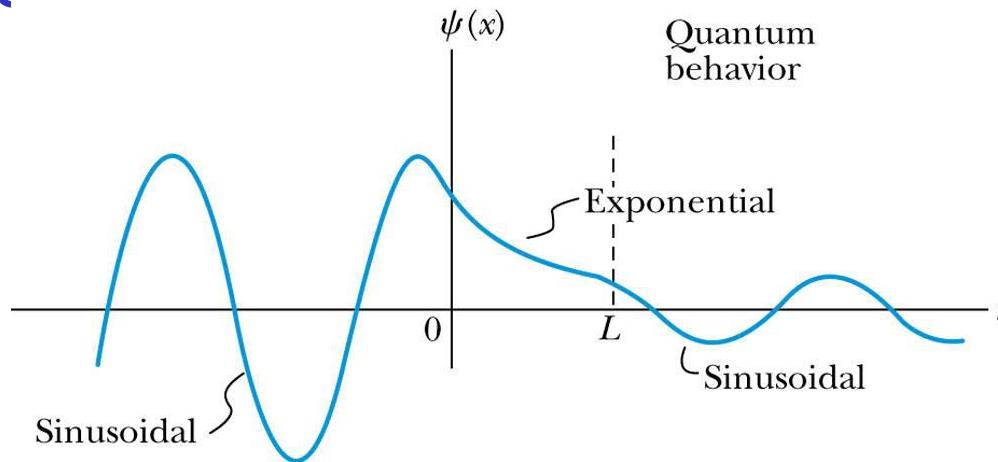
- Consider the potential energy function in which the potential energy of the system is zero everywhere except L .
- Square barrier – U is the barrier height
- If $E>U$, classically the particle reflected back.
- Classically if the particle exist in region II, its KE would be negative
- Classically region II and III are forbidden to the particle

Tunneling



- Acc to QM, all the regions are accessible regardless of its energy
- Schrodinger eqn has valid solutions for I and III which are sinusoidal, region II solution is exponential
- Applying the boundary conditions, the wave functions in the three regions must join smoothly at the boundaries
- Thus mathematically it is satisfied

Tunneling

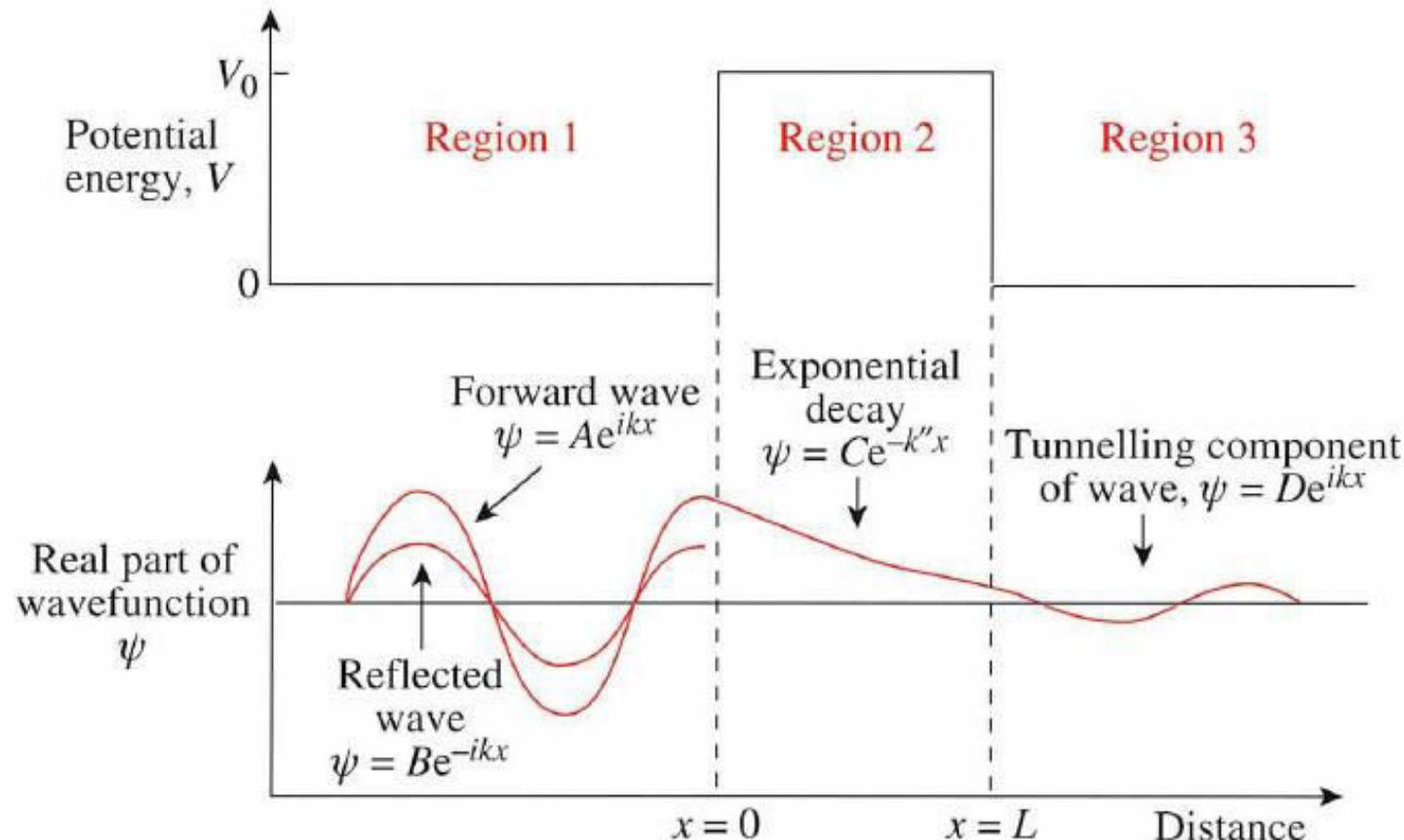


- Hence finding the particle beyond the barrier in region III is nonzero
- The movement of the particle to the far side of the barrier is called tunneling or barrier penetration

~~CLASSICAL~~

QUANTUM

Tunneling



Transmission and Reflection coefficient

- The transmission coefficient represents the probability that the particle penetrates to the other side of the barrier
- The reflection coefficient is the probability that the particle is reflected by the barrier

$$T + R = 1$$

$$T = e^{-2CL}, \text{ where } C = \frac{\sqrt{2m(U - E)}}{\hbar}$$

Periodic graphics

A collaboration between C&EN and Andy Brunning, author of the popular graphics blog **Compound Interest**

More online

 To see more of Brunning's work, go to compoundchem.com. To see all of C&EN's Periodic Graphics, visit cenm.ag/periodicgraphics.

EVERYDAY USES OF NANOTECHNOLOGY

National Nanotechnology Day (Oct. 9) is a yearly event in the U.S. to celebrate the tiny tech. Here, we take a look at various consumer products that utilize nanotechnology and the chemistry behind them.

WHAT IS NANOTECHNOLOGY?



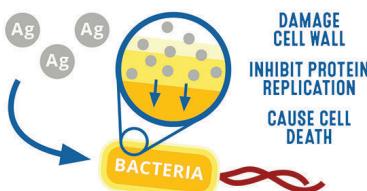
SALT GRAIN = 100,000 nm



NANOPARTICLES = 1–100 nm

Nanotechnology involves the applications of nanoparticles, which are collections of atoms or molecules less than 100 nm across. Because of their small size, the particles have properties that can differ from those of larger amounts of the same material.

ANTIMICROBIAL USES



SUNSCREENS



ZnO MAINLY BLOCKS UV-A



UV-A wavelength 320–400 nm



UV-B wavelength 290–320 nm

CLOTHES



SILVER ANTIMICROBIAL

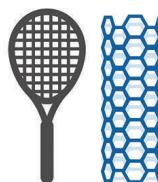
SILICA WATER-REPELLENT

TiO₂/ZnO UV-ABSORBING

ANTIMONY-DOPED TIN OXIDE ANTISTATIC

UV-absorbing titanium oxide and zinc oxide nanoparticles can be incorporated into clothes to prevent sunburn and sometimes to act as antistatic agents. Silicon dioxide nanoparticles can prevent stains and help clothing repel water.

SPORTS EQUIPMENT



CARBON NANOTUBES

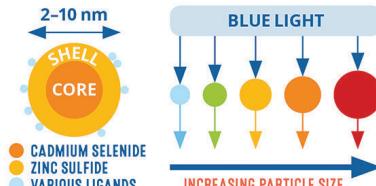
100 TIMES AS STRONG AS STEEL

ONE-SIXTH THE WEIGHT OF STEEL

AS STIFF AS DIAMOND

Sports equipment such as tennis rackets and bicycles are sometimes built using nanomaterials including carbon nanotubes. The nanotubes improve strength and durability and decrease weight. Titanium nanoparticles can also be used.

QUANTUM DOTS



Quantum dots, which are nanoparticles of semiconductors such as cadmium selenide, absorb light of one color, such as blue light, and emit it as another depending on particle size. The particles are more energy-efficient than light-emitting diodes.



© C&EN 2016 Created by Andy Brunning for *Chemical & Engineering News*

Did You Know?

- Carbon nanotubes, composed of interlocking carbon atoms, are 1000x thinner than an average human hair – but can be 200x stronger than steel.

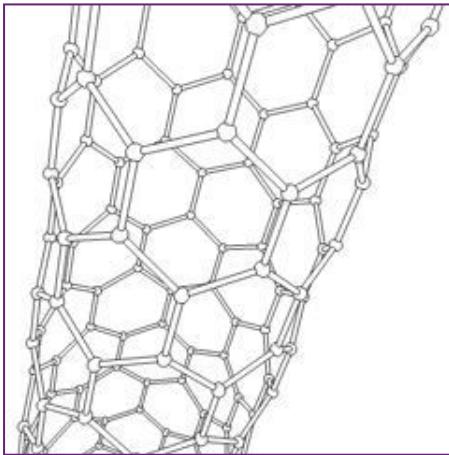


Image: Schwarzm, Wikipedia



CNT - Discovery

In 1991, Sumio Iijima discovers **multiwalled nanotubes** (MWNT) using the method of Krätschmer and Huffman.

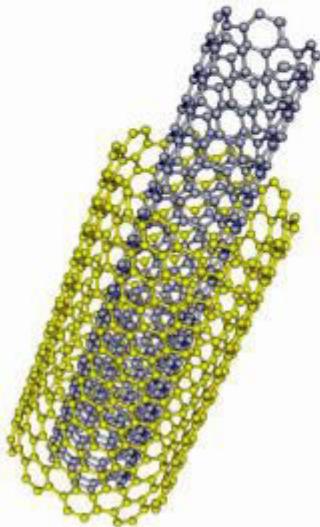
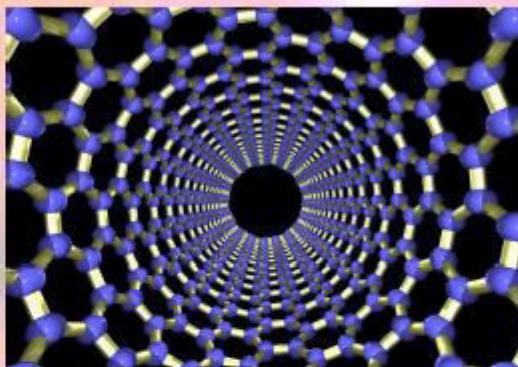


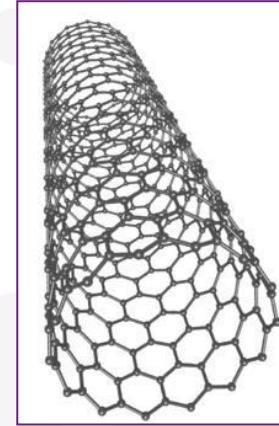
Image: Schwarzm, wikipedia



In 1993, Donald Bethune makes **single-walled** (SWNT) nanotubes by adding transition metals.



Which Of These Objects Are Made From Carbon?



Diamond

Graphite

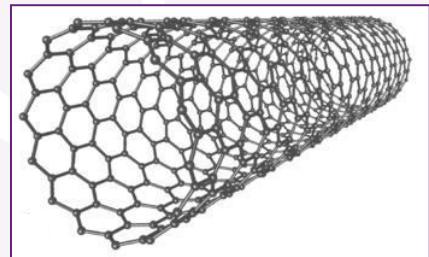
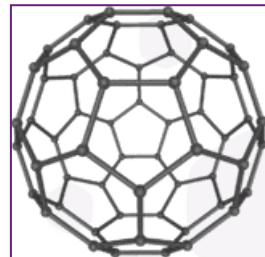
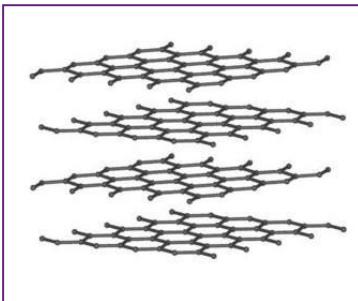
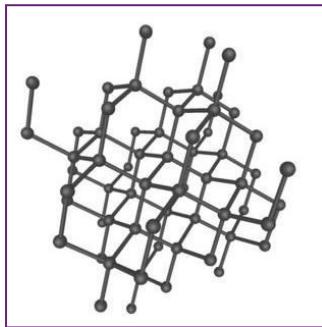
Nanotube

Coal



Did You Know?

Allotropes of carbon have different covalent bonding arrangements.



diamond

graphite

buckyball

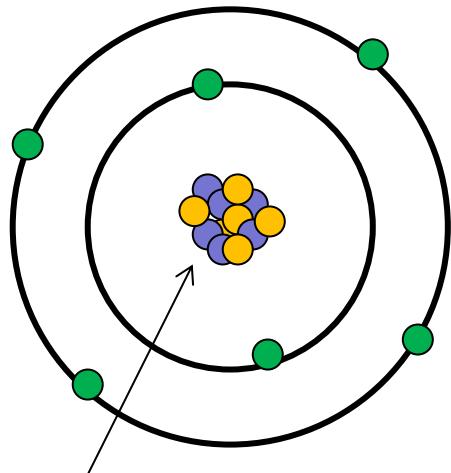
nanotube

- Carbon atoms form covalent bonds by sharing outer shell electrons with each other
- Diamond, graphite, buckyballs and carbon nanotubes all have different covalent arrangements of carbon atoms
- The differing covalent arrangements of carbon atoms lead to the different properties of carbon allotropes.



Covalent Bonding Sharing Electrons

- proton
- neutron
- electron



6 protons + 6
neutrons

A **covalent bond** is a form of chemical bonding that is characterised by the sharing of pairs of electrons between atoms

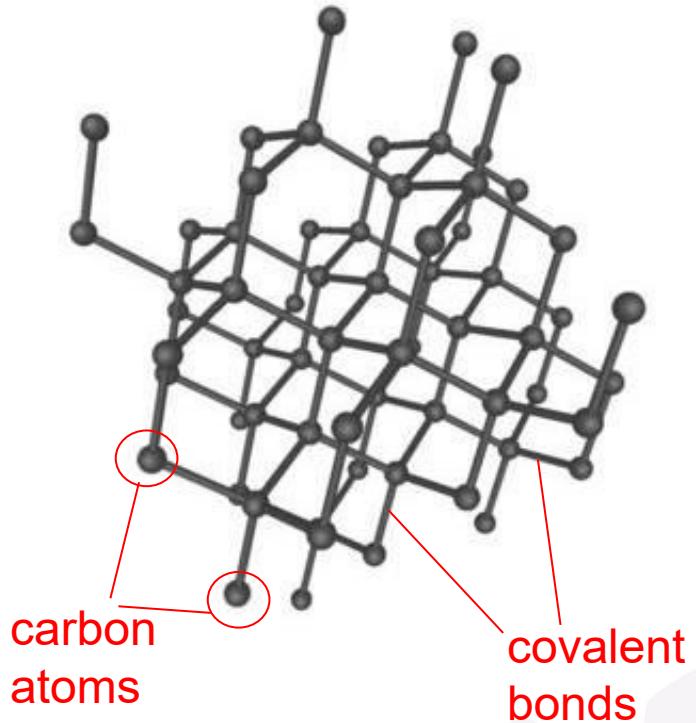
Valence electrons are the electrons in the outer shell or energy level of an atom that form covalent bonds

A carbon atom has 6 electrons, 4 of which are Valence electrons

Therefore, carbon atoms can form up to 4 Covalent Bonds



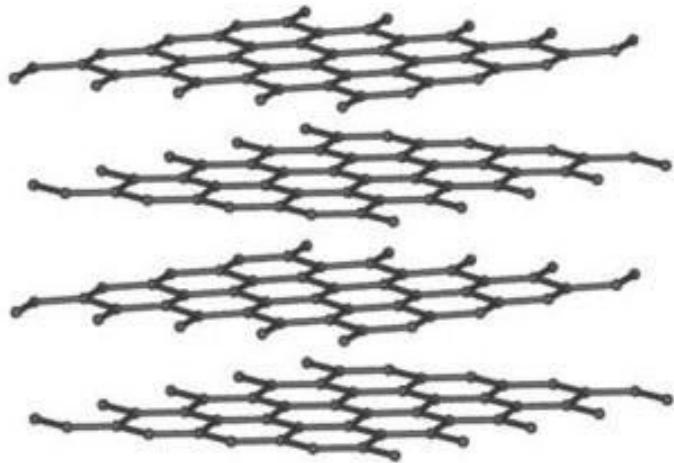
Covalent Bonds In Diamond



- Diamond is formed by a 3D box-like network of carbon atoms
- The continuous nature of the covalent arrangements forms a giant molecule
- Electrons are fixed.



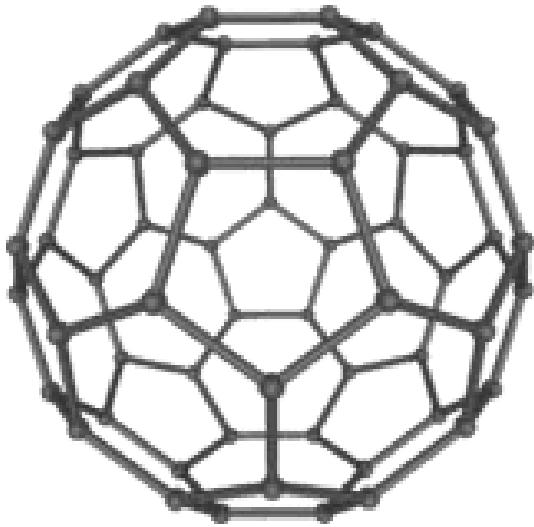
Covalent Bonds In Graphite



- Graphite is formed by hexagonally-arranged carbon molecules forming 2D layers of sheets
- Electrons are free to move between each carbon sheet.



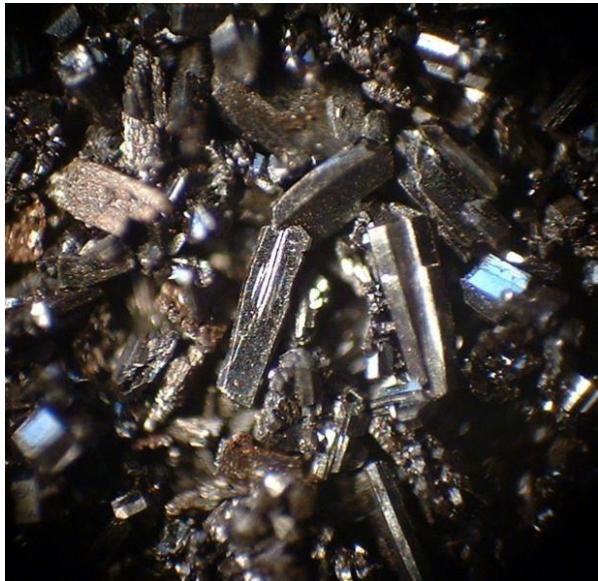
Covalent Bonds In Buckyballs



- Carbon atoms in buckyballs are arranged in a soccer ball shape
- C₆₀ Buckyballs have 20 regular hexagon faces and 12 regular pentagon faces
 - these faces come together at 60 carbon atom vertices
- Electrons are localised internally due to the curvature of the structure.



A Bit More About Buckyballs

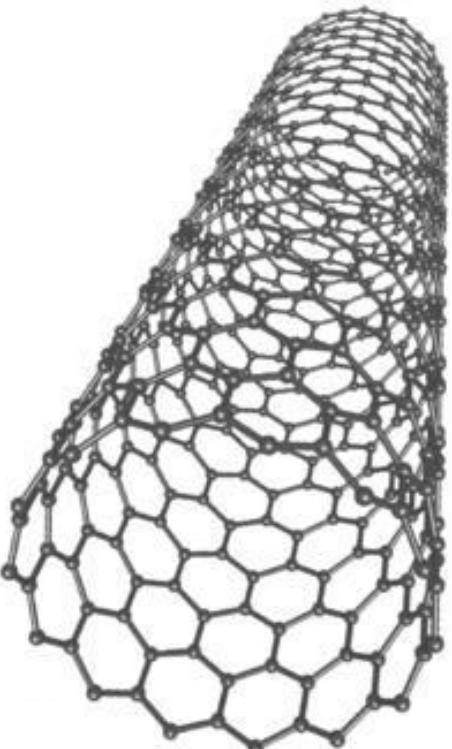


Buckyballs in crystalline form

- Buckyballs are also called fullerenes (after architect Richard Buckminster Fuller)
- Buckyballs were discovered in 1985 by Robert Curl, Harold Kroto and Richard Smalley
 - these scientists won the 1996 Nobel Prize in Chemistry for discovering this new allotrope of carbon.



Covalent Bonds In Carbon Nanotubes



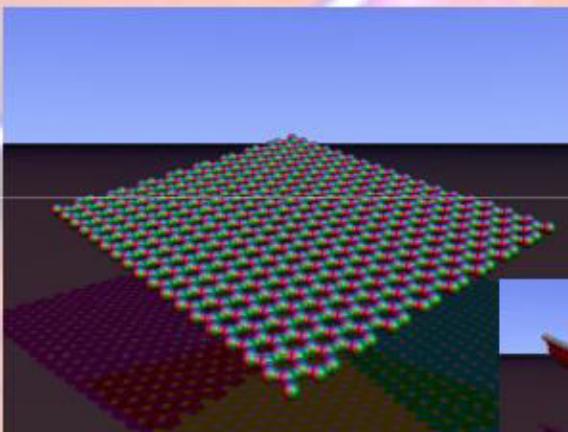
- Carbon nanotubes are formed by a layer of hexagonally-arranged carbon atoms rolled into a cylinder
 - usually have half buckyballs on one or both ends
- Electrons are localised internally, and some can move along the length of the tube
- Carbon nanotube diameter $\sim 1\text{nm}$
- Carbon nanotube length can be a million times greater than its width
- Nanotubes can be
 - single-walled ($d = 1\text{-}2\text{ nm}$), or
 - multi-walled ($d = 5\text{-}80\text{ nm}$).



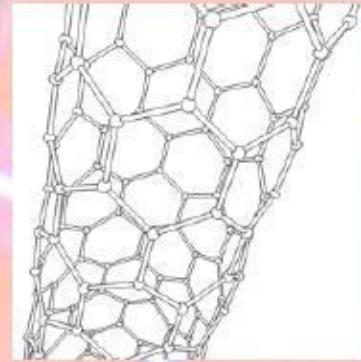
Properties of Carbon Allotropes

Carbon nanotubes

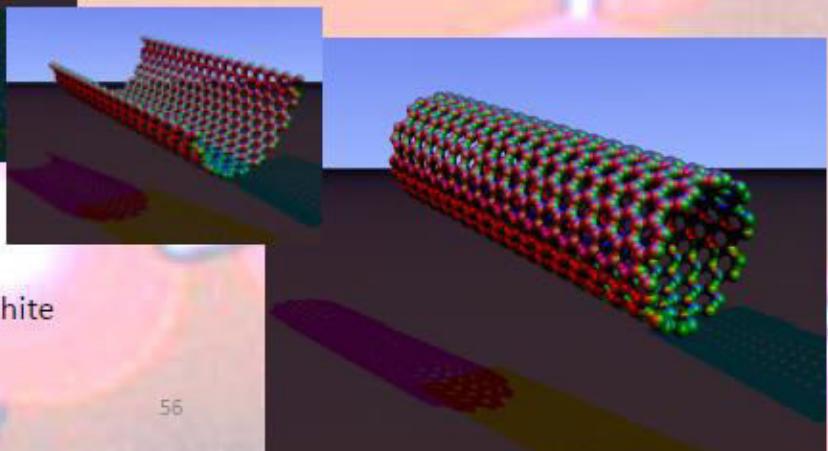
Carbon nanotubes (CNTs) are allotropes of **carbon** with a cylindrical nanostructure.



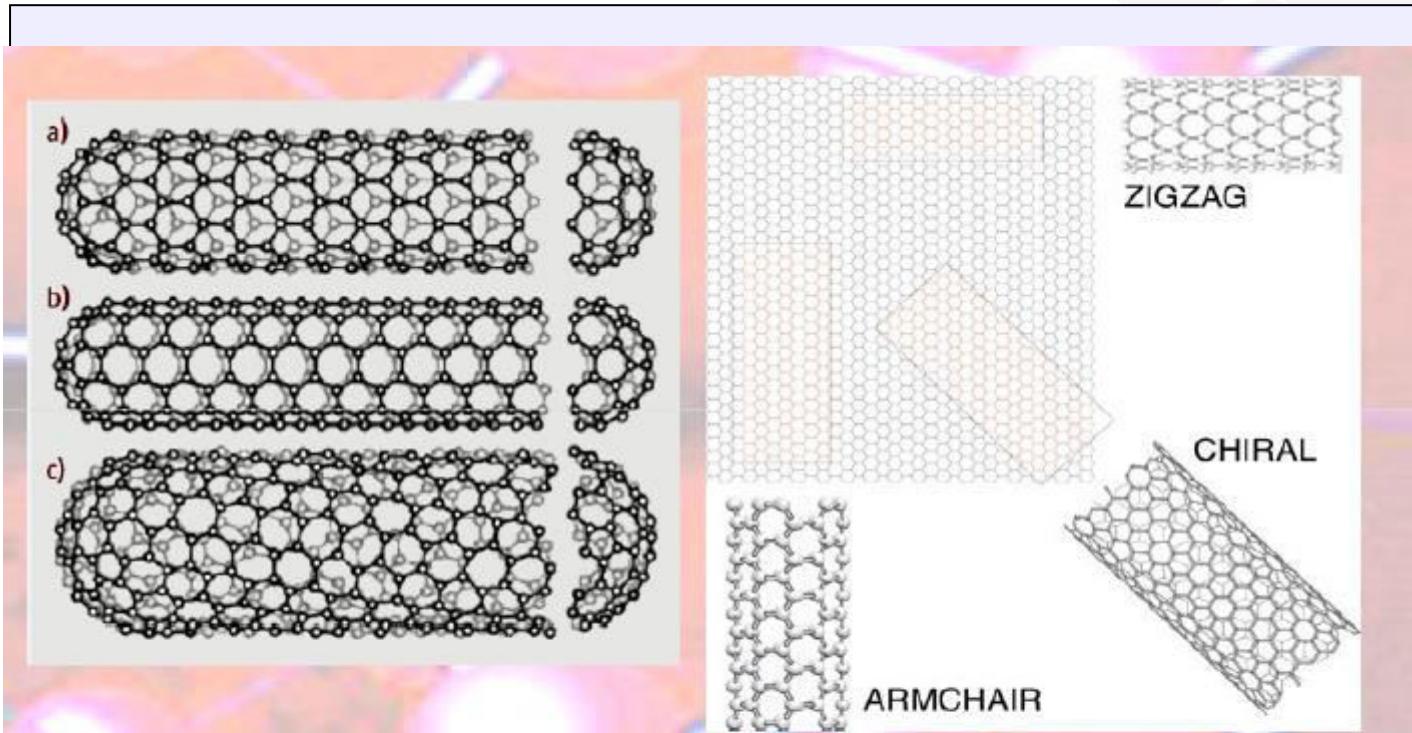
Structure is a single sheet graphite rolled into a tube.



Most **single-walled nanotubes (SWNT)** have a diameter of close to 1 nm, with a tube length that can be many millions of times longer.

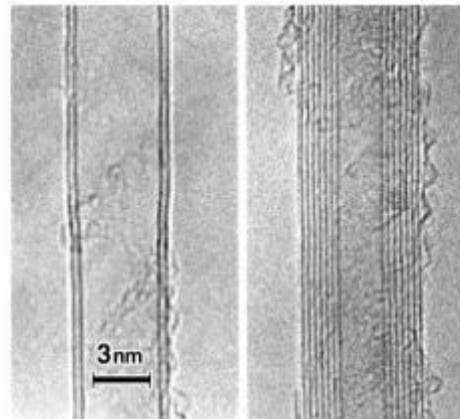
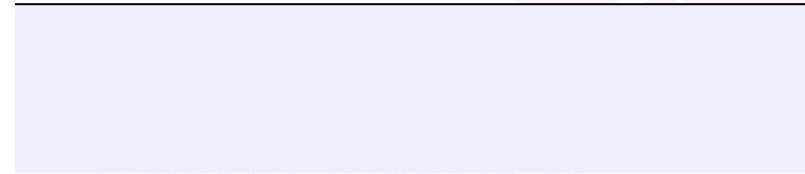
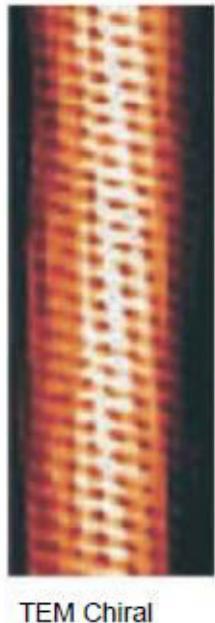
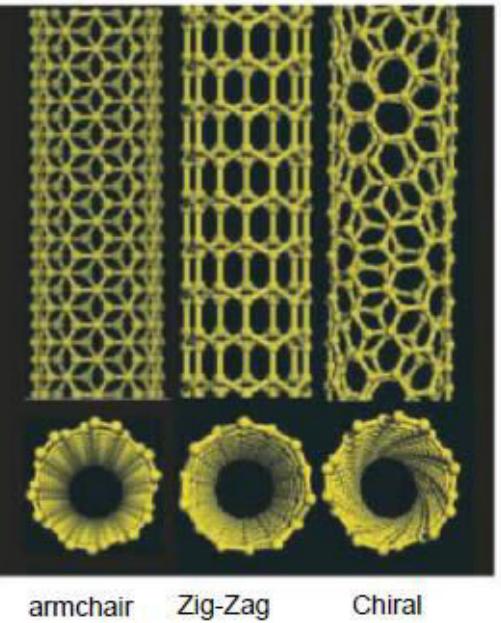


Types of Carbon Allotropes

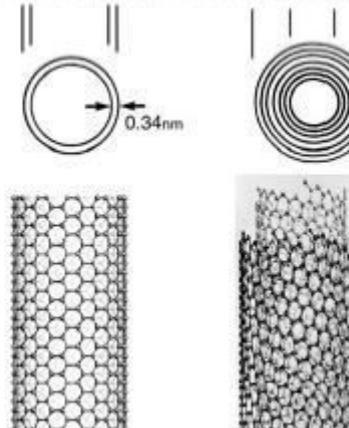


Sketch of three different SWNT structures as examples for
(a) a zig-zag-type nanotube, (b) an armchair type nanotube,
(c) a helical nanotube

Types of Carbon Allotropes

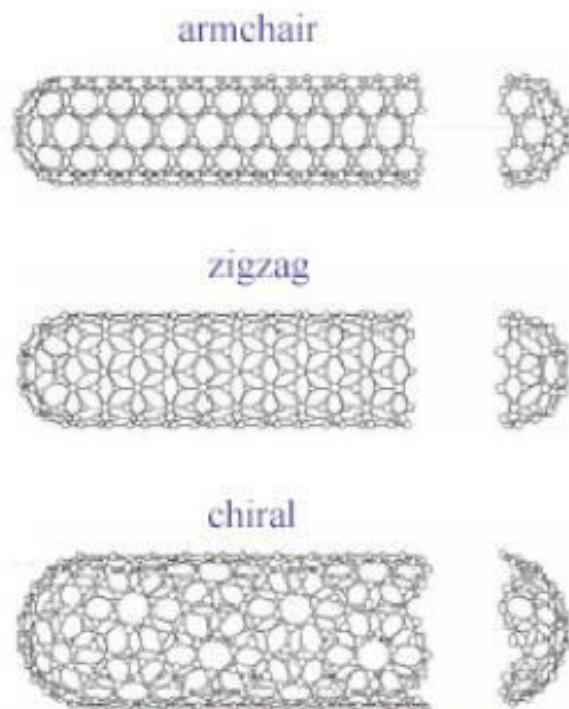
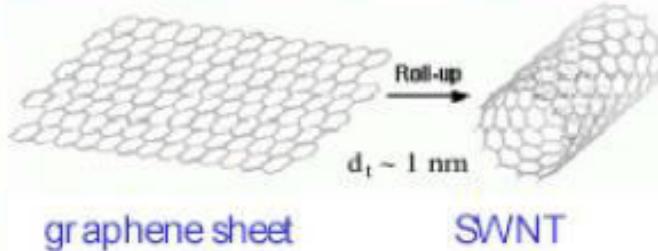


SWNT
(single
Wall nanotube)
Diameter ~ 1.4 nm



MWNT
(multiwall)
Diameter 10-20 nm

Properties of Carbon Allotropes



- **Size:** Nanostructures with dimensions of ~1 nm diameter (~10 atoms around the cylinder)
- **Electronic Properties:** Can be either metallic or semiconducting depending on diameter and orientation of the hexagons
- **Mechanical:** Very high strength, modulus, and resiliency. Good properties on both compression and extension.
- **Physics:** 1D density of electronic states
- Single molecule Raman spectroscopy and luminescence.
- Single molecule transport properties.
- Heat pipe, electromagnetic waveguide.

Properties of Carbon Allotropes

Allotrope	Hardness	Tensile strength	Conducts heat	Conducts electricity
Coal	+	+	+	no
Graphite	++	++	++++	++++
Diamond	++++	Not known	+++	no
Buckyballs	++++	+++	+	+
Carbon Nanotubes	+++++	++++	++++	++++

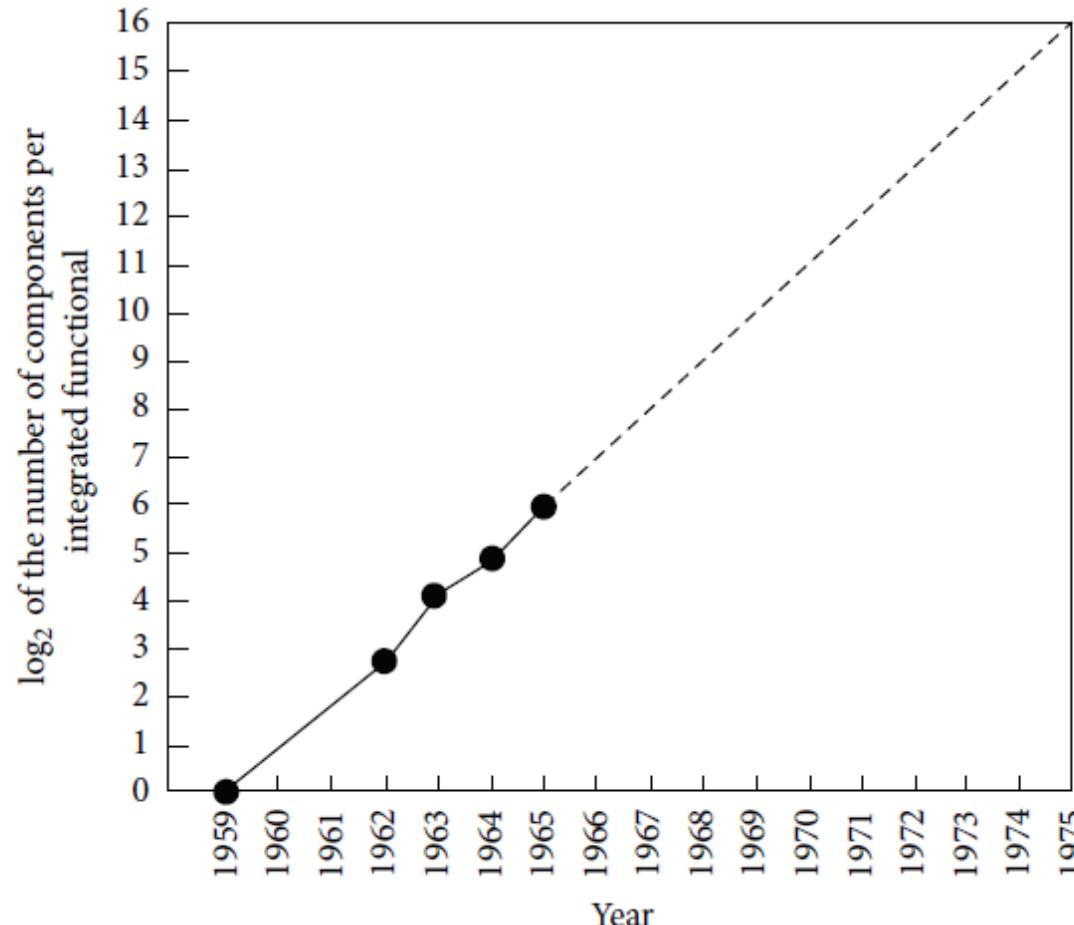


Moore's Law History

- Alan Turing in his **1950** paper "**Computing machinery and intelligence**" had predicted that by the turn of the millennium.
- we would have "**computers with a storage capacity of about 10^9** " what today we would call *128 megabytes* processing speed, memory capacity, sensors, and even the number and size of pixels in digital cameras, for example, smartphone.
- After him, **Gordon Earle Moore** (born January 3, 1929; UC Berkeley BS Chemistry, 1950; Caltech PhD. major in Chemistry and minor Physics, 1954) is the *cofounder and Chairman Emeritus of Intel Corporation*.
- In 1965, Moore, a founder of *Fairchild Semiconductor* (later Intel), observed in his famous paper that
 - "**The number of transistors incorporated in a chip will approximately double every 24 months.**"
 - "**the complexity for minimum component costs has increased at a rate of roughly a factor of two per year**"

Moore's Law History

- Extrapolating this trend for a decade, Moore predicted that chips with 65,000 components would be available by 1975.



Science and Technology of Nano Materials

Introduction

Matter and Energy are manifestations of the universe they exist in a variety of forms and interact with each other in many ways.

Nano means 10^{-9} . (Nanometer is one thousand Millionth of a Meter)

To understand how small one nm is let us see few comparisons

1. A **Red blood cell** is approximately **7000nm** wide.
2. Water Molecule is almost **0.3nm** across.
3. Human hair which is about **80,000nm** wide.

Nano Material

03_Nano_-_Intro_Reference_Material_II_01-Feb-2021

1. Nano Materials could be defined as the materials with at least one of its dimensions in the range of a Nano meter.
2. Thus the material need not be so small that it cannot be seen, it can be a large surface or a long wire whose thickness is in the scale of Nanometers.
3. Materials that are Nano scale in one dimension are layers, such as a **Thin films** or **Surface coatings**.
4. Materials that are Nano Scale in two dimensions include **Nano wires** and **Nano tubes**.
5. Materials that are Nano scale in three dimensions are particles for example **precipitates**, **colloids** and **quantum dots** (Small particles of Semiconductor Materials)

Nano Science

Nano Science can be defined as the study of phenomena and manipulation of materials at Atomic, Molecular and Macromolecular scales where properties differ significantly from those at a larger scale.

Nano Science is the study and understanding of properties of Nano Particles.

Nano Technology

Nanotechnology can be defined as the design, characterization, production and application of structures devices and systems by controlling shape and size at a Nano meter Scale.

Why properties of Nano Materials are different ?

The properties of Nano Materials are very much different from those at a larger scale.

Two principal factors cause the properties of Nano Materials to differ significantly from other materials.

- 1.Increased relative surface area.
- 2.Quantum confinement effect.

These factors can change or enhance properties such as reactivity , strength and electrical characteristics.

Increase in a Surface Area to Volume ratio

Nano Materials have a relatively larger Surface area when compared to the same volume or mass of the material produced in a larger form.

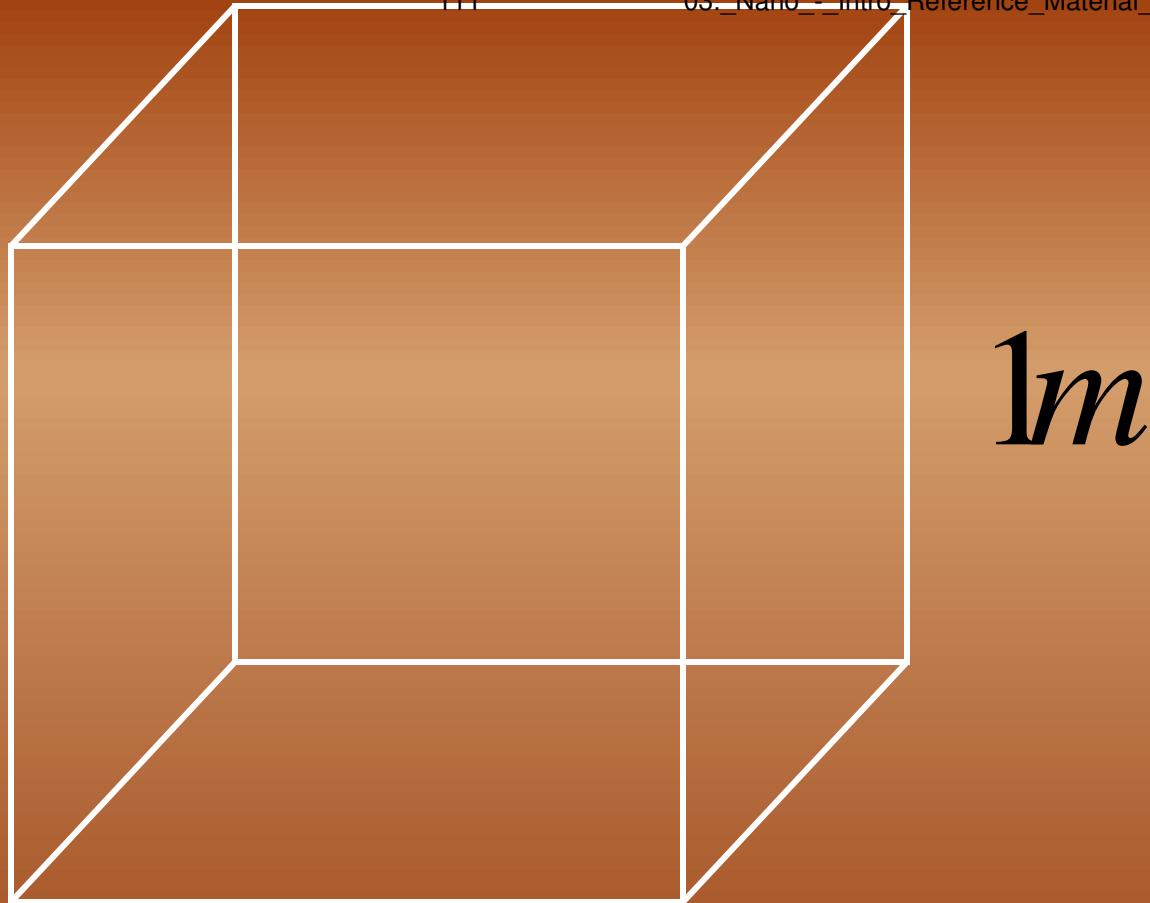
Let us consider a Sphere of radius “r”.

Its Surface Area = $4\pi r^2$.

Its volume= $\frac{4}{3}\pi r^3$

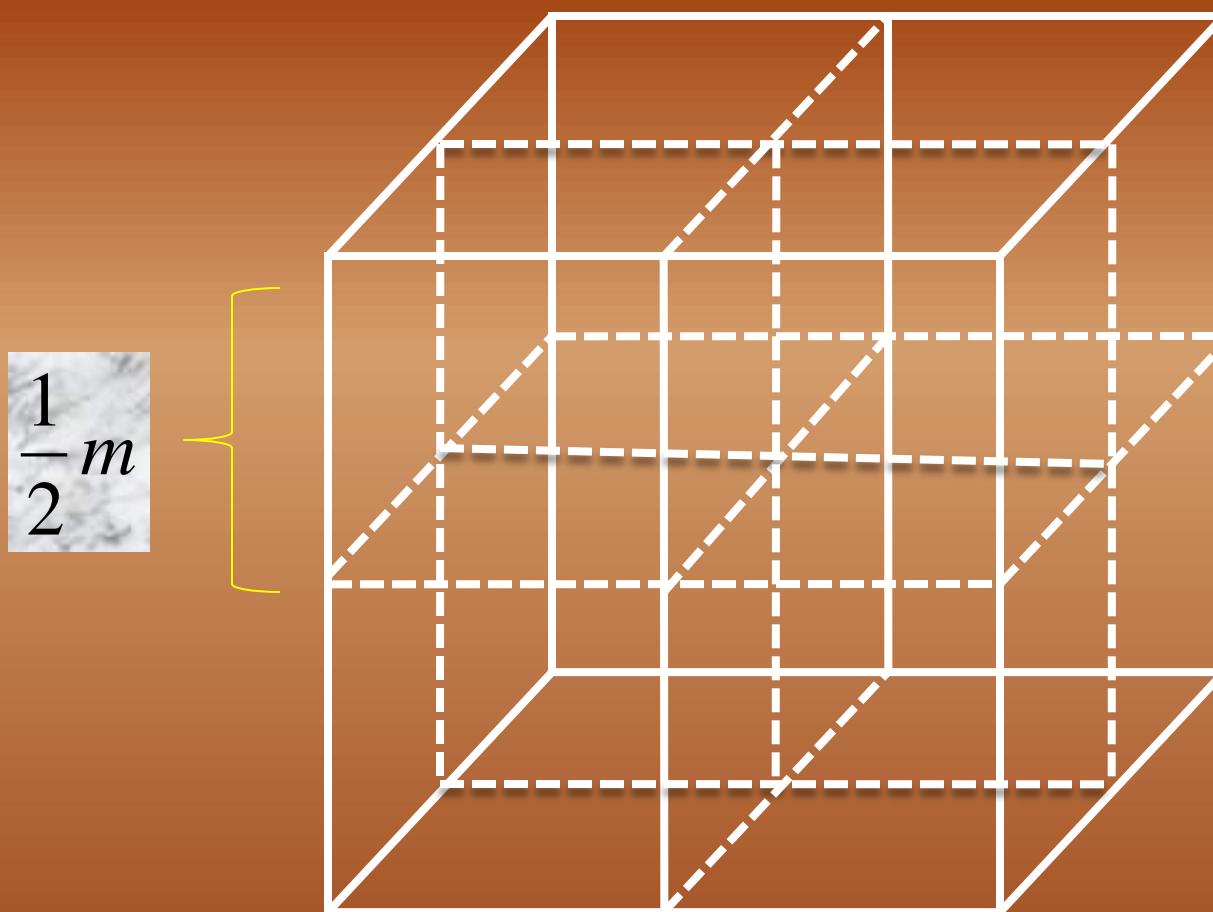
Surface Area to Volume Ratio= $3/r$.

Thus when the radius of the Sphere decreases , its Surface to Volume ratio increases.



$$\text{surface area} = 6 \times 1m^2 \Rightarrow 6m^2$$

- Let us consider one Cubic Volume shown in figure its the Surface Area is 6m^2 .
- When it is divided into eight pieces its Surface Area becomes 12m^2 , similarly When the same volume is divided into 27 pieces its Surface Area becomes 18m^2 .
- Thus we find that when the given volume is divided into smaller pieces the Surface Area increases.
- Hence as particle size decreases a greater proportion of atoms are found at the surface compared to those inside.
- Nano particles have a much greater surface area per given volume compared with larger particles. It makes materials more Chemically reactive.



$$\text{surface area} = 6 \times \left(\frac{1}{2} m\right)^2 \times 8 = 12m^2$$

Quantum Confinement

In Nano Crystals, the Electronic energy levels are not continuous as in the bulk but are discrete (finite density of states), because of the confinement of the electronic Wave function to the physical dimensions of the particles. This phenomenon is called Quantum confinement and therefore Nano Crystals are also referred to as quantum dots (QDs).

Properties of Nano Materials

Nano Materials have properties that are different from those of bulk materials.

Most Nano structure materials are Crystalline in nature and they posses unique properties.

Physical Properties of Nano Particles

- Crystal structure of Nano particles is same as bulk structure with different lattice parameters.
- The inter atomic spacing decreases with size and this is due to long range electrostatic forces and the short range core-core repulsion.
- The Melting point of Nanoparticles decreases with size.

Chemical Properties

- The Electronic structure of Nanoparticles is dependent on its size and the ability of Nano cluster to react, depends on cluster size.
- The large Surface area to volume ratio the variations in geometry and the electronic structure of Nano particles have a strong effect on catalytic properties.

Electrical properties

- The electronic structure of Nano materials is different from its bulk material.
- The density of the energy states in the conduction band changes.
- When the energy spacing between two energy levels is more than $K_B T$, energy gap is created.
- Nano clusters of different sizes will have different electronic structures and different energy level separations.
- The Ionization potential at Nano sizes are higher than that for the bulk materials

Magnetic Properties

- ⦿ The Magnetic Moment of Nano particles is found to be very less when compared them with its bulk size.

- ⦿ Nanoparticles made of semiconducting materials Germanium , Silicon and Cadmium are not Semiconductors.

Applications of Nanomaterials

Chemical Industry:

- Fillers for point systems
- Coating Systems based on Nano composites.
- Magnetic fluids.

Automotive Industry:

- Light weight construction
- Painting
- Catalysts
- Sensors

Medicine

- Drug delivery systems
- Active agents
- Medical rapid tests
- Antimicrobial agents and coatings.
- Agents in cancer therapy.

Electronic Industry:

- Data memory
- Displays
- Laser diodes
- Glass fibers
- Filters
- Conductive, antistatic coatings.

Energy Sources

- Fuel cells
- Solar cells
- Batteries
- capacitors.

Cosmetics

- Sun protection creams
- Tooth paste

Preparation of Nanomaterials

Physical Methods

Ball Milling

Sputtering/
Evaporation

Chemical Methods

Electro-
Deposition

Chemical
reactions

Chemical vapor
Deposition

Consider a homogeneous solid material of compact shape (let us say spherical) and macroscopic dimensions (let us say millimetric). Most of its properties will be related to its chemical composition and crystal structure. This is what is traditionally studied in the physics and chemistry of solids. When the size of the object is reduced to the nanometric range, the properties of materials may become totally different from what is observed in the bulk solid system.

Nanoscience is a new discipline concerned with the unique properties associated with **nanomaterials**, which are assemblies of atoms or molecules on a **nano scale**. ‘Nano’ refers a scale of size in the metric system. It is used in scientific units to denote one-billionth of the base unit - approximately 100,000 times smaller than the diameter of a human hair. ‘Scale’ refers an order of magnitude - of size or length- reference to objects that are sized on a scale that is relevant to nanometer. A nanometer is 10^{-9} meter ($1\text{nm} = 10^{-9}$ meter), a dimension in the world of atoms and molecules (the size of H atom is 0.24 nm and for instance, 10 hydrogen atoms lined up measure about 1 nm).

Nanoscience is actually the study of objects/particles and phenomena at very small scale, roughly 1 to 100 nanometers. **Nanoparticles** are those particles which contain from hundred to ten thousand of atoms (100 to 10,000 atoms). So the size of the particle range roughly from 1 to 100 nm. They are the building block of **Nanomaterials**.

What are nanomaterials? The materials which are created from blocks of nanoparticles or they are defined as a set of substances where at least one dimension is less than approximately 100 nanometers. Nanomaterials are of interest because at this scale unique optical, magnetic, electrical, and other properties emerge. These emergent properties have the potential for great impacts in electronics, medicine, and other fields. Nano carbons such as fullerenes and carbon nanotubes are excellent examples of nanomaterials.

Classification of Nanomaterials: Nanomaterials can be nanoscale in one dimension (e.g., surface films), two dimensions (e.g., strands or fibers), or three dimensions (e.g., precipitates, colloids). They can exist in single, fused, aggregated or agglomerated forms with spherical, tubular, and irregular shapes.

Nanostructures (dimensions of the nanomaterials) are the ordered system of one dimension, two dimensions or three dimensions of nanomaterials, assembled with nanometer scale in certain pattern which includes nanosphere, nanotubes, nanorod, nanowire, and nanobelt. Nano-structured materials are classified as zero dimensional, one dimensional, two dimensional, three dimensional nanostructures.

(a) 0D spheres and clusters	(b) 1D films, plates, and networks	(c) 2D nanofibers, wires, and rods,	(d) 3D nanomaterials or nanocrystallite

The **Clusters** are particles containing a very small number of atoms or simply few hundred atoms or smaller. A **Nanocrystallite** is generally understood to possess crystalline order in addition to nanoscale size. **Note:** If one dimension of the three dimensional nanostructure is at nano scale, then it is called a **Quantum Well**. If two dimension of the three dimensional nanostructure is at nano scale, then it is called a **Quantum Wire**. If all the three dimension of the nanostructure is at nano scale, then it is called a **Quantum Dot**. Nanocrystallites are also called quantum dot.

There is considerable variety in the types of nanoparticle systems that have been fabricated and studied. Aside from differences in their size and shape one important variable is their composition. Almost every element in the periodic table, together with various alloys and compounds, can form nanoparticles. They can be metallic, semiconducting, or insulating and typically their properties are very different to those of the corresponding bulk material.

The **seven main nanomaterial categories** are (1) Carbon based nanomaterials (2) Nano-composites (3) Nano-metals & Nano-alloys (4) Biological nanomaterials (5) Nano-polymers (6) Nano-glasses (7) Nano-ceramics.

Nanotechnology is defined as the creation of functional materials, devices, and systems through control of matter on the nanometer length scale and the utilization of new properties and phenomena developed at that scale. It is a field of science whose goal is to control individual atoms and molecules to create devices that are thousands of times smaller than present technologies permit. An important idea that underpins much of nanotechnology is that by controlling composition, size, and structure at the nanoscale one can engineer almost any desired properties.

For an object of the macroscopic size, the **surface atoms** comprise a negligible proportion of the total number of atoms and will therefore play a negligible role in the bulk properties of the material. When the size of the object is reduced to the nanometric range the proportion of surface atoms no longer negligible. So a large fraction of the atoms are located at the surface of the object in the nanomaterials, it will modify its properties. The properties of nanomaterials become totally different from what is observed in the bulk solid system.

The properties of nanomaterials are very much different from Reference Materials large 01s Fall 2021
Two principal factors cause the properties of Nano Materials to differ significantly from other materials (1) Increased relative surface area and (2) Quantum confinement effect. These factors can change or enhance properties such as reactivity, strength and electrical characteristics.

Increase in a Surface Area to Volume ratio: Nanomaterials have a relatively larger surface area when compared to the same volume or mass of the material produced in a larger form. When the given volume is divided into smaller pieces the surface area increases. So the particle size decreases a greater proportion of atoms are found at the surface compared to those inside. Hence Nanoparticles have a much greater surface area per given volume compared with larger particles. It makes materials more chemically reactive.

Quantum Confinement: In nanocrystals, the electronic energy levels are not continuous as in the bulk but are discrete (finite density of states), because of the confinement of the electronic Wave function to the physical dimensions of the particles. This phenomenon is called Quantum confinement. If one length of three dimensional nanostructures is at nano-dimension, then it is called a **Quantum Well**. If two sides of three dimensional nanostructures are at nano-dimension, then it is called a **Quantum Wire**. If all three dimensional nanostructures is at nano-dimension (**Nano Crystals**), are referred as **Quantum Dots (QDs)**.

Properties of Nanomaterials: Nanomaterials have properties that are different from those of bulk materials. Most nanostructure materials are crystalline in nature and they have unique properties. Filling polymers with nanoparticles or nanorods and nanotubes, respectively, leads to significant improvements in their mechanical properties.

Physical Properties: Crystal structure of nanoparticles is same as bulk structure with different lattice parameters. The inter-atomic spacing decreases with size and this is due to long range electrostatic forces and the short range core-core repulsion. The melting point of nanoparticles decreases with size.

Chemical Properties: A large fraction of the atoms are located at the surface of the nanomaterial which increase its reactivity and catalytic activity. The large surface area to volume ratio, the variations in geometry and the electronic structure of nano particles have a strong effect on catalytic properties.

Electrical properties: The energy band structure and charge carrier density in the materials can be modified quite differently from their bulk and in turn will modify the electronic properties of the materials. Nanoparticles made of semiconducting materials like Germanium, Silicon and Cadmium are not semiconductor. Nanoclusters of different sizes will have different electronic structures and different energy level separations. So they show diverse electronic properties which depend on its size.

Magnetic Properties: The magnetic moment of ~~nano~~ particles is Reference Material les 01 Feb 2021 compared them with its bulk size. Actually, it should be possible that non-ferromagnetic bulk exhibit ferromagnetic-like behavior when prepared in nano range. Bulk Gold and Pt are non-magnetic, but at the nano size they are magnetic.

Applications of Nanomaterials: (1) **Chemical Industry:** Fillers for point systems. Coating systems based on nano composites, Magnetic fluids (2) **Automotive Industry:** Light weight construction, Painting, Catalysts, Sensors (3) **Medicine:** Drug delivery systems, Active agents, Medical rapid tests, Antimicrobial agents and coatings, Agents in cancer therapy (4) **Electronic Industry:** Data memory, Displays, Laser diodes, Glass fibers, Filters, Conductive, antistatic coatings (5) **Energy Sources:** Fuel cells, Solar cells, Batteries, Ultracapacitors (6) **Cosmetics:** Sun protection creams, Tooth paste

Applications of Nanomaterials in Medicine: Medical application of nanomaterials include
(a) Fluorescent biological labels (b) Drug and gene delivery (c) Bio-detection of pathogens
(d) Detection of proteins (e) Probing of DNA structure (f) Tissue engineering (g) Tumour destruction
(h) Separation and purification of biological molecules and cells.

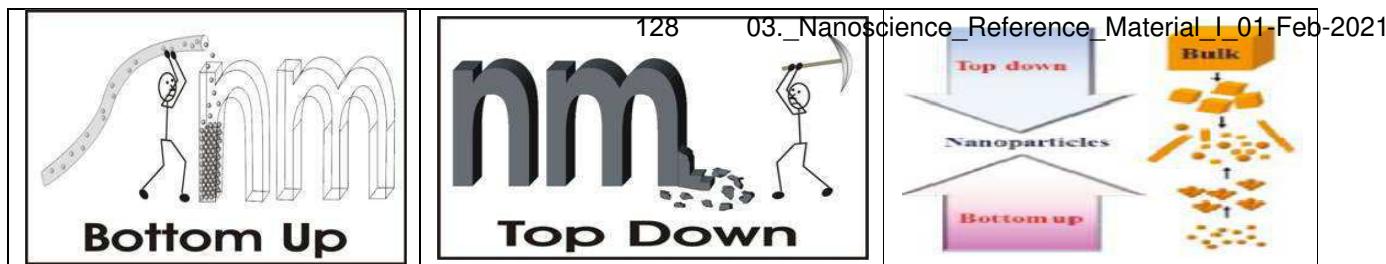
Future application of Nanomedicine is as follows: (1) The elimination of bacterial infections in a patient within minutes (2) The ability to perform surgery at the cellular level, removing individual diseased cells and even repairing defective portions of individual cells (3) **Qdots:** that identify the location of cancer cells in the body, here nanomaterials that deliver chemotherapy drugs directly to cancer cells to minimize damage to healthy cells (4) **Nanocells:** that concentrate the heat from infrared light to destroy the cancer cells with minimal damage to surrounding health cells.

Applications of Nanomaterials in Catalysis: The extremely small size of the nanomaterials maximizes the surface area exposed to the reactants, allowing more reactions to occur. The application ranges from fuel cell to catalytic converters (nano-Pt), photocatalytic devices (e.g., nano-TiO₂) and for the production of chemicals (e.g.. nano-ZrO₂ in CO hydrogenation in isobutene synthesis)

There are two ways of approaching the properties of nanoscale objects: the **Bottom-up** approach and the **Top-down** approach.

In **Bottom-up** approach (or **Atom-Atom Assembly**), one assembles atoms and molecules into objects whose properties vary discretely with the number of constituent entities, and then increases the size of the object until this discretisation gives way in the limit to continuous variation. The relevant parameter becomes the size rather than the exact number of atoms contained in the object.

In **Top-down** approach (Chisel away atoms), one considers the evolution of the properties of a sample as its size is whittled down from macroscopic toward nanometric lengths.



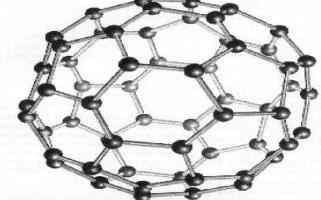
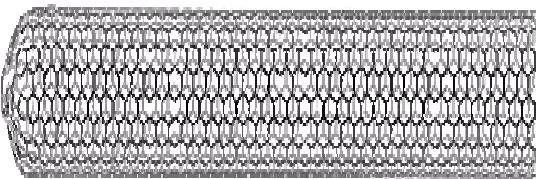
Carbon based Nanomaterials (or Nano carbons): They are defined as materials in which the nanocomponent is pure carbon. Fullerenes and carbon nanotubes are excellent examples of it.

When graphite was vapourized using a laser beam, H. Kroto, R. F. Curl and R. E. Smalley (in 1985) discovered a molecule that contained 60 carbon atoms as the major product called **Fullerenes**. They realized that the molecule was a closed cage based on rings of 5 and 6 carbon atoms, looking just like a soccer ball. **Carbon nanotubes (CNTs)** are molecular tubes having diameters of a few nanometers (as low as 1 nm) made up of lattices of carbon atoms. CNTs were discovered in 1991 by the Japanese electron microscopist Sumio Iijima who was studying the material deposited on the cathode during the arc-evaporation synthesis of fullerenes.

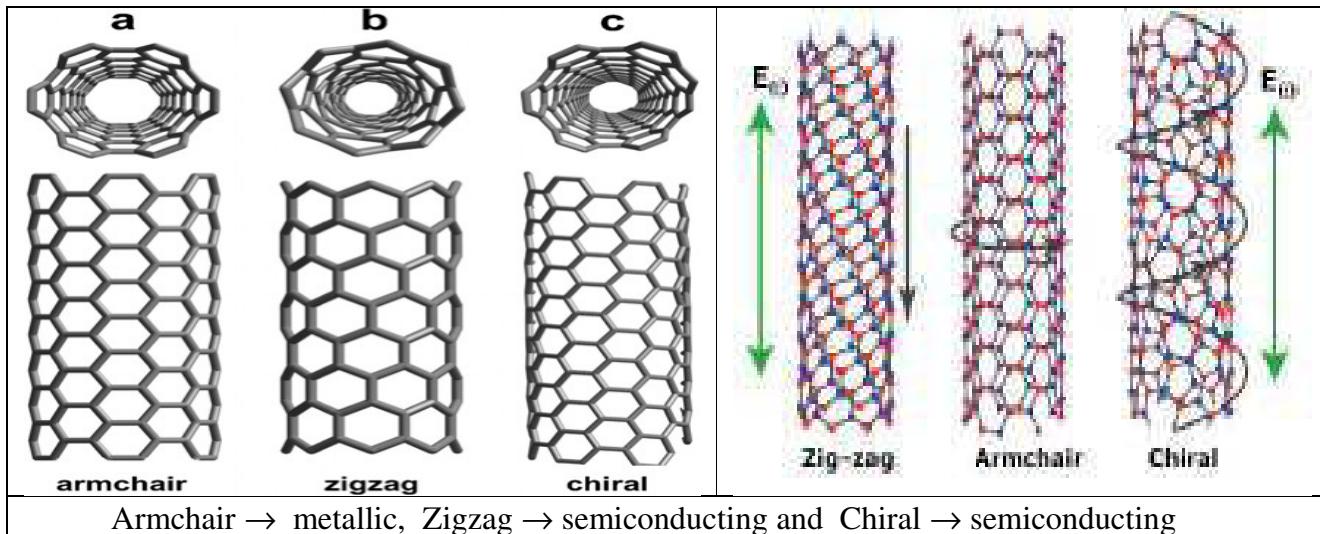
Fullerenes are made by the heating of graphite in an electric arc in the presence of inert gases such as helium or argon. The sooty material formed by condensation of vapourized C_n small molecules consists of mainly C_{60} with smaller quantity of C_{70} and traces of fullerenes consisting of even number of carbon atoms up to 350 or above. Fullerenes are the only pure form of carbon because they have smooth structure without having ‘dangling’ bonds. Fullerenes are cage like molecules. C_{60} molecule has a shape like soccer ball and called **Buckminsterfullerene**. Spherical fullerenes are also called **Bucky balls** in short.

Fullerenes are a class of cage-like carbon compounds that contained 60 carbon atoms, composed of fused pentagonal and hexagonal carbon rings, looking just like a soccer ball. It contains twenty six- membered rings and twelve five membered rings. A six membered ring is fused with six or five membered rings but a five membered ring can only fuse with six membered rings. C_{60} can be used as excellent microscopic ball bearings, lubricant and catalyst. They can cage other molecules, so in the future this may be used to deliver drugs in small amounts for slow release, e.g., cancer treatment.

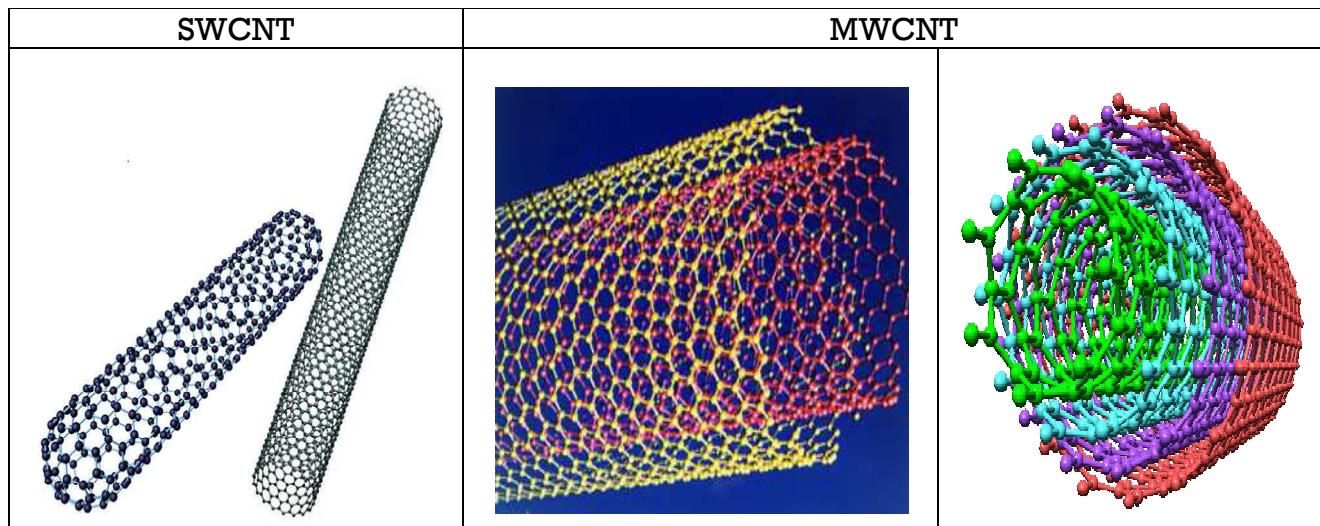
Carbon Nanotubes (CNT) are sheets of graphite rolled to make a tube. They are molecular tubes (cylindrical nanostructure) having diameters of a few nanometers made up of lattices of carbon atoms. CNTs are at least 100 times stronger than steel, but only one-sixth as dense. In addition, they conduct heat and electricity far better than copper. So, CNTs can be used in tiny, physically strong conducting devices. CNTs have been filled with potassium atoms, making them even better electrical conductors.

Fullerene (C ₆₀)	¹²⁹ Carbon Nanotube (CNT)
	

There are three kinds of nanotubes: (a) Armchair nanotubes containing a line of hexagons parallel to the axis of the nanotube, (b) Zigzag nanotubes where there is a line of carbon bonds down the centre and (c) Chiral nanotubes which exhibit a twist or spiral around the nanotube.



Carbon nanotubes exist as nanotubes within nanotubes, leading to the categorization as **Single-Walled Carbon Nanotubes (SWCNT)** and **Multi-Walled Carbon Nanotubes (MWCNT)**. An SWNT contains only a single cylinder whereas a MWCNT contains multiple concentric nanotube cylinders. They have diameter close to 1 nm. The structure SWNT/MWCNT can be conceptualized by wrapping a one-atom-thick layer/multi-atom-thick layer of graphite (graphene) into a seamless cylinder.

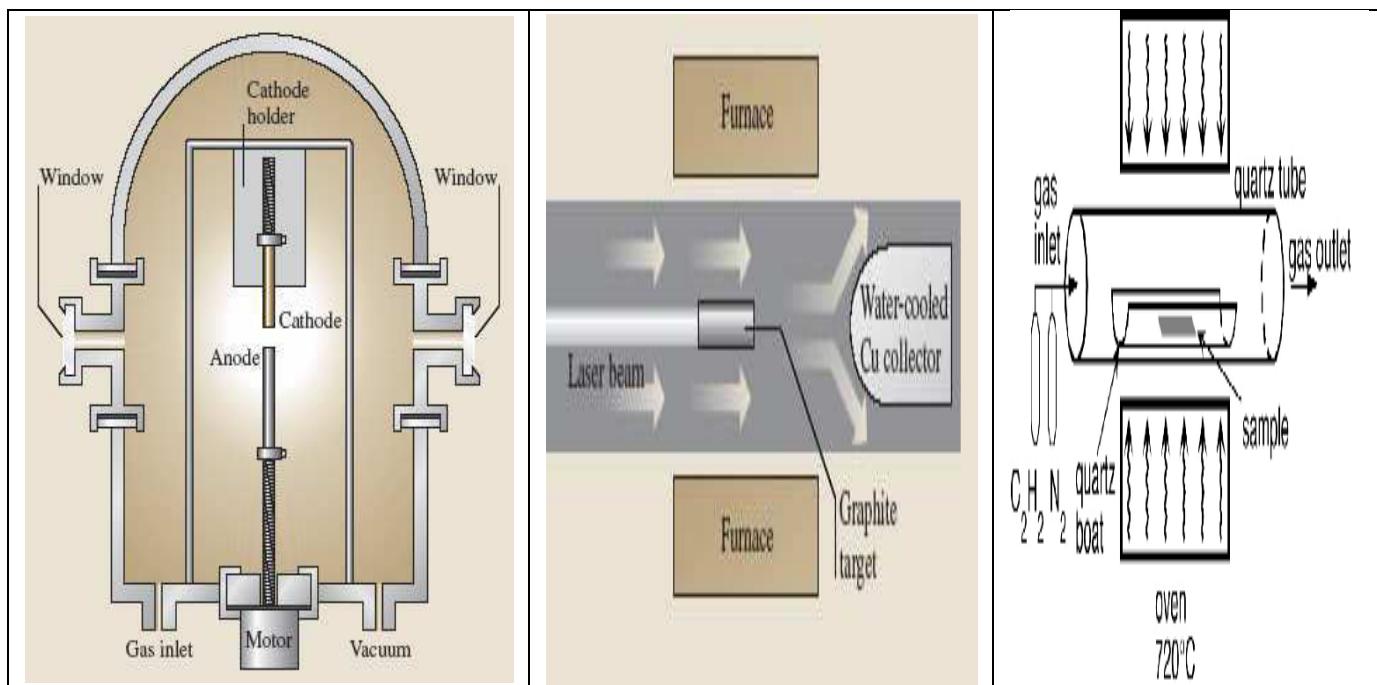


techniques; (1) **Arc discharge** (2) **Laser ablation** and (3) **Chemical vapour deposition**.

In **Arc Discharge Method**, a vapour of carbon atoms is created by an arc discharge between two graphite electrodes with or without catalyst. This vapour of carbon atoms self-assembles to form the carbon nanotubes. When pure graphite rods are used, the anode evaporates and it is deposited on the cathode, which contains CNTs. These CNTs are MWNTs. When a graphite rod containing metal catalyst (Fe, Co, etc.) is used as the anode with a pure graphite cathode, SWNTs are generated in the form of soot.

In **Laser Ablation Method**, a high power laser beam impinges on a graphite sheet, it vaporizes graphite and the vaporized carbon will then condense on a cooled surface, where carbon nanotubes will collect. High-quality SWNTs are produced from this method.

In **Chemical Vapor Deposition (CVD) Method**, it can be achieved by breaking the carbon-based gaseous molecules (methane, acetylene and carbon monoxide) into reactive atomic carbon in a high temperature furnace and sometime assisted by plasma to increase the generation of atomic carbon. **To start the process**, a substrate is layered with metal catalyst (Ni or Co). The intended diameter of the carbon nanotubes is related to the size of the metal catalyst deposited or etched on the substrate. The substrate is heated to 700°C, and carbon-based gases are then slowly delivered into the substrate's chamber. Two gases are used – one is called a “process gas”, which would be ammonia, N₂, H₂, while the other is a carbon based gas (methane, acetylene and carbon monoxide). The carbon nanotubes grow on the catalyst because the carbon-based gas breaks apart at the catalyst surface. It is the one of the most common methods of carbon nanotube synthesis.

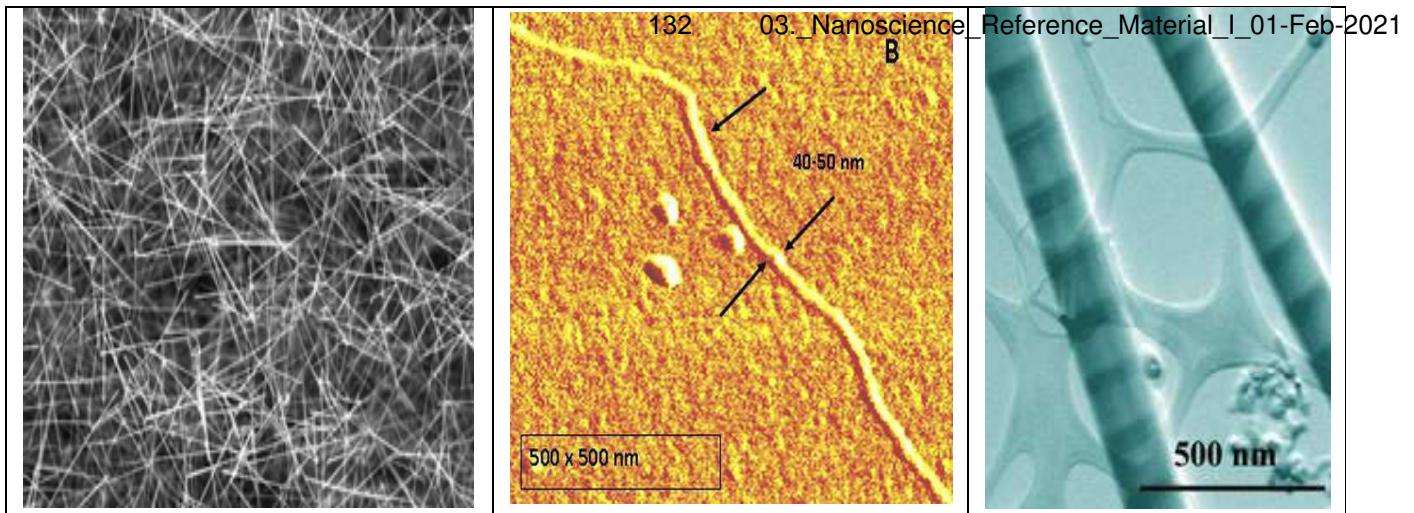


and stiffest materials. [2] **Hardness:** SWNTs are used for synthesizing super-hard material, by compressing it at room temperature. [3] **Kinetic:** Multiple concentric nanotubes precisely nested within one another, exhibit telescoping property. [4] **Electrical:** Their electrical conductivity is better than metals. Electron travelling through a CNT behaves like a wave travelling through smooth channel - **Ballistic Transport.** MWNTs with interconnected inner shells show superconductivity with relatively high temperatures. [5] **Thermal:** Their thermal conductivity also is better than metals. All nanotubes are expected to be very good thermal conductors along the tube but good insulators laterally to the tube axis - **Blastic Conduction.** The thermal stability of carbon nanotubes to be upto 2800°C in vacuum and about 750°C in air.

Application of Carbon Nanotubes: It includes [1] **Structural:** Its superior mechanical properties, uses ranging from everyday items like clothes and sports gear to combat jackets and space elevators. [2] In **Electrical Circuits:** Nanotube based transistors have been made that operates at room temperature. Carbon nanotubes are used for miniaturizing electronic devices. [3] As **Paper Batteries:** A paper battery is a battery engineered to use a paper thin sheet of cellulose infused with aligned carbon nanotubes. It gives a steady power output. This battery also functions as a super-capacitor which give a quick explode of high energy. [4] **Solar Cells:** It is formed by a mixture of carbon nanotubes and fullerenes. Electrons trapped inside fullerenes, add sunlight lead to the flow of this electrons which produce the current. [5] They are used for making **ultra-capacitors** which provide a large surface to store electrical charge. [6] They act as molecular size test tubes or capsules for drug delivery. [7] They are used as tips for analysis of DNA and proteins by Atomic Fourse Microscopey (AFM).

Nanowire: Nanowires are microscopic wires that have a width measured in nanometers. A nanowire is a connecting structure that has a diameter of 10^{-9} meters, which is extremely small. Typically, the diameter of nanowires range from 40 to 50 nanometers, but their length is not so limited. Since they can be lengthened by simply attaching more wires end to end or just by growing them longer, they can be as long as desired. Nanowires are also known as **Quantum Wires**, these connectors are used to connect tiny components together into very small circuits.

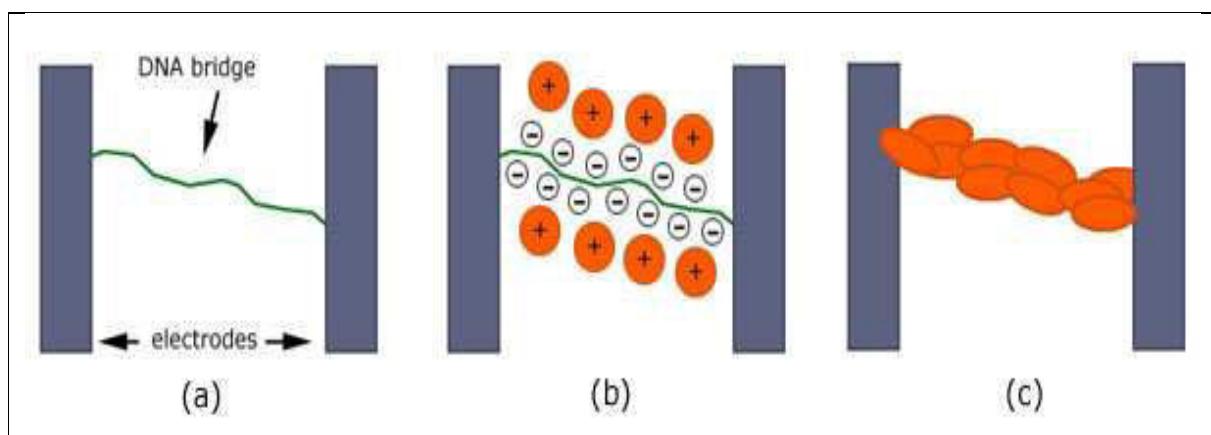
The electrical properties of nanowires are peculiar mainly because of their size and its conductivity is strongly influenced by edge effects. The edge effects arise from atoms that lie at the nanowire surface which are not fully bonded to neighbouring atoms as that in the bulk of the nanowire. These nonbonded atoms often become source of defects within the nanowire causing it to conduct electricity.



Four Different types of Nanowires: They are (a) **Metal nanowires** are made from nickel, platinum or gold (b) **Semi-conducting** wires are comprised of silicon, indium phosphide or gallium nitride, (c) **Insulating** wires are made of silicon dioxide or titanium dioxide (d) To create a **Molecular** nanowire, the process involves repeating organic or inorganic molecular units in a particular format.

Striped nanowires are capable of performing more than one task along the same wire. They are striped with different materials that posses different properties, an attribute which allows different operations to be performed at the same time. This also enables devices to be more compacted because fewer wires are needed; each nanowire is serving multiple functions.

Production of nanowires: A solution containing the desired metal is mixed with DNA and then exposed to UV light. When exposed, the metal in the mixture bonds to the DNA and forms a microscopic wire, a nanowire. Its width is dependent upon how concentrated the solution of the metal is. The more concentrated the metal solution, the wider the nanowire; likewise, the less concentrated, the thinner the wire will be.



Nanowires are not being heavily manufactured because they are still in the development stage and are only produced in the laboratory and have not moved to manufacturing plants. Until production has been streamlined, made easier and faster, they will not be heavily manufactured for commercial purposes.

greatly reduce the size of electronic devices while allowing us to increase the efficiency of those devices. The most apparent impact this will have on society would be the increase of storage space for MP3 players, computers, and phones without increasing size.

(1) Data storage/transfer - transfer data up to 1,000 times faster, and store data for as long as 100,000 years without degradation (2) Batteries/generators - tiny, efficient solar panels, turning light into energy, able to hold 10 times the charge of existing batteries (3) Transistors, LED's, Optoelectronic devices, Biochemical sensors and Heat-pumping thermoelectric devices.

Nanowires are also being developed for prototype sensors. These sensors will be used on gases and biological molecules. They will be used to detect harmful agents by scanning each gas or chemical on a molecular level. This is possible due to how small these wires can be. They will be made out of materials that react to harmful agents, thus alerting to the presence of harmful agents.

The conductivity and tiny size make them ideal for future computer processors and connectors. The use of nanowire in a transistor is an ideal method of producing smaller and faster microprocessor components for the computer and electronic industry. Although nanowire transistors function better than the current transistors, the high costs required to create them is a barrier to wider manufacturing.

Silicon Nanowires (SiNW): The most promising building blocks for future nanoscale electronic devices. Atomic configurations of silicon nanowires created through assembly of silicon clusters or created through etching of crystalline silicon. The small sizes of SiNWs make their electronic and electrical properties strongly dependent on growth direction, size, morphology and surface reconstruction. A well-known example is the size dependence of the electronic band gap width of SiNWs irrespective of wire direction. As the wire diameter decreases, the band gap of the nanowire widens and deviates from that of bulk silicon gradually

Applications: Silicon nanowire (SiNW) is a very attractive two-dimensional nanomaterial for future nanoelectronic applications.

- 1) Silicon nanowire solar cells which consists of arrays of radial p-n junction nanowires
- 2) Silicon nanowires can help to reduce the size of microchips
- 3) Scientists have fabricated a memory device that combines silicon nanowires with a more traditional type of data-storage. The device is a type of “non-volatile” memory, meaning stored information is not lost when the device is without power

Syllabus:

Introduction to Nanoscience – carbon nanotubes and nanowires – applications

PROPERTIES OF NANOMATERIALS

- **Nanomaterials have properties that are different from those of bulk materials.**

- **Most nanostructure materials are crystalline in nature and they have unique properties.**

PROPERTIES OF NANOMATERIALS

- ❑ Certain structures exist only in the nanoscale (i.e. there are no bulk counterparts to these structures).
 - Examples ◊ carbon nanotubes (CNTs), Fullerenes, carbon onions etc.
- ❑ Certain properties arise only in the nanoscale (i.e if the size of the relevant unit is made larger the property under consideration would not be observed).
 - Examples ◊ super-hydrophobicity, super- catalytic activity, superparamagnetism, giant magneto resistance etc.
- ❑ Certain combination of properties can only be obtained with nanomaterials and nanostructures
 - Example ◊ Abalone shell has a fracture toughness reaching more than 1000 times that of calcium carbonate the chief ingredient in the shell; while still retaining the hardness of nacre.
- ❑ Drastic change in properties may be observed on approaching the nanoscale
 - Example ◊ Fracture strength of Ni has shown to increase from 100 MPa to 900 MPa once the nanometer-sized grains are obtained.
- ❑ The performance of some systems depend on a functional entity in the nanoscale
 - Example ◊ A devise sensing a signal from a single DNA strand has to be in the nanoscale so as to extract the local signal from one strand. Nanoscale entity (such as CNTs) can be utilized for measuring the electrical signal in response to a stimulus.
- ❑ In multi-lengthscale structures (with special properties) the fundamental unit has to be nanoscale for the other lengthscales to be effective
 - Example ◊ non-buckling nano-hairs are at the heart of the hierarchical structure on a lotus leaf, which gives it superhydrophobicity.

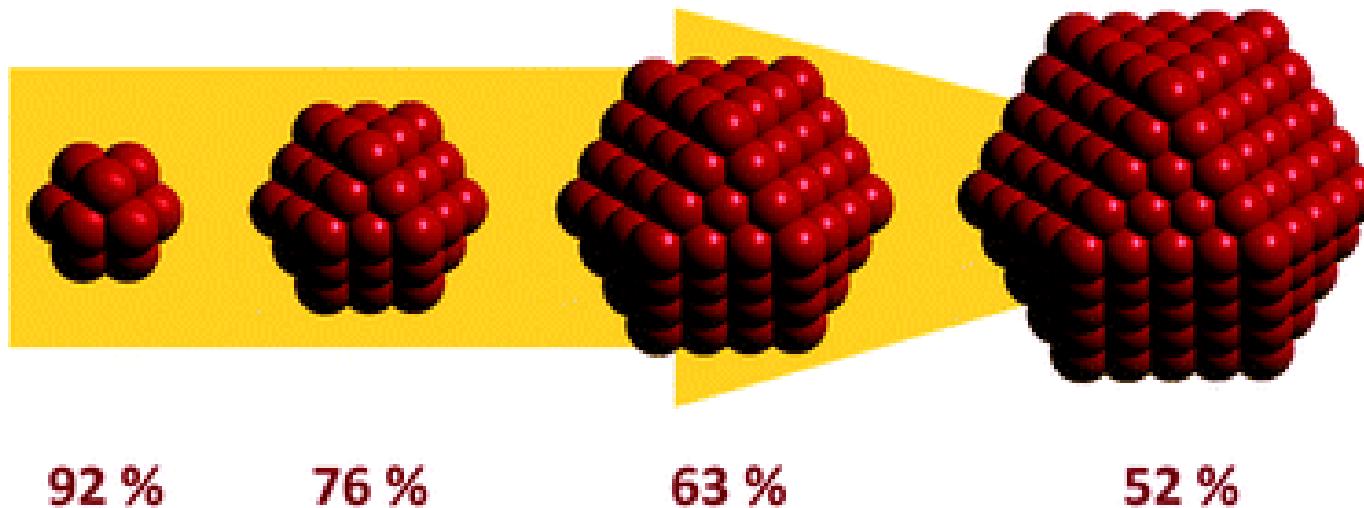
PHYSICAL PROPERTIES

- Crystal structure of nanoparticles is same as bulk structure with different lattice parameters.
- The inter-atomic spacing decreases with size and this is due to long range electrostatic forces and the short range core-core repulsion.
- The melting point of nanoparticles decreases with size.

CHEMICAL PROPERTIES

- A large fraction of the atoms are located at the surface of the nanomaterial which increase its reactivity and catalytic activity.
- The large surface area to volume ratio, the variations in geometry and the electronic structure of nano particles have a strong effect on catalytic properties.

Decrease of surface-to-volume ratio



ELECTRICAL PROPERTIES

- The **energy band structure** and **charge carrier density** in the materials can be modified quite differently from their bulk and in turn will **modify the electronic properties of the materials**.
- Nanoparticles made of **semiconducting materials like Germanium, Silicon and Cadmium** are **not semiconductor**.
- **Nanoclusters of different sizes will have different electronic structures and different energy level separations.** So they show diverse electronic properties which depend on its size.

- ❑ In bulk materials conduction electrons are delocalized and travel ‘freely’ till they are scattered by phonons, impurities, grain boundaries etc.
- ❑ In nanoscale conductors two effects become important:
 - Quantum effect: Continuous ('nearly') bands are replaced with discrete energy states
 - Classical effect: mean free path (MFP) for inelastic scattering becomes comparable to the size of the system (can lead to reduction in scattering events).
- ❑ In metals: change in DOS on reduction of size of the system plays a major role (along with change in electronic and vibrational energy levels).
- ❑ In semiconductors quantum confinement of both the electron and hole leads to an increase in the effective band gap of the material with decreasing crystallite size.
- ❑ These effects can lead to altered conductivity in nanomaterials.

MECHANICAL PROPERTIES

CLASSIFICATION OF DEFECTS BASED ON DIMENSIONALITY

0D (Point defects)

Vacancy

Impurity

Frenkel defect

Schottky defect

1D (Line defects)

Dislocation

Disclination

Dispiration

2D (Surface / Interface)

Surface

Interphase boundary

Grain boundary

Twin boundary

Stacking faults

Anti-phase boundaries

3D (Volume defects)

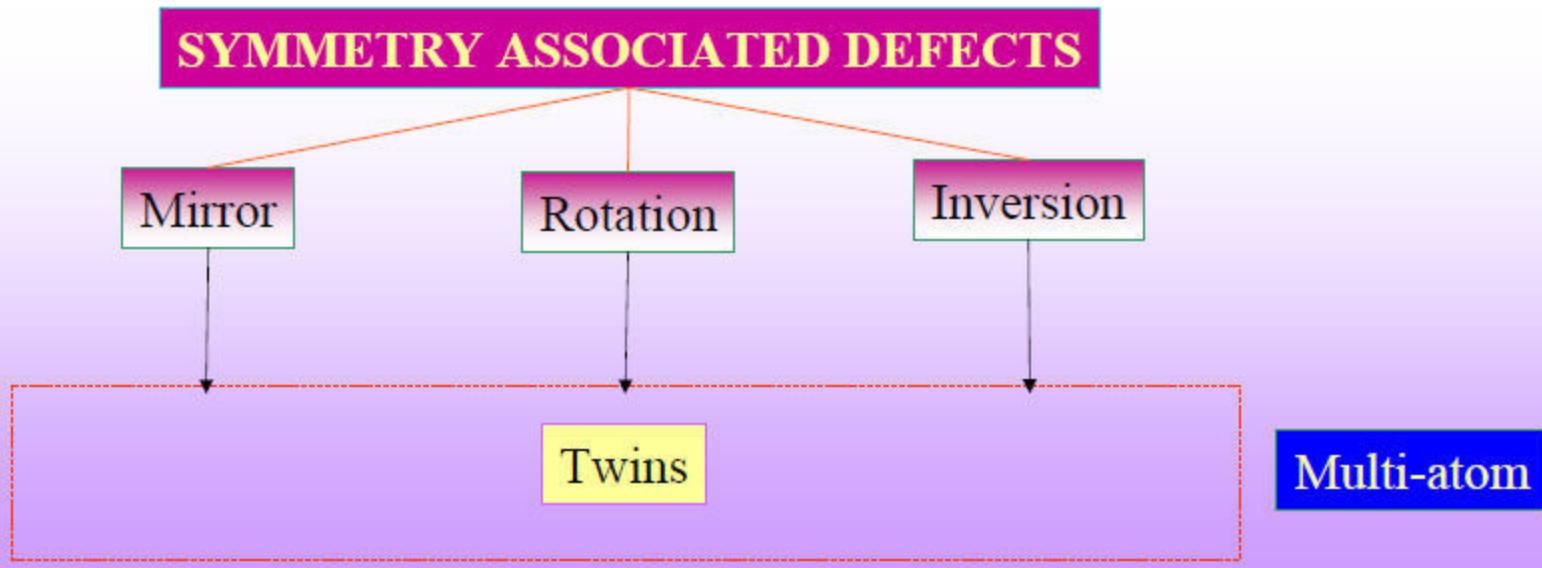
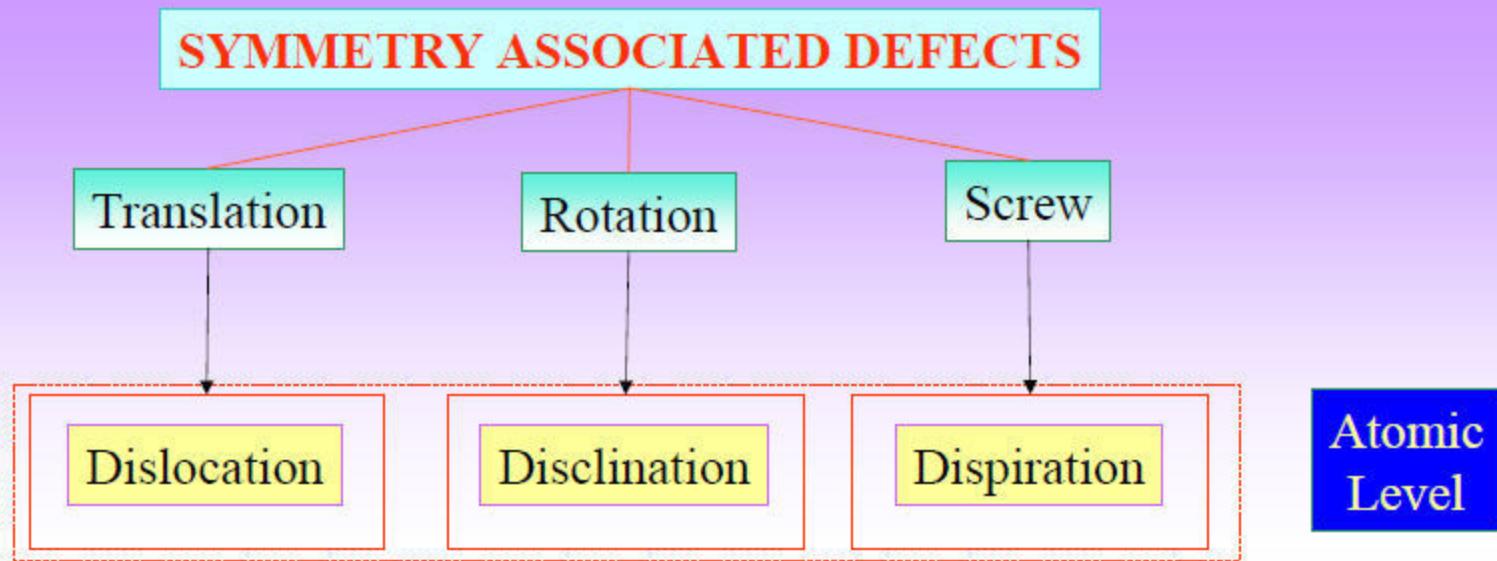
Twins

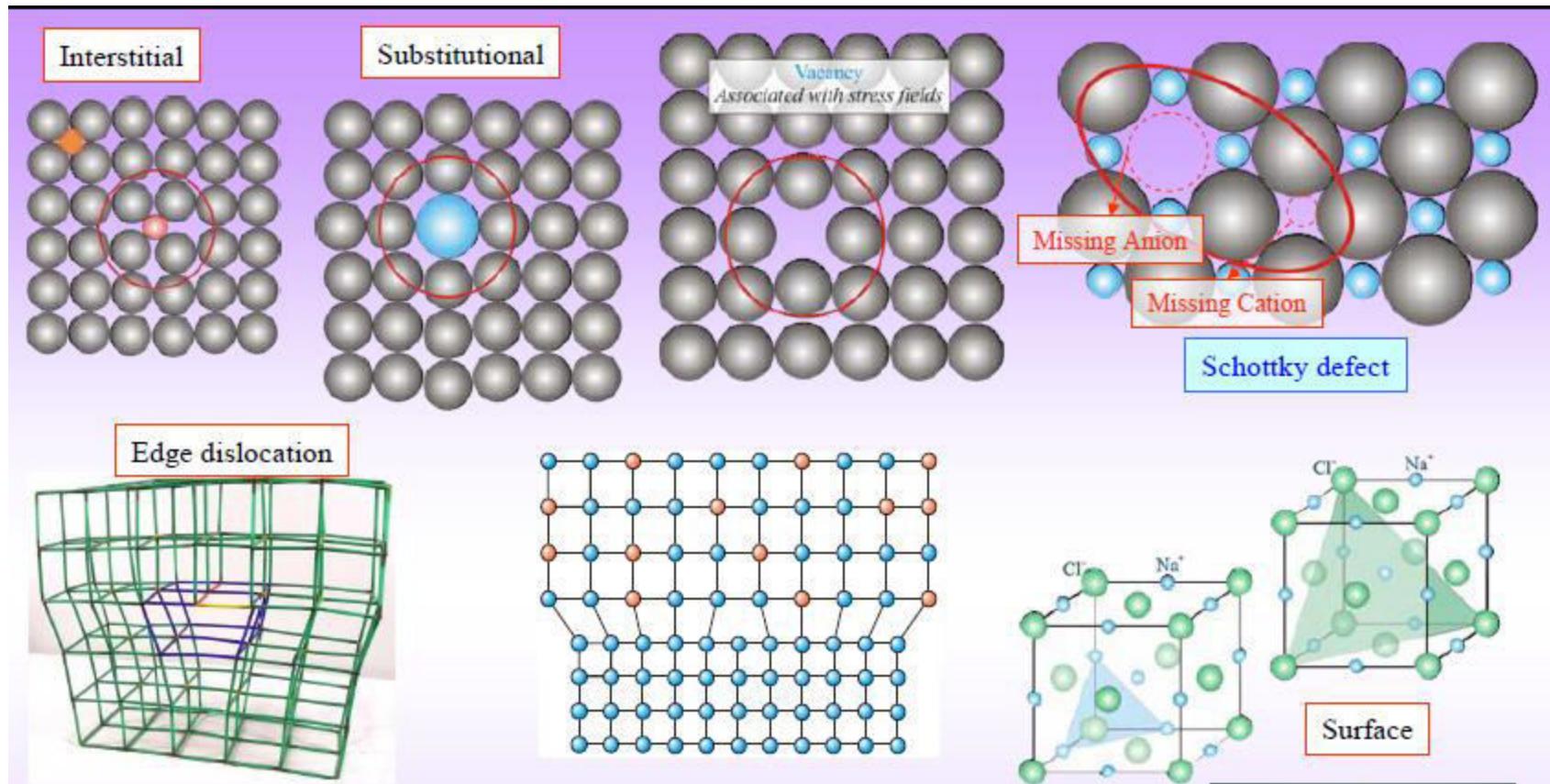
Precipitate

Faulted region

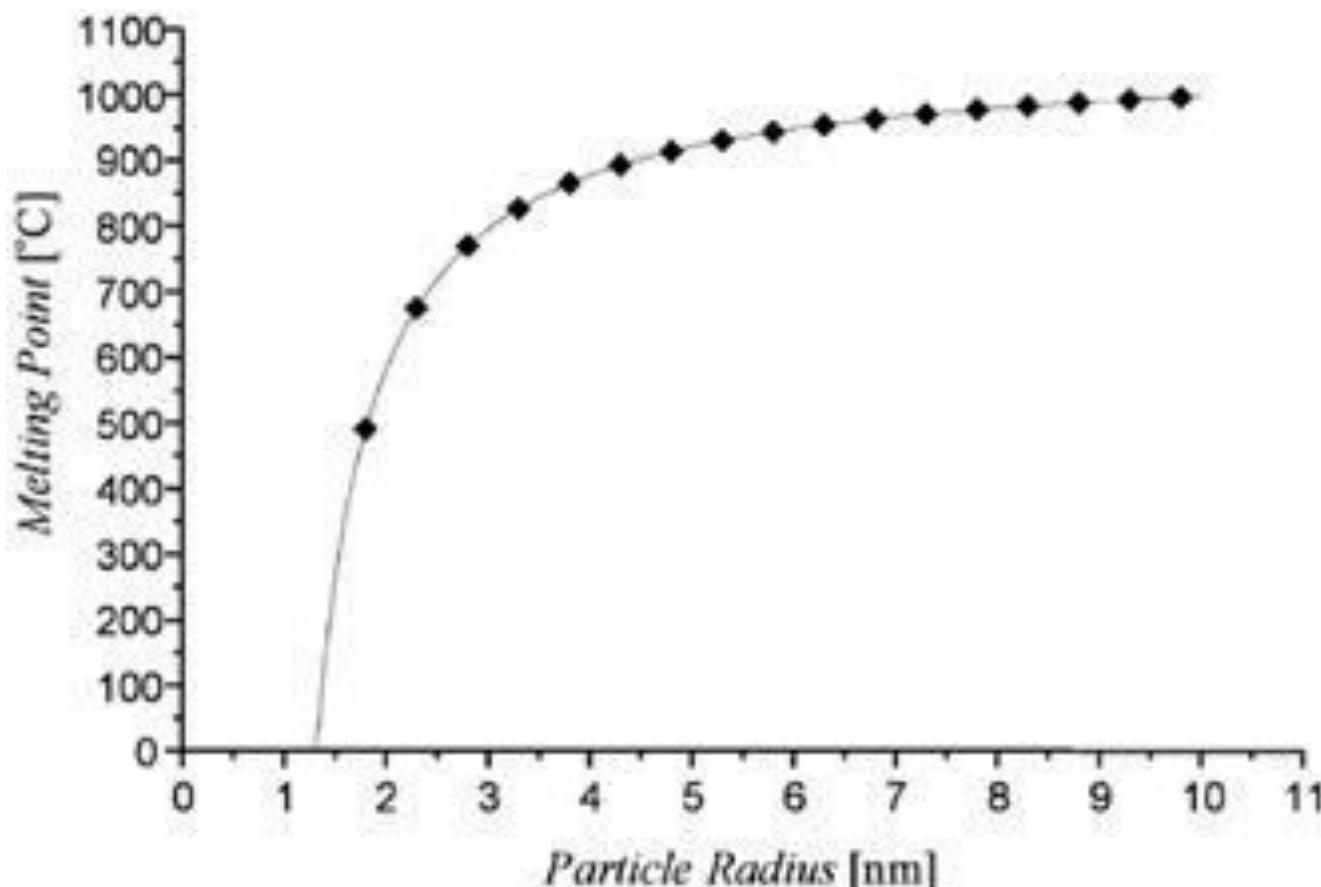
Voids / Cracks

Thermal vibration





- ☐ Melting point of a material is an indicator of the bond strength of a material (though boiling point is better reference as it is structure independent).
- ☐ In bulk systems surface to volume ratio is small and curvature effects can be ignored.
- ☐ In nanocrystals the number of surface atoms is a considerable fraction of the total number of atoms. These atoms have a higher energy and higher freedom for vibrations.
 - This implies that surface atoms *are expected* to melt below the melting point of the bulk.



MAGNETIC PROPERTIES

- The **magnetic moment** of nano particles is found to be **very less** when compared them with its bulk size.
- It should be possible that **non-ferromagnetic** bulk exhibit **ferromagnetic-like behavior** when prepared in nano range.
- Bulk Gold and Pt are **non-magnetic**, but at the nano size they are **magnetic**.

Even in bulk magnetic materials some structures can be in the nanoscale:

- Domain walls in a ferromagnet (~60nm for Fe).
- Some domains (especially those in the vicinity of the surface or grain boundaries), could themselves be nanosized.
- Spin clusters above paramagnetic Curie temperature (θ_p) could be nano-sized.

When we go from bulk to 'nano' only the structure sensitive magnetic properties (like coercivity) is expected to change significantly.

Some of the possibilities when we go from bulk to nano are:

- Ferromagnetic particles becoming **single domain**
- Superparamagnetism** in small ferromagnetic particles (i.e. particles which are ferromagnetic in bulk)

Quantum Confinement

145

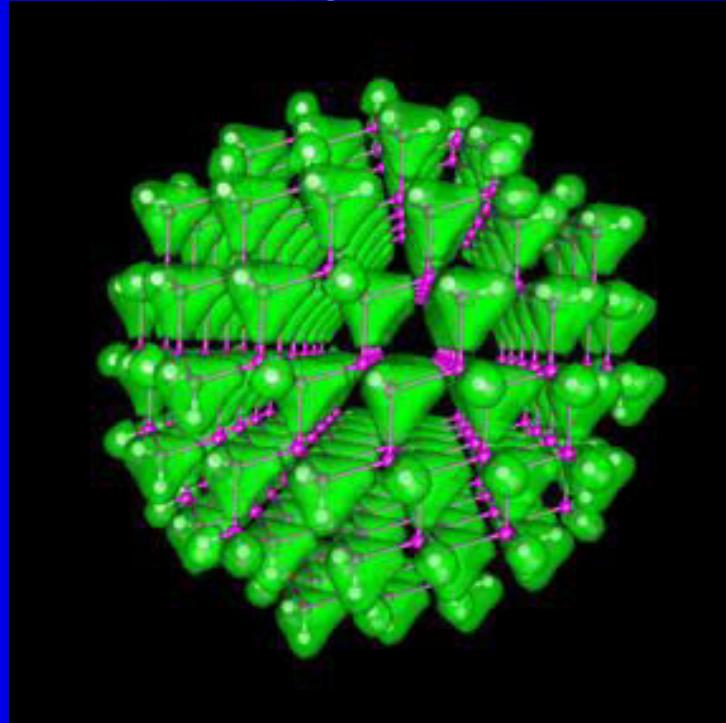
03 - Quantum dots - Reference - Material_1_08-Feb-2021

William E. Buhro (2003)

When the size of the material becomes in the order of the de Broglie wavelength of an electron, then the charge carriers are confined within the small region. The state of the system is known as quantum confinement.

Particle confined in this way is similar to a particle in a box of infinite height. It has only quantized energy levels. It introduces new quantum effects at nanosize which are used to make new electronic devices in telecommunication and optoelectronics.

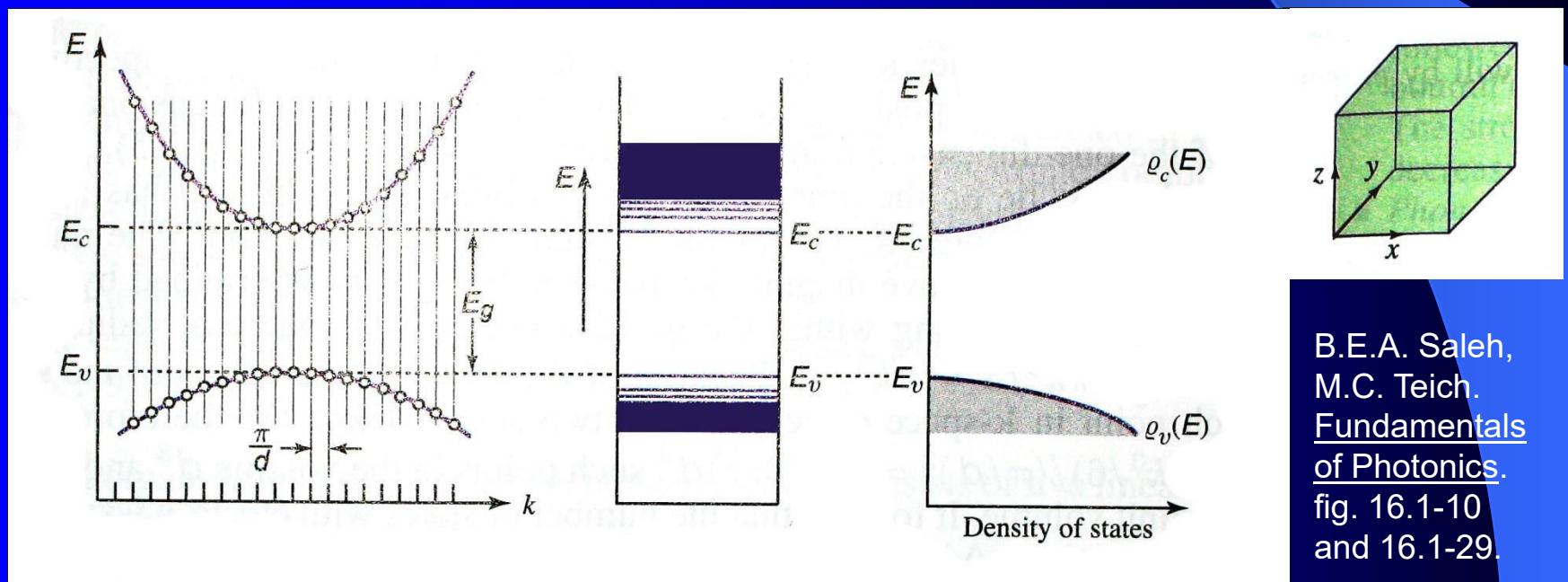
Quantum Well, wire and dot



Quantum effect describes the phenomenon resulting from electrons and electron holes being squeezed into a dimension that approaches a critical quantum measurement, called the exciton Bohr radius. In current application, a quantum dot such as a small sphere confines in three dimensions, a quantum wire confines in two dimensions, and a quantum well confines only in one dimension. These are also known as zero-, one- and two-dimensional potential wells, respectively. In these cases they refer to the number of dimensions in which a confined particle can act as a free carrier.

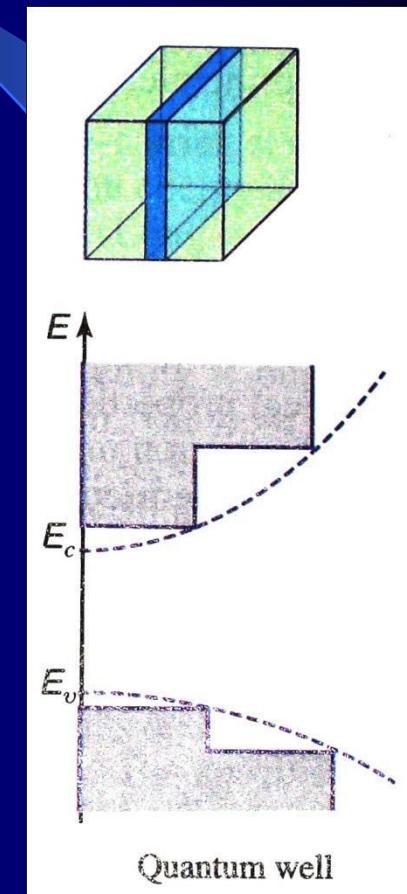
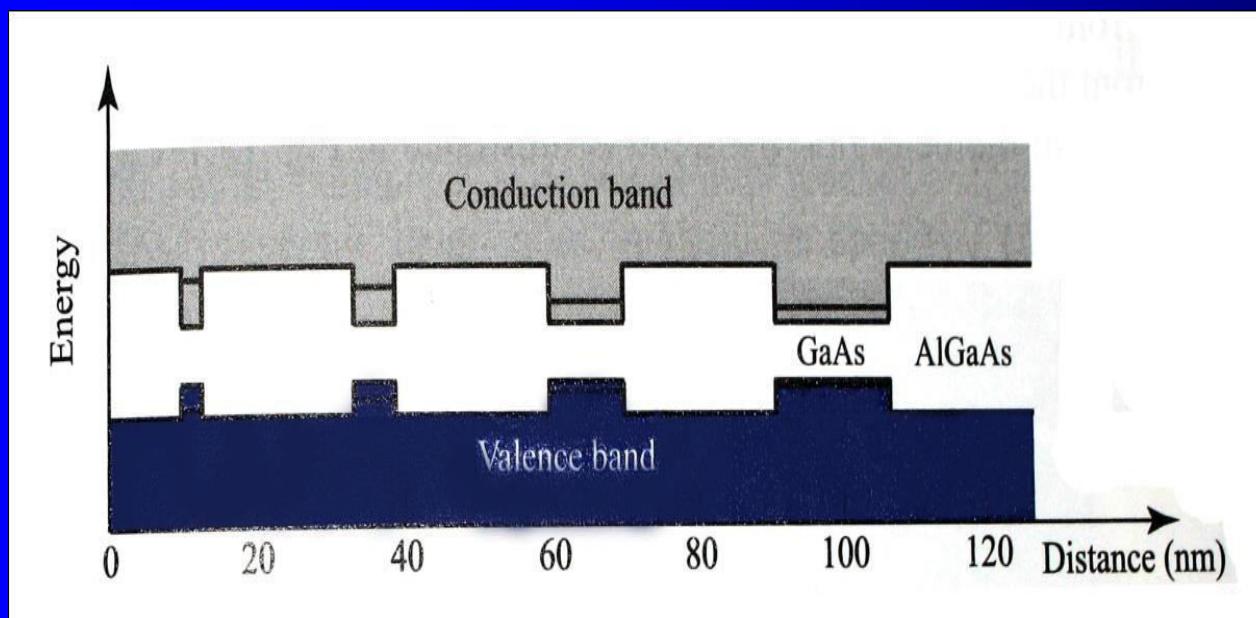
Bulk Semiconductors

- Electrons in conduction band (and holes in the valence band) are free to move in all three dimensions of space.



Thin Film Semiconductors

- Electrons in conduction band (and holes in the valence band) are free to move in two dimensions.
- Confined in one dimension by a potential well.
 - Potential well created due to a larger bandgap of the semiconductors on either side of the thin film.



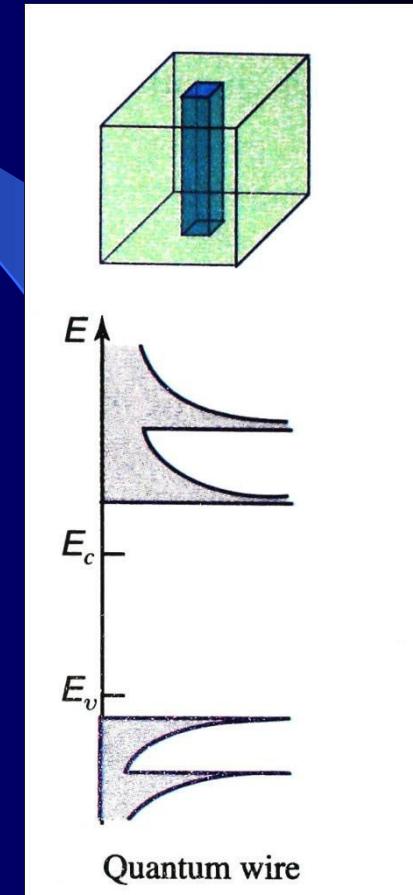
Electrons in quantum wells have a density of states as a function of energy that has distinct steps, versus a smooth square root dependence that is found in bulk materials.

The effective mass of holes in the valence band is changed to more closely match that of electrons in the conduction band.

Quantum wells are in wide use in diode lasers, including red lasers for DVDs and laser pointers, infra-red lasers in fiber optic transmitters, or in blue lasers. They are also used to make HEMTs (High Electron Mobility Transistors), which are used in low-noise electronics. Quantum well infrared photodetectors are also based on quantum wells, and are used for infrared imaging.

Quantum Wire

- Thin semiconductor wire surrounded by a material with a larger bandgap.
 - Surrounding material confines electrons and holes in two dimensions (carriers can only move in one dimension freely) due to its larger bandgap.
 - Quantum wire acts as a potential well.

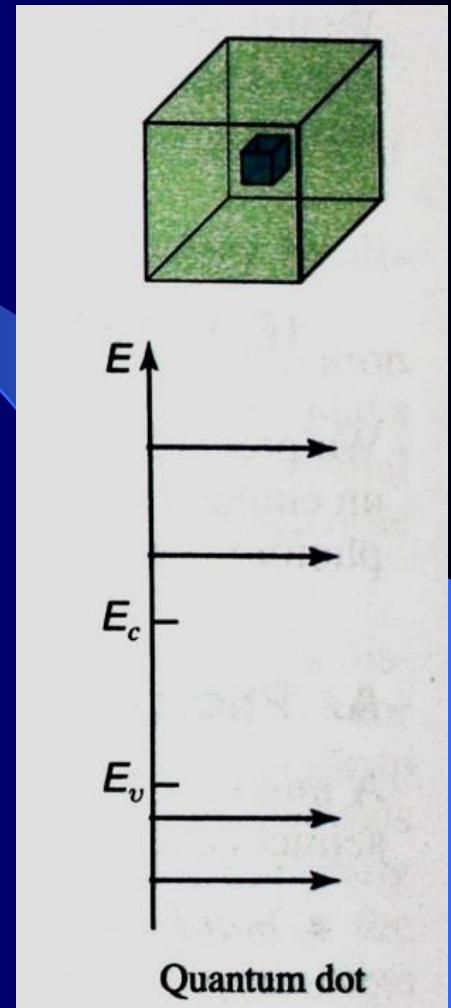


The carbon nanotube is an example of a quantum wire. A metallic single-walled carbon nanotube that is sufficiently short to exhibit no internal scattering has a conductance that approaches two times the conductance quantum $4e^2/h$

The factor of two arises because carbon nanotubes have two spatial channels

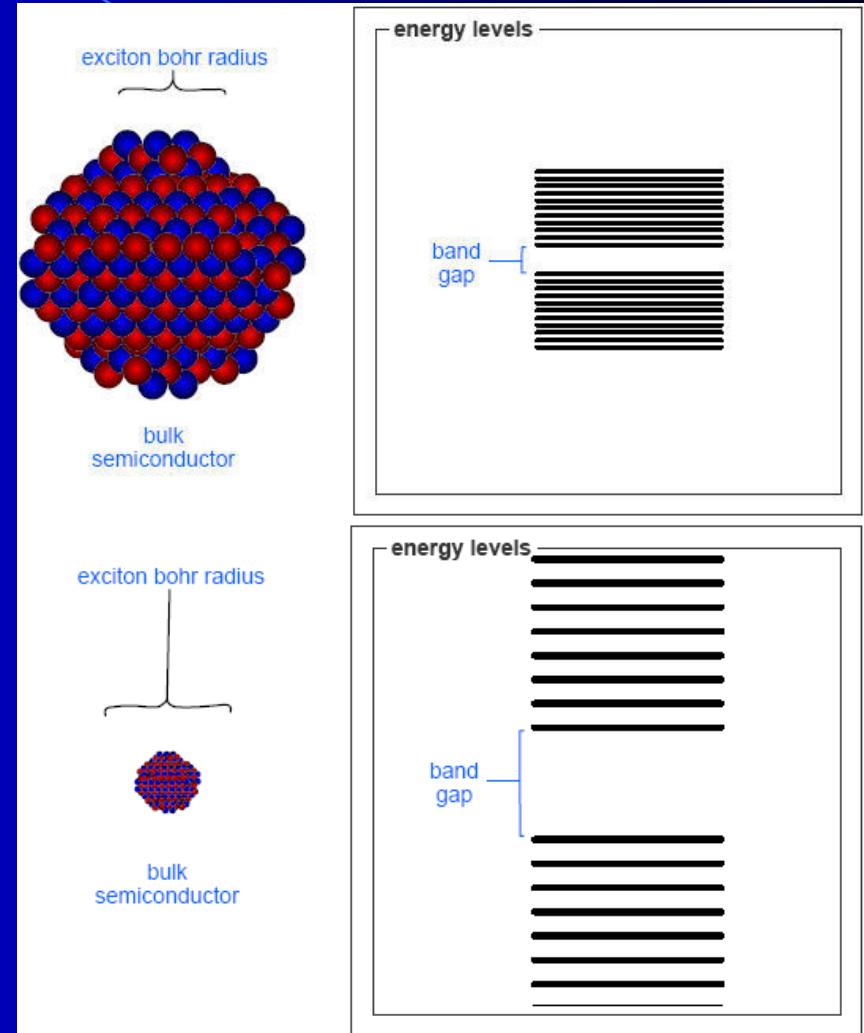
Quantum Dot

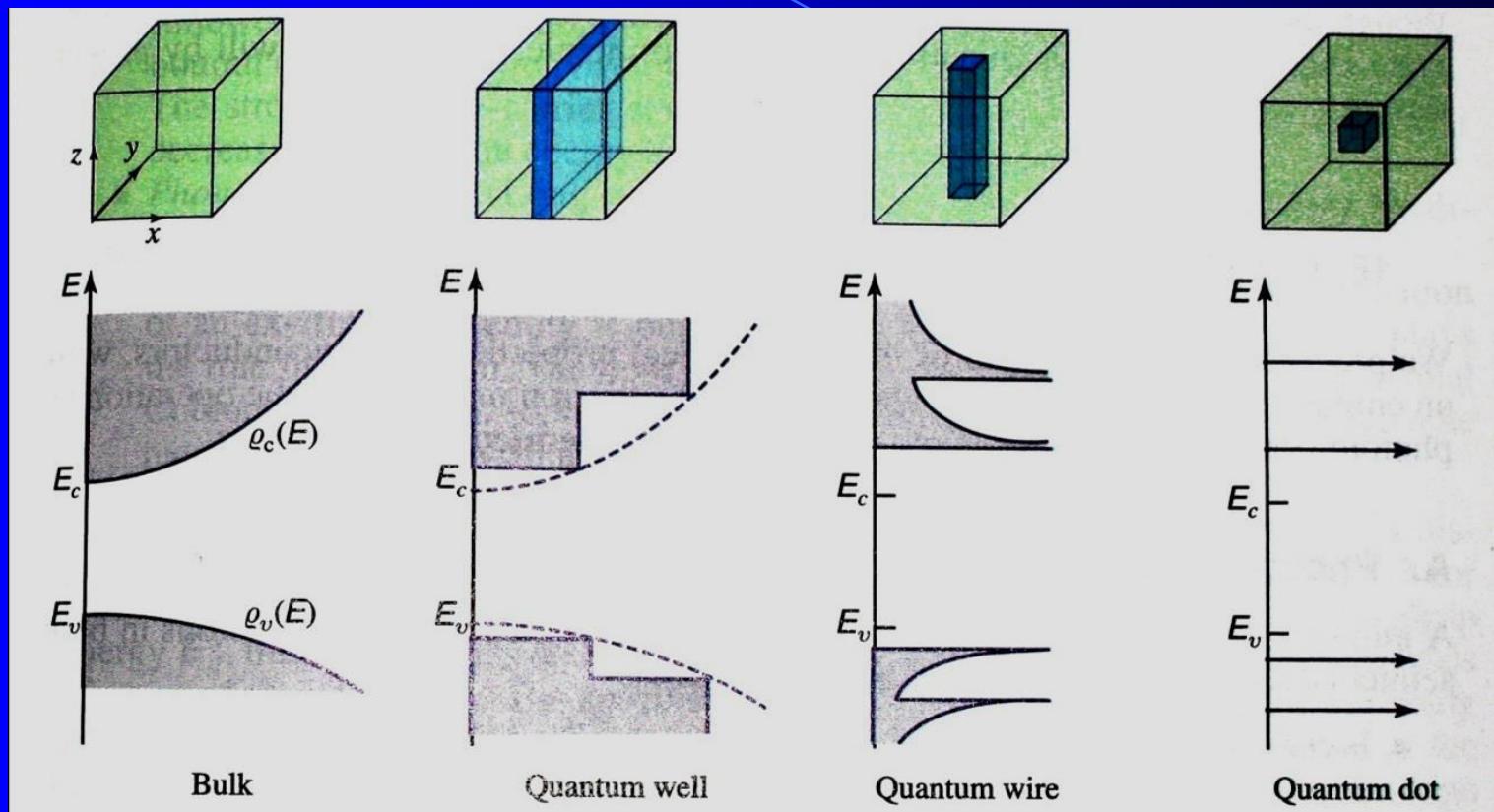
- Electrons and holes are confined in all three dimensions of space by a surrounding material with a larger band gap.
- Discrete energy levels (artificial atom).
- A quantum dot has a larger bandgap.
- Like bulk semiconductor, electrons tend to make transitions near the edges of the bandgap in quantum dots.



Discrete Energy Levels

- The energy levels depend on the size, and also the shape, of the quantum dot.
- Smaller quantum dot:
 - Higher energy required to confine excitons to a smaller volume.
 - Energy levels increase in energy and spread out more.
 - Higher band gap energy.





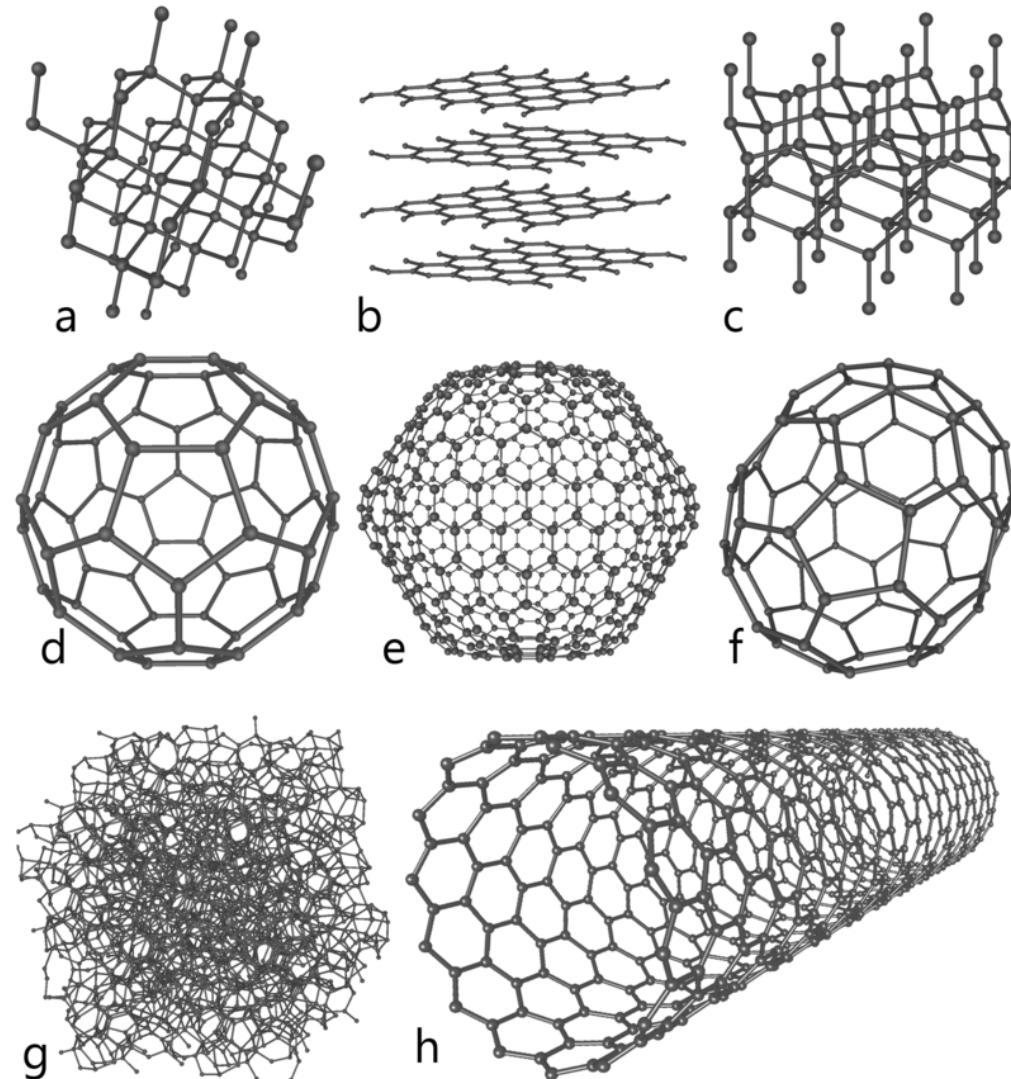
B.E.A. Saleh, M.C. Teich. Fundamentals of Photonics. fig. 16.1-29.

Carbon NanoTubes (CNT)

Carbon NanoTube (CNT)

Allotropes of Carbon

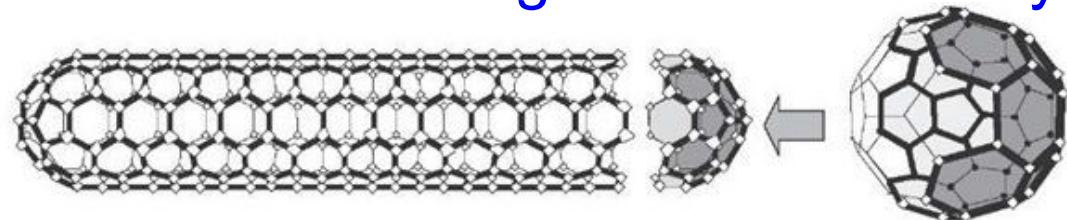
- (a). Diamond
- (b). Graphite
- (c). Lonsdaleite
- (d). C_{60} Fullerene
- (e). C_{540} Fullerite
- (f). C_{70}
- (g). Amorphous carbon
- (h). Carbon nanotube



Thanks wikipedia

Carbon NanoTube (CNT)

- Graphite: Each carbon linked to other by (sp^2 hybridization) covalent bonds and forms hexagonal rings in the planar structure (Graphene).
- Carbon nanotube is imagined to be obtained in the cylindrical form by rolling graphene sheet and closing both the ends by fullerene hemispheres.

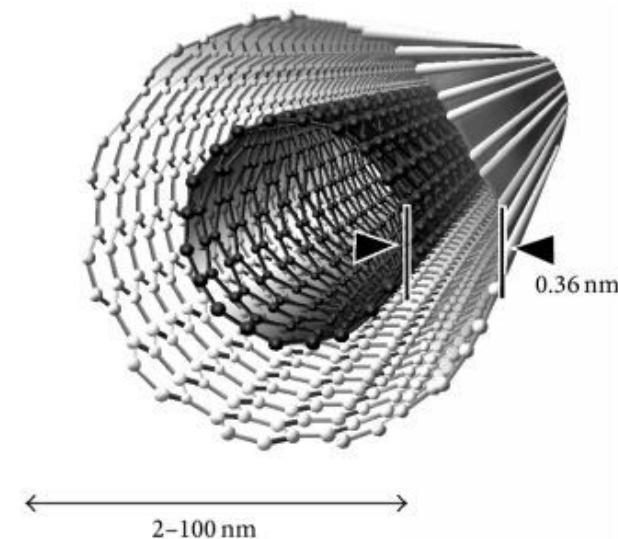
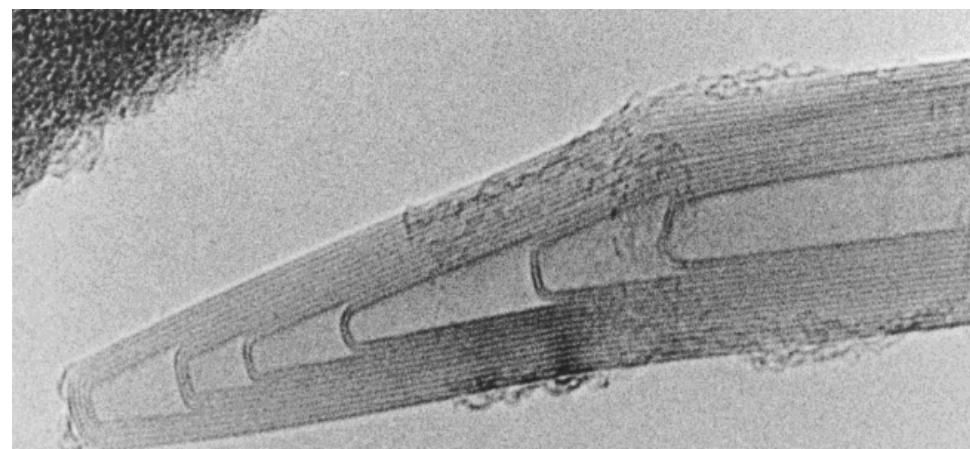
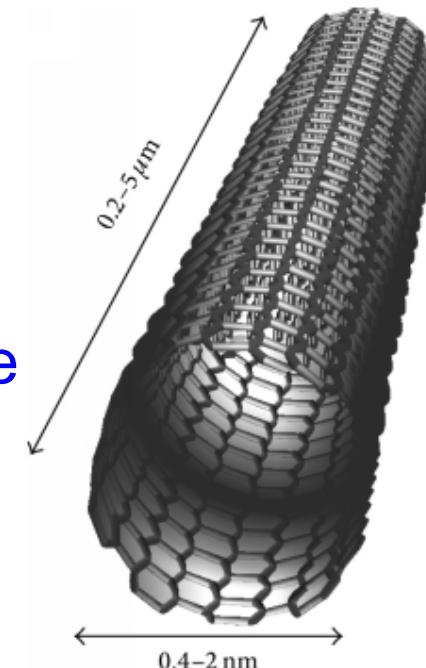


- The circular curvature will cause $\sigma - \pi$ rehybridization in which the 3 σ bonds are slightly out of plane; and for compensation, π bond is more delocalized outside the tube.
- This makes the nanotubes mechanically stronger, electrically and thermally more conductive, chemically and biologically more reactive than graphite.

Types of Carbon NanoTubes

Two types of CNTs

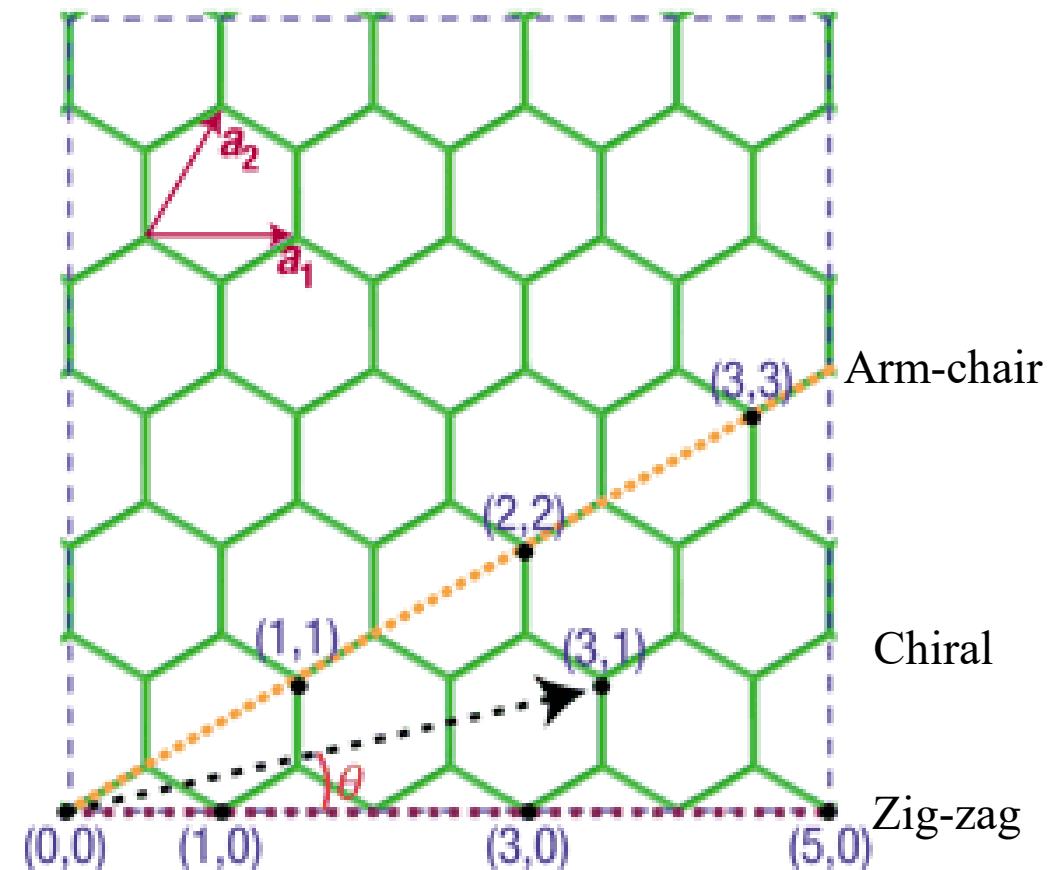
- Single walled (SWNT): A single-atom thick graphite (graphene) sheet rolled into cylinder and capped with fullerene hemisphere.
- Multi-walled (MWNT): Multiple rolled layers (concentric tubes) of graphene.



Structure of Carbon NanoTube

- SWNTs can have 3 different structures based on the direction of rolling graphene sheet.
- Nanotube lattice vectors are represented by Chiral vector

$$\vec{C}_h = n\vec{a}_1 + m\vec{a}_2$$
- a_1 and a_2 are basis vectors defined in terms of nearest-neighbour distance a_{cc} bond length of carbon atoms, $a_{cc} = 0.144 \text{ nm}$



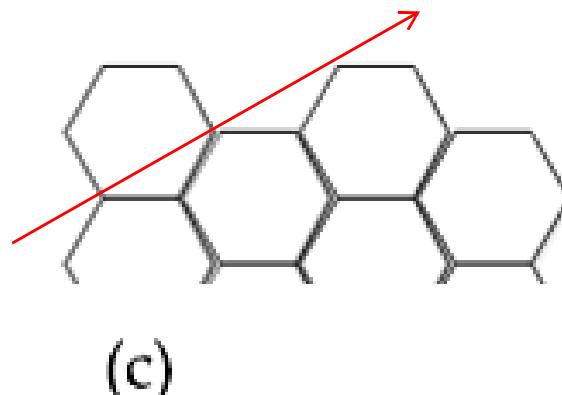
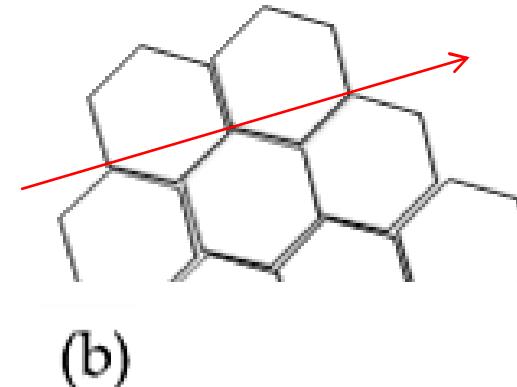
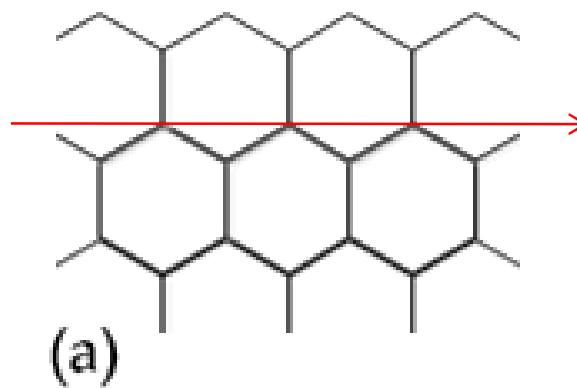
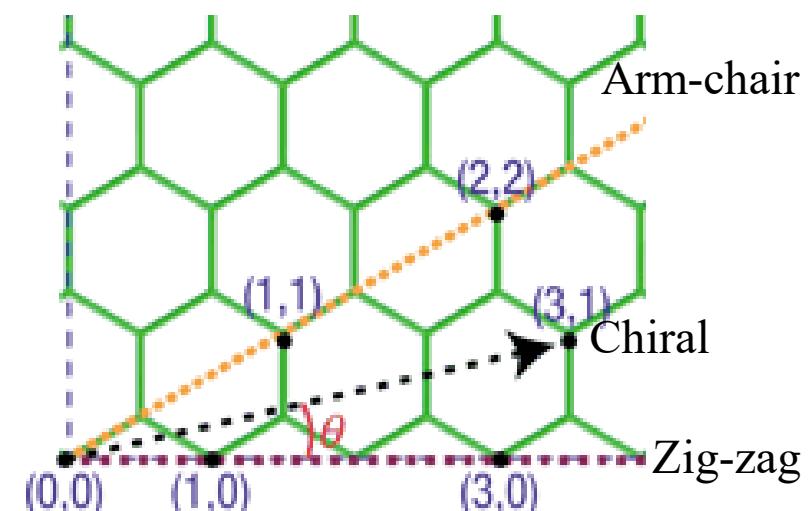
- Structure of nanotube depends on the chiral angle θ , defined as angle between chiral vector and zig zag direction.

$$\theta = \tan^{-1} \left(\frac{\sqrt{3}m}{m+2n} \right)$$

(a). Zig-zag ($\theta = 0^\circ$ for $m = 0$)

(b). Chiral ($0 < \theta < 30^\circ$ for $n \neq m$)

(c). Arm chair ($\theta = 30^\circ$ for $n = m$)



Potential Applications

**Terabit
memory device**



Conductive composite



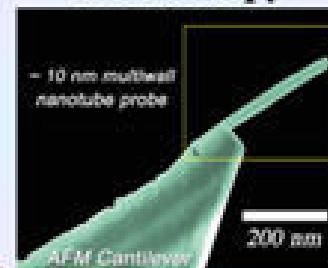
**Fuel cell &
Hydrogen storage**



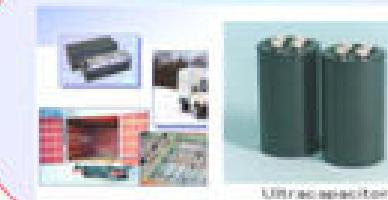
**Field
emission display**



Microscopy



Super capacitor



Did You Know?

- Carbon nanotubes, composed of interlocking carbon atoms, are 1000x thinner than an average human hair – but can be 200x stronger than steel.

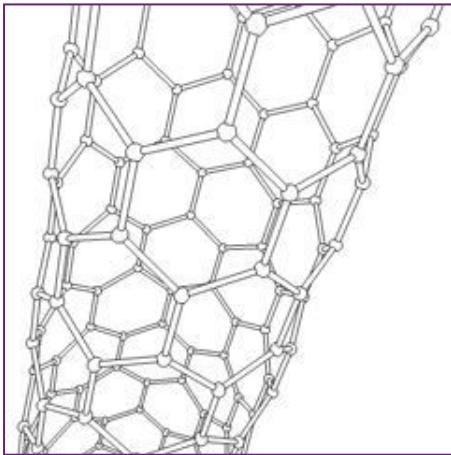


Image: Schwarzm, Wikipedia



CNT - Discovery

In 1991, Sumio Iijima discovers **multiwalled nanotubes** (MWNT) using the method of Krätschmer and Huffman.

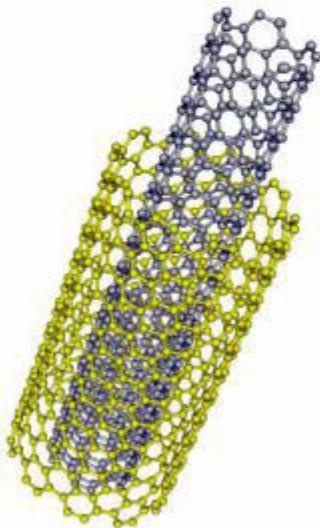
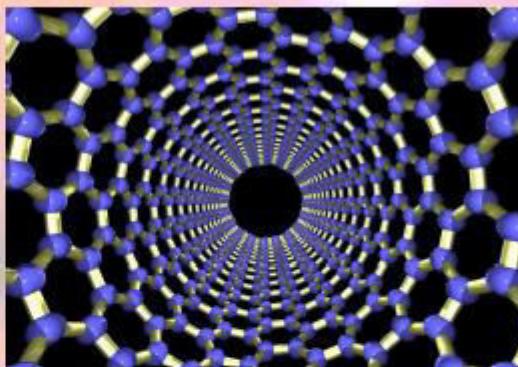


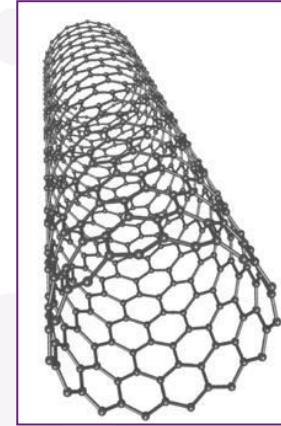
Image: Schwarzm, wikipedia



In 1993, Donald Bethune makes **single-walled** (SWNT) nanotubes by adding transition metals.



Which Of These Objects Are Made From Carbon?



Diamond

Graphite

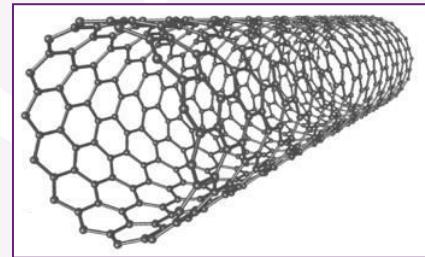
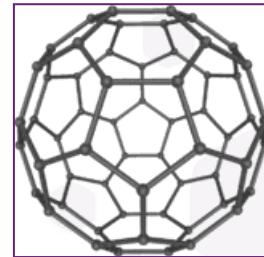
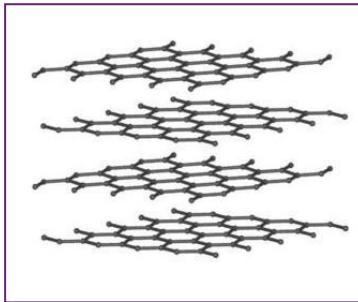
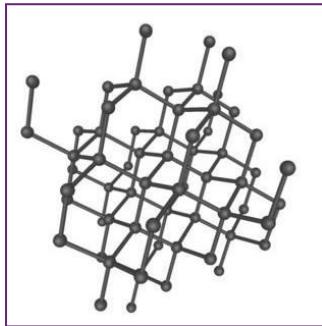
Nanotube

Coal



Did You Know?

Allotropes of carbon have different covalent bonding arrangements.



diamond

graphite

buckyball

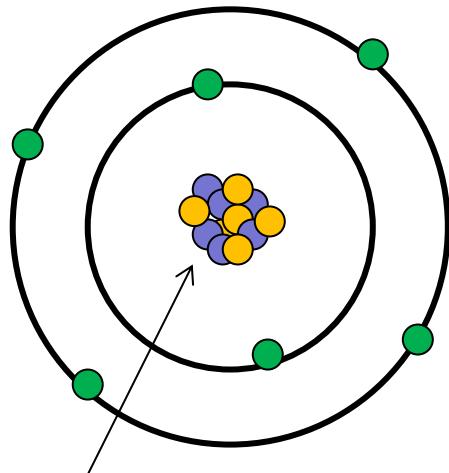
nanotube

- Carbon atoms form covalent bonds by sharing outer shell electrons with each other
- Diamond, graphite, buckyballs and carbon nanotubes all have different covalent arrangements of carbon atoms
- The differing covalent arrangements of carbon atoms lead to the different properties of carbon allotropes.



Covalent Bonding Sharing Electrons

- proton
- neutron
- electron



6 protons + 6
neutrons

A **covalent bond** is a form of chemical bonding that is characterised by the sharing of pairs of electrons between atoms

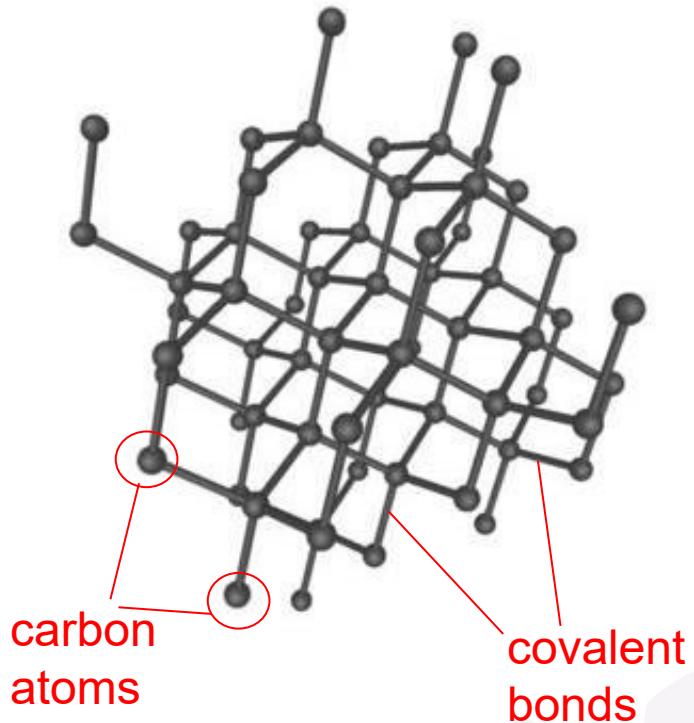
Valence electrons are the electrons in the outer shell or energy level of an atom that form covalent bonds

A carbon atom has 6 electrons, 4 of which are Valence electrons

Therefore, carbon atoms can form up to 4 Covalent Bonds



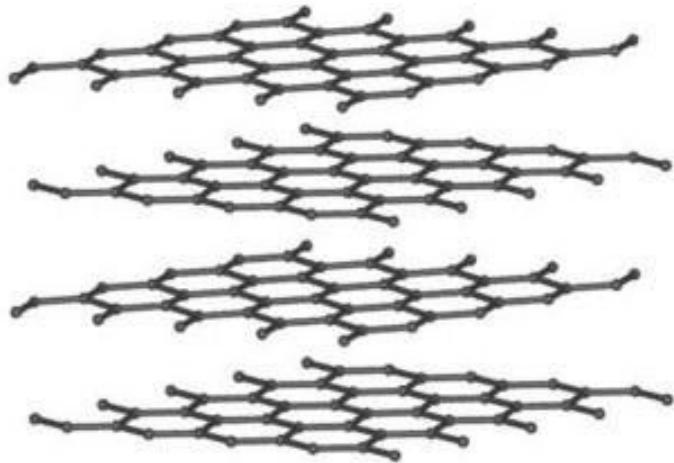
Covalent Bonds In Diamond



- Diamond is formed by a 3D box-like network of carbon atoms
- The continuous nature of the covalent arrangements forms a giant molecule
- Electrons are fixed.



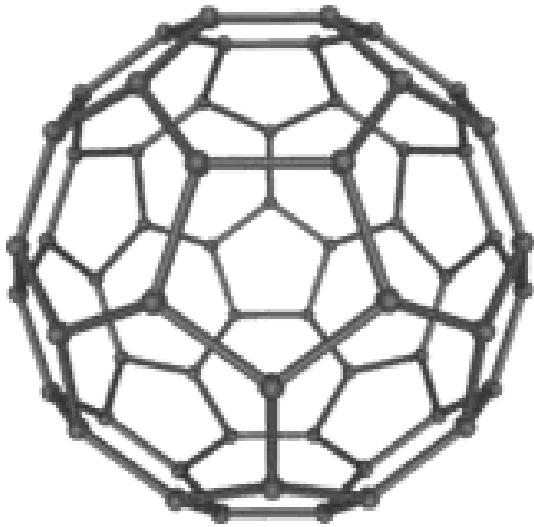
Covalent Bonds In Graphite



- Graphite is formed by hexagonally-arranged carbon molecules forming 2D layers of sheets
- Electrons are free to move between each carbon sheet.



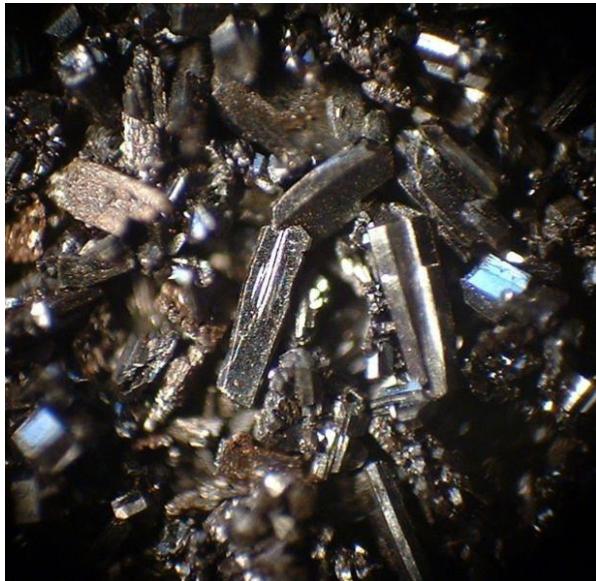
Covalent Bonds In Buckyballs



- Carbon atoms in buckyballs are arranged in a soccer ball shape
- C₆₀ Buckyballs have 20 regular hexagon faces and 12 regular pentagon faces
 - these faces come together at 60 carbon atom vertices
- Electrons are localised internally due to the curvature of the structure.



A Bit More About Buckyballs

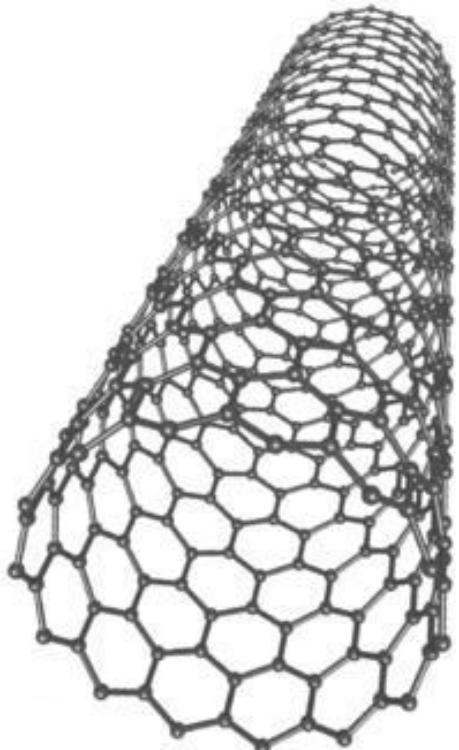


Buckyballs in crystalline form

- Buckyballs are also called fullerenes (after architect Richard Buckminster Fuller)
- Buckyballs were discovered in 1985 by Robert Curl, Harold Kroto and Richard Smalley
 - these scientists won the 1996 Nobel Prize in Chemistry for discovering this new allotrope of carbon.



Covalent Bonds In Carbon Nanotubes



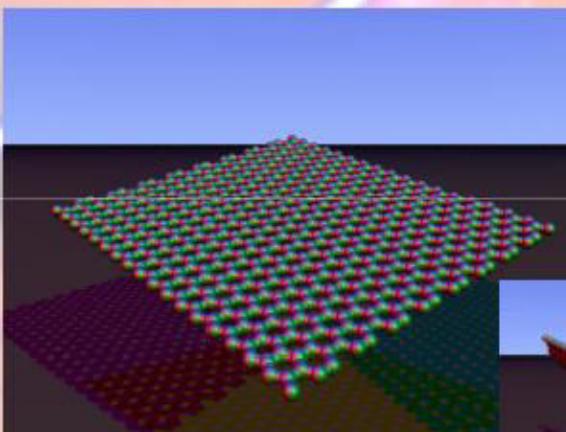
- Carbon nanotubes are formed by a layer of hexagonally-arranged carbon atoms rolled into a cylinder
 - usually have half buckyballs on one or both ends
- Electrons are localised internally, and some can move along the length of the tube
- Carbon nanotube diameter $\sim 1\text{nm}$
- Carbon nanotube length can be a million times greater than its width
- Nanotubes can be
 - single-walled ($d = 1\text{-}2\text{ nm}$), or
 - multi-walled ($d = 5\text{-}80\text{ nm}$).



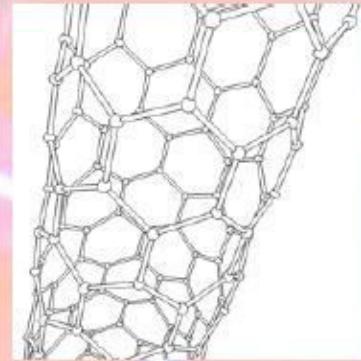
Properties of Carbon Allotropes

Carbon nanotubes

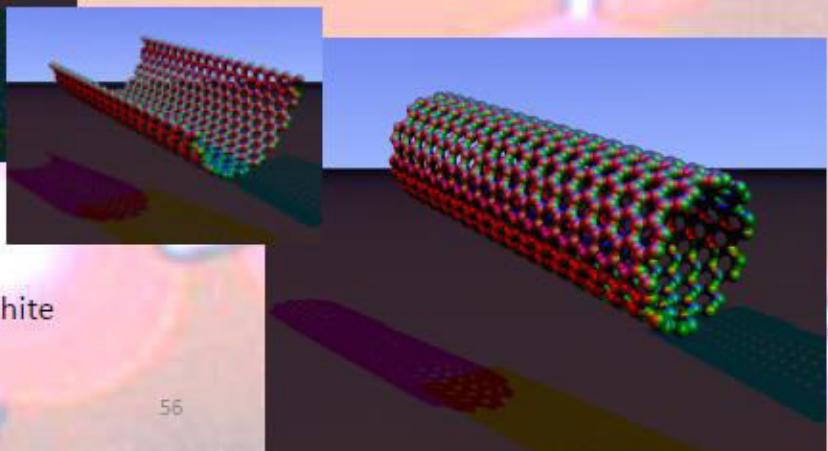
Carbon nanotubes (CNTs) are allotropes of **carbon** with a cylindrical nanostructure.



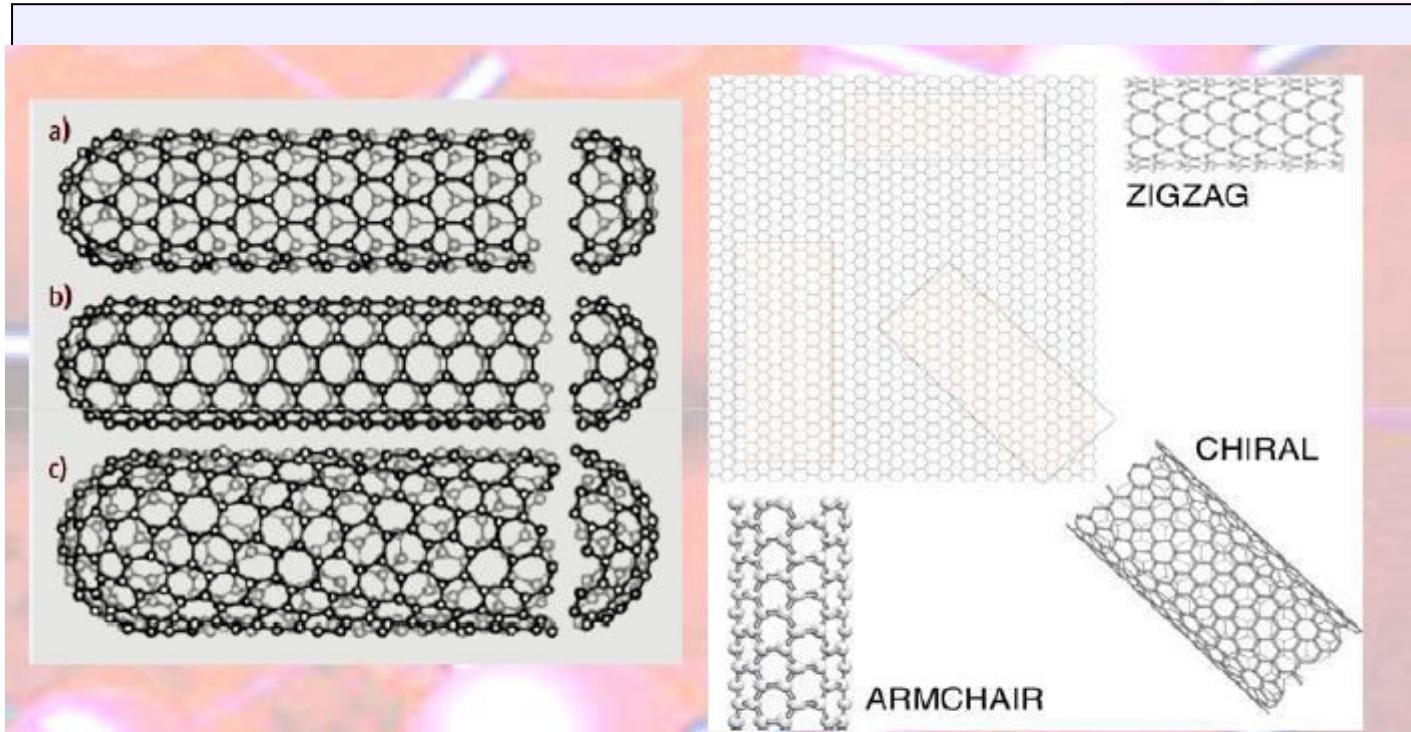
Structure is a single sheet graphite rolled into a tube.



Most **single-walled nanotubes (SWNT)** have a diameter of close to 1 nm, with a tube length that can be many millions of times longer.



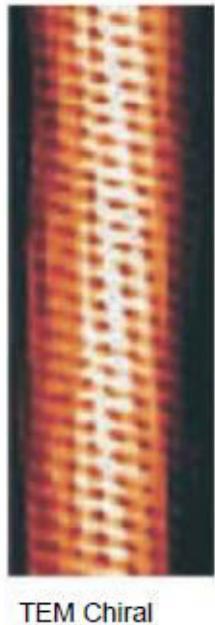
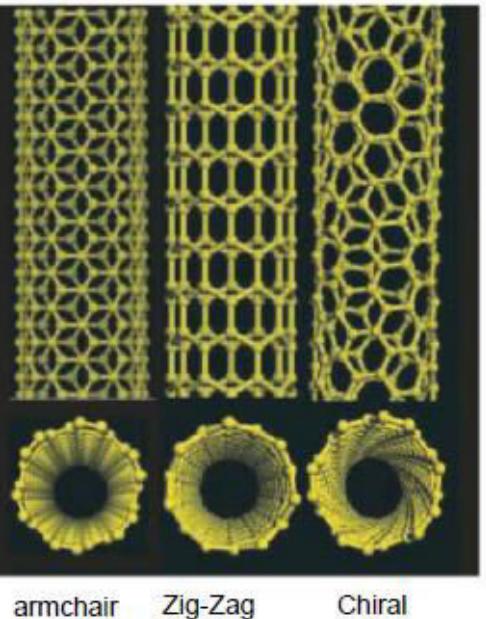
Types of Carbon Allotropes



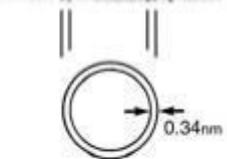
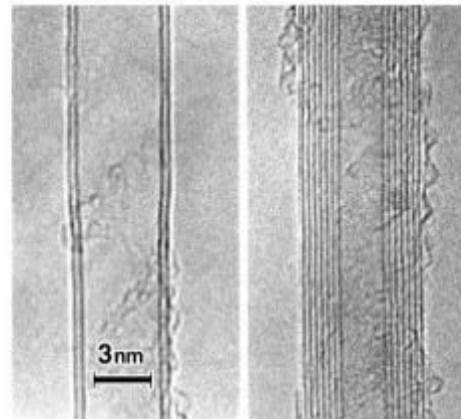
Sketch of three different SWNT structures as examples for
(a) a zig-zag-type nanotube, (b) an armchair type nanotube,
(c) a helical nanotube



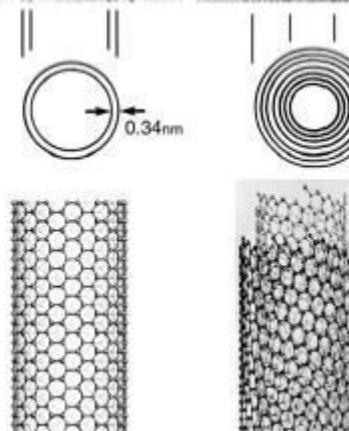
Types of Carbon Allotropes



TEM Chiral

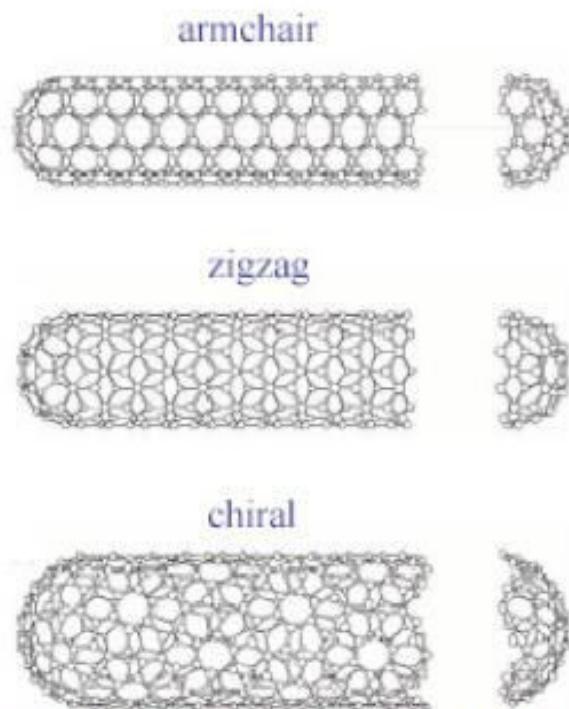
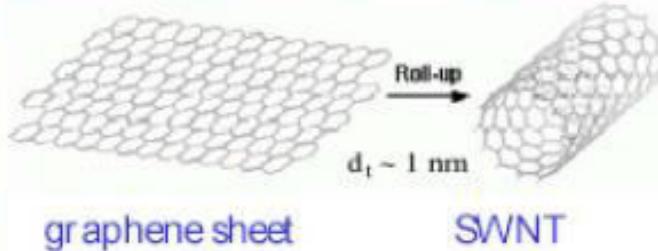


SWNT
(single
Wall nanotube)
Diameter ~ 1.4 nm



MWNT
(multiwall)
Diameter 10-20 nm

Properties of Carbon Allotropes



- **Size:** Nanostructures with dimensions of ~1 nm diameter (~10 atoms around the cylinder)
- **Electronic Properties:** Can be either metallic or semiconducting depending on diameter and orientation of the hexagons
- **Mechanical:** Very high strength, modulus, and resiliency. Good properties on both compression and extension.
- **Physics:** 1D density of electronic states
- Single molecule Raman spectroscopy and luminescence.
- Single molecule transport properties.
- Heat pipe, electromagnetic waveguide.

Properties of Carbon Allotropes

Allotrope	Hardness	Tensile strength	Conducts heat	Conducts electricity
Coal	+	+	+	no
Graphite	++	++	++++	++++
Diamond	++++	Not known	+++	no
Buckyballs	++++	+++	+	+
Carbon Nanotubes	+++++	++++	++++	++++



Periodic graphics

A collaboration between C&EN and Andy Brunning, author of the popular graphics blog **Compound Interest**

More online

 To see more of Brunning's work, go to compoundchem.com. To see all of C&EN's Periodic Graphics, visit cenm.ag/periodicgraphics.

EVERYDAY USES OF NANOTECHNOLOGY

National Nanotechnology Day (Oct. 9) is a yearly event in the U.S. to celebrate the tiny tech. Here, we take a look at various consumer products that utilize nanotechnology and the chemistry behind them.

WHAT IS NANOTECHNOLOGY?



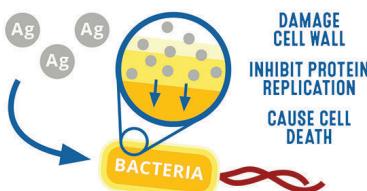
SALT GRAIN = 100,000 nm



NANOPARTICLES = 1–100 nm

Nanotechnology involves the applications of nanoparticles, which are collections of atoms or molecules less than 100 nm across. Because of their small size, the particles have properties that can differ from those of larger amounts of the same material.

ANTIMICROBIAL USES



SUNSCREENS



ZnO MAINLY BLOCKS UV-A



UV-A wavelength 320–400 nm



UV-B wavelength 290–320 nm

CLOTHES



SILVER ANTIMICROBIAL

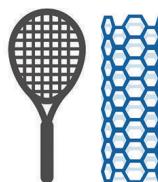
SILICA WATER-REPELLENT

TiO₂/ZnO UV-ABSORBING

ANTIMONY-DOPED TIN OXIDE ANTISTATIC

UV-absorbing titanium oxide and zinc oxide nanoparticles can be incorporated into clothes to prevent sunburn and sometimes to act as antistatic agents. Silicon dioxide nanoparticles can prevent stains and help clothing repel water.

SPORTS EQUIPMENT



CARBON NANOTUBES

100 TIMES AS STRONG AS STEEL

ONE-SIXTH THE WEIGHT OF STEEL

AS STIFF AS DIAMOND

Sports equipment such as tennis rackets and bicycles are sometimes built using nanomaterials including carbon nanotubes. The nanotubes improve strength and durability and decrease weight. Titanium nanoparticles can also be used.

QUANTUM DOTS



Quantum dots, which are nanoparticles of semiconductors such as cadmium selenide, absorb light of one color, such as blue light, and emit it as another depending on particle size. The particles are more energy-efficient than light-emitting diodes.



© C&EN 2016 Created by Andy Brunning for *Chemical & Engineering News*

LASER

4

4.1 INTRODUCTION

The word Laser is an abbreviation for ‘Light Amplification by Stimulated Emission of Radiation’. It is one of the most important discovery of 20th century. The word Laser is also used for a device that emits a narrow intense beam of light which differs from ordinary light and has very special applications. Output of a laser can either be a continuous beam of low to medium power or pulses of intense radiation.

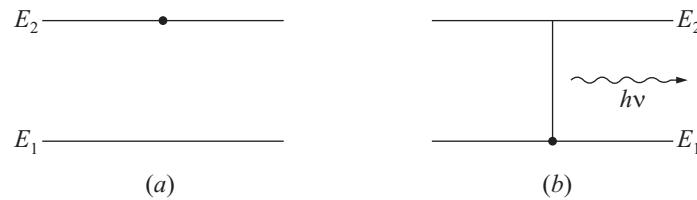
The first laser was built in 1960 by Theodore Maiman and other scientists in California (U.S.A.), although the essential ingredients for lasers were provided by Einstein in 1917. At present various types of lasers have been developed using liquids and solids. Before describing the operation of a particular laser, it is important to discuss the following three transition phenomenon given by Einstein.

4.2 SPONTANEOUS EMISSION

It is well known that there are various energy levels in an atom. Ground state of the atom is the minimum energy state and it is the most stable state. When the atom gets suitable thermal energy, its valence electron (say of energy E_1) jumps to higher energy level (say to energy E_2) called excited level. Electrons in this level are also called atoms in its excited state. Life time of electron in the excited state is very small, of the order of 10^{-8} sec, hence the electron within this time falls back to lower energy level E_1 by emitting a radiation. This process is called spontaneous emission. The frequency, v of the emitted radiation is given by

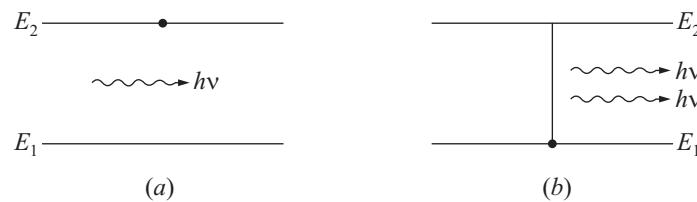
$$v = \frac{E_2 - E_1}{h} \quad \dots(4.1)$$

where h is the Planck constant. We also say that in this transition a photon of energy hv is emitted. If there are large number of atoms in the upper energy level then the emitted radiations will have randomly different initial phases and directions and the emitted radiations will be incoherent.

**Fig. 4.1.** (a) Initial state, (b) Final state

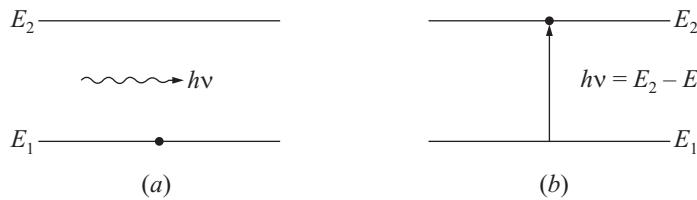
4.3 STIMULATED EMISSION

If an atom is in an excited state of energy E_2 and an incident photon of suitable energy ($h\nu = E_2 - E_1$) may cause the atom (electron) to jump to lower energy state ($\approx E_1$) emitting an additional photon of same frequency, same phase and in the same direction. Thus now two photons each of energy equal to $E_2 - E_1$ are present. This kind of transition is called stimulated emission of radiation. This is shown in Fig. 4.2.

**Fig. 4.2.** (a) Initial state (b) Final state

4.4 ABSORPTION OF RADIATION

If an atom in its ground state of energy E_1 and radiation of suitable energy ($h\nu = E_2 - E_1$) is given such that the atom goes to excited state E_2 i.e., its electron jumps from E_1 level to higher energy state E_2 by absorbing a quantum of radiation or photon. This kind of transition is called the absorption of radiation. This is shown in Fig. 4.3.

**Fig. 4.3.** (a) Initial state (b) Final state

4.5**RELATION BETWEEN EINSTEIN'S A AND B COEFFICIENTS**

Let us consider an enclosure containing atoms which are in thermal equilibrium or in steady state. Let N_1 and N_2 are the number of atoms per unit volume called population in energy levels E_1 and E_2 , respectively. Here E_2 is greater than E_1 . In thermal equilibrium three processes of transition described above will take place.

1. Spontaneous Emission: According to Einstein the probability of spontaneous emission from energy level E_2 to energy level E_1 per unit time is denoted by

$$(P_{21})_{\text{spontaneous}} = A_{21} \quad \dots(4.2)$$

A_{21} is called the Einstein's A coefficient of spontaneous emission of radiation. Thus the number of photons of energy $E_2 - E_1$ emitted per second by spontaneous emission in the system is equal to $N_2 A_{21}$.

2. Induced Emission: According to Einstein the probability of induced emission from energy level E_2 to energy level E_1 per unit time can be written as

$$(P_{21})_{\text{induced emission}} = B_{21} u(v) \quad \dots(4.3)$$

Here B_{21} is called the Einstein's B coefficient of induced emission of radiation and $u(v)$ is the energy density of the radiation of frequency v . Then the number of photons of energy $h\nu$ emitted per second by induced emission in the system is equal to $N_2 B_{21} u(v)$.

3. Absorption of Radiation: According to Einstein the probability of absorption of energy for transition from energy level E_1 to energy level E_2 per unit time can be written as

$$(P_{12})_{\text{absorption}} = B_{12} u(v) \quad \dots(4.4)$$

Here B_{12} is called the Einstein's B coefficient of absorption of radiation and $u(v)$ is the energy density of the radiation of frequency v . Then the number of photons of energy $h\nu$ absorbed per second in the system is equal to $N_1 B_{12} u(v)$.

In the thermal equilibrium state (*i.e.*, steady state) total number of photons absorbed per second should be equal to the total number of photons emitted per second. It can be written as

$$N_1 B_{12} u(v) = N_2 B_{21} u(v) + N_2 A_{21}$$

$$\text{or} \quad u(v) [N_1 B_{12} - N_2 B_{21}] = N_2 A_{21}$$

$$\text{or} \quad u(v) = \frac{A_{21}}{B_{21}} \frac{1}{\left(\frac{N_1}{N_2} \frac{B_{12}}{B_{21}} - 1 \right)} \quad \dots(4.5)$$

According to Maxwell-Boltzmann distribution the number of atoms N_1 and N_2 in the energy states E_1 and E_2 , respectively, in the steady state at temperature T are given by

$$N_1 \propto e^{-E_1/kT} \quad \text{and} \quad N_2 \propto e^{-E_2/kT} \quad \dots(4.6)$$

Here k is the Boltzmann constant. Therefore

$$\frac{N_1}{N_2} = e^{+(E_2 - E_1)/kT}$$

but $E_2 - E_1 = h\nu$ (energy of the photon emitted or absorbed).

$$\therefore \frac{N_1}{N_2} = e^{+hv/kT} \quad \dots(4.7)$$

Substituting the value of $\frac{N_1}{N_2}$ in Eqn. (4.5), we get

$$u(v) = \frac{A_{21}}{B_{21}} \frac{1}{\left(e^{hv/kT} \frac{B_{12}}{B_{21}} - 1 \right)} \quad \dots(4.8)$$

Planck derived an expression for $u(v)$ in the following form

$$u(v) = \frac{8\pi hv^3}{c^3} \frac{1}{e^{hv/kT} - 1} \quad \dots(4.9)$$

Comparing Eqns. (4.8) and (4.9) for $u(v)$, we get

$$B_{12} = B_{21} \quad \dots(4.10)$$

and
$$\frac{A_{21}}{B_{21}} = \frac{8\pi hv^3}{c^3} \quad \dots(4.11)$$

Relation (4.10) shows that transition probability of absorption is equal to transition probability of induced emission. Relation (4.10) has also been proved in semiclassical theory of radiation by using perturbation theory. Equation (4.11) shows that A_{21}/B_{21} is proportional to v^3 . This is the ratio of Einstein's A coefficient of spontaneous emission and Einstein's B coefficient of induced emission.

4.6 POPULATION INVERSION

It is well-known that the process of spontaneous emission is independent of external radiations. Einstein proved theoretically that transition probability of absorption is equal to transition probability of induced emission between same two levels. If the number of atoms N_1 in lower state (energy E_1) is more than the number of atoms N_2 in the upper state (energy E_2) then in the presence of external radiations of frequency $(E_2 - E_1)/h$, the absorption dominates over emission. On the other hand, if N_1 is less than N_2 then the emission dominates. In general at room temperature N_1 is greater than N_2 , so the absorption is more than emission. In some materials if the life time of any upper state (level) is of the order of 10^{-3} sec, called metastable state, then it is possible to have that material with $N_2 > N_1$. If this happens, then the population inversion takes place in the medium. These materials are called active materials and they can be used for lasing transition. Thus for laser action population inversion is necessary.

Population of a particular energy level of atoms is given by Maxwell-Boltzmann distribution function as

$$N \propto e^{-E/kT} \quad \dots(4.12)$$

where N is the number of atoms in the state whose energy is E , T is the temperature of the material and k is the Boltzmann constant. Using this relation we get

$$\frac{N_2}{N_1} = e^{-(E_2 - E_1)/kT} = e^{-hv/kT} \quad \dots(4.7)$$

Since $E_2 > E_1$, therefore $N_2 < N_1$ under normal conditions. As mentioned above for laser action N_2 should be more than N_1 .

The process by which atoms are raised from lower level to upper level is called the pumping. Following are the commonly used methods of laser pumping:

- (a) Optical pumping
- (b) Electrical pumping
- (c) Chemical pumping.

(a) Optical Pumping: In this process the population inversion is achieved by using strong source of light such as gaseous discharge, flash lamp or arc lamp. This process keeps more atoms in the upper excited (metastable) state. This process is suited to solid state (*e.g.* Ruby) and liquid (*e.g.*, dye) lasers.

(b) Electrical Pumping: In this process the population inversion is achieved by using an intense electrical discharge in the medium which is in the gaseous form. This process is suited to gas lasers and semiconductors. The discharge converts the gas into plasma and the population of the upper level also increases.

(c) Chemical Pumping: In this process the population inversion is achieved by using suitable chemical reaction in the absence of any other source of energy. For example, in the reaction $A + B \rightarrow AB^*$, AB^* is the excited vibrational state of AB molecule.

4.7

LASER ACTION

Following are the three main things required for laser action:

(1) An active material is required in which population inversion can be attained.

(2) An specially designed cylindrical tube or rod fitted with mirrors at the two end surfaces which help in increasing the light intensity by multiple reflections. One end is completely silvered while the other end is partially silvered so that an intense beam can emerge out of it.

(3) Pumping system is required to achieve population inversion.

At first sight it appears that a system with two levels only can be used for lasing action and population inversion can be achieved through interaction of the material with a sufficiently strong *e.m.* field of frequency $\nu = (E_2 - E_1)/h$. Such type of interaction does not work. At thermal equilibrium level 1 is more populated than level 2, hence absorption will dominate over induced emission. With passage of time the population of both the levels will become equal and then the rate of absorption will be equal to rate of emission and the material will then become transparent. This situation is often referred to as two level saturation. Thus with the use of just two levels 1 and 2, it is impossible to produce a population inversion.

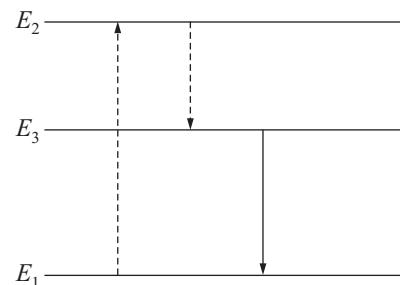


Fig. 4.4

In order to understand laser action, let us consider three energy levels E_1 , E_2 and E_3 as shown in Fig. 4.4. Here E_1 is the ground state, E_2 is the normal excited state of life time about 10^{-8} sec and E_3 is the meta stable state of life time of the order of 10^{-3} sec. With the help of pumping system active material in the ground state E_1 is continuously excited to E_2 state. From this state they immediately relax down to the intermediate energy level E_3 which is metastable state so the atoms stay there for a long time. Net effect of it is that the population of E_1 decreases and population of E_3 increases continuously i.e., population inversion between states E_1 and E_3 takes place. This is the necessary condition for laser action. Now if an atom from level E_3 jumps down to E_1 level by spontaneous emission then a

photon of energy $E_3 - E_1$ (or radiation of frequency $\nu = \frac{E_3 - E_1}{h}$) is released. When these photons of

frequency ν are incident on the atoms in the excited state, the atoms are forced to undergo transitions from excited state E_3 to their ground state E_1 . These forced photons (radiations) are in the same phase and direction as the incident photons (radiations). This emission is called induced emission or stimulated emission. These emitted radiations amplify the light which is named as laser. The word LASER stands for ‘Light Amplification by Stimulated Emission of Radiation’.

In actual laser instrument the active material is placed between a pair of reflecting mirrors (e.g., plane parallel mirrors) forming what is known as ‘Resonator Cavity’. Usually one of the two mirrors is made partially transparent so as to get a useful output in the form of intense laser beam.

4.8 MAIN FEATURES OF LASER

The most striking features of a laser are the following:

(a) Directionality: A conventional source of light (like a sodium lamp or an incandescent bulb) emit radiations in all directions. An aperture is used in front of such a source to get a narrow beam of light, whereas laser emits radiations only in one direction. Thus laser has high degree of beam directionality.

For a typical laser, the beam divergence is found to be less than one milliradian. It means that laser beam spreads less than one millimeter for every meter travelled by it. For example, the spread of laser beam sent from earth to moon (384400 km) is just few kilometers.

(b) Intensity: The light from a conventional source spreads out uniformly in all the directions.

At distance of 30 cm from a 100 watt bulb, power entering* the eye is less than $\frac{1}{1000}$ watt. On the other hand the laser gives out light into a narrow beam and its energy is concentrated in a small region. Therefore even a laser of 1 watt would appear thousand times more intense than 100 watt ordinary bulb.

*Light per cm^2 from a bulb of 100 W at 30 cm away

$$= \frac{100}{4\pi r^2} = \frac{100}{4\pi \times 30^2} = \frac{1}{36\pi} = \frac{1}{113} \text{ watt/cm}^2$$

If the area of the eye's lens is about $\left(\frac{1}{3}\right)^2 \text{ cm}^2$, then the power entering the eye is

$$\frac{1}{9 \times 113} = \frac{1}{1017} \text{ W} < \frac{1}{1000} \text{ W.}$$

(c) Monochromaticity: The light emitted by a laser is extraordinary monochromatic. Monochromaticity (single wavelength) of a laser light is much more than that of any conventional monochromatic source. Broadening of the emitted light from a source is the result of non-monochromaticity.

An inspection of a line emitted by an ordinary monochromatic source reveals that it has a spread (or width) over a frequency range of thousands of mega Hertz ($\sim 10^{10}$ Hz) whereas the spread (or width) of laser light is almost negligible (~ 500 Hz). This is shown in Fig. 4.5 for the monochromatic light of

$$\text{frequency } v_0 \text{ (say } \simeq 5 \times 10^{14} \text{ Hz} = \frac{3 \times 10^8}{\lambda} \text{ m.}$$

or

$$\lambda = \frac{3 \times 10^8}{5 \times 10^{14}} \text{ m} = 6000 \text{ \AA}$$

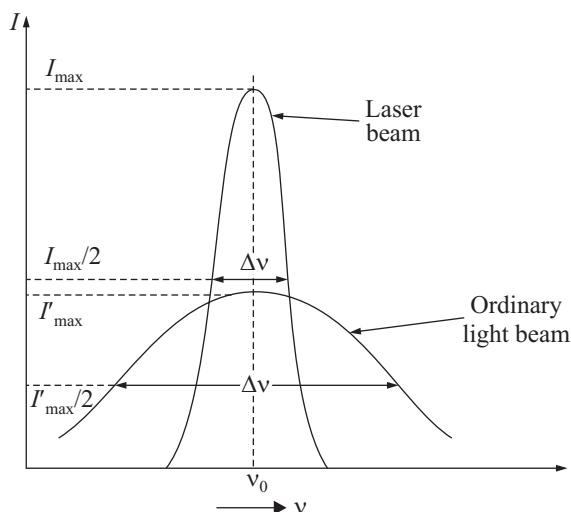


Fig. 4.5. Band widths for laser and ordinary light beam

The spread in laser beam is largely due to the presence of ‘spontaneous’ emission in addition to stimulated emission.

(d) Coherence: Coherence is a measure of the degree of phase correlation that exists in the radiation field of a light source at different locations and different times. Due to stimulated emission in laser the photons are emitted in the same phase and there is very high degree of coherence in laser as compared to the conventional coherent sources. The wave front of the light emitted by conventional monochromatic light source change from one point to the other and varies from instant to instant. Corresponding to these there are two concepts of coherence, namely spatial coherence (on lateral coherence) and temporal coherence (or longitudinal coherence).

(i) Spatial Coherence or Lateral Coherence: To understand spatial coherence consider two points P_1 and P_2 which, at time $t = 0$, lie on the same wave front of some given electromagnetic wave and let $E_1(t)$ and $E_2(t)$ be the corresponding electric fields at these points. According to definition the difference between the phases of the two fields at time $t = 0$ is zero. If this difference in phase remains zero at any time $t > 0$, we would say that there is perfect coherence between these two points *i.e.*, we will say that the wave has perfect spatial coherence. Thus if the electric fields at any two different

points on the *e.m.*, wave front has constant phase difference over any time *t*, then perfect coherence is said to occur.

The concept of spatial coherence can be understood by Young's double slit experiment as shown in Fig. 4.6. Here *S* is a point source, *S*₁ and *S*₂ are equally spaced pinholes. Now on the screen *T* near the central position *O* one would get interference fringes of good contrast. Let us now consider another point source *S'* placed near *S* and there is no phase relationship between the waves from *S* and *S'*. Now the interference pattern on the screen *T* will be due to superposition of intensity coming from sources *S* and *S'* through the pinholes *S*₁ and *S*₂. If *S'* is moved slowly away from *S*, the contrast of fringes becomes poorer. For a particular distance *SS'* such that $S'S_2 - S'S_1 = \frac{\lambda}{2}$, the interference maximum produced by *S* falls on the interference minimum produced by *S'* and vice versa. At this position of *S'* the interference fringes due to *S* will just disappear and uniform illumination on the screen is observed. From the geometry of the Fig. 4.6.

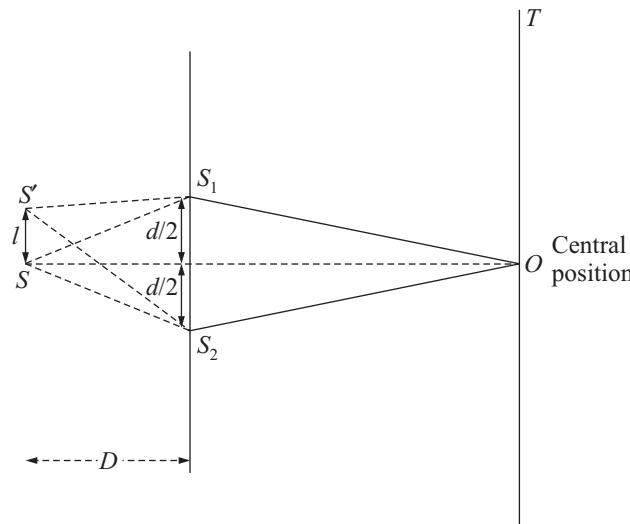


Fig. 4.6. Young's double slit experiment to study the spatial coherence of a light source

$$S'S_2 = \sqrt{D^2 + \left(\frac{d}{2} + l\right)^2} = D \left[1 + \frac{1}{2D^2} \left(\frac{d}{2} + l \right)^2 \right]$$

and

$$S'S_1 = \sqrt{D^2 + \left(\frac{d}{2} - l\right)^2} = D \left[1 + \frac{1}{2D^2} \left(\frac{d}{2} - l \right)^2 \right]$$

$$\therefore S'S_2 - S'S_1 = \frac{D}{2D^2} (dl + dl)$$

or

$$S'S_2 - S'S_1 = \frac{ld}{D} = \frac{\lambda}{2} \quad \dots(4.13)$$

where *l* = *SS'*, *d* is the distance between pinholes *S* and *S*₂ and *D* is the distance of the pinholes from the sources *S* and *S'*. Equation (4.13) gives

$$l = \frac{D\lambda}{2d} \quad \dots(4.14)$$

Therefore for an extended source of width l , fringe pattern will be observed if

$$l \ll \frac{\lambda D}{d} \quad \dots(4.15)$$

Equivalently, for an extended source (of width l) the interference pattern of good contrast will be formed if the separation of the two point sources, d is

$$d \ll \frac{\lambda D}{l} \quad \dots(4.16)$$

Now l/D is the angle subtended by the source at the slit as shown in Fig. 4.7 i.e.,

$$\theta = \frac{l}{D} \quad \dots(4.17)$$

Substituting the value of $\frac{l}{D}$ in Eqn. (4.16), we get

$$d \ll \frac{\lambda}{\theta} \quad \dots(4.18)$$

In practice the pinhole S_1 is fixed and pinhole S_2 is moved to see the reduction in fringe contrast. The area over which the pinhole S_2 shows the interference fringes is called the 'Coherent Area' of light wave. The distance d_t between the pinholes for which the fringes just disappear is called transverse, lateral or spatial coherence length, Thus

$$\text{spatial coherence length} = d_t = \frac{\lambda D}{l} = \frac{\lambda}{\theta}. \quad \dots(4.19)$$

(ii) Temporal Coherence or Longitudinal Coherence: Temporal coherence refers to the correlation between electric fields at a point at two different times.

To understand temporal coherence, let us consider electric field of the electromagnetic wave at a given point P at time t and $t + \tau$. If, for a given time delay τ , the phase difference between the two field values remains the same for anytime t ($t < \tau$), we say that there is 'Temporal Coherence' over time τ . If this occurs for any value of τ , the e.m. wave is said to have perfect time coherence. If this occurs for a time delay τ such that $0 < \tau < \tau_c$, the e.m. wave is said to have partial time coherence with coherence time equal to τ_c as shown in Fig. 4.8, which shows a sinusoidal electric field undergoing phase jumps at time interval equal to τ_c .

The distance travelled by the wave train during time τ_c is called temporal or longitudinal coherence length i.e., longitudinal coherence length $= l = C \tau_c \quad \dots(4.20)$

where C is the speed of light. In fact the band width $\Delta\nu$ is related to the coherence time τ_c of the e.m. wave by the relation

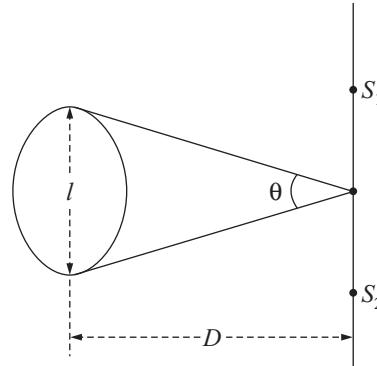


Fig. 4.7

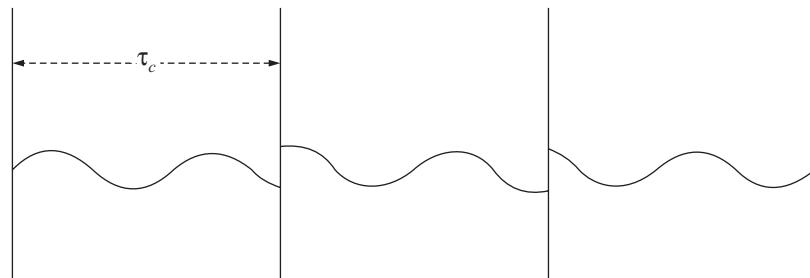


Fig. 4.8. Example of an e.m. wave with a coherence time τ_c

$$\Delta\nu = \frac{1}{\tau_c} \quad \dots(4.21)$$

Thus the concept of temporal coherence is directly connected to monochromaticity.

Temporal coherence can be studied by using Michelson interferometer. For mercury vapour lamp and He-Ne (Helium-Neon) laser, the typical temporal or longitudinal coherence lengths are 0.50 m and 300 m, respectively.

4.9 HELIUM-NEON LASER

The helium-neon (He-Ne) laser is the first gas laser operated successfully in 1961, by Ali Javan and his co-workers at Bell Telephone Laboratories in U.S.A. It consists of the following main components.

1. Active Material: Mixture of helium and neon gases in the ratio 7 : 1 at total pressure of 1 torr (1 torr = 1 mm of Hg column) is used as an active material. However, this ratio 5 : 1 to 10 : 1 may be taken. In the mixture helium is used for population inversion only whereas levels of neon are involved in laser action.

2. Resonator Cavity: Resonator cavity is made of quartz tube of about 80 cm length and 1 cm diameter. The gas mixture is enclosed between a pair of mirrors. One mirror is completely reflecting (99.99% reflectivity) and the other mirror called the output mirror is partially transmitting (about 90% reflectivity).

3. Exciting Source: Exciting source for creating discharge in the tube is the high voltage radio frequency such as Tesla coil.

As is shown below its operation involves four energy levels—three of neon and one of helium.

Operation: The working of the He-Ne laser is based on the fact that neon has excitation energy levels very close to meta-stable energy levels of helium. Few energy levels of He and Ne excited states are shown in Fig. 4.10.

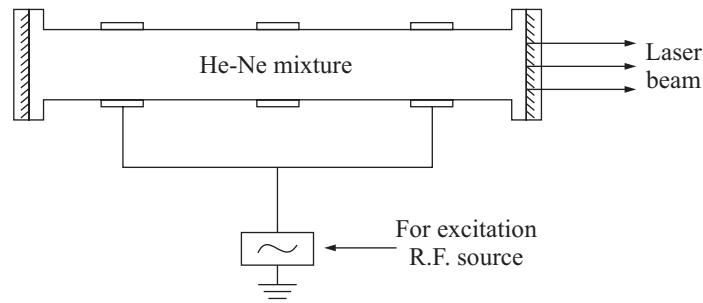


Fig. 4.9. He-Ne laser

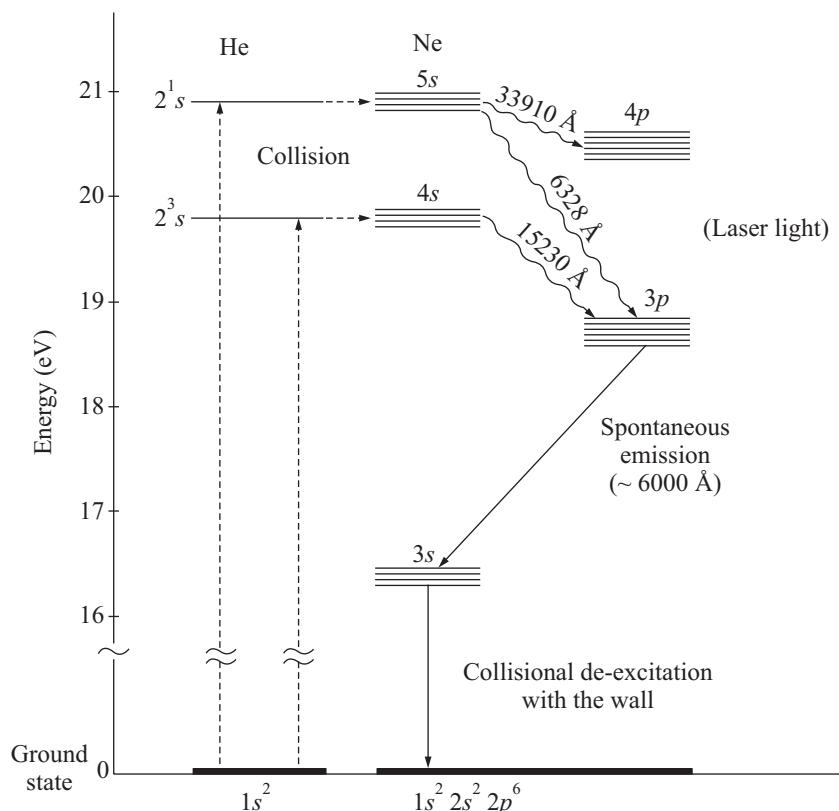


Fig. 4.10. Few energy levels of He and Ne excited states

Ground state of helium is $1s^2$. In the first excited state one of the electron goes to $2s$ level or the excited states are 2^3s (triplet) and 2^1s (singlet). Both the states (2^3s and 2^1s) of He are metastable states because $2^3s \rightarrow 1^1s$ is forbidden due to change in multiplicity or flip in the spin and $2^1s \rightarrow 1^1s$ transition is also forbidden by electric dipole transition. Ground state of Ne is $1s^2 2s^2 2p^6$ and the first few excited states are $[1s^2 2s^2 2p^5] 3s^1, 4s^1, 5s^1, 3p^1$ or $4p^1$. When electric discharge is passed through the gas the electrons collide with the helium and neon atoms and excite them to the higher levels of both helium and neon. The helium excited states (2^3s and 2^1s) being metastable states accumulate population whereas neon has radiative decay and they come to ground state. These excited He atoms collide with Ne atoms

and transfer the energy to excite Ne to $4s$ and $5s$ levels which have nearly the same energy as 2^3s and 2^1s levels, respectively, of He atoms. Thus there is resonant energy transfer. Since significant population can be built up in $4s$ and $5s$ levels, they prove suitable candidates as upper levels for laser transition. Taking account of the selection rules possible transitions are those to p -states. Transition from $5s$ to $4p$, $4s$ to $3p$ and $5s$ to $3p$ result in the emission of radiation having wavelengths 33910 \AA , 15230 \AA and 6328 \AA , respectively. Out of these 6328 \AA corresponds to visible light of He-Ne laser. The terminal $3p$ level decays radiatively with a life time of 10^{-8} sec to the $3s$ level of long life time. Because of long life of $3s$ state, atoms in this state tend to collect with time. These atoms after collision with electrons in the discharge may be excited back to $3p$ state which thus reduces the inversion and can even quench it. This situation may be avoided by taking tube of small diameter and then neon atoms in $3s$ state are de-excited by colliding with the wall of the tube.

Typical power output in He-Ne lasers lie between 1 and 50 mW of continuous wave for input of about 5 to 10 W. The output shows a strong dependence on the discharge current. The power supply is thus matched with the tube.

4.10 SEMICONDUCTOR LASER

The semiconductor laser is also sometimes called as the junction laser or the diode laser. These lasers use semiconductors as the lasing material. A number of semiconductors can be used for $p-n$ junction. When this junction is forward biased then the electrons drift from n -region into the p -region and the holes drift from the p -region into the n -region. They soon recombine near the junction and in this process they release* excess energy in the form of heat or light. In the case of silicon and germanium this recombination energy is released in the form of heat so these materials are of no use for laser action. In gallium arsenide most of this recombination energy is released in the form of light, therefore gallium-arsenide is usually used in semiconductor lasers.

First doped semiconductor was fabricated in 1962, by using gallium-arsenide in the form of diffused $p-n$ junction. $p-n$ junction is formed by diffusing aluminium (acceptor impurity) to wafers of n -type Ga-As. Just like in other laser systems, there can be three interaction processes.

(i) An electron in the valence band can absorb the incident radiation and can be excited to the conduction band. This process would create electron-hole pair.

* If the energy gap or band gap is equal to 1 e.v., then the wavelength of the emitted radiation due to pair annihilation is

$$\begin{aligned}\lambda &= \frac{hc}{E_g} \text{ because } E_g = h\nu = \frac{hc}{\lambda} \\ &= \frac{6.63 \times 10^{-34} \times 3 \times 10^8}{1 \times 1.6 \times 10^{-19}} \text{ m} = 12400 \text{ \AA}\end{aligned}$$

These radiations are in the far infrared region. If the band gap is equal to 1.6 e.v. (say), then λ of emitted radiation is

$$\lambda = \frac{6.63 \times 10^{-34} \times 3 \times 10^8}{1.6 \times 1.6 \times 10^{-19}} \text{ m} \approx 7700 \text{ \AA}, \text{ which is in the visible region.}$$

(ii) An electron from the conduction band can jump spontaneously to valence band and in this process it emits radiation.

(iii) Incident radiation may induce electrons of the conduction band to jump into the valence band. In this process it emits radiation.

If now by some mechanism a large density of electrons is created in the bottom of the conduction band and simultaneously in the same region of space a large density of holes is created at the top of the valence band then an optical beam with a frequency slightly greater than E_g/h will cause a larger number of stimulated emission as compared to absorptions and thus can be amplified. This can be achieved by cutting wafers, the Al diffused *n*-type Ga-Al into small chips with optically flat and polished parallel ends perpendicular to the plane of junction. These surfaces are coated to increase the reflectivity. As shown in Fig. 4.11 the remaining faces are made rough to avoid the leakage of laser beam. At low currents it acts as LED (light emitting diode). When current reaches the threshold value, a population inversion is achieved between filled electron levels near the bottom of conduction band. It is found that when the forward current density is more than the threshold current density ($\sim 50,000$ Amp/cm 2) the stimulated emission rate will exceed the absorption rate and the amplification will overcome the losses in the cavity and laser will begin to emit coherent radiation. Taking the typical cavity length 300 $\mu\text{m} \times 100 \mu\text{m}$ the required threshold current is 15 Amp.

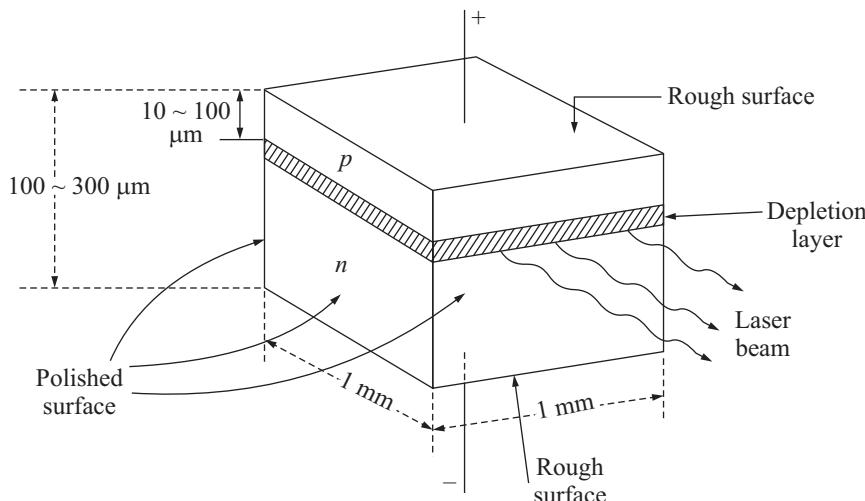


Fig. 4.11. Semiconductor laser

Now-a-days semiconductor lasers are based on the double heterojunction in which thin active layer of a semiconductor with a narrow band gap is sandwiched between two larger band gap semiconductors. Fortunately the refractive index of semiconductor decreases with an increase in bandgap. Thus the refractive index of the central active layer of GaAs is higher than the surrounding region. This change in refractive index helps in the confinement of the emitted optical radiation to the active region because of total internal reflections at the boundaries. The double heterojunction used here also reduces the value of the threshold current density to 2000–4000 Amp/cm 2 or the current required is only about 60 mA because the current is restricted within small lateral size of the active region.

4.11 APPLICATIONS OF LASER

Lasers are widely used for the following purposes:

1. Measurement of distances: With the help of laser beam, distances can be accurately measured such as the distance of moon from earth.
2. Communication: Lasers play the essential role in using thin strands of glass fibres to transmit light signals that can be received and translated into communication format.
3. Material Processing: Lasers are used in cutting, welding and drilling with more accuracy and precision in even very hard materials because of their very high intensity.
4. Medical Science: Lasers are used to burn up brain tumour and to remove tattoos. Laser welding is also used reconnect blood vessel. Eye surgeons also use laser for ophthalmic operations like treatment of cataract, glaucoma and diabetic retinopathy.
5. Applications in Physics and Chemistry: Laser can be used to study the non-linear optics with special mention of harmonic generation and stimulated scattering. Lasers are also used for producing irreversible chemical change *i.e.*, laser photochemistry.
6. Lasers are used in printing technology.
7. Lasers are used effectively and efficiently in holography to analyse holograms with more precision.
8. Laser beam being of very high energy, has been mentioned as potential “death ray” type of incendiary weapon for use against energy missiles.

SOLVED EXAMPLES

Example 1: Calculate the coherence length of a laser beam for which the band width $\Delta v = 3000 \text{ Hz}$. Speed of light $c = 3 \times 10^8 \text{ m/sec}$. (M.D.U., 2001, 2003, 2006)

Solution: Given band width $\Delta v = 3000 \text{ Hz}$

$$\therefore \text{Coherence time } \tau_c = \frac{1}{\Delta v} \quad (\text{See Eqn. 4.21})$$

$$= \frac{1}{3000} = \frac{10^{-3}}{3} \text{ sec.}$$

$$\begin{aligned} \therefore \text{Coherence length } l &= c\tau_c \\ &= 3 \times 10^8 \times \frac{10^{-3}}{3} \text{ m} \\ &= 10^5 \text{ m} = 100 \text{ km.} \end{aligned}$$

Example 2: For an ordinary source, the coherence time $\tau_c = 10^{-10} \text{ second}$. Obtain the degree of non-monochromaticity for wavelength $\lambda_o = 5400 \text{ \AA}$. (M.D.U., 2002)

Solution: Coherence time $\tau_c = 10^{-10} \text{ sec.}$

$$\therefore \text{Band width } \Delta v = \frac{1}{\tau_c} = \frac{1}{10^{-10}} = 10^{10} \text{ Hz}$$

For wavelength $\lambda_0 = 5400 \text{ \AA} = 5400 \times 10^{-10} \text{ m}$

$$\text{Frequency } v_0 = \frac{3 \times 10^8}{5400 \times 10^{-10}} \text{ Hz} = \frac{1 \times 10^{16}}{18} \text{ Hz}$$

Degree of non-monochromaticity is given by

$$\frac{\Delta\nu}{v_0} = \frac{10^{10}}{\frac{1}{18} \times 10^{16}} = 18 \times 10^{-6} = 1.8 \times 10^{-5}.$$

QUESTIONS

1. Explain the terms: Spontaneous and stimulated emission, population inversion, optical pumping. (M.D.U., 2009)
2. Discuss the essential requirements for producing laser action. Describe a He-Ne laser. (M.D.U., 2009)
3. Describe the principle, construction and working of He-Ne gas laser. (M.D.U., 2008)
4. Discuss Einstein's coefficients. Derive relation between them. (M.D.U., 2008)
5. What are the specialities of laser light? Give description of semiconductor laser. (M.D.U.; Dec. 2009)
6. Write a note on He-Ne laser. (M.D.U.; Dec. 2008)
7. Describe the construction and working of a semiconductor laser. (M.D.U., 2007)
8. Discuss the characteristic features of a laser beam. (M.D.U., 2007)
9. What do you mean by lasers? Describe laser action. What are the characteristics of a laser beam? (K.U., B.T., 2007)
10. Discuss the salient characteristics of a laser beam. (K.U.; B.T., 2005; N.I.T.K.U., 2007)

PROBLEMS

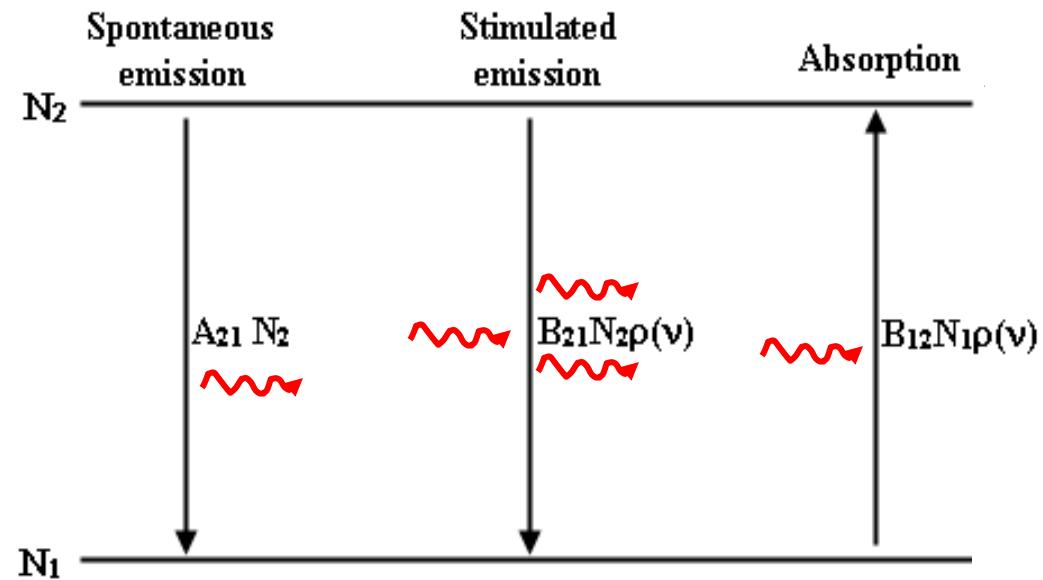
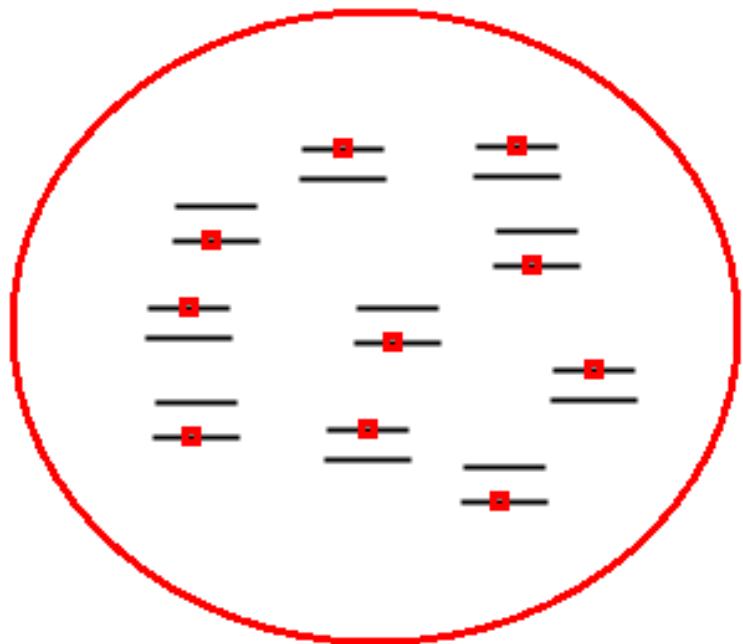
1. A laser beam has band width of 1200 Hz. Obtain its coherence length. [Ans. 250 km] (M.D.U., Dec. 2006)
2. For an ordinary source, the coherence time $\tau_c = 10^{-10} \text{ s}$. Calculate the degree of non-monochromaticity for wavelength $\lambda_0 = 6000 \text{ \AA}$. [Ans. } 2 \times 10^{-5}]

Relation between Einstein's A and B Coefficient

Relation between Einstein's A and B Coefficient

- In thermal equilibrium at temperature T , with radiation frequency n and energy density Q . Let N_1 and N_2 be the number of atoms in energy states 1 and 2 respectively at any instant. The number of atoms in state 1 absorb a photon and give rise to absorption per unit time
- Under thermal equilibrium transitions from E_1 to E_2 should be equal to transitions from E_2 to E_1
- No. of atoms absorbing photons per unit volume = the number of atoms emitting photons per second

- Einstein's A and B coefficients



- Under thermal equilibrium transitions from E_1 to E_2 should be equal to transitions from E_2 to E_1
No. of atoms absorbing photons per unit volume = the number of atoms emitting photons per second

$$B_{12}\rho(v)N_1 = A_{21}N_2 + B_{21}\rho(v)N_2$$

04_Einstein_Coefficient_Reference_Material_II_11-Feb-2021

$$B_{12}\rho(v)N_1 - B_{21}\rho(v)N_2 = A_{21}N_2$$

$$\rho(v)[B_{12}N_1 - B_{21}N_2] = A_{21}N_2$$

$$\rho(v) = A_{21}N_2 / [B_{12}N_1 - B_{21}N_2]$$

Taking $B_{21}N_2$ in the denominator as common

$$\rho(v) = A_{21}N_2 / B_{21}N_2 \{ 1 / [(B_{12}N_1 / B_{21}N_2) - 1] \}$$

$$\rho(v) = (A_{21} / B_{21}) \{ 1 / [(B_{12}N_1 / B_{21}N_2) - 1] \}$$

It was proved thermodynamically that $B_{12} = B_{21}$

Then

$$\rho(v) = (A_{21} / B_{21}) \{ 1 / [(N_1 / N_2) - 1] \}$$

Distribution of atoms under equilibrium among diff. states is given by Boltzmann's law according to which

$$\frac{N_2}{N_1} = \frac{e^{-E_2 / kT}}{e^{-E_1 / kT}} \quad \frac{N_2}{N_1} = e^{-(E_2 - E_1) / kT} \quad \frac{N_2}{N_1} = e^{-(hv) / kT}$$

Therefore

$$\rho(v) = (A_{21}/B_{21}) [1 / (e^{hv/kT} - 1)]$$

The above equation is the energy density of photon of frequency v in equilibrium with atoms in energy states 1 and 2 at temperature T comparing with Plank's radiation formula

$$\rho(v) = (8\pi h v^3/c^3) [1 / (e^{hv/kT} - 1)]$$

We get

$$A_{21}/B_{21} = 8\pi h v^3/c^3$$

The above equation gives the relation between Einstein's A and B coefficients

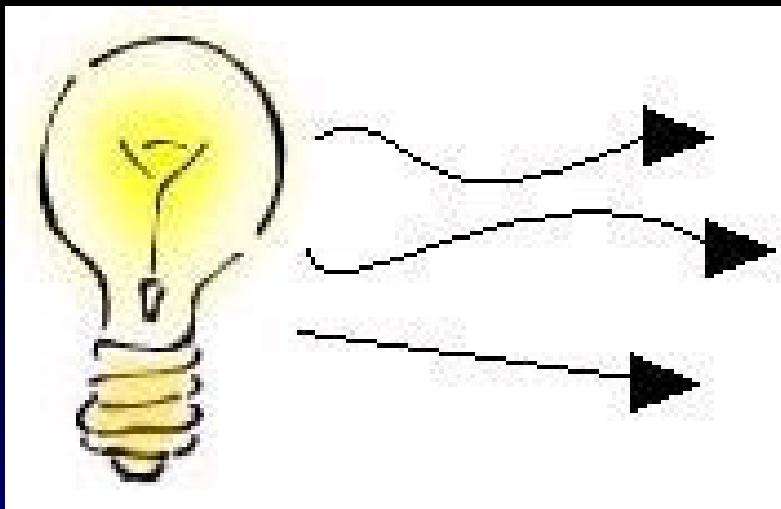
Significance

- $B_{12} = B_{21}$, The probability of spontaneous emission is same as that of induced absorption. This means that if these two processes will occur at equal rates, so that no population inversion can be attained in a two-level system.
- The ratio of spontaneous emission and stimulated emission is proportional to v^3 . This implies that the probability of spontaneous emission dominates over induced emission more and more as the energy difference between the two states increases.

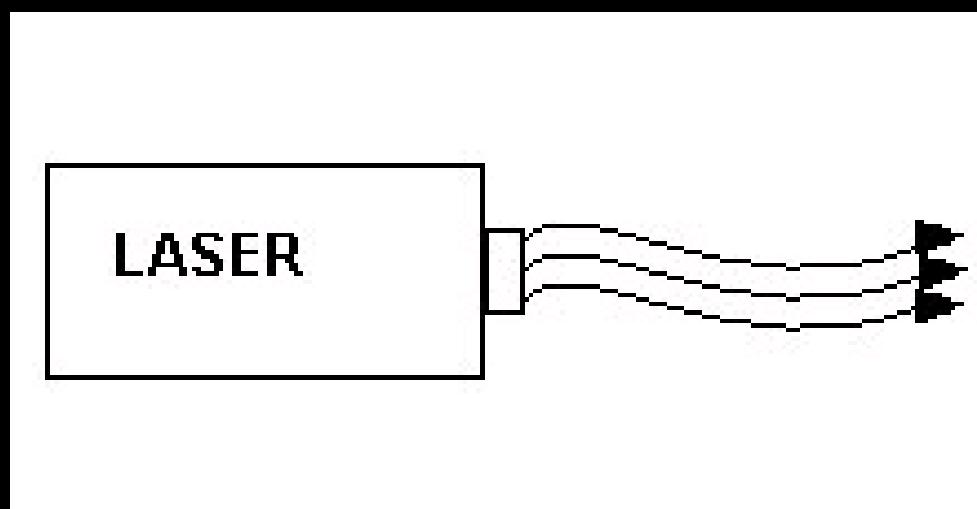
LASER

- **LIGHT AMPLIFICATION by STIMULATED EMISSION of RADIATION**
- Characteristics
 - Monochromaticity

Incandescent vs. Laser Light

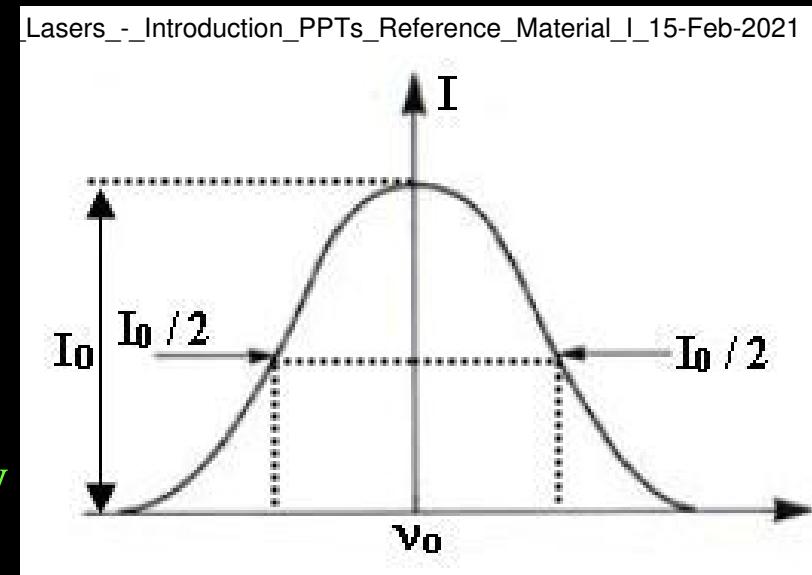


1. Many wavelengths
2. Multidirectional
3. Incoherent

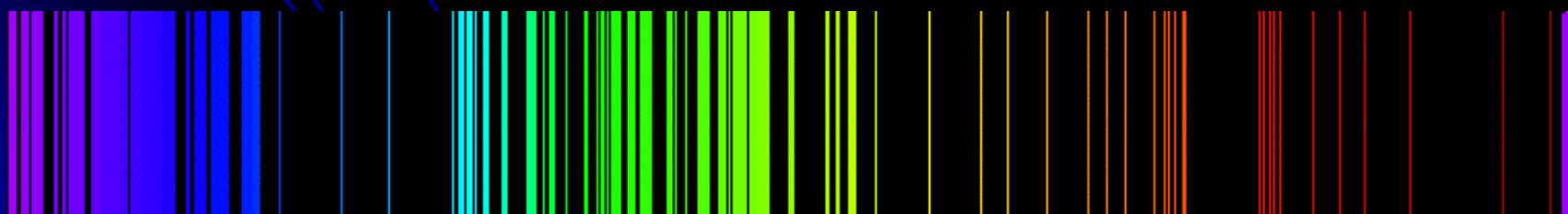


1. Monochromatic
2. Directional
3. Coherent

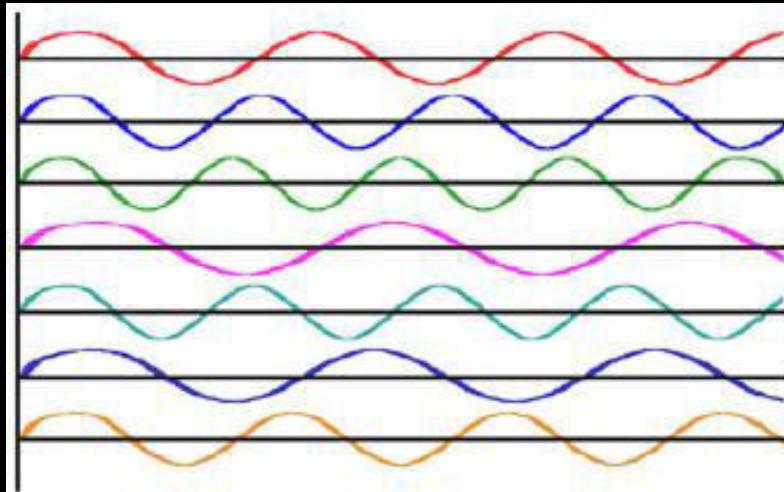
- Monochromatic light – having one frequency of oscillation
- Practically not possible
- Consist a band of frequencies closely spaced around a central frequency v_0



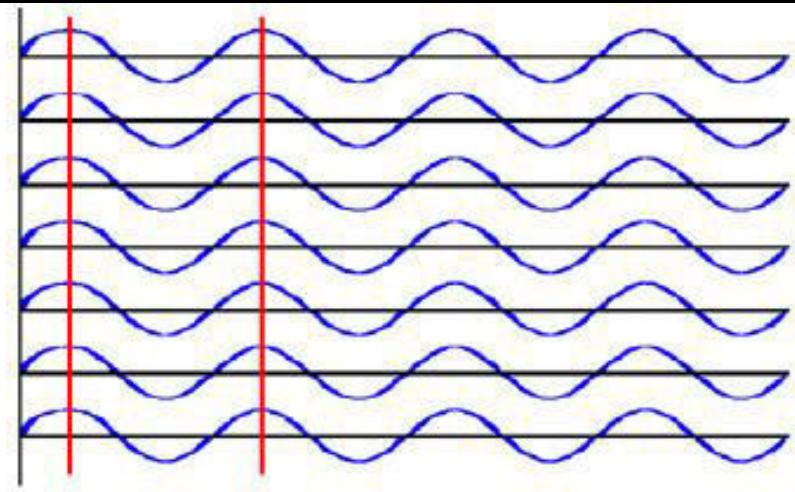
- Band frequency Δv is called line width or band width
- Light from conventional source – line width 10^{10} Hz or higher
- Laser source has a line width of 100 Hz.



- Higher degree of coherence



Incoherent light waves



Coherent light waves

- Coherent if in phase – maintain crest to crest and trough to trough correspondence
- Things necessary for light waves to be coherent
 - Start at same point with same phase
 - Wavelengths must be the same
- Laser light is a resultant of number of identical photons which are in phase, hence exhibiting higher degree of coherence

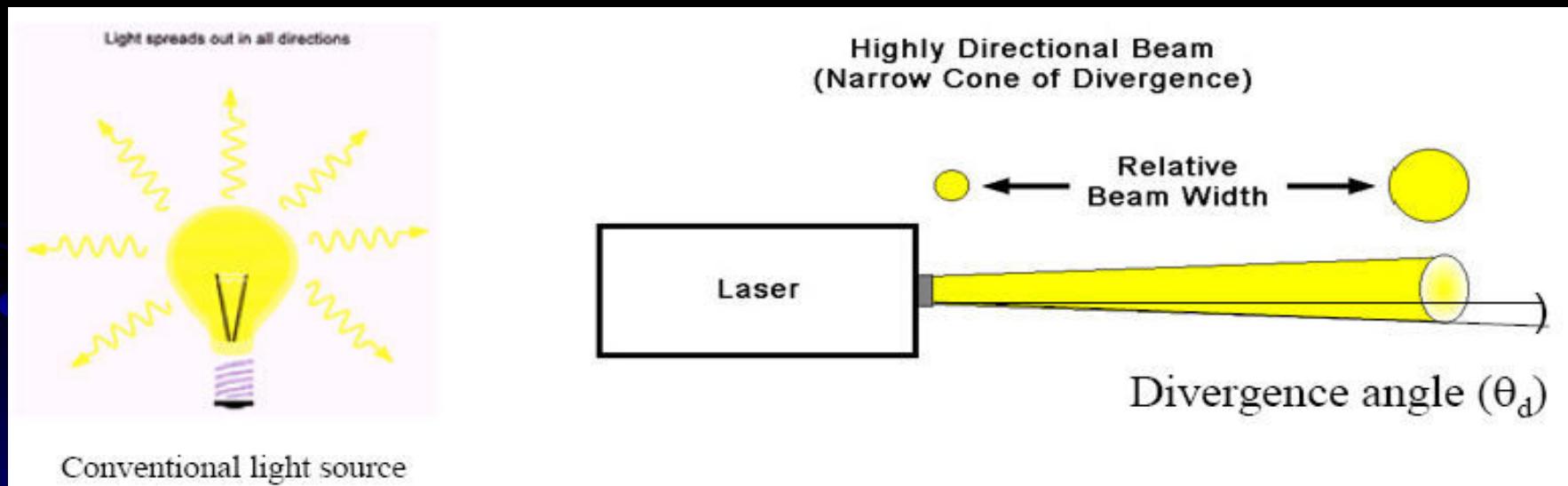
- **High intensity**

- Power o/p varies from few mW to few kW.
- This o/p power is concentrated by a very small CS
- Intensity of laser beam is given by

$$I = (10 / \lambda)^2 P$$

P is the power radiated by the laser

- **Directionality**

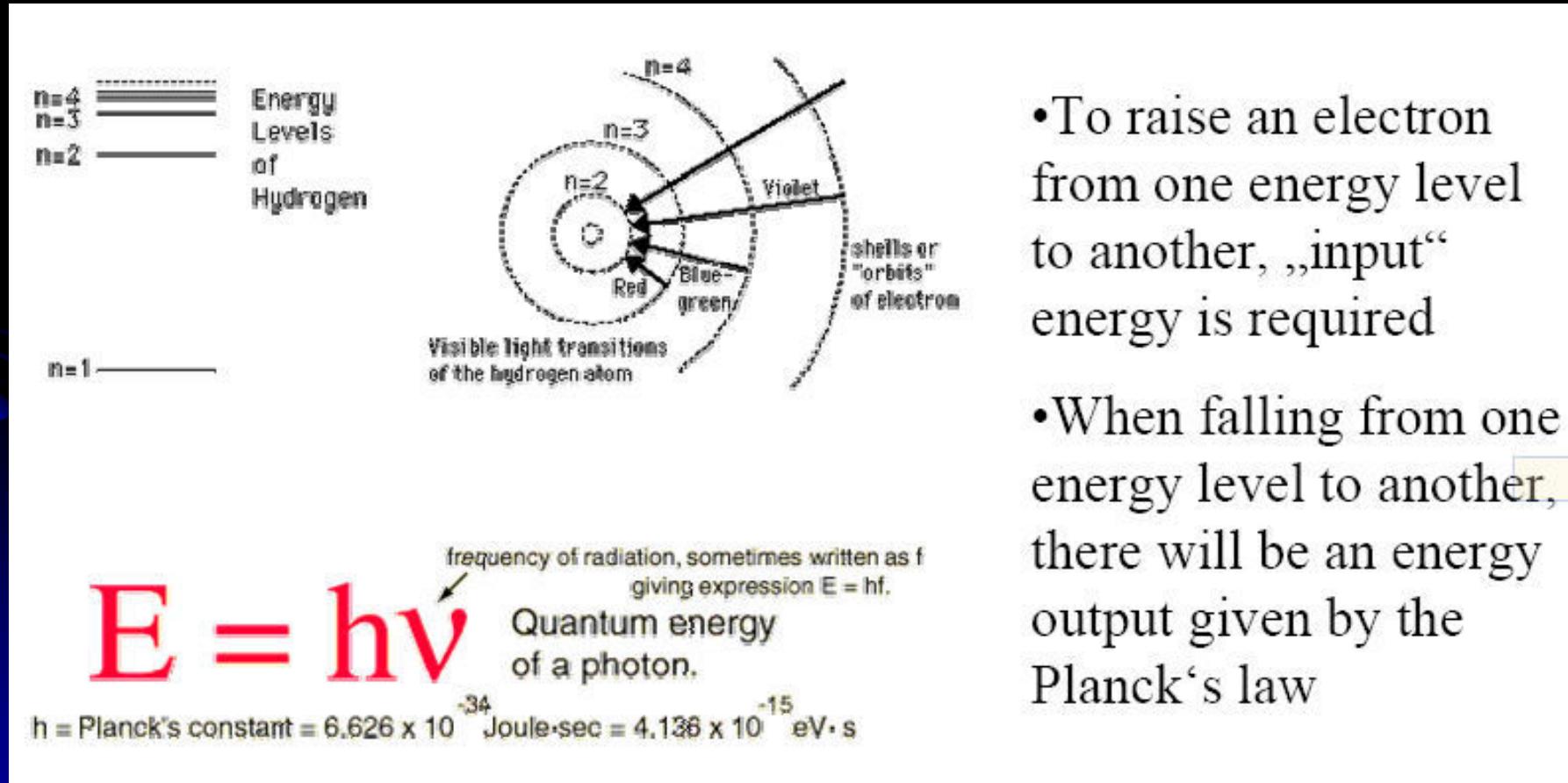


- **Light from laser diverges very little**

- Light beam can travel as a parallel beam up to a distance of d^2 / λ

d – diameter of the aperture, λ wavelength of light

- After the distance d^2 / λ light beam spreads radially
- Ordinary light beam angular spread is given by $\Delta\theta = \lambda / d$
- For laser angular spread is $1\text{mm} / \text{m}$ (ordinary light $1\text{m} / \text{m}$)
- Einstein's Prediction



- An excited atom tend to return randomly to ground state.
- As gnd. State population is high, more atoms are excited and a state will reach where all the electrons from the gnd. state are excited – violation of thermal equilibrium.
- Hence Einstein suggested there could be some emission mechanism by which the excited atom comes to the gnd. State.
- He predicted that photons in the light field induce the excited atom to fall to lower state and give up their excess energy in the form of photons. He called this type of second emission as stimulated emission.

- Almost all electronic transitions that occur in atoms that involve photons fall into one of **three categories**:
- Stimulated absorption or absorption**

- Transition from E_1 to E_2 due to the absorption of energy from the incident photon is called as stimulated absorption.
- The process can be represented as
atom + Photon \rightarrow atom*

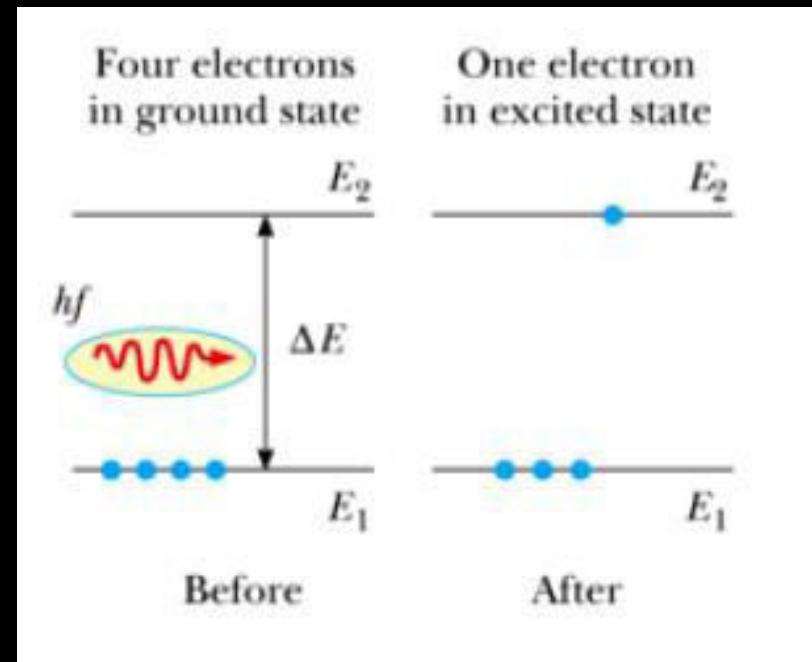
No. of absorption transition occurring in the Material will be α to the population in the lower level and number of photons / unit Volume in the incident beam.

- The rate of absorption may be expressed as

$$R_{\text{abs}} = B_{12} \rho(v) N_1 = B_{12} Q N_1 \quad \text{where } Q = \rho(v)$$

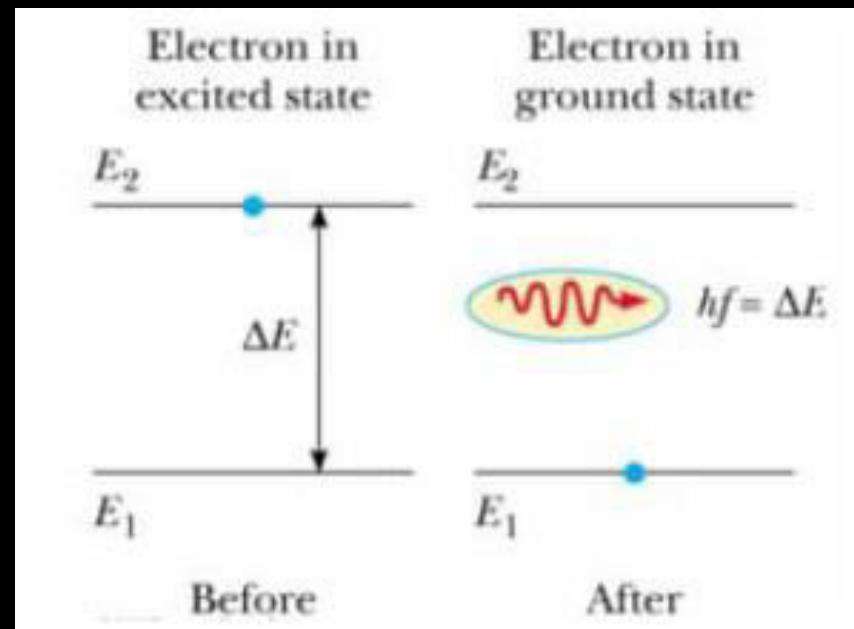
$\rho(v)$ – energy density of incident light

B_{12} – Einstein's coefficient for absorption



● Spontaneous emission

- Excited atom stays in the excited state for a period of 10^{-8} s.
- If not stimulated undergoes transition to gnd. State of its own emitting energy in the form of photons
- Known as spontaneous emission
- The process is represented as
 $\text{atom}^* \rightarrow \text{atom} + \text{photon}$



No. of photons generated is \propto to population in the excited level

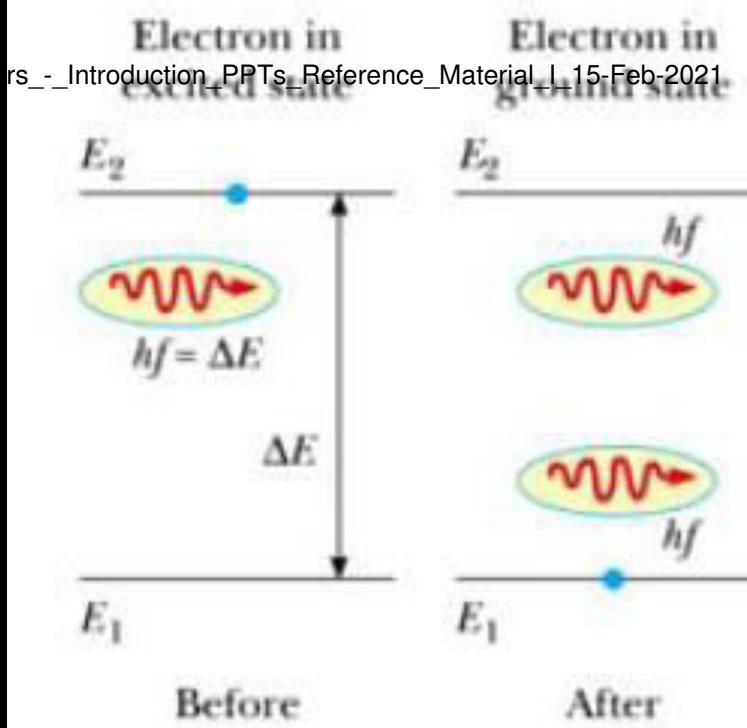
$$R_{SP} = A_{21} N_2$$

A_{21} – Einstein co efficient for spontaneous emission

Thus spontaneous emission is independent of energy density

- Stimulated emission

- Atoms in the excited state is stimulated by the photon with appropriate energy then transition to gnd. Level occurs with emission of additional photon.
- Hence this is called as stimulated emission.



- The process may be expressed as

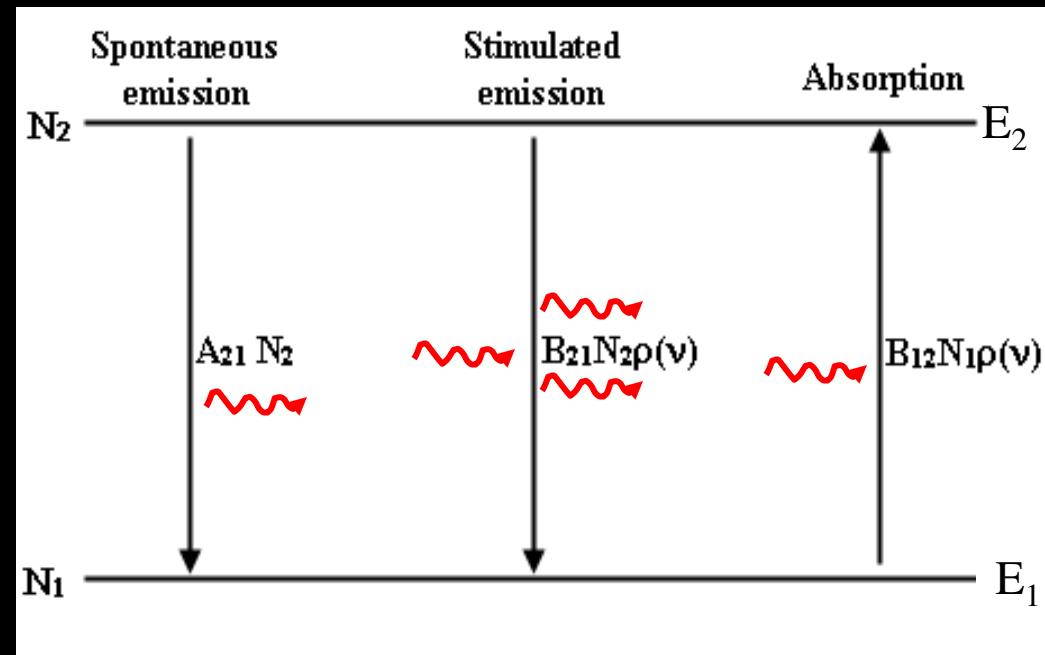
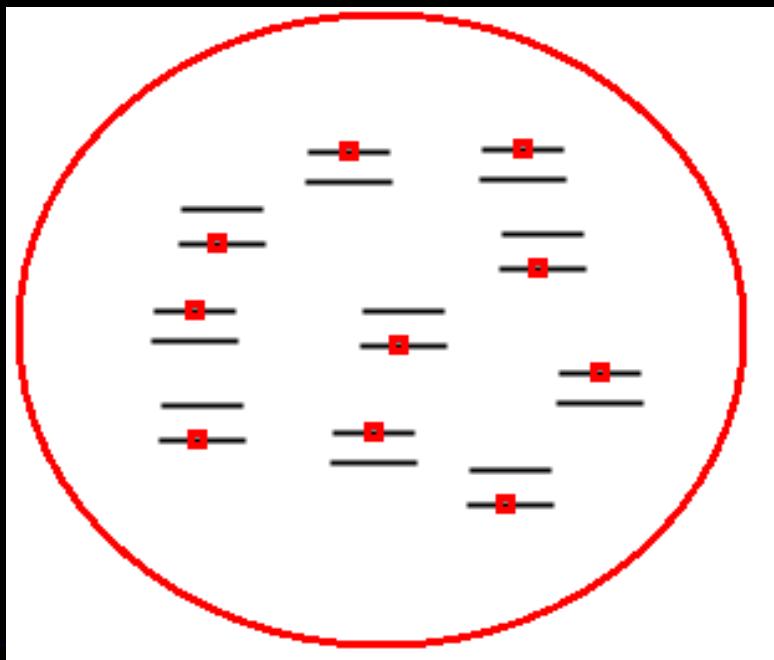
atom* + Photon = atom + 2 Photons

The rate of stimulated emission of photons is given by

$$R_{St} = B_{21} \rho(v) N_2 = B_{21} Q N_2 \text{ where } Q = \rho(v)$$

B_{21} – Einstein coefficient of stimulated emission

- Einstein's A and B coefficients



- Under thermal equilibrium transitions from E_1 to E_2 should be equal to transitions from E_2 to E_1

No. of atoms absorbing photons per unit volume =
the number of atoms emitting photons per second

$$B_{12}\rho(v)N_1 = A_{21}N_2 + B_{21}\rho(v)N_2$$

$$B_{12}\rho(v)N_1 - B_{21}\rho(v)N_2 = A_{21}N_2$$

$$\rho(v)[B_{12}N_1 - B_{21}N_2] = A_{21}N_2$$

$$\rho(v) = A_{21}N_2 / [B_{12}N_1 - B_{21}N_2]$$

Taking $B_{21}N_2$ in the denominator as common

$$\rho(v) = A_{21}N_2 / B_{21}N_2 \{ 1 / [(B_{12}N_1 / B_{21}N_2) - 1] \}$$

$$\rho(v) = (A_{21} / B_{21}) \{ 1 / [(B_{12}N_1 / B_{21}N_2) - 1] \}$$

It was proved thermodynamically that $B_{12} = B_{21}$

Then

$$\rho(v) = (A_{21} / B_{21}) \{ 1 / [(N_1 / N_2) - 1] \}$$

Distribution of atoms under equilibrium among diff. states is given by Boltzmann's law according to which

$$\frac{N_2}{N_1} = \frac{e^{-E_2 / kT}}{e^{-E_1 / kT}}$$

$$\frac{N_2}{N_1} = e^{-(E_2 - E_1) / kT}$$

$$\frac{N_2}{N_1} = e^{-(hv) / kT}$$

Therefore

$$\rho(v) = (A_{21}/B_{21}) [1 / (e^{hv/kT} - 1)]$$

The above equation is the energy density of photon of frequency v in equilibrium with atoms in energy states 1 and 2 at temperature T comparing with Plank's radiation formula

$$\rho(v) = (8\pi hv^3/c^3) [1 / (e^{hv/kT} - 1)]$$

We get

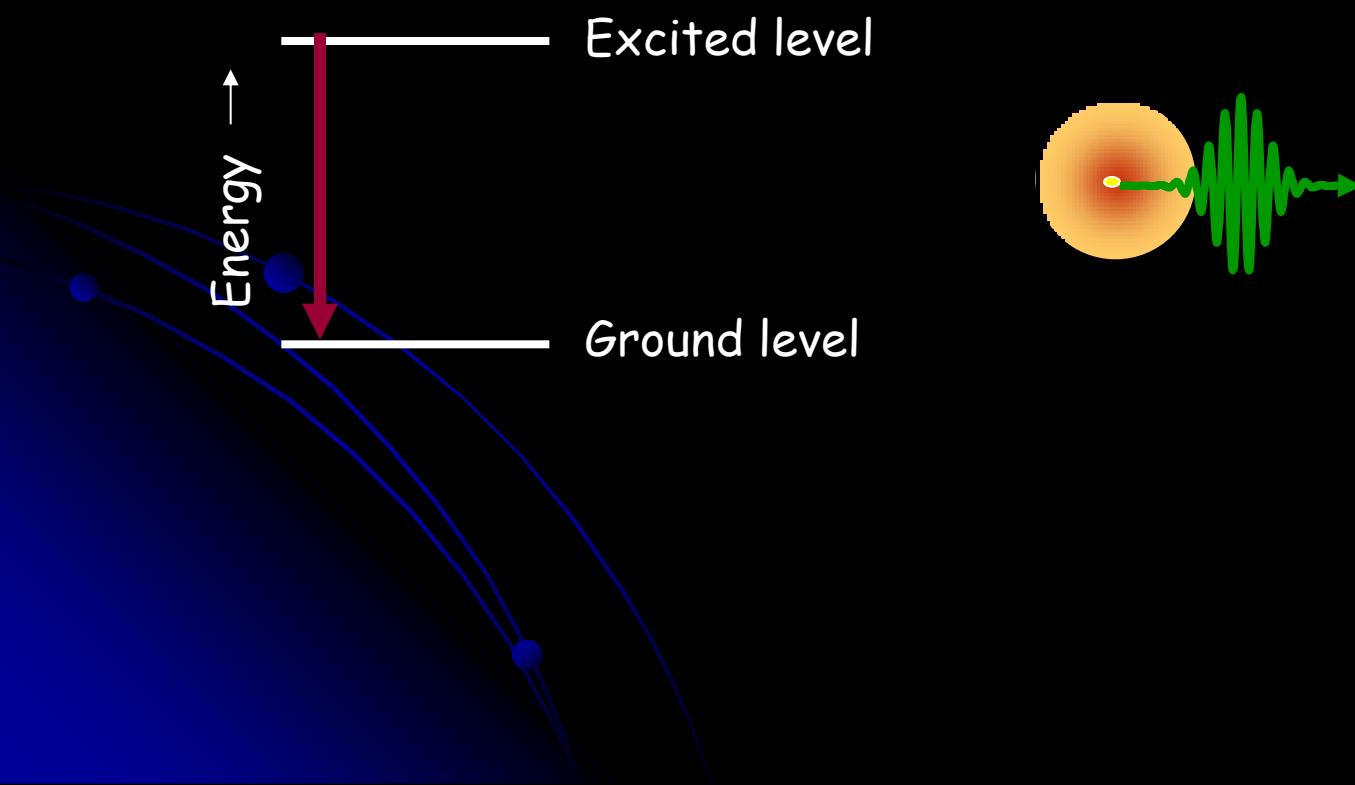
$$A_{21}/B_{21} = 8\pi hv^3/c^3$$

The above equation gives the relation between Einstein's A and B coefficients

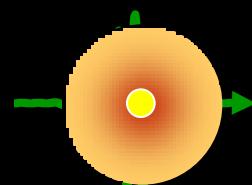
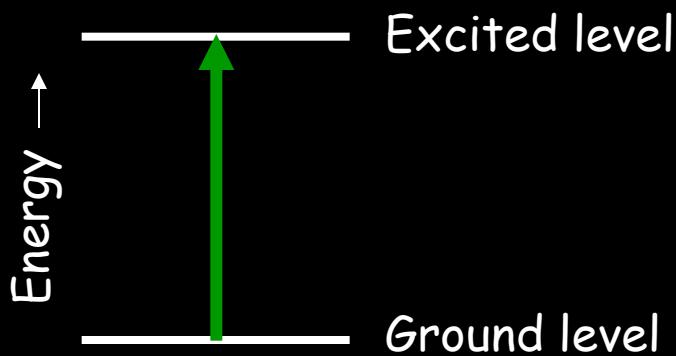
- Significance of Einstein's coefficients
 - Coefficients A_{21} , B_{21} and B_{12} are interrelated and can be calculated if one is known
 - Stimulated emission and absorption coefficients are equal at least for non-degenerate energy states
 - Since B_{21}/A_{21} is proportional to reciprocal of cube of ν , higher the frequency smaller the B_{21}

Excited atoms emit photons spontaneously.

When an atom in an excited state falls to a lower energy level, it emits a photon of light.

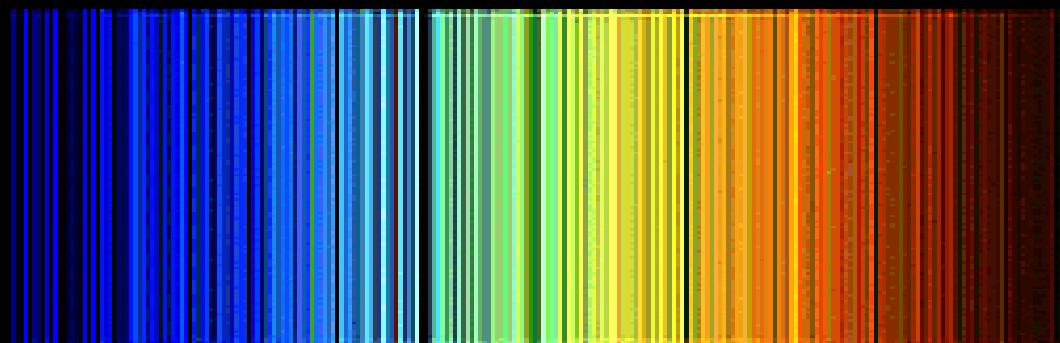


Atoms and molecules can also absorb photons, making a transition from a lower level to a more excited one.



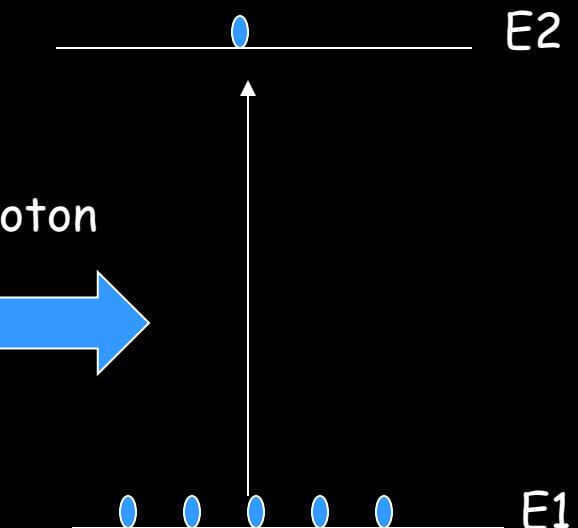
This is, of course, absorption.

Absorption lines in an otherwise continuous light spectrum due to a cold atomic gas in front of a hot source.



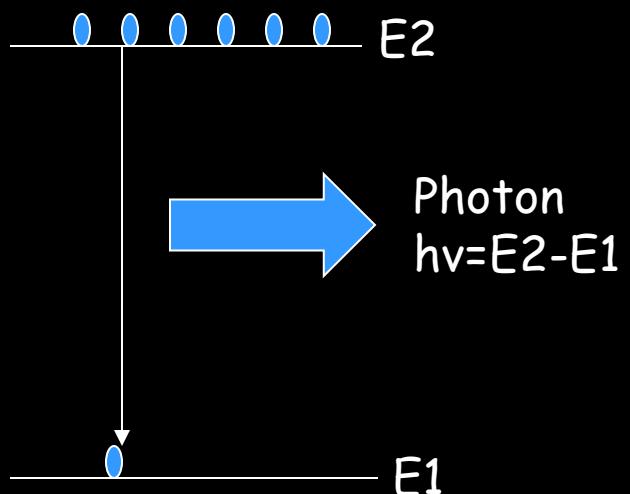
Spontaneous absorption

- Let us consider two energy level having energy E_1 & E_2 resp.
- The atom will remain in ground state unless some external stimulant is applied to it.
- When an EM wave i.e photon of particular freq fall on it , there is finite probability that atom will jump form energy state E_1 to E_2 .



Spontaneous emission

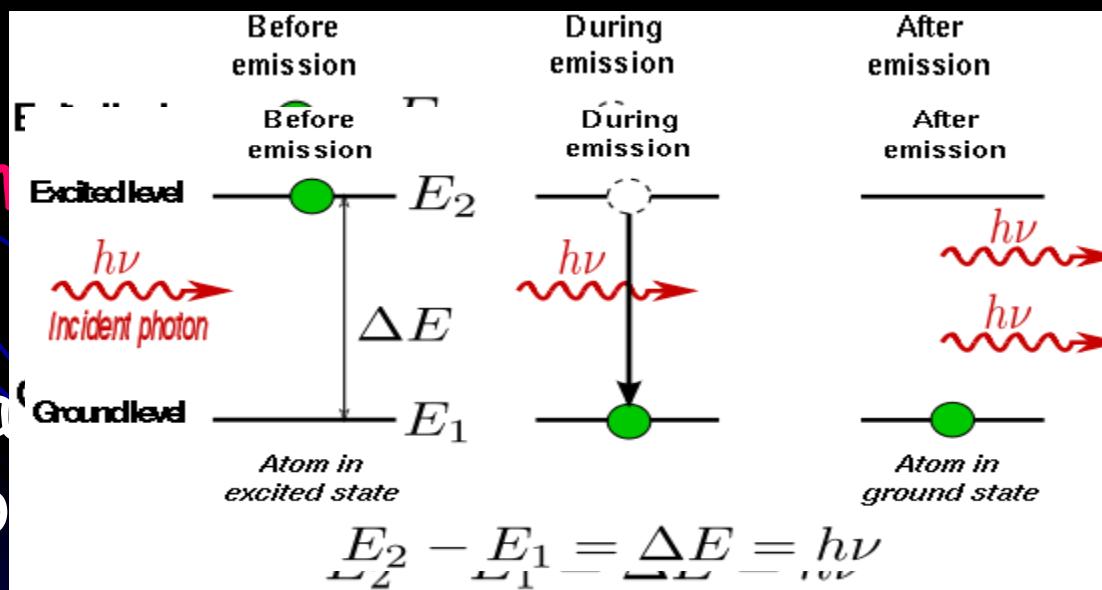
- Consider an atom in higher state (E_2).
- It can decay to lower energy level by emitting photon.
- Emitted photon have energy $h\nu = E_2 - E_1$.
- Life time of excited state is 10^{-9} sec.



Stimulated Emission

The **stimulated photons** have unique properties:

- In phase with the incident photon



- Same

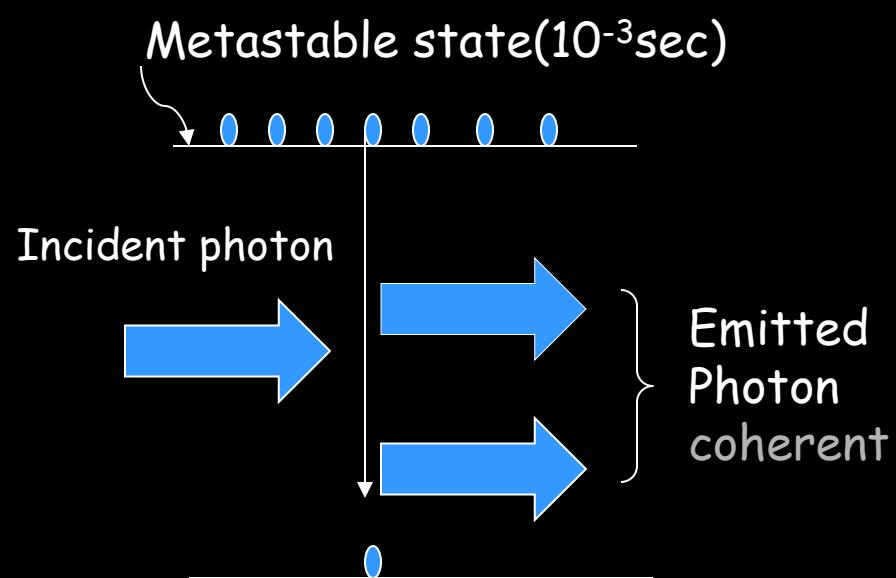
photon

- Transient

nt

Stimulated emission

- There are metastable state i.e. transition from this state is not allowed acc to selection rule.
- There life time is 10^{-3} sec.



- When an photon of suitable freq arrive it make the atom in metastable unstable.
- The emitted photon is in coherence with incident photon.

Stimulated vs Spontaneous Emission

Stimulated emission requires the presence of a photon. An “incoming” photon stimulates a molecule in an excited state to decay to the ground state by emitting a photon. The stimulated photons travel in the same direction as the incoming photon.

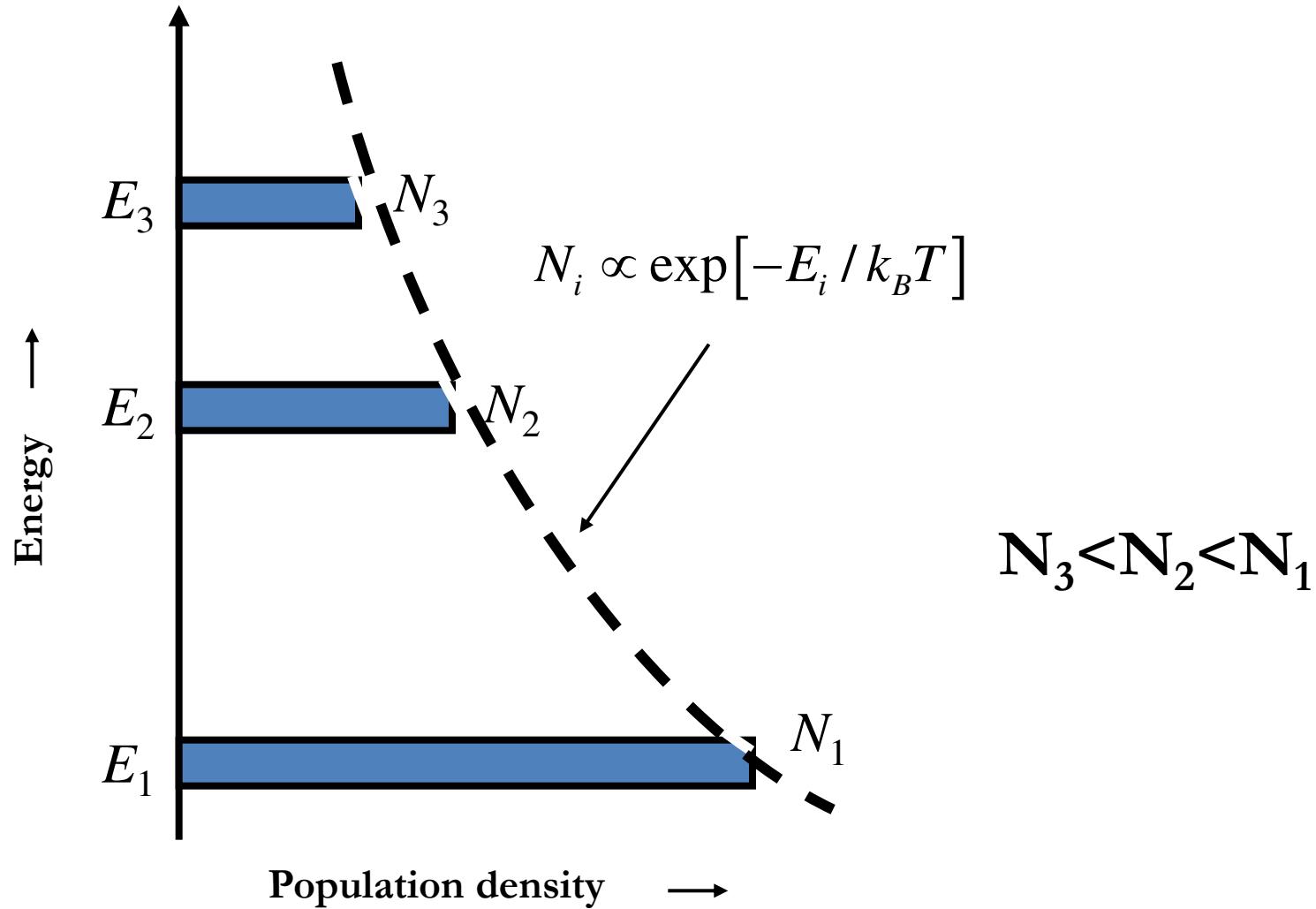
Spontaneous emission does not require the presence of a photon. Instead a molecule in the excited state can relax to the ground state by spontaneously emitting a photon. Spontaneously emitted photons are emitted in all directions.

Population Inversion

- At **equilibrium**, absorption and spontaneous emission is **simultaneous**, but $N_1 > N_2$
- A state in which a substance has been **energized**, or **excited** to specific energy levels.
- More atoms or molecules are in a higher excited state to obtain high percentage of stimulated emission.
- Non-equilibrium state in which population in N_2 **exceeds** to **that of N_1** is generally known as **Population inversion**
- The **process of producing** a population inversion is called **pumping**.

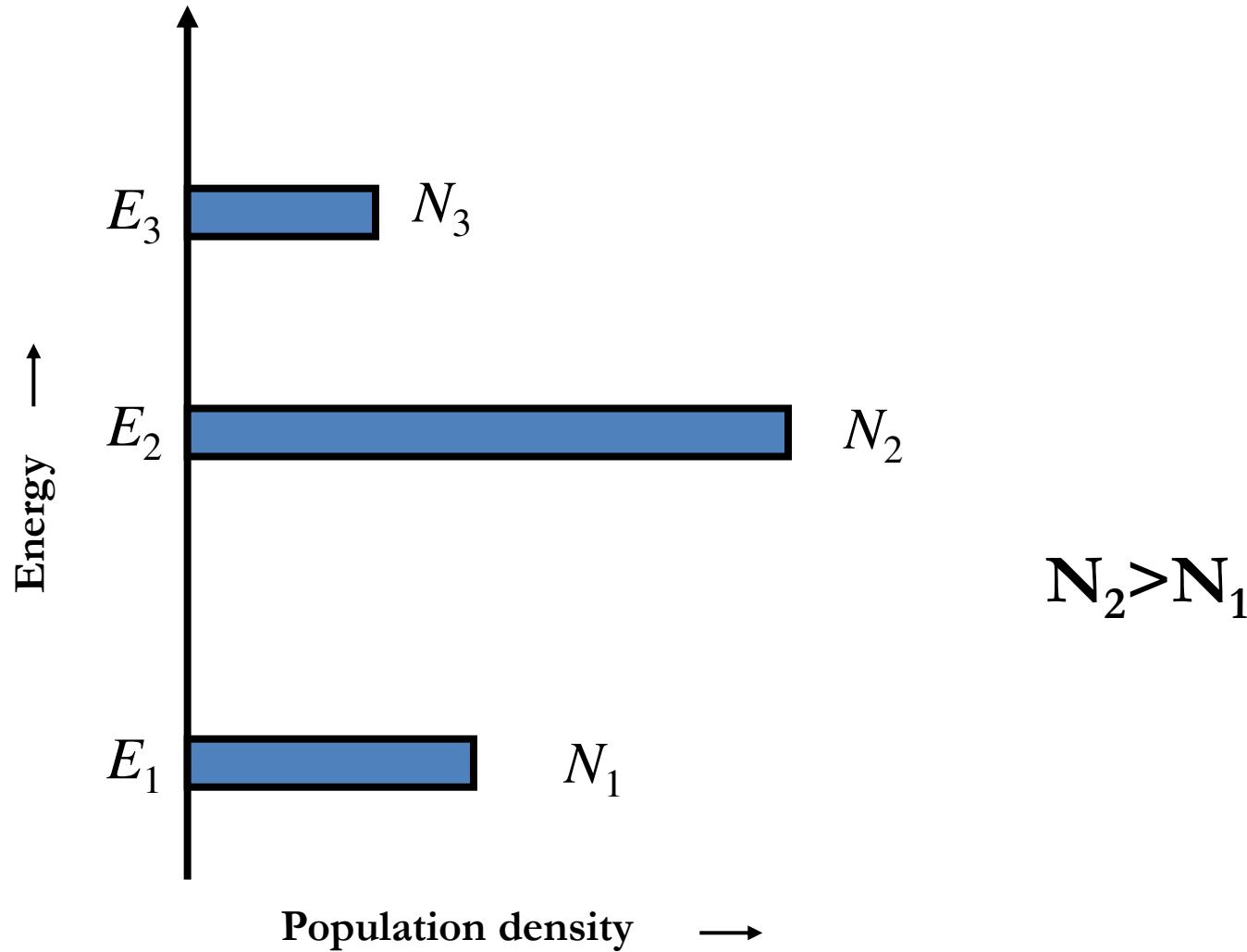
Population Inversion

Some of the Boltzmann population factors,



Population Inversion

Some of the Boltzmann population factors,



Pumping

- A medium in which **population inversion** is achieved is called **active medium**.
- The population inversion can be achieved usually by **exciting the active medium** with suitable form of energy – pumping.
- Most commonly used methods for pumping are
 - Optical pumping
 - Direct electron excitation
 - Inelastic atom-atom collisions
 - Chemical reactions
 - Direct conversion

Pumping

Optical pumping

- **Light source is used** – this energy comes in the form of short flash of light
- Suitable for any laser medium which is transparent to pump light.

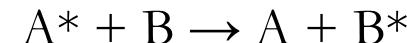
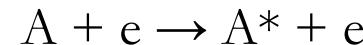
Electric discharge

- Preferred in **gases** – high voltage causes **electrons** emitted by the cathode to be **accelerated** towards anode. These accelerated electrons **collide** with the atoms in the active medium, **ionize** the medium and raise it to excited state causing population inversion.

Pumping

Inelastic atom-atom collisions

- Excitation by **electric discharge** provides the initial excitation which raises **one type of atoms** to their excited state.
- These atoms **collide inelastically** with **another type of atom** the later atoms provides population inversion.



Direct conversion

- In semiconductor lasers direct conversion of **electrical energy** into **light energy** takes place.

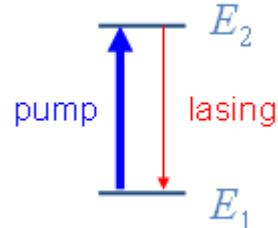
Chemical reactions

- The molecules undergo **chemical changes** in which one of the product of the reaction is a molecule or an atom that is **left** in an **excited state** under appropriate conditions.

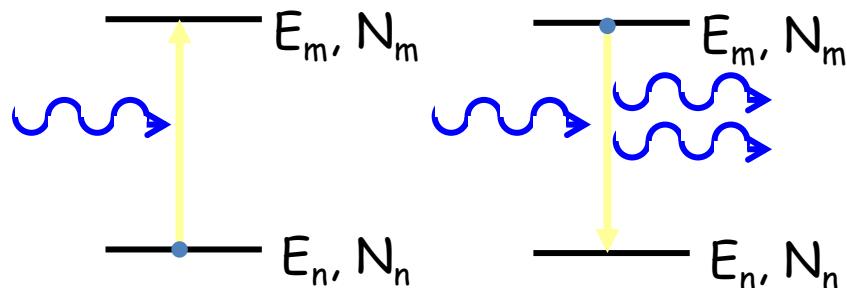
Types of Lasers

Lasers can be classified based on the factor, namely **number of levels involved** and the **type of active medium**.

Two – Level Laser



- **Not suitable** for attaining population inversion
- **Difficult to keep** the collection of atoms in the **excited state** until they are stimulated to emit photons

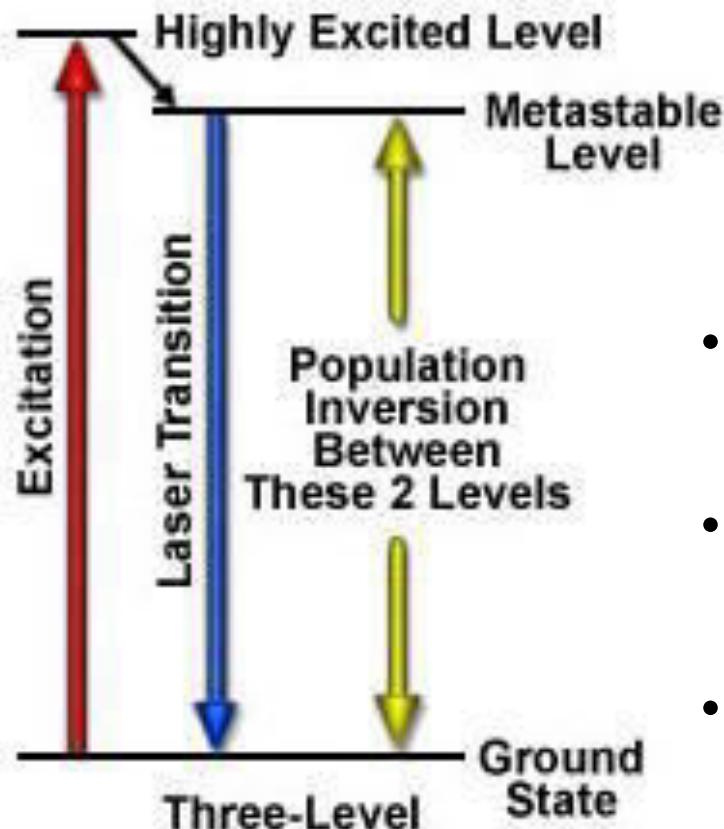


Even with very a intense pump source, the best one can achieve with a two-level system is

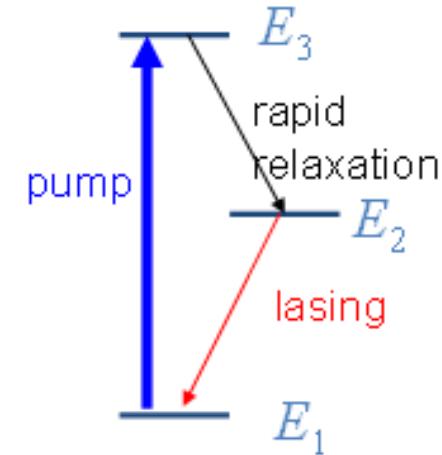
excited state population = ground state population

Types of Lasers

Three – Level Laser

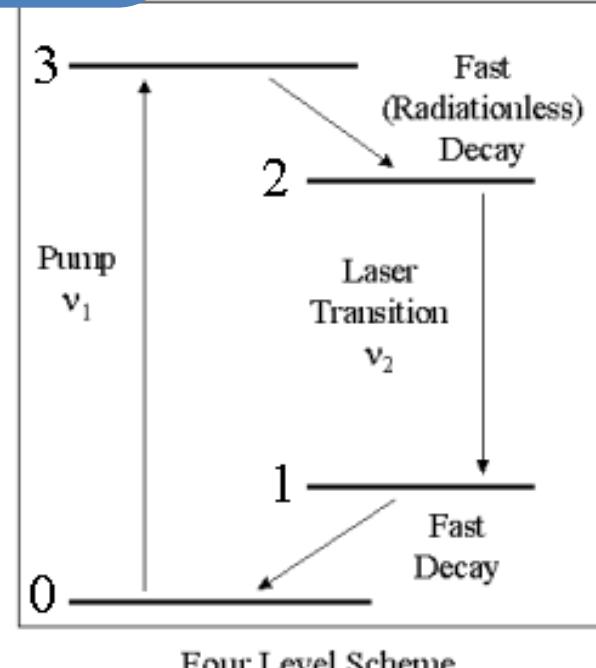
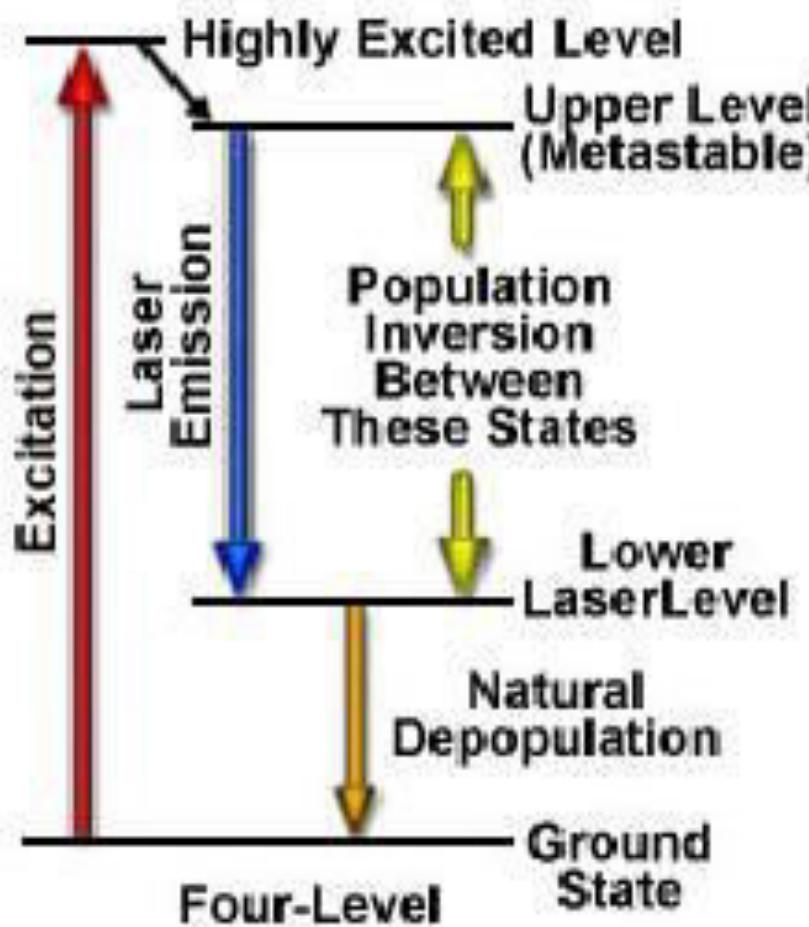


- Initially excited to a **short-lived high-energy state** .
- Then **quickly decay** to the intermediate **metastable level**.
- Population inversion is created between **lower ground state** and a **higher-energy metastable state**.



Types of Lasers

Four – Level Laser



- Laser transition takes place between the **third and second excited states**.
- Less pumping energy than three level laser
- Can operate in a **continuous mode**.

Types of Laser

Lasers can be classified based on

number of levels involved

- (i) Two-level Lasers
- (ii) Three-level Lasers
- (iii) Four-level Lasers

and types of active medium

- (i) Solid state Lasers
- (ii) Gas Lasers
- (iii) Liquid Lasers

Main Components

Three main components of **ANY** lasers are

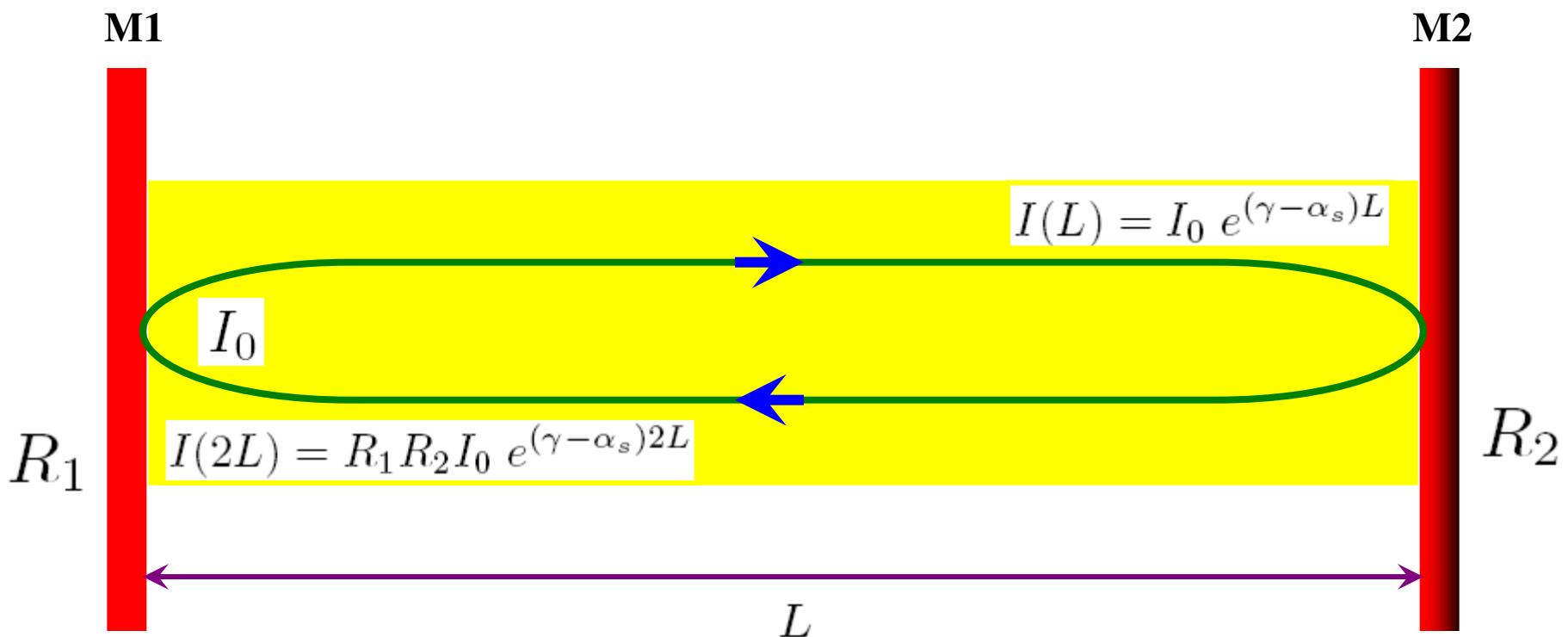
- (i) The active medium
- (ii) The pumping source
- (iii) The optical resonator

- The active medium acts as an **amplifier** for light waves
- For amplification, **medium** should be in a state of **Population inversion**
- Population inversion – **metastable levels** – lifetime is bit longer as compared to excited state
- The active medium placed inside an optical resonator – acts as an **oscillator**
- A pair of mirrors + active medium - optical resonators

230 04_Gain_Coefficient_Reference_Material_I_18-Feb-2021

Threshold condition for LASING action

- Output of the active medium bouncing back and forth in the optical resonator.
- During amplification it suffers various losses



1. Transmission at the output mirror
 2. Scattering, diffraction and absorption of light within the active medium
- Proper build up of laser oscillation: The amplification between two constructive reflections of light from rear end mirror must balance the losses.
 - Assume that the active medium of the laser fills the space between the mirrors M1 and M2.
 - The reflectivity of both mirrors are R_1 and R_2
 - Mirrors be separated by a distance L
 - Let the intensity of light beam at M1 be I_0
 - Traveling from M1 to M2, the beam intensity increases from

$$I(L) = I_0 e^{(\gamma - \alpha_s)L}$$

Where γ is the gain of the laser medium

α_s losses due to scattering, diffraction and absorption in the medium

After reflection at mirror M2, the beam intensity will be

$$I(L) = R_2 I_0 e^{(\gamma - \alpha_s)L}$$

After complete the round trip the final intensity will be

$$I(2L) = R_1 R_2 I_0 e^{(\gamma - \alpha_s)2L}$$

The amplification obtained during the round trip

$$G = \frac{I(2L)}{I_0} = R_1 R_2 e^{(\gamma - \alpha_s)2L}$$

The product $R_1 R_2$ represents the losses at the mirror

233 04_Gain_Coefficient_Reference_Material | 18-Feb-2021

The losses are balanced by gain, when the amplification factor

$$G \geq 1$$

It requires that

$$R_1 R_2 e^{(\gamma - \alpha_s)2L} \geq 1$$

$$e^{(\gamma - \alpha_s)2L} \geq \frac{1}{R_1 R_2}$$

$$(\gamma - \alpha_s)2L \geq \ln\left(\frac{1}{R_1 R_2}\right)$$

$$\gamma - \alpha_s \geq \frac{1}{2L} \ln \left(\frac{1}{R_1 R_2} \right)$$

$$\gamma \geq \alpha_s + \frac{1}{2L} \ln \left(\frac{1}{R_1 R_2} \right)$$

The above equation is the condition for the lasing action

This equation is determined the threshold value of pumping energy for lasing action

As the pump power is slowly increased, a value of γ^{th} called threshold value is reached and the laser starts oscillations

The threshold value γ^{th} is given by

$$\gamma_{\text{th}} = \alpha_s + \frac{1}{2L} \ln \left(\frac{1}{R_1 R_2} \right)$$

Nd:YAG Lasers (Neodymium: Yttrium aluminium garnet $\text{Y}_3\text{Al}_5\text{O}_{12}$)

First successful working laser belongs to solid-state medium was ruby laser built by T.H. Maiman in 1960

Pure material acts as host and the dopant acts as guest material responsible for the lasing action

Material

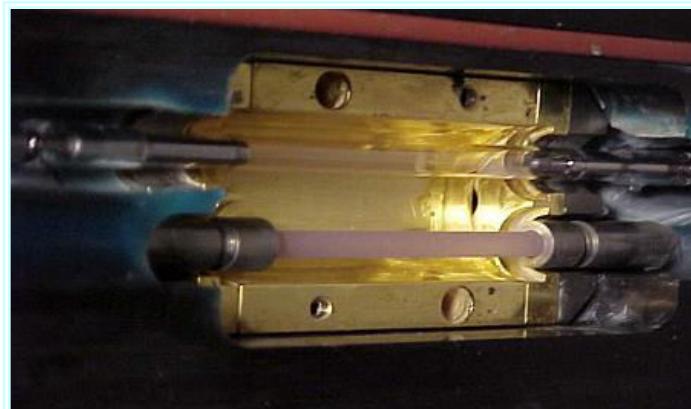
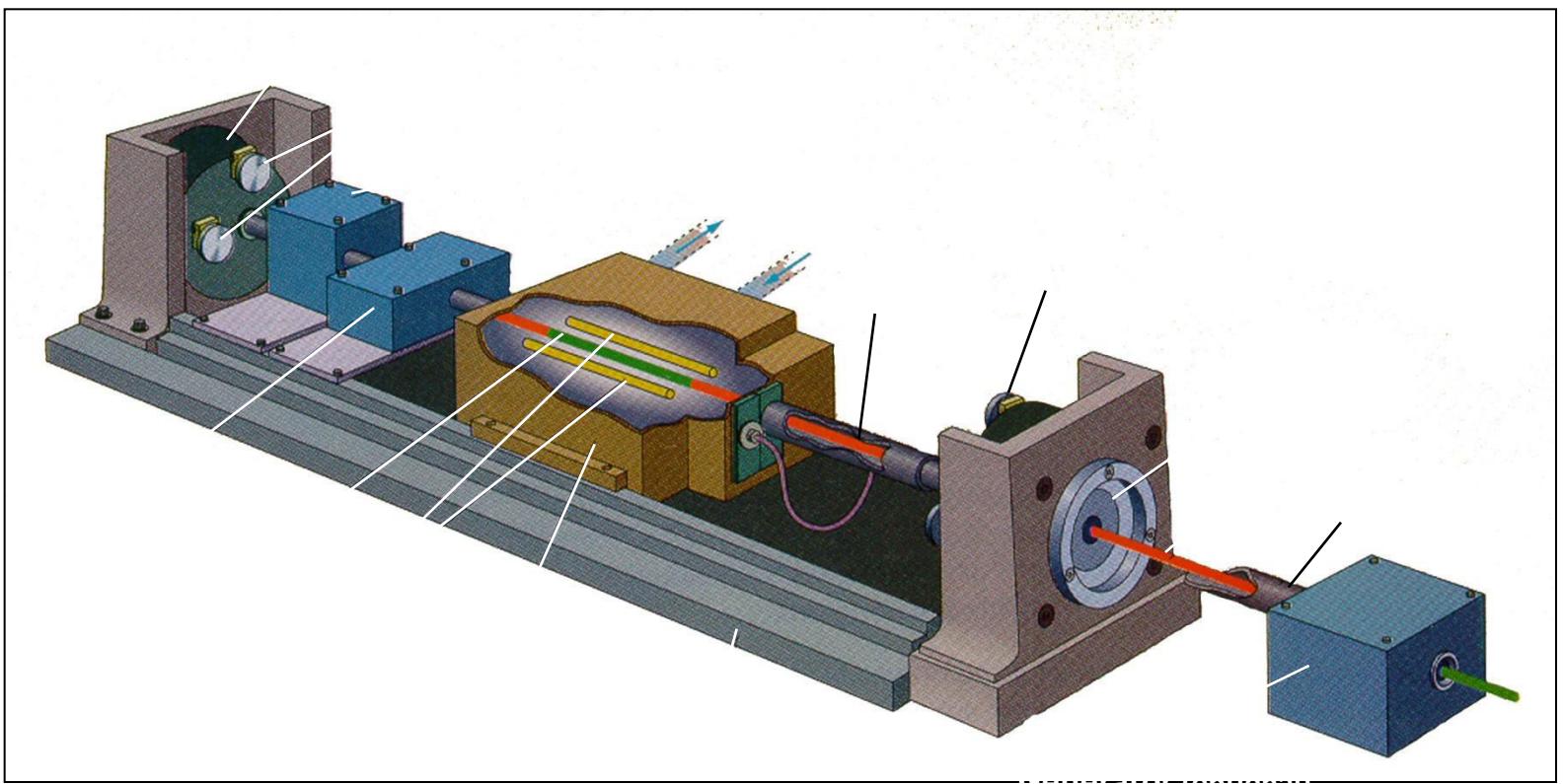
It is a four level laser

$\text{Y}_3\text{Al}_5\text{O}_{12}$ commonly called as YAG is an optical isotropic crystal

Some of Y^{3+} ions in the crystal is replaced by neodymium ions, Nd^{3+}

Crystal atoms do not participate in lasing action-but as host for active centers

where Nd^{3+} resides

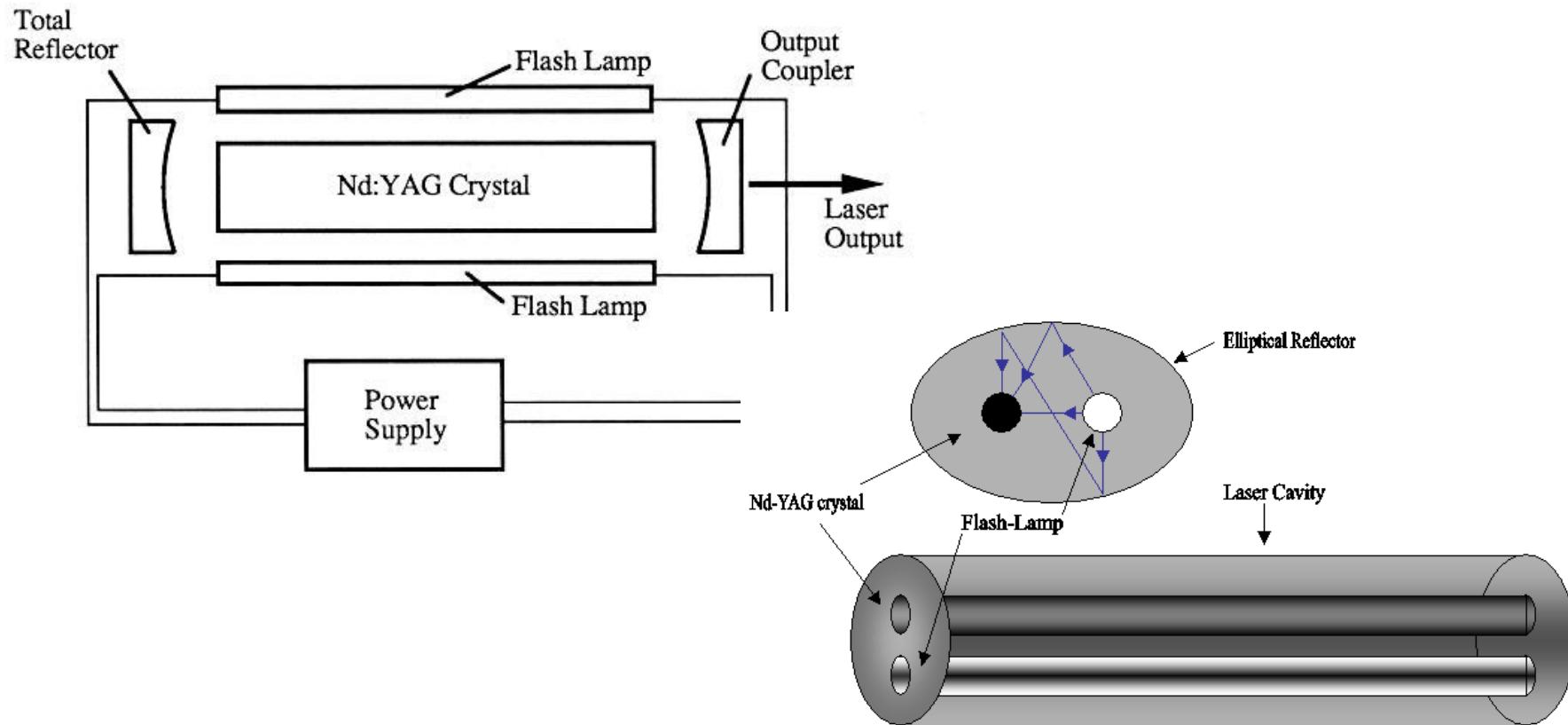


• Construction

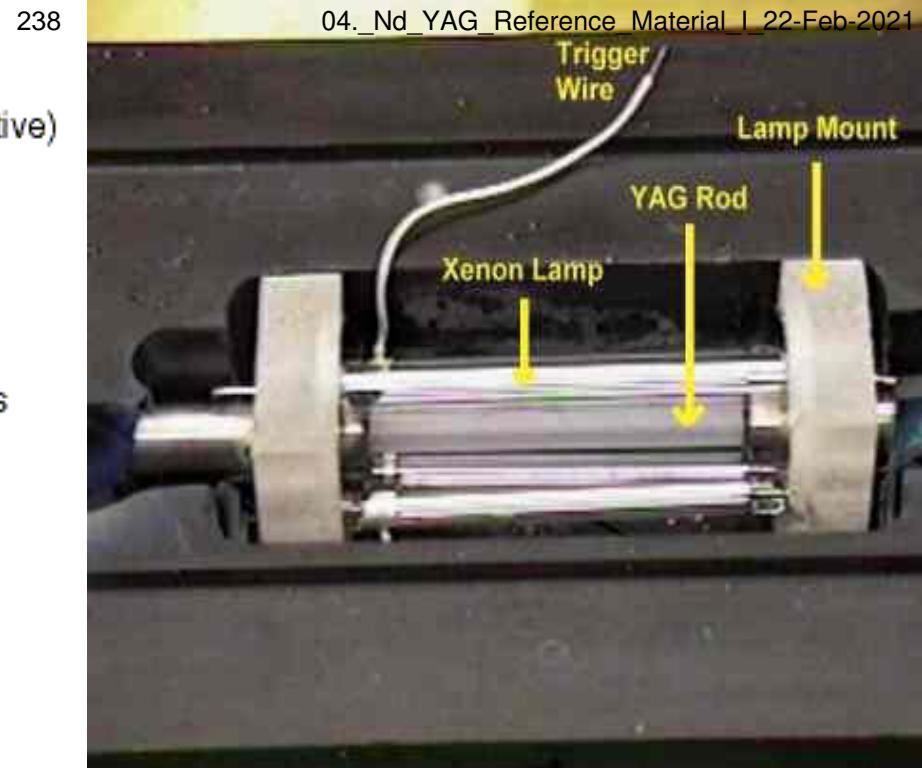
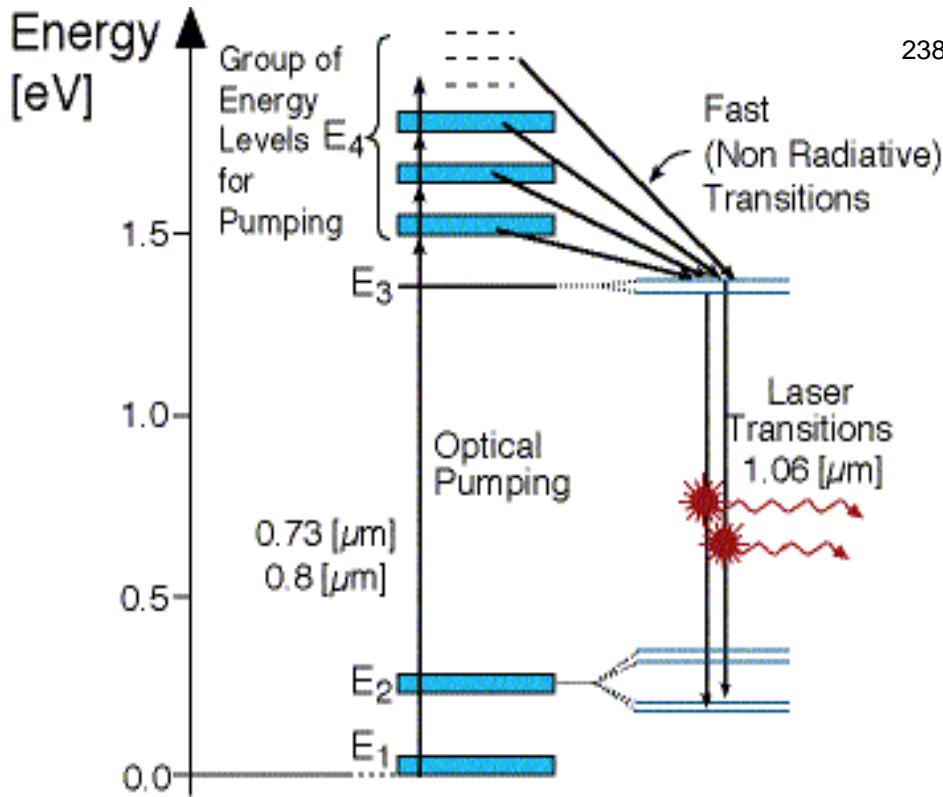
237

04._Nd_YAG_Reference_Material_I_22-Feb-2021

- Nd: YAG laser is made up of elliptical cylindrical reflector



- One end is fixed with focus krypton lamp acting as pumping device
- Other focus is silvered – flash from the krypton lamp after reflection concentrate at YAG rod placed at the other end
- Ends of the laser rod is polished with silver to achieve the resonance mechanism of lasing action



- Nd ion has two absorption band – the excitation is done by optical pumping
- Nd ions transfer to upper laser level by non-radiative transition – then stimulated emission to its lower level – from the lower level it is by non-radiative transition to ground level.

Some important features of $\text{Nd}^{239}\text{:YAG}$ laser are:

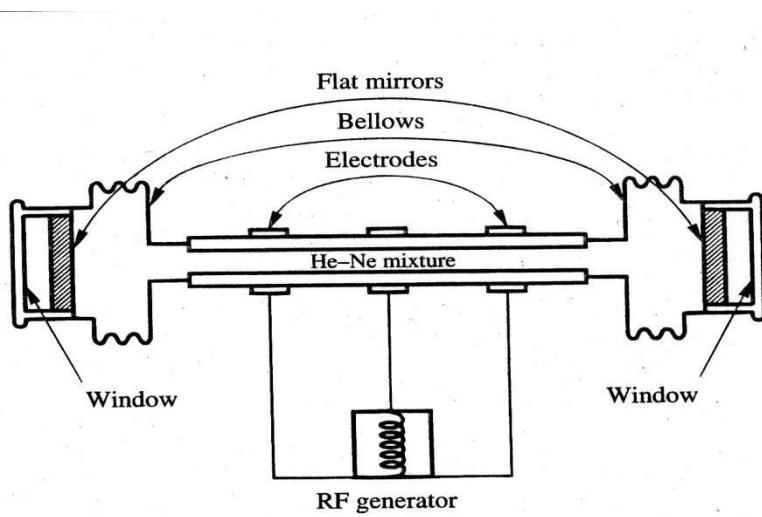
- It is a four-level laser
- RE like Nd, Er, Dy can be used
- Nd ion concentration corresponds to ground state population of 6×10^{19} ions/cc
- The metastable level $4F_{3/2}$ has a lifetime of 0.23×10^{-3} s
- The Nd:YAG laser is a quasi-continuous wave laser
- Cr ions can be dopped in addition to Nd ions. Xenon flash lamp can be used for pumping

Helium-neon Laser

- The **Helium-Neon** laser was the first **continuous laser**
- It was invented by **Ali Javan** et. al. in **1961**.
- Is an atomic laser employs **four level pumping** scheme.
- Its usual **operation wavelength** is **632.8 nm**, in the **red** portion of the visible spectrum.
- It operates in **Continuous Working (CW)** mode.
- Active medium is a mixture of **10 parts** of **Helium** and **one part** of **Neon**.
- Ne atoms are **active centers** and have energy levels suitable for laser transitions while, He atoms **help in efficient excitation** of Ne atoms.

Construction of He-Ne laser

- The setup consists of a discharge tube of **length 35 cm** and bore **diameter of 1 cm**.
- The medium - **mixture** of helium and neon gases, in a **5:1** to **20:1** ratio, contained at low pressure (an average **50 Pa per cm of cavity length**) in a glass envelope.
- The energy or pump source of the laser is provided by an electrical discharge of few kilo volts through an anode and cathode at each end of the glass tube. A current of 5 to 100 mA is typical for CW operation.



Schematic arrangement of the first gas laser.

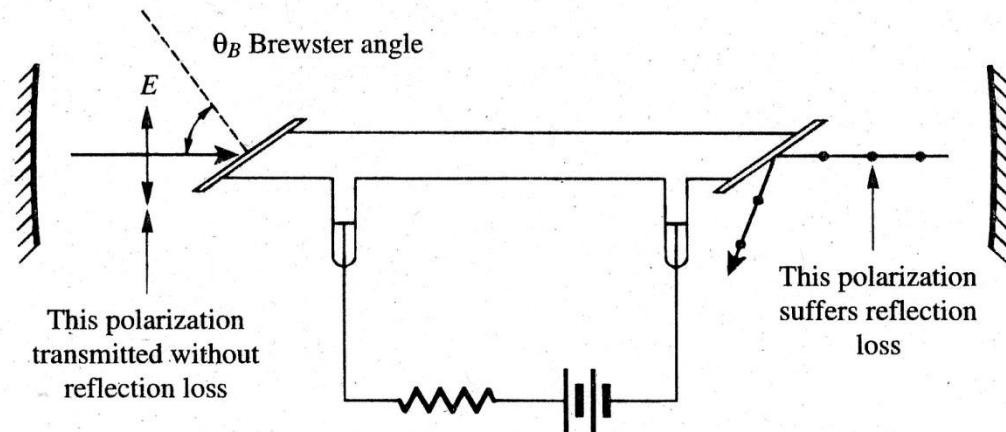
Helium-Neon Lasers

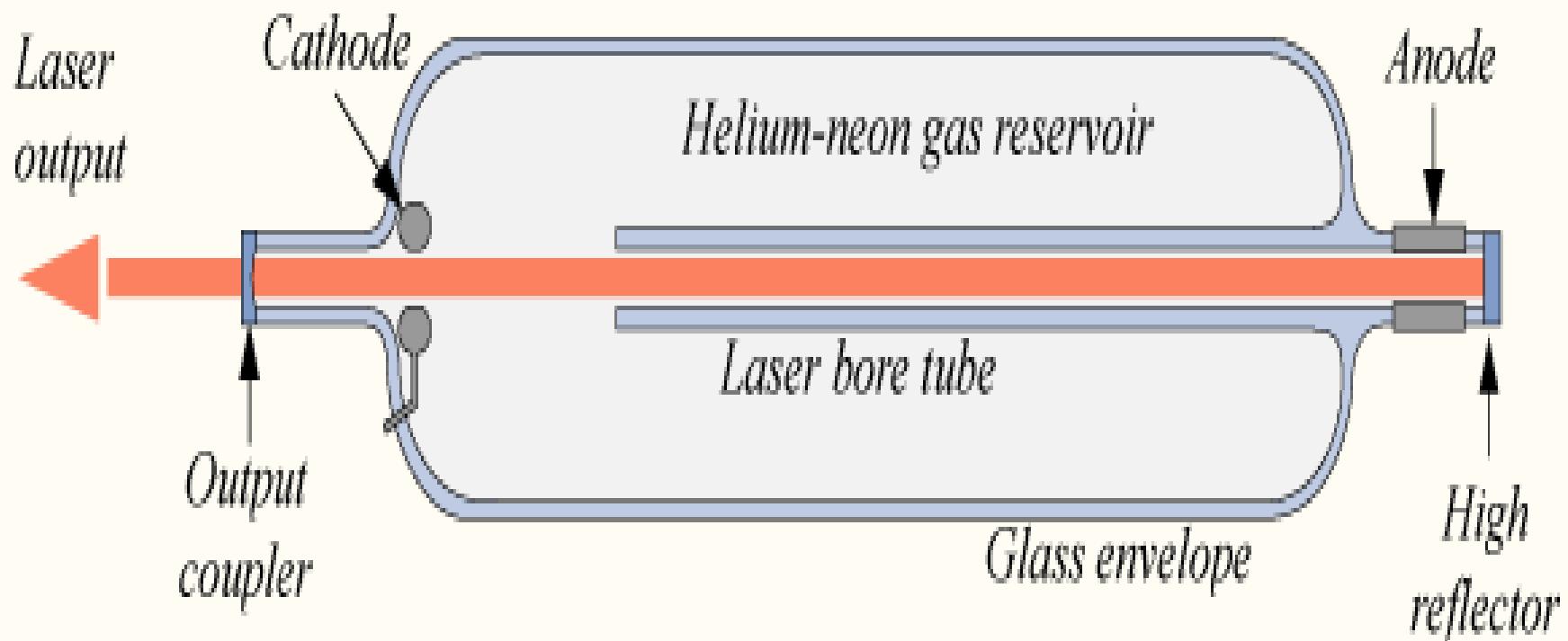
243

04_Helium_Neon_Laser_Reference_Material_I_23-Feb-2021

- The optical cavity of the laser typically consists of a plane, high-reflecting mirror at one end of the laser tube, and a concave output coupler mirror of approximately 1% transmission at the other end.
- Since the cavity window is outside the tube, Brewsters windows may be used at the ends of tube to minimize reflection loss.
- HeNe lasers are normally small, with cavity lengths of around 15 cm up to 0.5 m, and optical output powers ranging from 1 mW to 100 mW.

Typical schematic design of a modern laser.

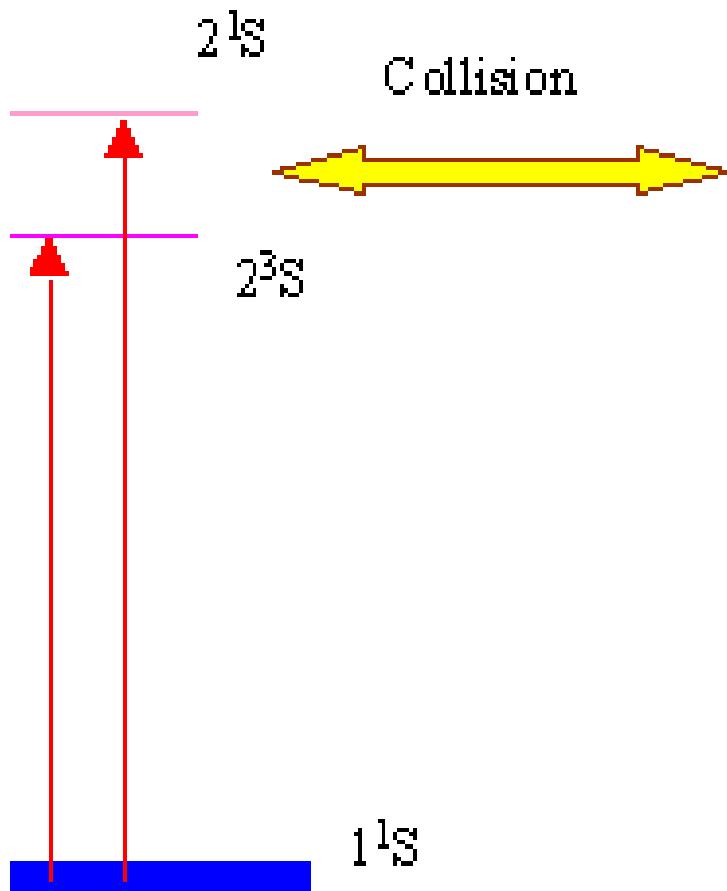




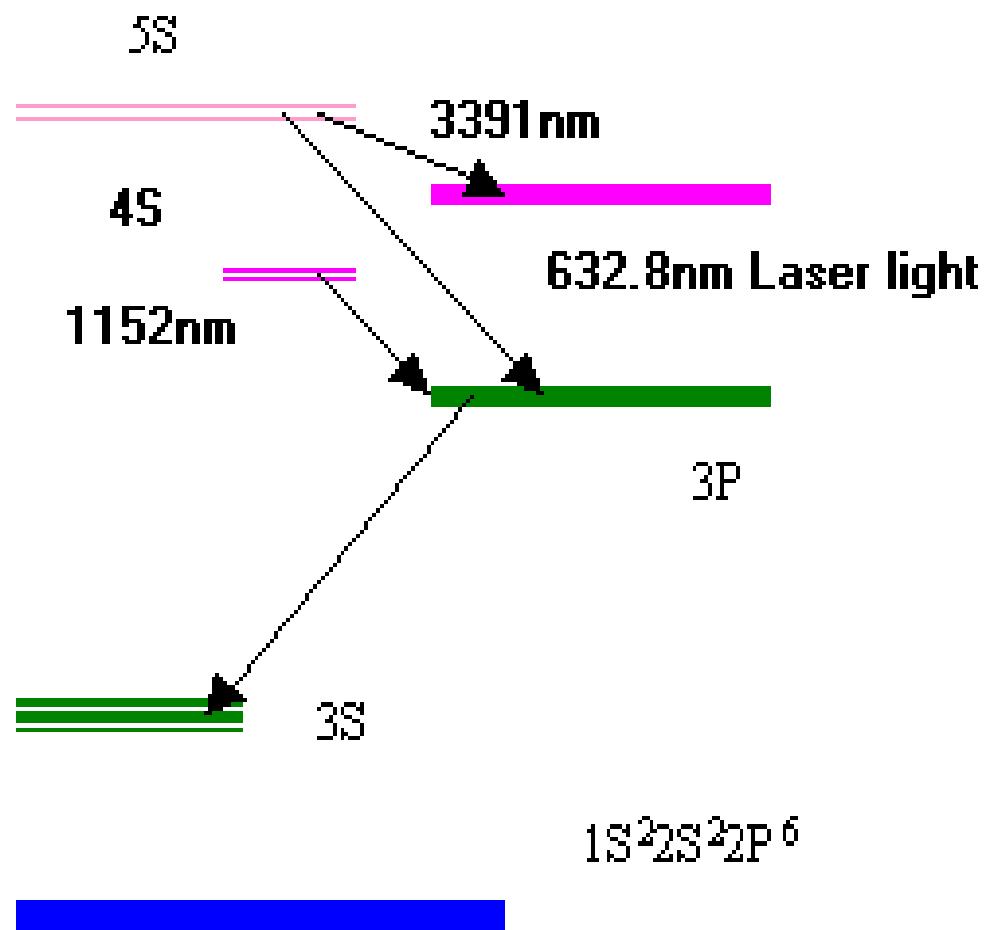
Working of He-Ne laser

- When **voltage** is applied to the **electrodes** it **ionizes** the **gas**, the **electrons** and **ions** thus **produced** are accelerated towards anode and cathode respectively.
- Electrons acquire higher velocity due to its smaller mass when compared to the others. They **transfer K.E to He atoms** through **inelastic atom-atom collision**.
- **He atoms** are readily **excited by electrons** impact because of its **fairly light mass**.
- Thus He atoms are excited to **F₂** and **F₃** states which lie at **19.81** and **20.61 eV** respectively.
- These are **meta stable states** and these atoms **cannot return** to **ground state** by spontaneous emission.
- These atoms **returns to ground state** by **transferring energy to Ne atom** in the state which has identical energy. Such transfer is called **resonant transfer of energy**.
- Ne energy levels **E₆** and **E₄** **nearly coincide with F₃** and **F₂** so resonant transfer can occur.
- This energy exchange process occurs with high probability only because of the accidental near equality of the two excitation energies of the two levels in these atoms. Thus, the purpose of population inversion is fulfilled.

Helium



Neon



Ground State

- When Ne atom in the **E₆ level** and **E₄ level** emits a photon parallel to the axis of the tube
- In reality neon energy levels E₆, E₅, E₄, E₃, E₂ are not single but a **group of lines**. Consequently several laser transitions are possible.
- Three main laser transitions are
 - **E₆ to E₃** – generates laser beam of **red** color at **6328Å**
 - **E₄ to E₃** – IR beam at wavelength of **1.15 mm**
 - **E₆ to E₅** – light in Far IR region at **3.39 mm**
- Level E₂ is a metastable state, the neon atoms tend to accumulate at this level, if they are not removed the population in the ground state continuously decreases restricting lasing action.
- E₂ to E₁ transition can be induced by collision with the walls of the discharge tube.

Applications of He-Ne laser

- The Narrow red beam of He-Ne laser is used in supermarkets to read bar codes.
- The He- Ne Laser is used in Holography in producing the 3D images of objects.
- He-Ne lasers have many industrial and scientific uses, and are often used in laboratory demonstrations of optics.

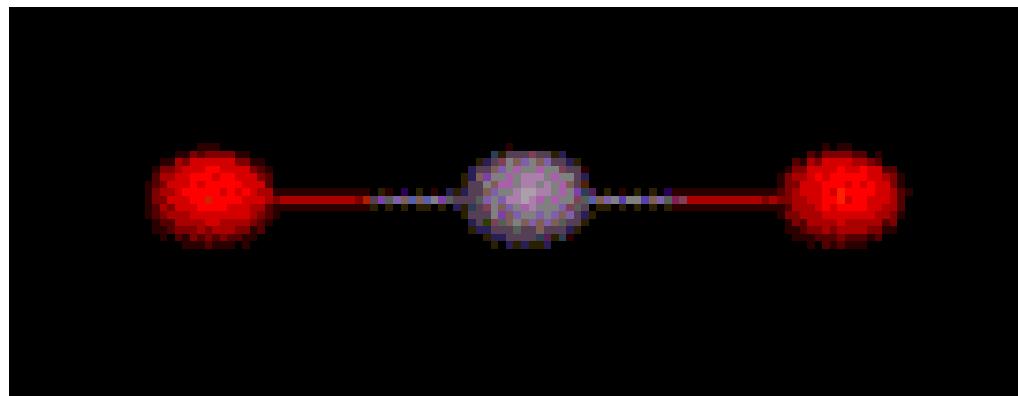
CO₂ Laser

Carbon-di-oxide laser

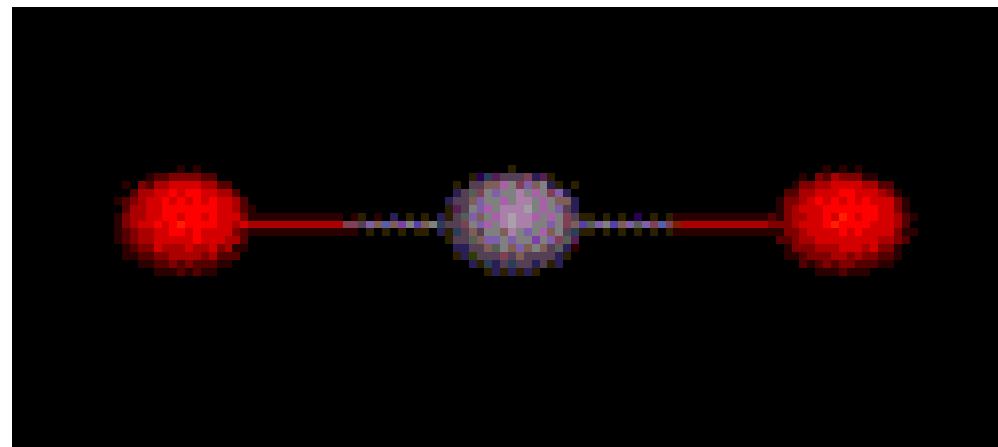
- CO₂ laser was one of the earliest gas laser to be developed in BELL Labs in 1964 by **Indian born CKN Patel.**
- It is a **molecular gas laser**
- o/p of these lasers are more **powerful**
- In CO₂ laser the transition occurring between **different vibrational states** are responsible for laser effect
- CO₂ molecule has a central carbon and two Oxygen one at either ends
- Such a molecule can vibrate in **three different modes of vibration**, and in each mode of vibration the center of gravity remains fixed
 - Three modes are
 1. Sym. Stretching mode
 2. Bending mode
 3. Assym. Stretching mode.

- **Symmetric stretching mode**

- Carbon atom fixed in its position and each oxygen atom can vibrate along the axis of the molecule simultaneously departing or approaching the fixed carbon atom.
- The frequency corresponding to this mode of vibration is called symmetric stretching frequency



- Bending mode
 - Oxygen atom and carbon atom may vibrate at right angles to the line passing through the centre of gravity.
 - The frequency corresponding to this mode of vibration is called as bending frequency

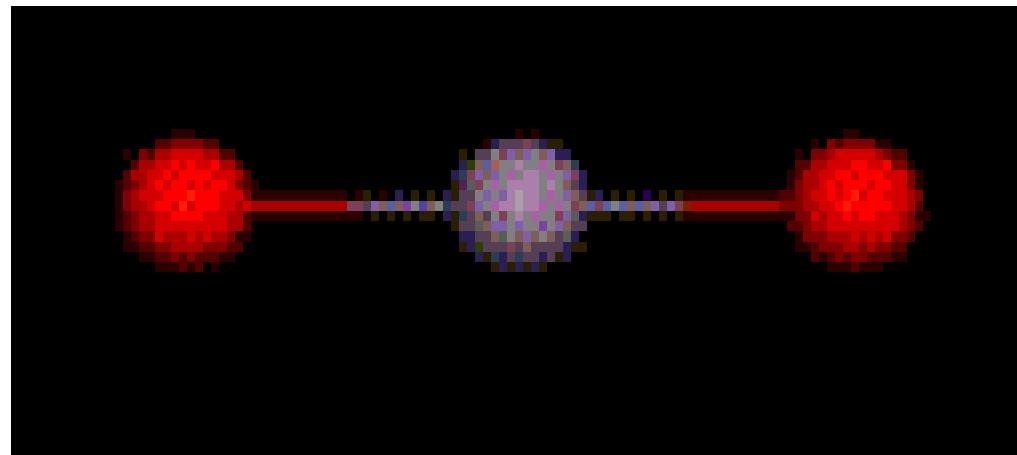


- Asymmetric stretching mode

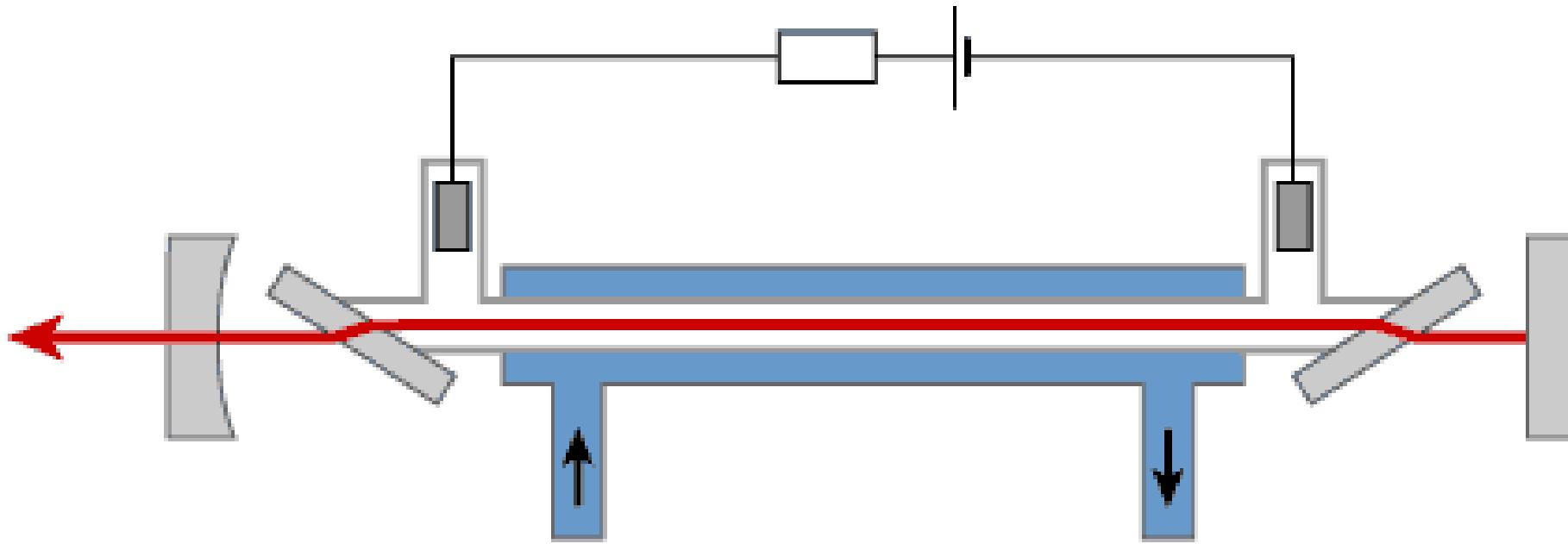
253

04._Helium_Neon_Laser_Reference_Material_I_23-Feb-2021

- The two oxygen atom and the carbon atom at the centre vibrates asymmetrically or they vibrate in an opposite directions with respect to its mean position.
- The corresponding frequency is called asymmetric stretching frequency.



- The experimental setup is as shown



- Consist of discharge tube made of fused Quartz of **2.5 cm** in **diameter** and **5 m long**.
- Special feature of the CO₂ laser is the o/p is **dependent** on the **diameter** of the discharge tube.
- Active medium consists of He, N₂ and CO₂, **Brewster's windows** at the ends,
- Near confocal silicon mirror coated with aluminium formed at the resonant cavity

• Working

255

04._Helium_Neon_Laser_Reference_Material_I_23-Feb-2021

- In CO₂ lasers N₂ plays the **similar role** of He in He-Ne laser
- N₂ goes to excited state by **collision with electrons**



- The excited N₂ transfers energy to CO₂ and CO₂ gets excited



- The lowest vibrational level of N₂ has nearly as much energy as the asymmetric stretching mode of CO₂ molecule and so the excited N₂ readily transfer energy to CO₂ in resonant collisions.

• Merits and demerits

- Has very high o/p power
- o/p power can be increased by increasing the length of the gas tube

Energy cm^{-1}

1000

2000

3000

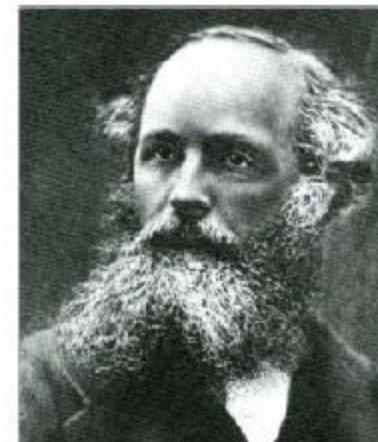
 $\frac{1388 \text{ cm}^{-1}}{100}$ 10.4 μm

020

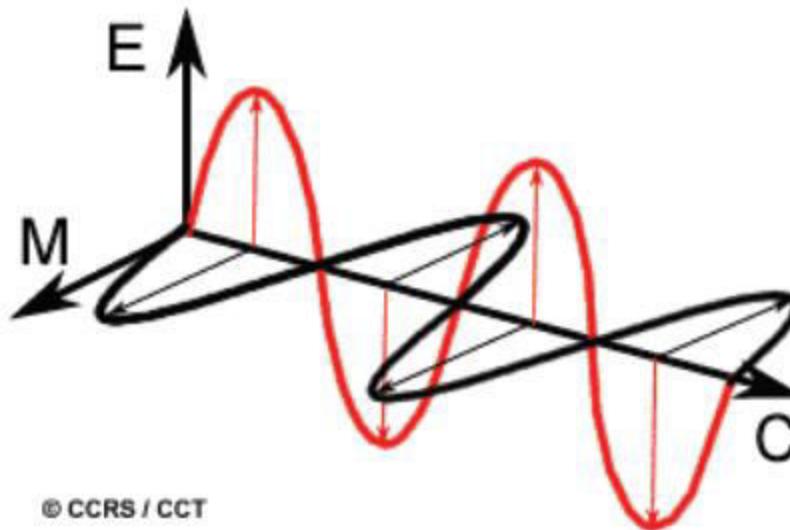
 $\frac{667 \text{ cm}^{-1}}{010}$ CO₂
Symmetric
StretchCO₂
Bending
ModeCO₂
Asymmetric
StretchCollisional
excitation
transfer $\frac{2349 \text{ cm}^{-1}}{001}$ 9.4 μm N₂

Electromagnetic radiation: wave model

- James Clerk Maxwell (1831-1879) – Scottish mathematician and physicist
- Wave model of EM energy
 - Unified existing laws of electricity and magnetism (Newton, Faraday, Kelvin, Ampère)
 - Oscillating electric field produces a magnetic field (and vice versa) – propagates an EM wave
 - Can be described by 4 differential equations
 - Derived speed of EM wave in a vacuum



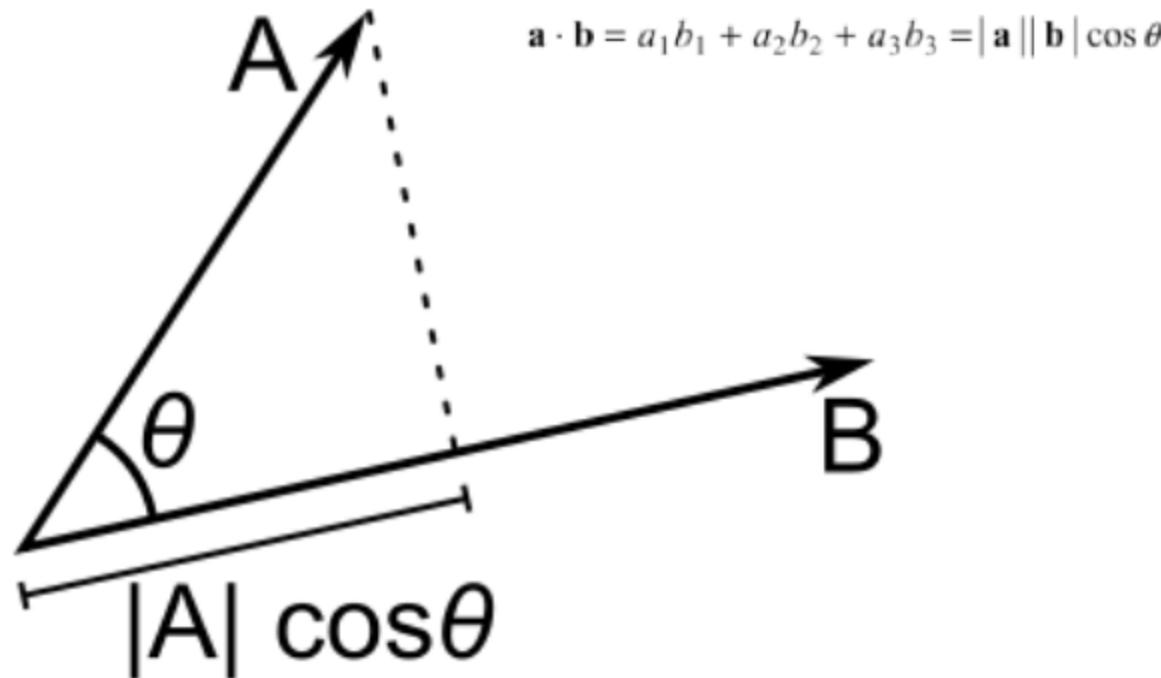
Electromagnetic radiation



© CCRS / CCT

- EM wave is:
- Electric field (E) perpendicular to magnetic field (M)
- Travels at velocity, c (3×10^8 ms⁻¹, in a vacuum)

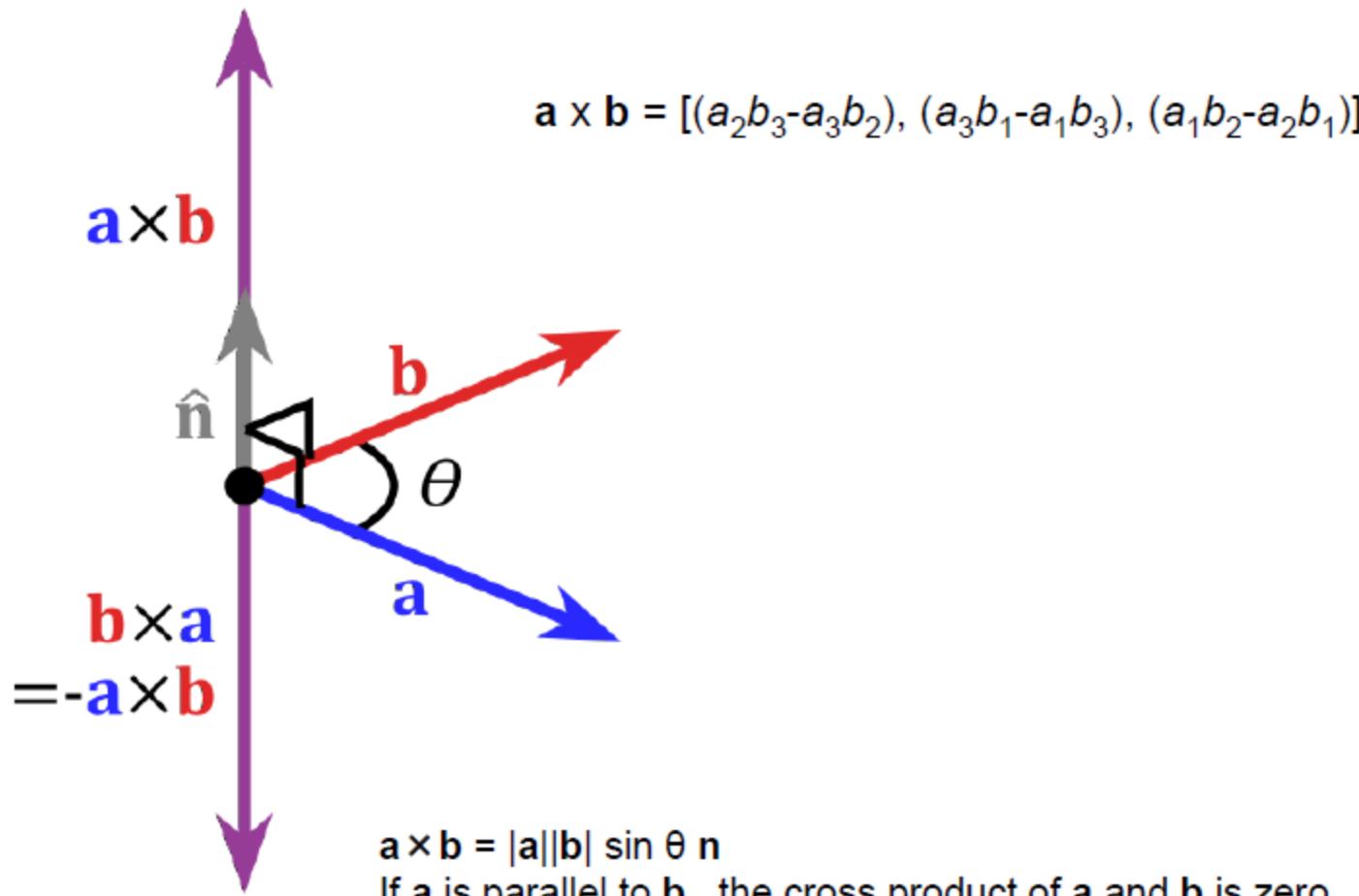
Dot (scalar) product



$$\mathbf{A} \cdot \mathbf{B} = |\mathbf{A}| |\mathbf{B}| \cos \theta$$

If \mathbf{A} is perpendicular to \mathbf{B} , the dot product of \mathbf{A} and \mathbf{B} is zero

Cross (vector) product



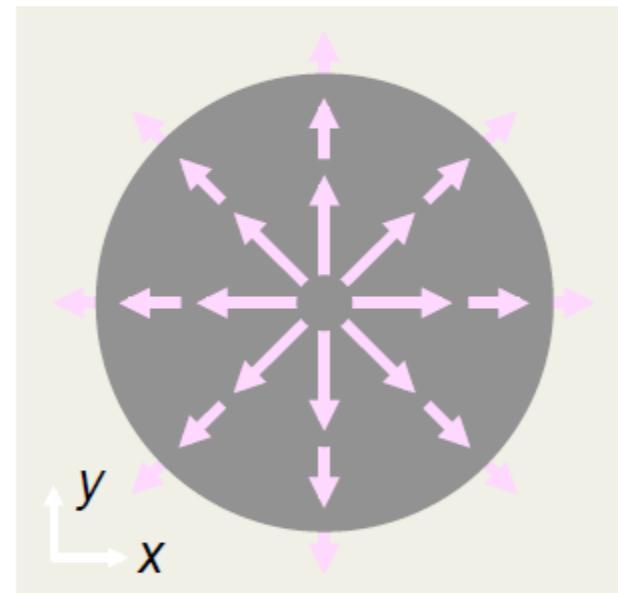
Div, Grad, Curl

The **Divergence** of a vector function (scalar):

$$\nabla \cdot f = \frac{\partial f_x}{\partial x} + \frac{\partial f_y}{\partial y} + \frac{\partial f_z}{\partial z}$$

The **Divergence** is nonzero if there are sources or sinks.

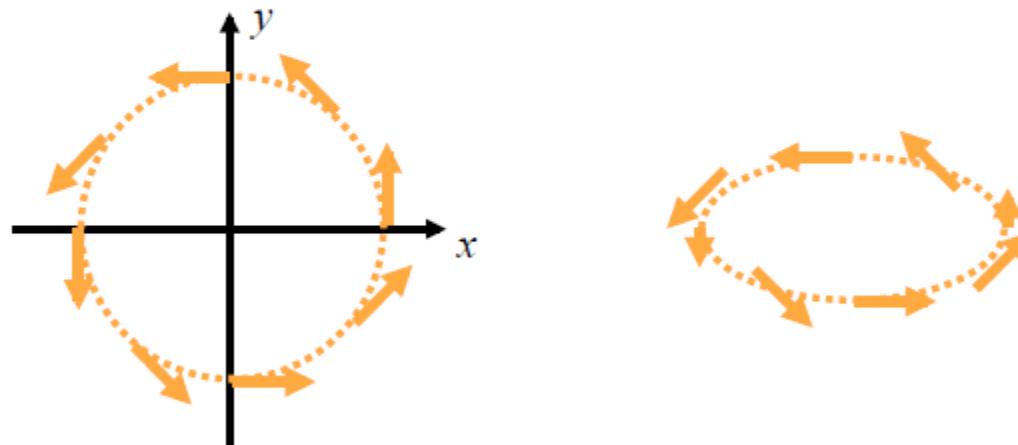
A 2D source with a large divergence:



Div, Grad, Curl

The **Curl** of a vector function \vec{f} :

$$\vec{\nabla} \times \vec{f} \equiv \left(\frac{\partial f_z}{\partial y} - \frac{\partial f_y}{\partial z}, \frac{\partial f_x}{\partial z} - \frac{\partial f_z}{\partial x}, \frac{\partial f_y}{\partial x} - \frac{\partial f_x}{\partial y} \right)$$



Functions that tend to **curl around** have large curls.

Div, Grad, Curl

The **Laplacian** of a scalar function :

$$\begin{aligned}\nabla^2 f &\equiv \vec{\nabla} \cdot \vec{\nabla} f = \vec{\nabla} \cdot \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right) \\ &= \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}\end{aligned}$$

The **Laplacian of a vector** function is the same,
but for each component of f :

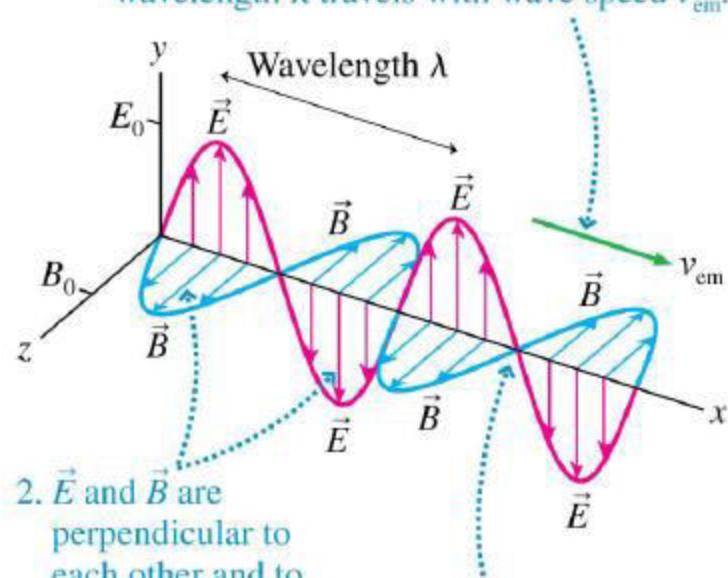
$$\nabla^2 \vec{f} = \left(\frac{\partial^2 f_x}{\partial x^2} + \frac{\partial^2 f_x}{\partial y^2} + \frac{\partial^2 f_x}{\partial z^2}, \frac{\partial^2 f_y}{\partial x^2} + \frac{\partial^2 f_y}{\partial y^2} + \frac{\partial^2 f_y}{\partial z^2}, \frac{\partial^2 f_z}{\partial x^2} + \frac{\partial^2 f_z}{\partial y^2} + \frac{\partial^2 f_z}{\partial z^2} \right)$$

The Laplacian tells us the curvature of a vector function.

The propagation direction of a light wave

FIGURE 35.19 A sinusoidal electromagnetic wave.

1. A sinusoidal wave with frequency f and wavelength λ travels with wave speed v_{em} .



2. \vec{E} and \vec{B} are perpendicular to each other and to the direction of travel. The fields have amplitudes E_0 and B_0 .

3. \vec{E} and \vec{B} are in phase. That is, they have matching crests, troughs, and zeros.

$$\vec{v} = \vec{E} \times \vec{B}$$

Right-hand screw rule

Maxwell's Equations

- Four equations relating electric (**E**) and magnetic fields (**B**) – vector fields
- ϵ_0 is **electric permittivity of free space** (or vacuum permittivity - a constant) – *resistance to formation of an electric field in a vacuum*
- $\epsilon_0 = 8.854188 \times 10^{-12}$ Farad m⁻¹
- μ_0 is **magnetic permeability of free space** (or magnetic constant - a constant) – *resistance to formation of a magnetic field in a vacuum*
- $\mu_0 = 1.2566 \times 10^{-6}$ T.m/A (T = Tesla; SI unit of magnetic field)

$$\nabla \bullet E = \frac{\rho}{\epsilon_0}$$

$$\nabla \bullet B = 0$$

$$\nabla \times E = -\frac{\partial B}{\partial t}$$

$$\nabla \times B = \mu_0 J + \epsilon_0 \mu_0 \frac{\partial E}{\partial t}$$

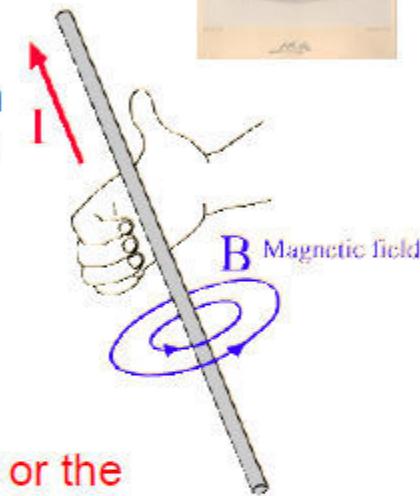
Note: $\nabla \bullet$ is 'divergence' operator and $\nabla \times$ is 'curl' operator

Biot-Savart Law (1820)



- Jean-Baptiste Biot and Felix Savart (French physicist and chemist)
- The magnetic field \mathbf{B} at a point a distance R from an infinitely long wire carrying current I has magnitude:

$$B = \frac{\mu_0 I}{2\pi R}$$



- Where μ_0 is the **magnetic permeability of free space or the magnetic constant**
- Constant of proportionality linking magnetic field and distance from a current
- Magnetic field strength decreases with distance from the wire
- $\mu_0 = 1.2566 \times 10^{-6} \text{ T.m/A}$ (T = Tesla; SI unit of magnetic field)

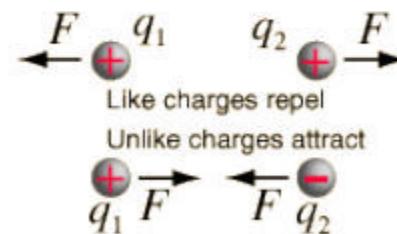
Coulomb's Law (1783)



- Charles Augustin de Coulomb (French physicist)
- The magnitude of the electrostatic force (F) between two point electric charges (q_1, q_2) is given by:

$$F = \frac{q_1 q_2}{4 \pi \epsilon_0 r^2}$$

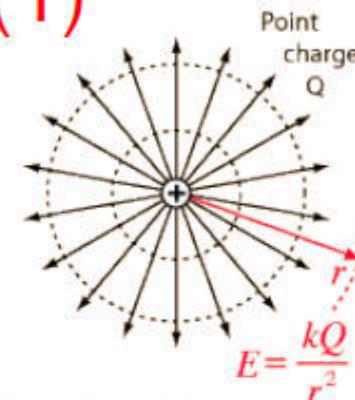
- Where ϵ_0 is the **electric permittivity or electric constant**
- Like charges repel, opposite charges attract
- $\epsilon_0 = 8.854188 \times 10^{-12}$ Farad m⁻¹





Maxwell's Equations (1)

$$\nabla \cdot E = \frac{\rho}{\epsilon_0}$$

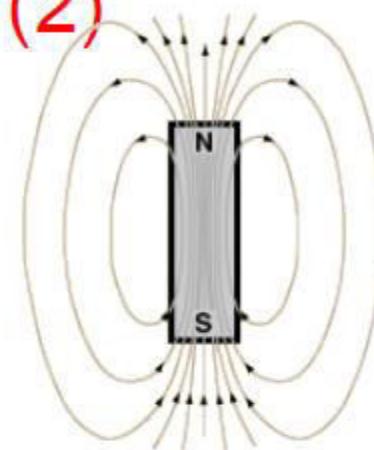


- **Gauss' law for electricity:** the electric flux out of any closed surface is proportional to the total charge enclosed within the surface; i.e. a charge will radiate a measurable field of influence around it.
- \mathbf{E} = electric field, ρ = net charge inside, ϵ_0 = vacuum permittivity (constant)
- Recall: divergence of a vector field is a measure of its tendency to converge on or repel from a point.
- Direction of an electric field is the direction of the force it would exert on a positive charge placed in the field
- If a region of space has more electrons than protons, the total charge is negative, and the direction of the electric field is negative (inwards), and vice versa.

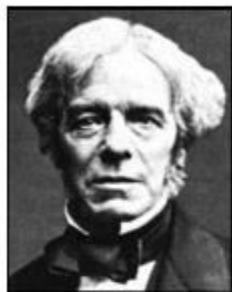


Maxwell's Equations (2)

$$\nabla \cdot \mathbf{B} = 0$$

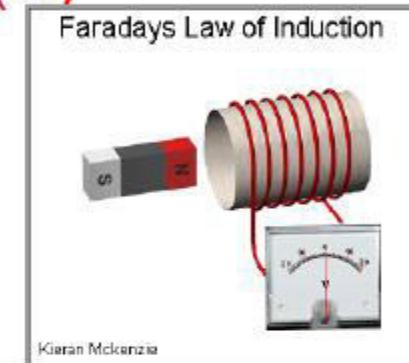


- **Gauss' law for magnetism:** the net magnetic flux out of any closed surface is zero (i.e. magnetic monopoles do not exist)
- \mathbf{B} = magnetic field; magnetic flux = $\mathbf{B}\mathbf{A}$ (\mathbf{A} = area perpendicular to field \mathbf{B})
- Recall: divergence of a vector field is a measure of its tendency to converge on or repel from a point.
- Magnetic sources are dipole sources and magnetic field lines are loops – we cannot isolate N or S 'monopoles' (unlike electric sources or point charges – protons, electrons)
- Magnetic monopoles *could* exist, but have never been observed



Maxwell's Equations (3)

$$\nabla \times E = -\frac{\partial B}{\partial t}$$



- **Faraday's Law of Induction:** the curl of the electric field (**E**) is equal to the negative of rate of change of the magnetic flux through the area enclosed by the loop
- **E** = electric field; **B** = magnetic field
- Recall: curl of a vector field is a vector with magnitude equal to the maximum 'circulation' at each point and oriented perpendicularly to this plane of circulation for each point.
- Magnetic field weakens → curl of electric field is positive and vice versa
- Hence changing magnetic fields affect the curl ('circulation') of the electric field – basis of electric generators (moving magnet induces current in a conducting loop)



Maxwell's Equations (4)

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \epsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t}$$

- **Ampère's Law:** the curl of the magnetic field (\mathbf{B}) is proportional to the electric current flowing through the loop

AND to the rate of change of the electric field. ← added by Maxwell

- \mathbf{B} = magnetic field; \mathbf{J} = current density (current per unit area); \mathbf{E} = electric field
- The curl of a magnetic field is basically a measure of its strength
- First term on RHS: in the presence of an electric current (\mathbf{J}), there is always a magnetic field around it; \mathbf{B} is dependent on \mathbf{J} (e.g., electromagnets)
- Second term on RHS: a changing electric field generates a magnetic field.
- Therefore, generation of a magnetic field does not require electric current, only a changing electric field. An oscillating electric field produces a variable magnetic field (as $d\mathbf{E}/dT$ changes)

Putting it all together....

- An oscillating electric field produces a variable magnetic field. A changing magnetic field produces an electric field....and so on.
- In 'free space' (vacuum) we can assume current density (J) and charge (ρ) are zero i.e. there are no electric currents or charges
- Equations become:

$$\nabla \bullet E = 0$$

$$\nabla \bullet B = 0$$

$$\nabla \times E = -\frac{\partial B}{\partial t}$$

$$\nabla \times B = \epsilon_0 \mu_0 \frac{\partial E}{\partial t}$$

Maxwell Equations

Gauss Law

$$\oint_S \mathbf{E} \cdot \hat{\mathbf{n}} \, dS = \frac{q}{\epsilon_0} \quad \nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}$$

**Gauss Law
of magnetism**

$$\oint_S \mathbf{B} \cdot \hat{\mathbf{n}} \, dS = 0 \quad \nabla \cdot \mathbf{B} = 0$$

Faraday's Law

$$\oint_l \mathbf{E} \cdot d\mathbf{l} = -\frac{d}{dt} \int_S \mathbf{B} \cdot \hat{\mathbf{n}} \, dS \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

**Ampere-
Maxwell's Law**

$$\oint_l \mathbf{E} \cdot d\mathbf{l} = \mu_0 \left(I + \epsilon_0 \frac{d}{dt} \int_S \mathbf{E} \cdot \hat{\mathbf{n}} \, dS \right) \quad \nabla \times \mathbf{B} = \mu_0 \left(J + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right)$$

Electromagnetic Wave equation

- EM wave eqn can be derived by using Maxwell's equation for EM wave propagation through a homogeneous, isotropic dielectric medium
- As the dielectric medium offers infinite resistance to the electric current – conductivity ‘s’ is zero i.e., $J = 0$
($J = sE$)
- In homogeneous isotropic medium, there is no volume distribution of charge, thus the charge density is Zero.
hence $J = 0$; $r = 0$; $D = e_0 e_r E$ and $B = m_0 m_r H = mH$

- Hence Maxwell's equation for a dielectric medium becomes

275

05._Electromagnetic_Wave_equation_08-Mar-2021

$$\nabla \bullet E = 0 \quad \longrightarrow \quad (1)$$

$$\nabla \bullet B = 0 \quad \longrightarrow \quad (2)$$

$$\nabla_x E = -\frac{\partial B}{\partial t} \quad \longrightarrow \quad (3)$$

$$\nabla_x B = \mu \epsilon \frac{\partial E}{\partial t} \quad \longrightarrow \quad (4)$$

Taking Curl of eqn (4), we get

$$\nabla_x \nabla_x B = \nabla_x \mu \epsilon \frac{\partial E}{\partial t}$$

$$= \mu \epsilon \left(\nabla_x \frac{\partial E}{\partial t} \right)$$

$$= \mu \epsilon \frac{\partial}{\partial t} (\nabla_x E)$$

since $\nabla_x E = -\frac{\partial B}{\partial t}$

$$= \mu \epsilon \frac{\partial}{\partial t} \left(-\frac{\partial B}{\partial t} \right) = -\mu \epsilon \frac{\partial^2 B}{\partial t^2} \quad \longrightarrow \quad (5)$$

- We have

$$\nabla \times \nabla \times \mathbf{B} = \nabla (\nabla \cdot \mathbf{B}) - \nabla^2 \mathbf{B}_{\text{Electromagnetic Wave equation 08-Mar-2021}}$$

$$\nabla(0) - \nabla^2 \mathbf{B}$$

therefore

$$\nabla \times \nabla \times \mathbf{B} = -\nabla^2 \mathbf{B} \longrightarrow (6)$$

Substituting eqn (6) in (5) we get

$$-\nabla^2 \mathbf{B} = -\mu \epsilon \left(\frac{\partial^2 \mathbf{B}}{\partial t^2} \right)$$

or

$$\nabla^2 \mathbf{B} = \mu \epsilon \left(\frac{\partial^2 \mathbf{B}}{\partial t^2} \right) \longrightarrow (7)$$

Similarly from eqn (3) we can show that

$$\nabla^2 \mathbf{E} = \mu \epsilon \left(\frac{\partial^2 \mathbf{E}}{\partial t^2} \right) \longrightarrow (8)$$

Eqn (7 & 8) represents the relation between the space and time variation of magnetic field \mathbf{B} and electric field \mathbf{E} . –Wave equations for \mathbf{B} and \mathbf{E} resp.

- The above eqns are similar to general form of differential equation of wave motion –

given by

$$\nabla^2 y = \frac{1}{v^2} \frac{\partial^2 y}{\partial t^2} \quad \longrightarrow \quad (9)$$

where v is the velocity of the wave and y in its amplitude

Comparing eqn (8 & 9) $\mu\varepsilon$ and $1/v^2$ has the same significance

So we find that variations of E and B are propagated in homogeneous, isotropic medium with a velocity given by

$$1/v^2 = \mu\varepsilon \text{ or } v^2 = 1/\mu\varepsilon$$

Therefore

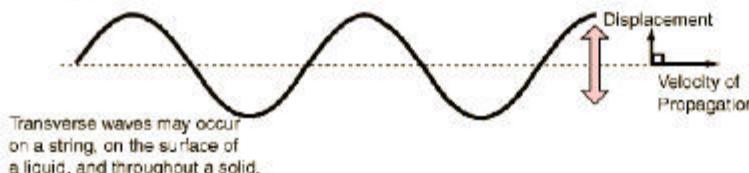
$$v = 1/(\mu\varepsilon)^{1/2}$$

Where μ and ε are permeability and permittivity of the medium

For free space

$$v = 1/(\mu_0\varepsilon_0)^{1/2} = 3 \times 10^8 \text{ m/s}$$

Why light waves are transverse



Suppose a wave propagates in the x -direction. Then it's a function of x and t (and not y or z), so all y - and z -derivatives are zero:

$$\frac{\partial E_y}{\partial y} = \frac{\partial E_z}{\partial z} = \frac{\partial B_y}{\partial y} = \frac{\partial B_z}{\partial z} = 0$$

In a charge-free medium,

$$\vec{\nabla} \cdot \vec{E} = 0 \text{ and } \vec{\nabla} \cdot \vec{B} = 0$$

that is,

$$\frac{\partial E_x}{\partial x} + \cancel{\frac{\partial E_y}{\partial y}} + \cancel{\frac{\partial E_z}{\partial z}} = 0 \quad \frac{\partial B_x}{\partial x} + \cancel{\frac{\partial B_y}{\partial y}} + \cancel{\frac{\partial B_z}{\partial z}} = 0$$

Substituting the zero values, we have:

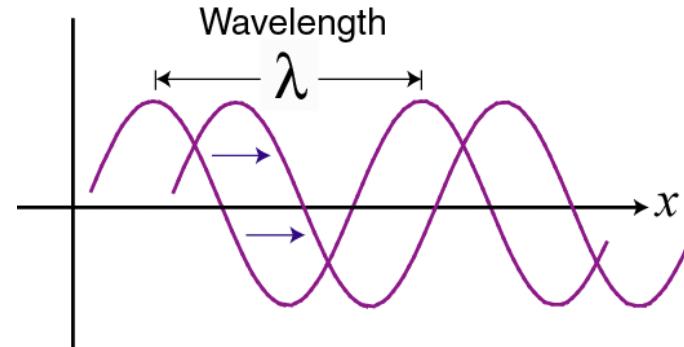
$$\frac{\partial E_x}{\partial x} = 0 \text{ and } \frac{\partial B_x}{\partial x} = 0$$

So the longitudinal fields (parallel to propagation direction) are at most **constant**, and not waves.

The Phase Velocity

How fast is the wave traveling?

Velocity is a reference distance divided by a reference time.



The phase velocity is the wavelength / period: $v = \lambda / \tau$

Since $f = 1/\tau$:

$$v = \lambda f$$

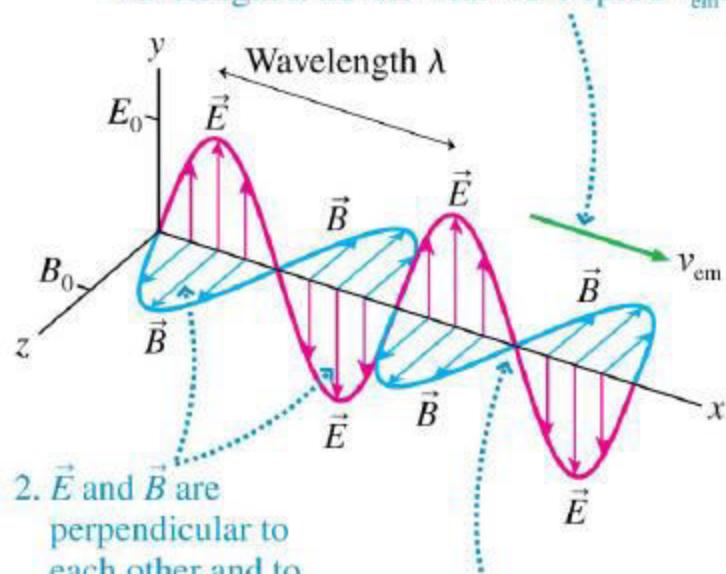
In terms of k , $k = 2\pi / \lambda$, and the angular frequency, $\omega = 2\pi / \tau$, this is:

$$v = \omega / k$$

The propagation direction of a light wave

FIGURE 35.19 A sinusoidal electromagnetic wave.

1. A sinusoidal wave with frequency f and wavelength λ travels with wave speed v_{em} .



2. \vec{E} and \vec{B} are perpendicular to each other and to the direction of travel. The fields have amplitudes E_0 and B_0 .

3. \vec{E} and \vec{B} are in phase. That is, they have matching crests, troughs, and zeros.

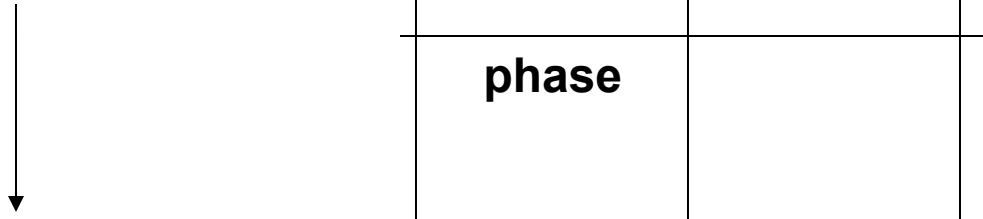
$$\vec{v} = \vec{E} \times \vec{B}$$

Right-hand screw rule

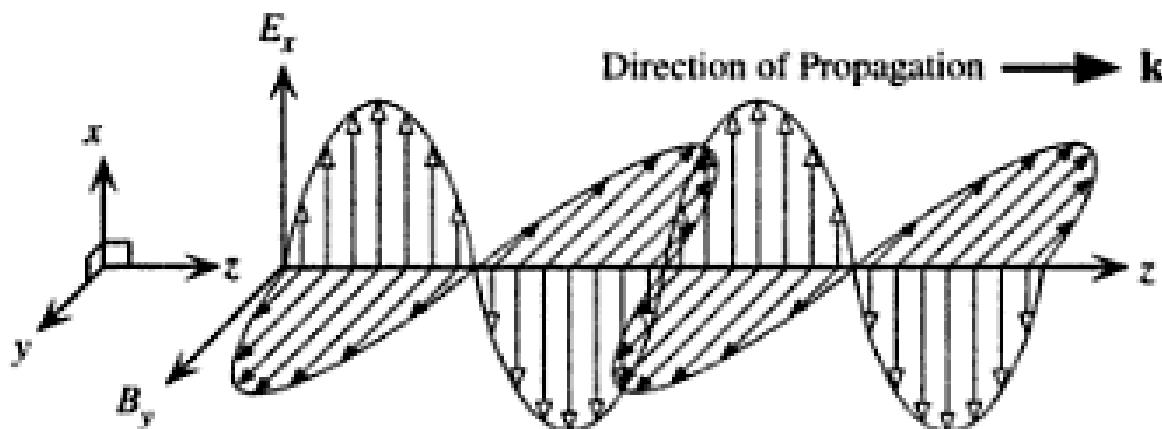
The solution of electric field wave equation is²⁸¹

05_Group_Velocity_09-Mar-2021

$$E(z, t) = E_o \cos(\omega t - kz + \phi_0)$$



Amplitud e
Angular frequency
Propagation constant or wave number
Phase constant



$$E(z,t) = E_o \cos(\overset{282}{\omega} t - kz + \phi_0)$$

05_Group_Velocity_09-Mar-2021

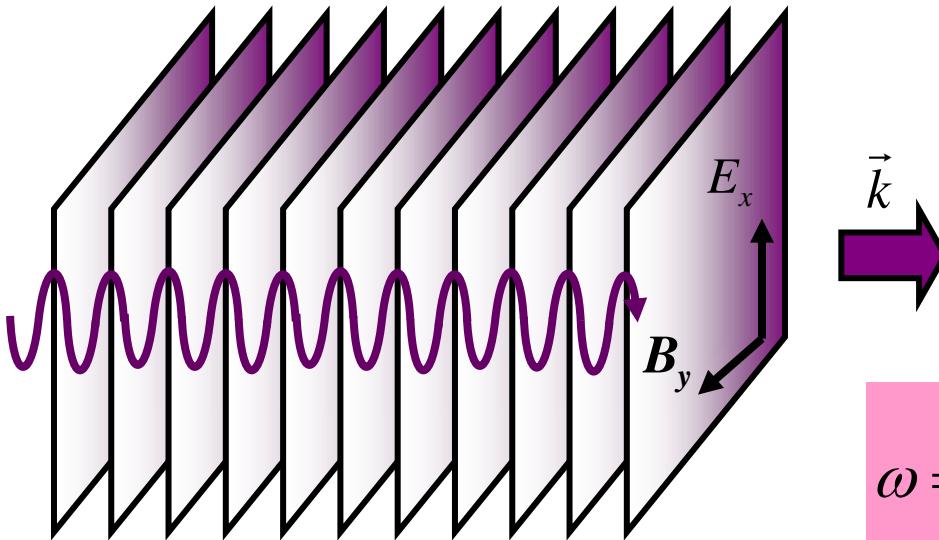
It represents a monochromatic plane wave of infinite extent traveling in the positive z direction.

Wave front: any plane perpendicular to the direction of propagation (z axis), the phase of the wave is constant. A surface over which the phase of the wave is constant is known as wave front.

The velocity with which the constant phase moves is called phase velocity

We can get the expression for phase velocity follows

$$\omega dt - kz = 0$$



Phase velocity,

$$V_p = \frac{dz}{dt} = \frac{\omega}{k}$$

$$\omega = \left| \frac{\partial \phi(z,t)}{\partial t} \right|$$

$$k = \left| \frac{\partial \phi(z,t)}{\partial z} \right|$$

Group velocity

Consider the superposition of two monochromatic waves with slightly different angular frequency

$$\omega_0 + \Delta\omega \text{ and } \omega_0 - \Delta\omega$$

Corresponding wave vectors

$$k_0 + \Delta k \text{ and } k_0 - \Delta k$$

Resulting signal

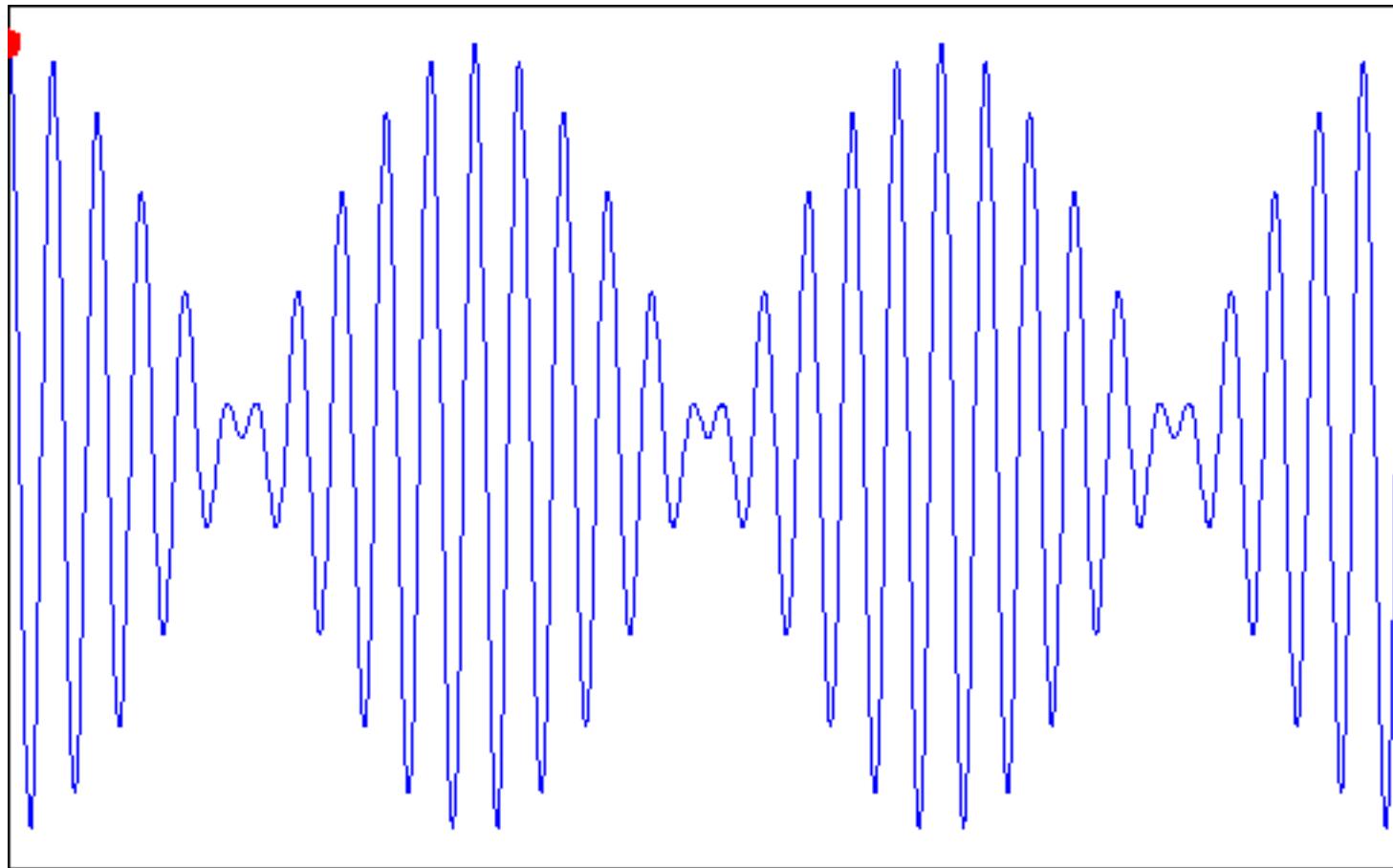
$$E_g = E_0 \cos[(\omega_0 + \Delta\omega)t - (k_0 + \Delta k)z] + E_0 \cos[(\omega_0 - \Delta\omega)t - (k_0 - \Delta k)z]$$

$$\cos[A] + \cos[B] = 2 \cos\left[\frac{A+B}{2}\right] \cos\left[\frac{A-B}{2}\right]$$

$$\begin{aligned}
& \cos[(\omega_0 + \Delta\omega)t - (k_0 + \Delta k)z]^{284} + \cos[(\omega_0 - \Delta\omega)t - (k_0 - \Delta k)z]^{05.(\text{Group Velocity})-09-Mar-2021} \\
&= 2 \cos \left[\frac{(\omega_0 + \Delta\omega)t - (k_0 + \Delta k)z + (\omega_0 - \Delta\omega)t - (k_0 - \Delta k)z}{2} \right] \\
&\quad \times \cos \left[\frac{(\omega_0 + \Delta\omega)t - (k_0 + \Delta k)z - (\omega_0 - \Delta\omega)t + (k_0 - \Delta k)z}{2} \right] \\
&= 2 \cos \left[\frac{2(\omega_0 t - k_0 z)}{2} \right] \times \cos \left[\frac{2(\Delta\omega t - \Delta k z)}{2} \right] \\
&= 2 \cos[\omega_0 t - k_0 z] \cos[\Delta\omega t - \Delta k z]
\end{aligned}$$

It is considered as a wave of frequency ω_0 . its amplitude is modulated by the function $\cos(\Delta\omega t - \Delta k z)$ having angular frequency $\Delta\omega$ and phase constant Δk .

The Group Velocity

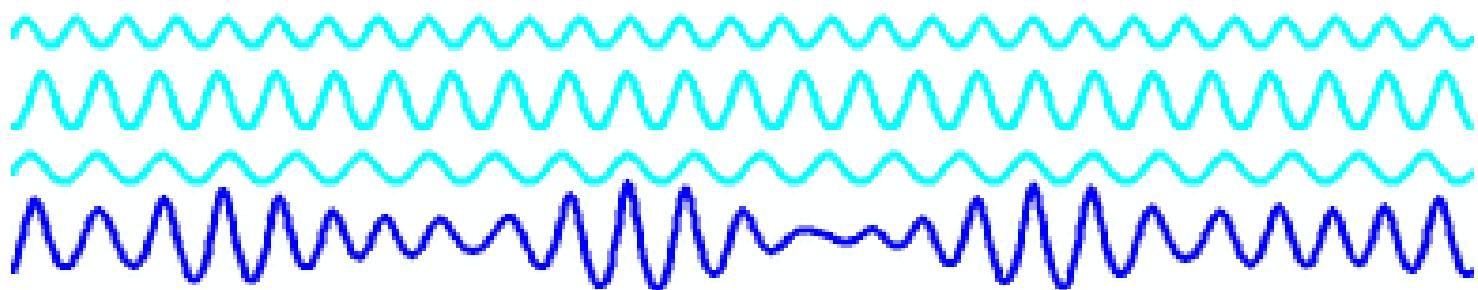
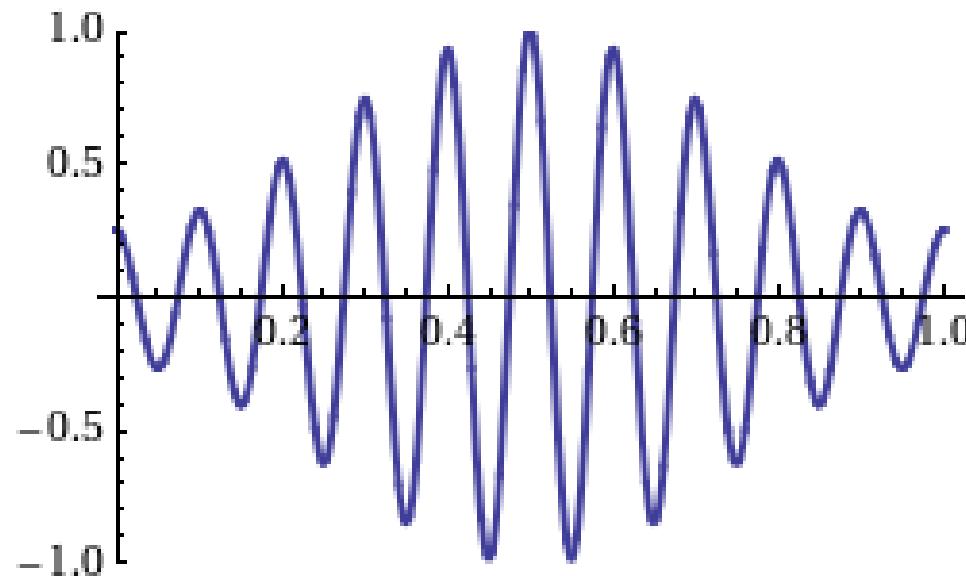


This is the velocity at which the overall shape of the wave's amplitudes, or the wave 'envelope', propagates. (= *signal velocity*)

Here, phase velocity = group velocity (the medium is *non-dispersive*)

Group velocity

Consider the superposition of two monochromatic wave with slightly different angular frequency

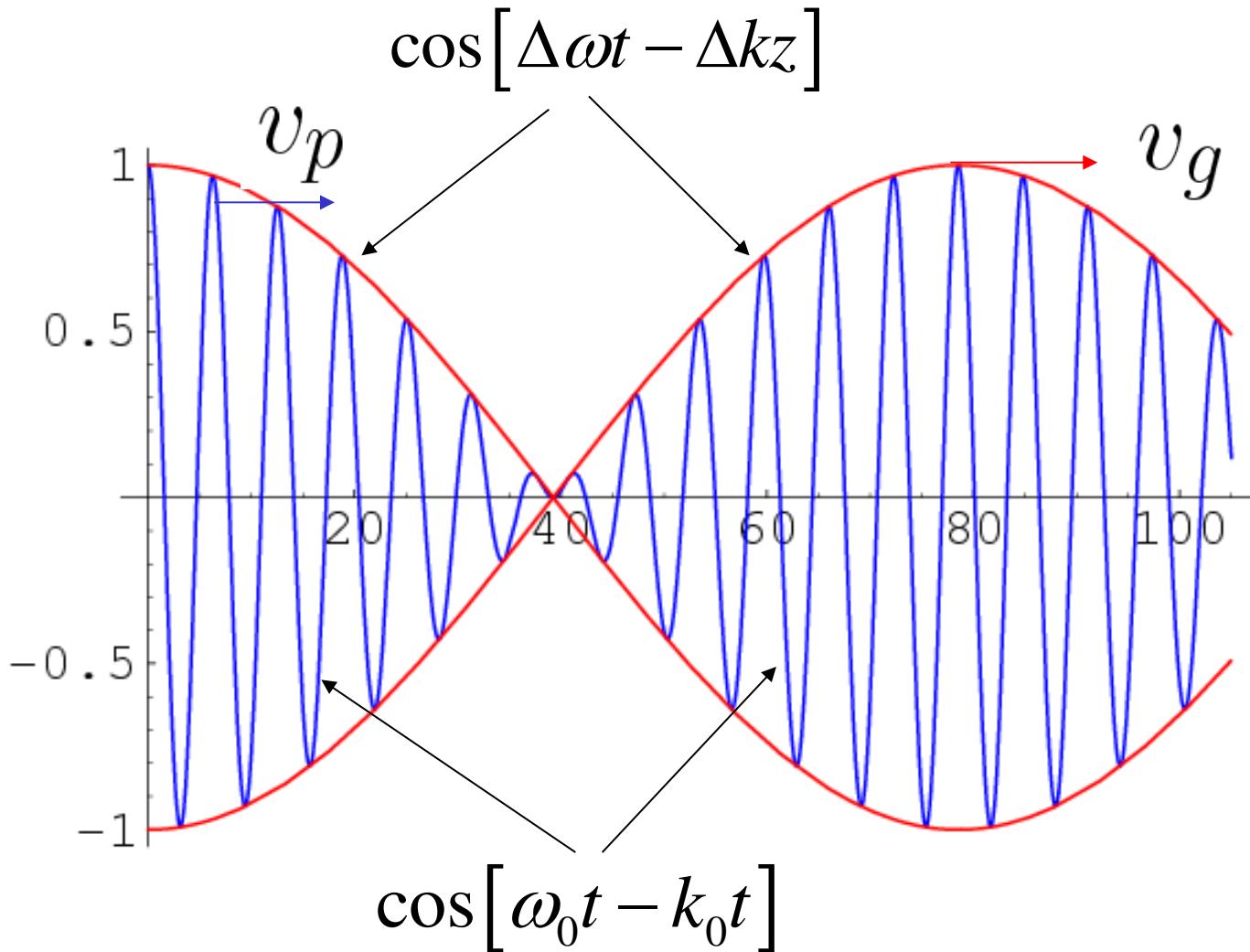


$$E_g = 2 \cos[\omega_0 t - k_0 t] \cos[\Delta\omega t - \Delta k z]$$

05_Group_Velocity_09-Mar-2021

The velocity with which the modulated function moves in space is called group velocity.

$$V_g = \frac{d\omega}{dk}$$



Refractive index is defined as

$$n = \frac{\text{Velocity of light}_{\text{vacuum}}}{\text{Velocity of light}_{\text{medium}}}$$

$$n = \frac{c}{V_p} \quad V_p = \frac{c}{n} = \frac{\omega}{k}$$

$$ck = n\omega$$

$$c \frac{d}{d\omega}(k) = \frac{d}{d\omega}(n\omega)$$

$$\frac{dk}{d\omega} = \frac{1}{c} \frac{d}{d\omega}(n\omega)$$

$$\text{Group velocity}, V_g = \frac{d\omega}{dk}$$

$$\frac{1}{V_g} = \frac{dk}{d\omega}$$

$$\frac{1}{V_g} = \frac{dk}{d\omega} = \frac{1}{c} \frac{d}{d\omega}(n\omega)$$

$$\frac{1}{V_g} = \frac{dk}{d\omega} = \frac{1}{c} \left[n \frac{d\omega}{d\omega} + \omega \frac{dn}{d\omega} \right]$$

$$\frac{1}{V_g} = \frac{dk}{d\omega} = \frac{1}{c} \left[n + \omega \frac{dn}{d\omega} \right]$$

$$\frac{1}{V_g} = \frac{dk}{d\omega} = \frac{1}{c} \left[n + \omega \frac{dn}{d\omega} \right]$$

$$\frac{1}{V_g} = \frac{dk}{d\omega} = \frac{\left[n + \omega \frac{dn}{d\omega} \right]}{c} = \frac{n_g}{c}$$

$$n_g = n + \omega \frac{dn}{d\omega}$$

**Where n_g – Group Index
In terms of wavelength
Group Index is written as**

$$\omega \frac{dn}{d\omega} = \frac{2\pi c}{\lambda} \frac{dn}{d\left(\frac{2\pi c}{\lambda}\right)}$$

$$= \frac{(2\pi c)}{\lambda} \frac{dn}{(2\pi c)d(\lambda^{-1})} = \frac{1}{\lambda} \frac{dn}{d(\lambda^{-1})}$$

$$\omega \frac{dn}{d\omega} = \frac{1}{\lambda} \frac{dn}{(-)\lambda^{-2} d\lambda}$$

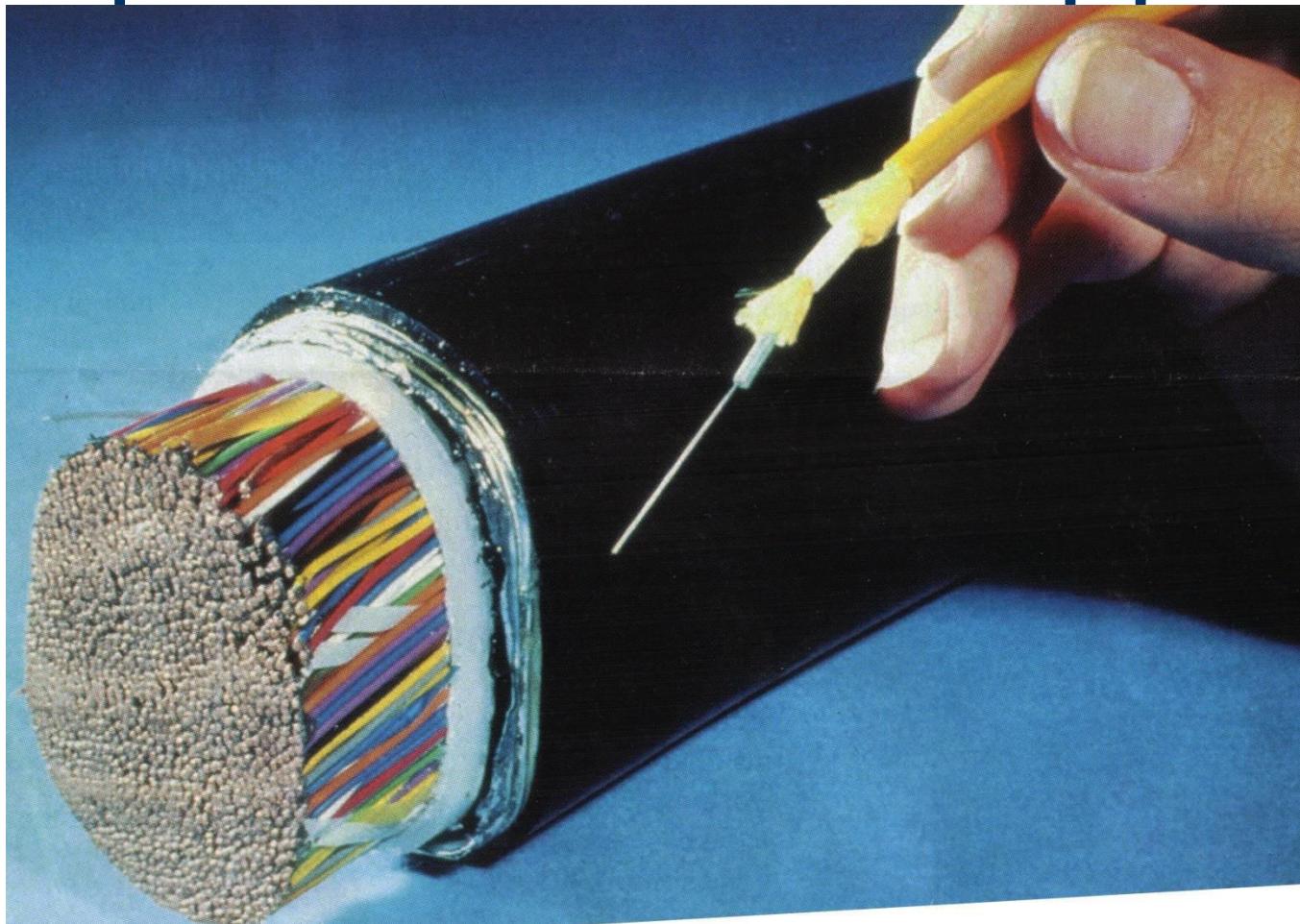
$$\omega \frac{dn}{d\omega} = -\lambda \frac{dn}{d\lambda}$$

$$n_g = n - \lambda \frac{dn}{d\lambda}$$

Introduction to Fiber Optics



Optical fiber vs Copper



What are optical fibers

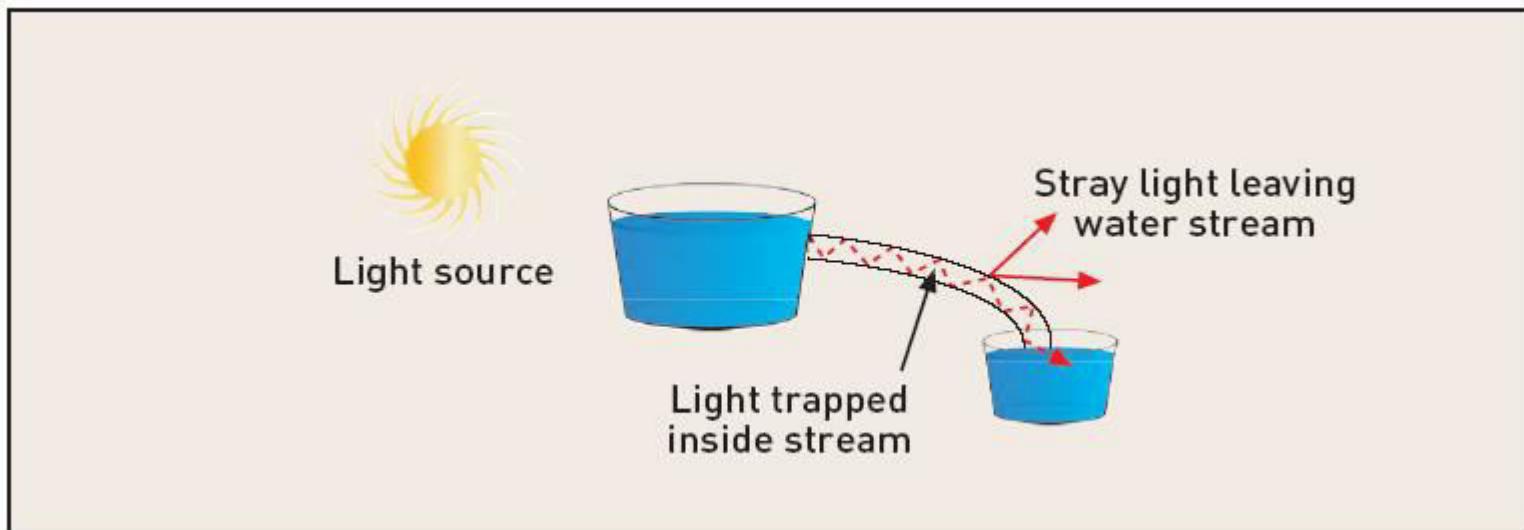
- Thin strands of pure glass
- Carry data over long distances
- At very high speeds
- Fiber can be bent or twisted

Background on Optical Communications

06._TIR_Chandar_15-Mar-2021

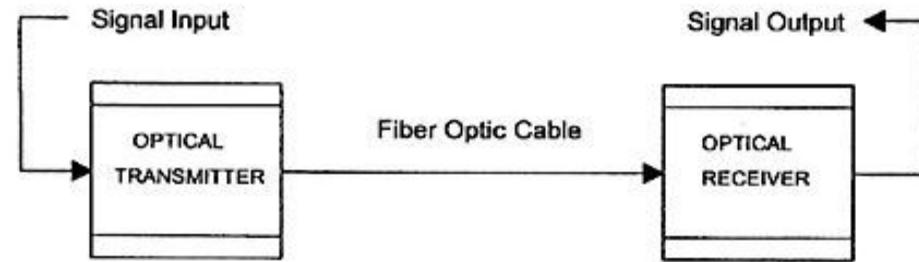
- Fiber-optic communication
 - Trapping light inside an optical fiber
 - Can carry any form of information
 - Fiber is an optical medium, which means it is capable of transmitting light
 - Based on total internal reflection (TIR)

Tyndall's Experiment



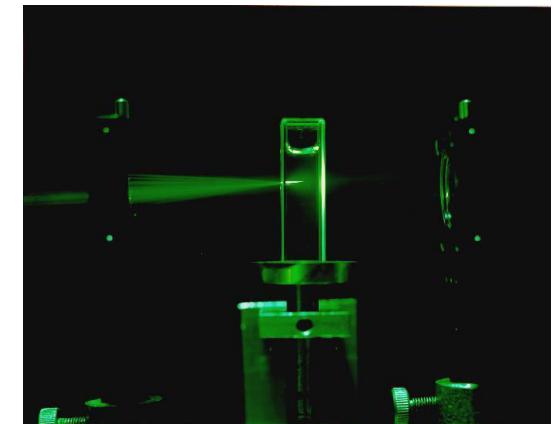
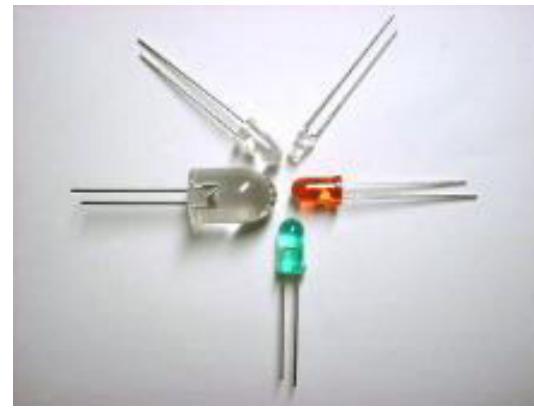
Fiber optic technology

- Sources
- Transmission medium
- Detectors



Sources of light

- Light emitting diodes
- Lasers

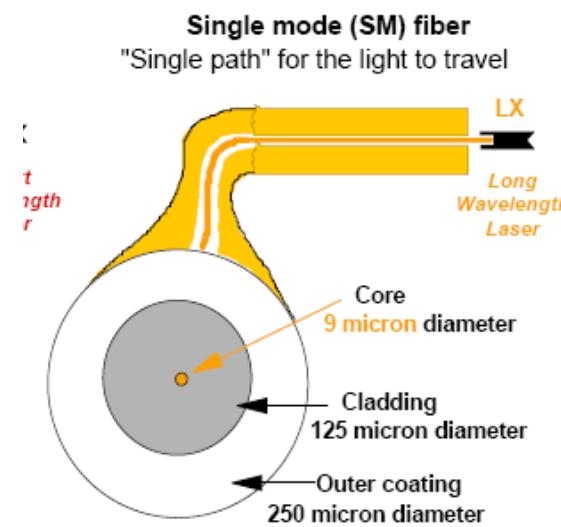
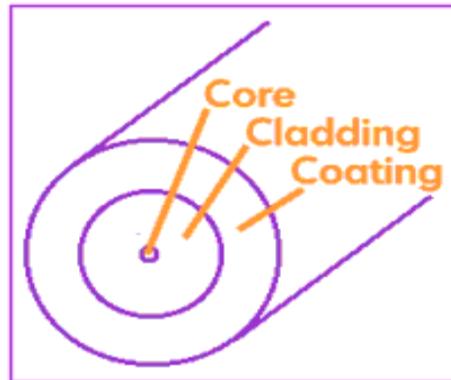


Sources

- Modulate electrical signals into optical signals
- Mostly modulate at 850nm, 1300nm and 1550 nm
- Lasers give high intensity, high frequency light
- LEDs are economical

Transmission medium

- Optical fiber is replacing copper
- Light is used as the carrier of information
- Much higher data rate



Structure of Fiber-Optic Cables – Cladding

- Cylindrical material made of glass or specialized plastic
- Central portion of the fiber
- Light signal carrying the information travels through the core
- The diameter of the core can range from a couple of micrometers (μm -one millionth of a meter) to a couple of millimeters (mm -one thousandth of a meter)

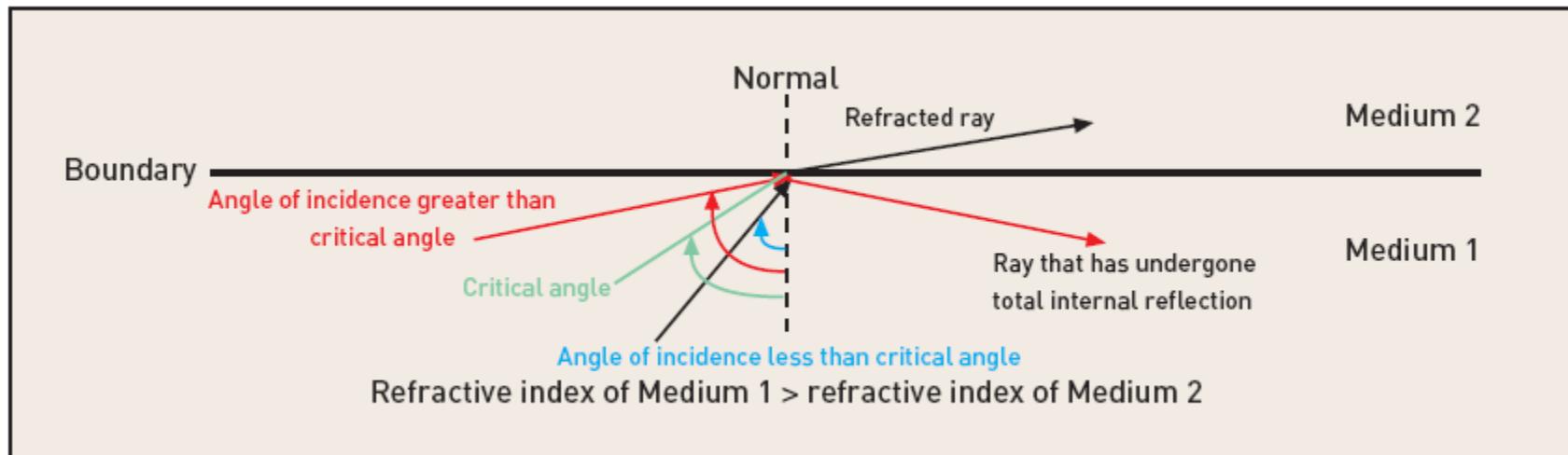
Structure of Fiber-Optic Cables – Jacket

- Surrounds the cladding
- Insulates and protects the fiber from physical damage and environmental effects, such as moisture, that might interfere with the inner workings of the cable
- Usually made of opaque plastic or another type of material

How Light Travels Through Fiber

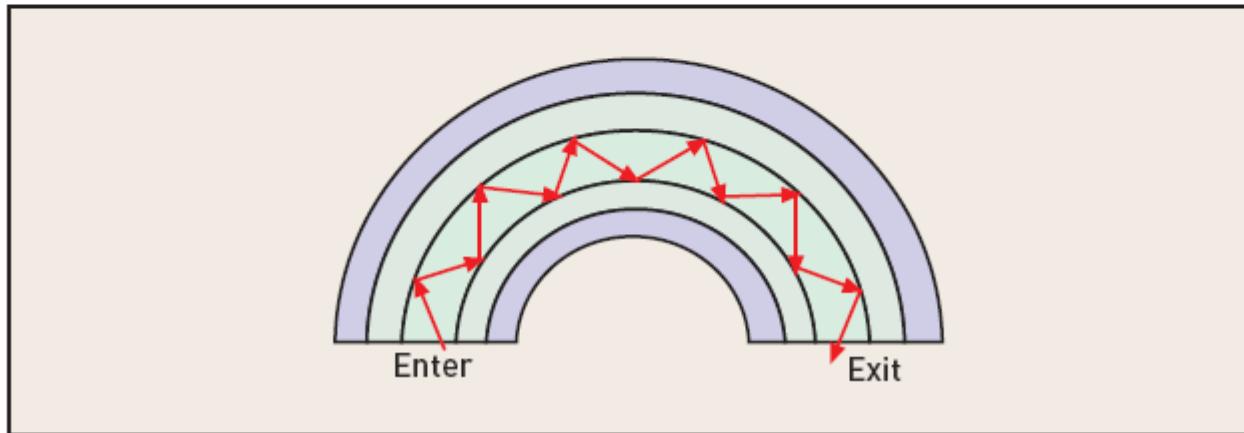
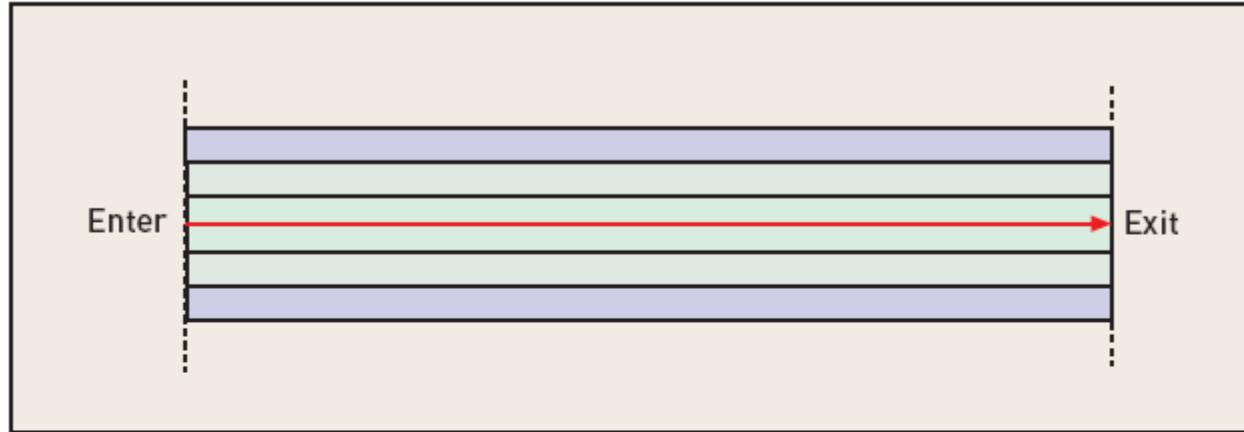
- TIR is the basis of fiber-optic communication
- TIR may be considered to be an extreme case of refraction
- When a light ray strikes a boundary of two materials with different RIs, it bends, or in other terms, refracts to an extent that depends on the ratio of the RIs of the two materials

Total Internal Reflection



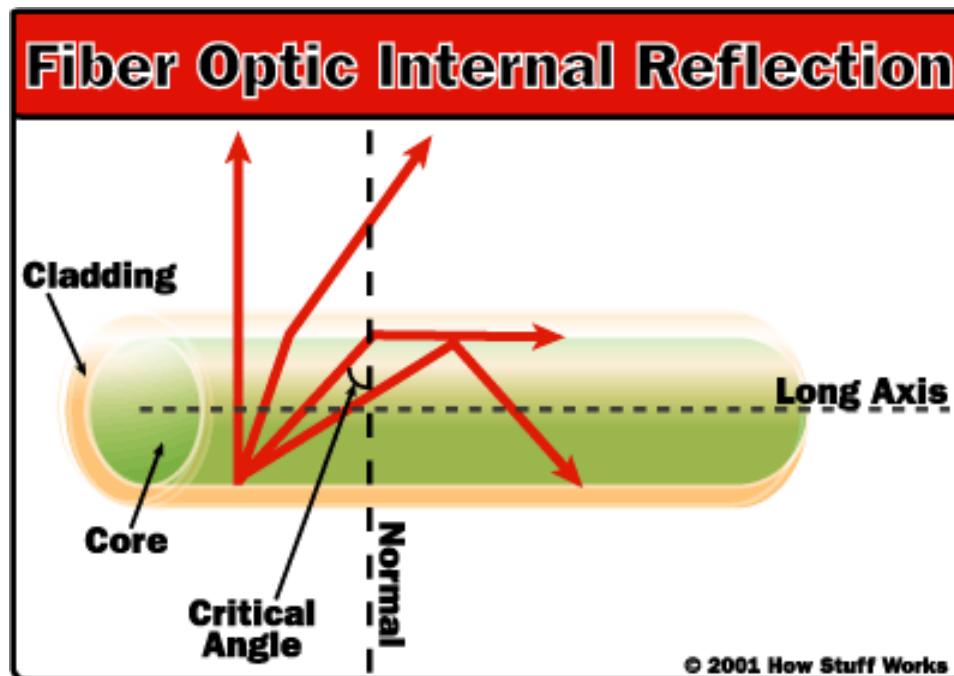
Total Internal Reflection³⁰³

(continued)



Total internal reflection

- Trapping light in the fiber



Fibers can be bent!!

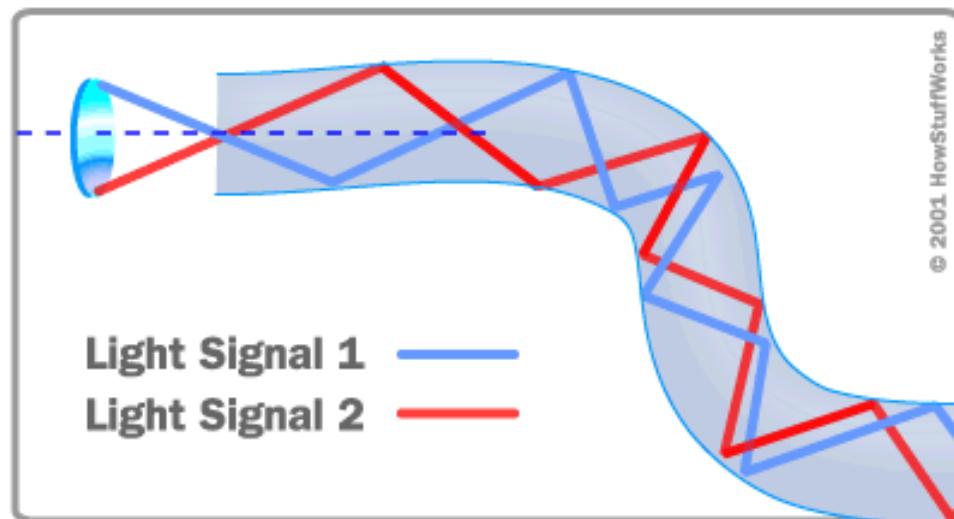


Fig: Illustration of total internal reflection

Types of optical fibers

■ Single mode

- only one signal can be transmitted
- use of single frequency

■ Multi mode

- Several signals can be transmitted
- Several frequencies used to modulate the signal

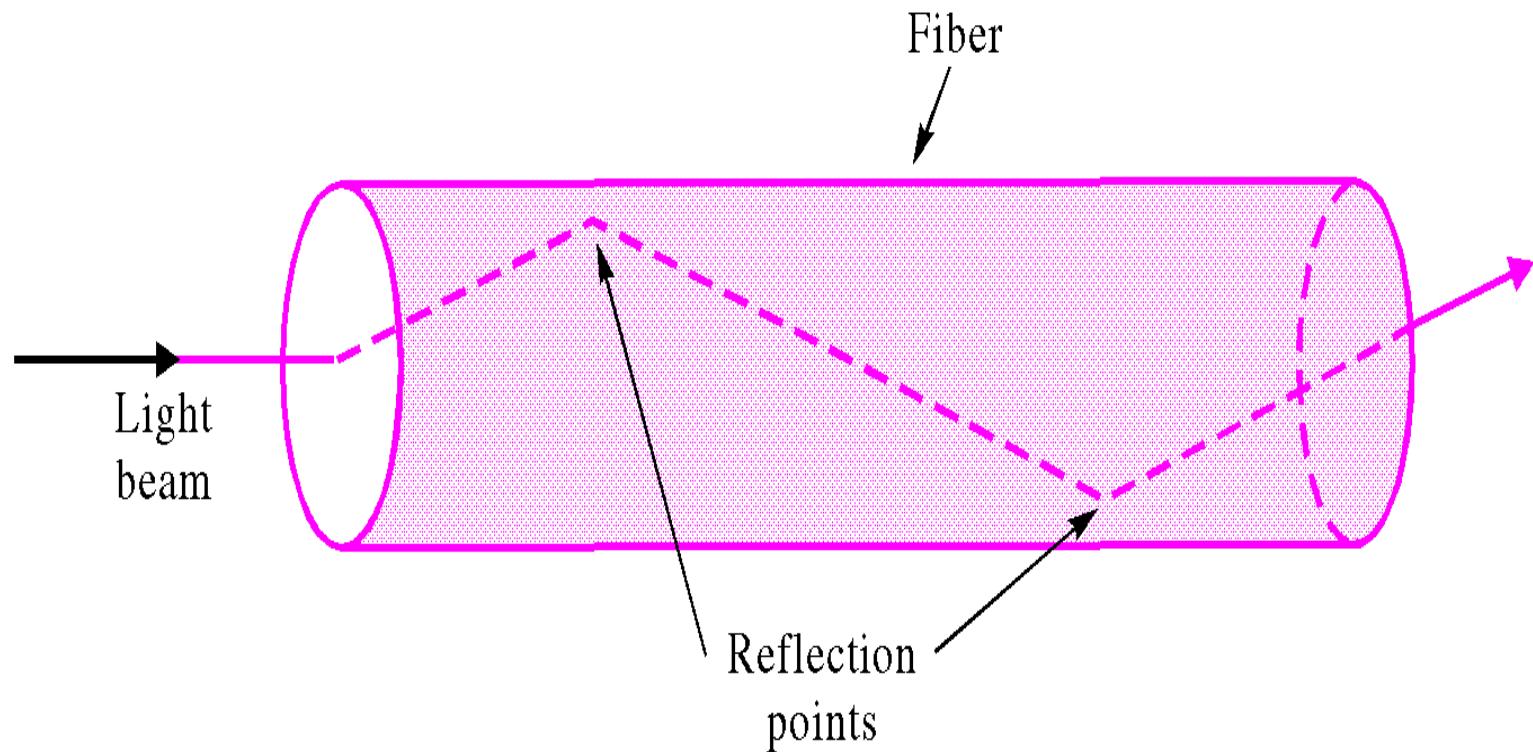
Losses in optical fibers

- Attenuation loss
- Dispersion loss
- Waveguide loss

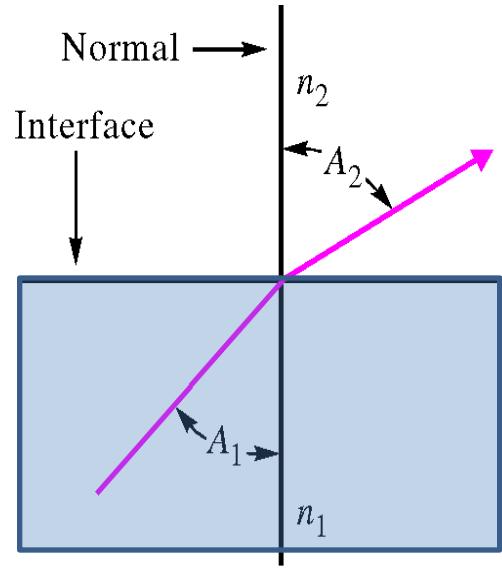
Advantages of optical fibers

- Can carry much more information
- Much higher data rates
- Much longer distances than co-axial cables
- Immune to electromagnetic noise
- Light in weight
- Unaffected by atmospheric agents

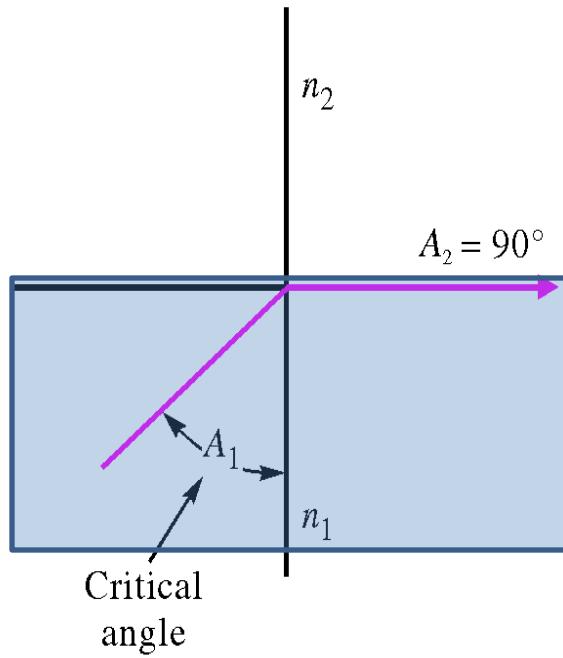
Reflection in Optical Fiber



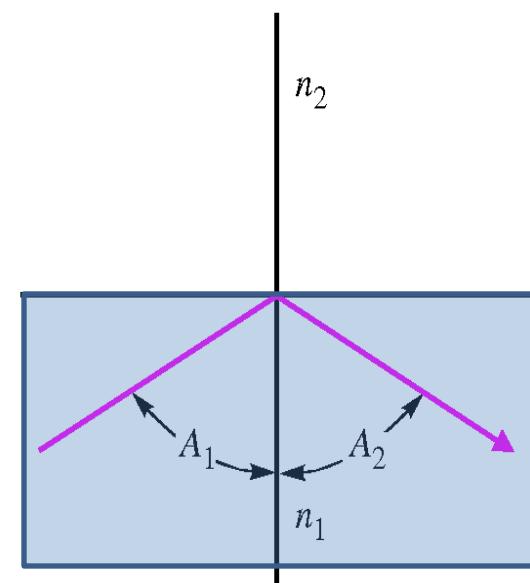
The *critical angle* is the angle of incidence that will produce a 90° angle of refraction.



(A)



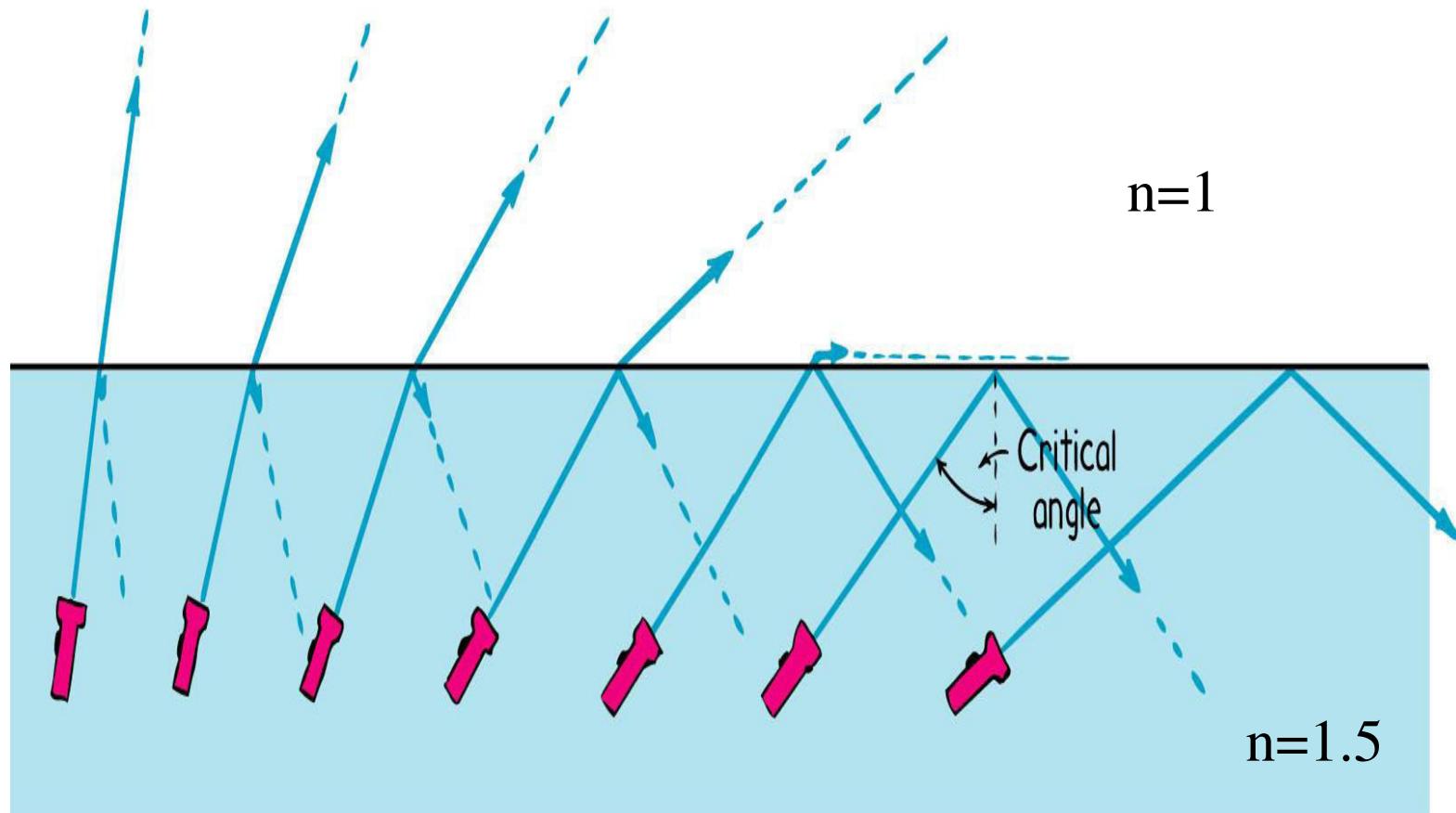
(B)



(C)

3 specific conditions are shown here. The angle of incidence, A_1 and the angle of refraction, A_2 .

Material 1 is more dense than material 2, so n_1 is greater than n_2 .

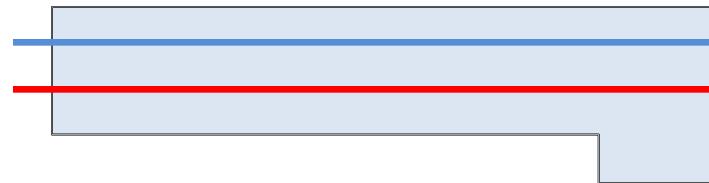


Total Internal Reflection in Fiber

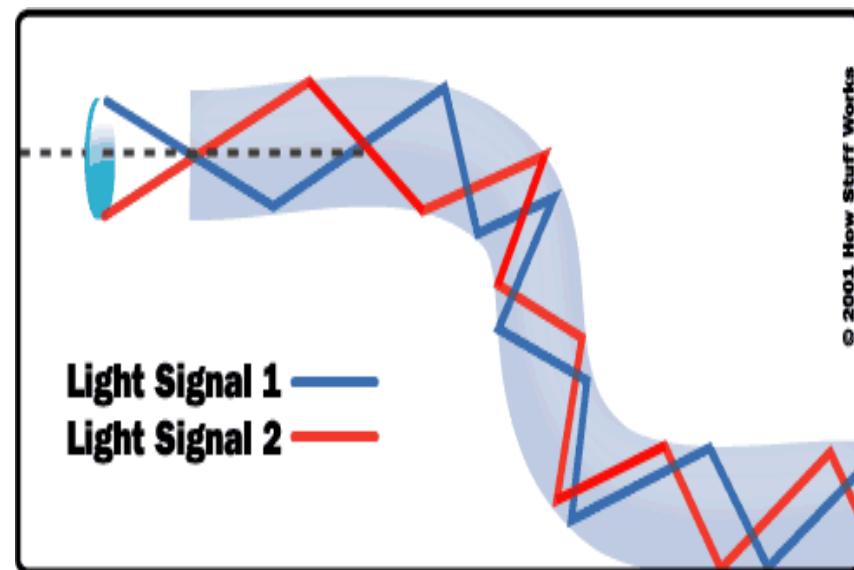
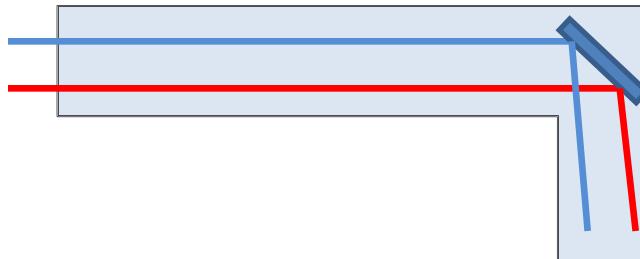
- Straight hallway



- Bent hallway



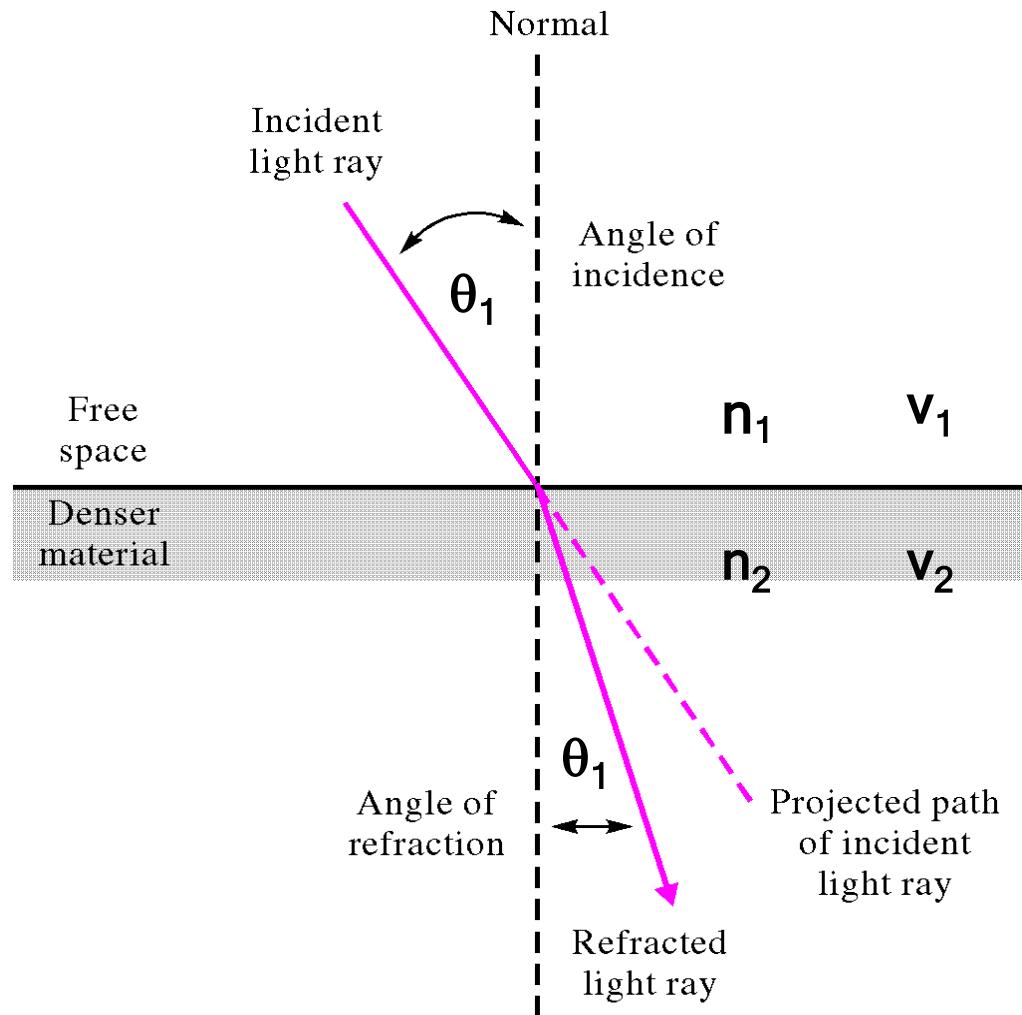
- Bent hallway with a mirror



Acceptance Angle

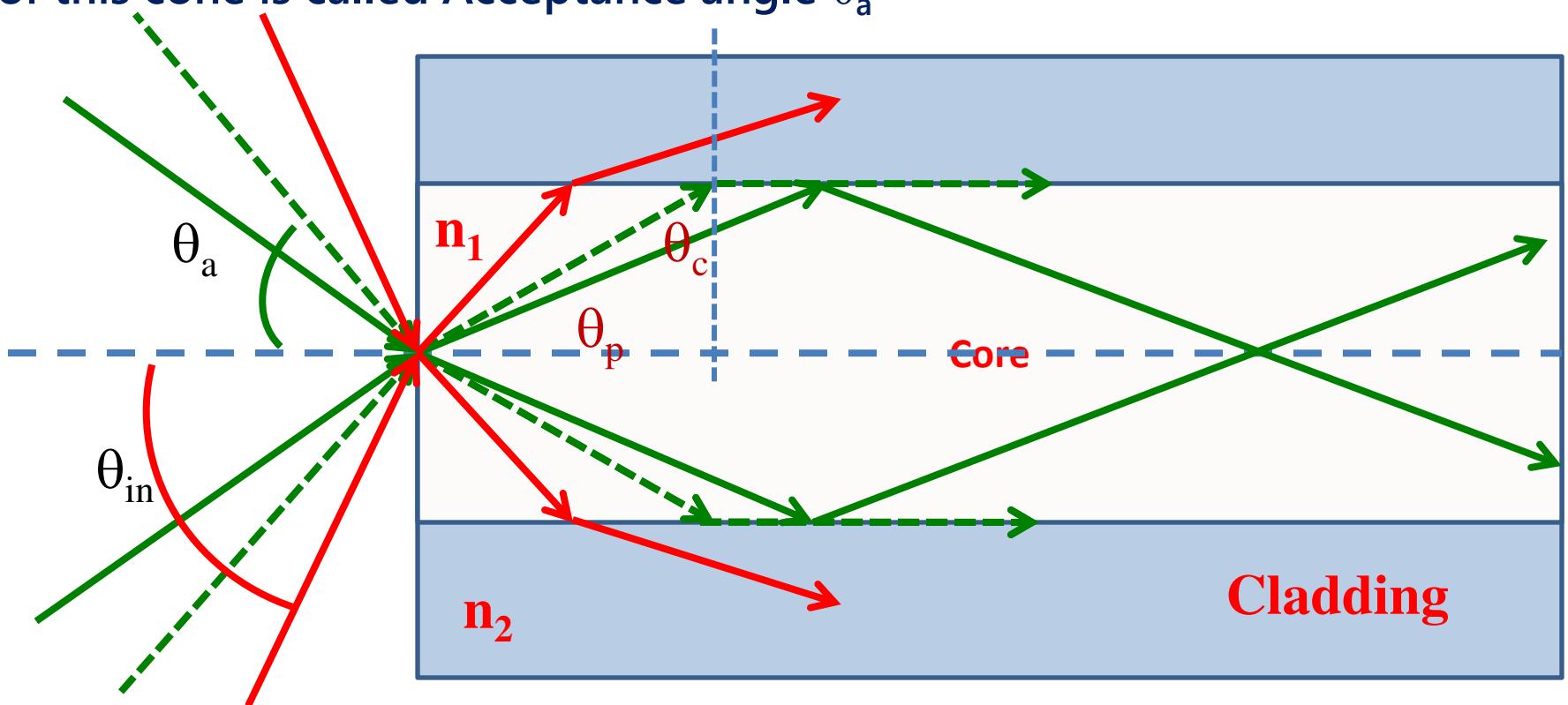
Snell's law states that the ratio of the sines of the angles of incidence and refraction is equivalent to the ratio of phase velocities in the two media, or equivalent to the opposite ratio of the indices of refraction:

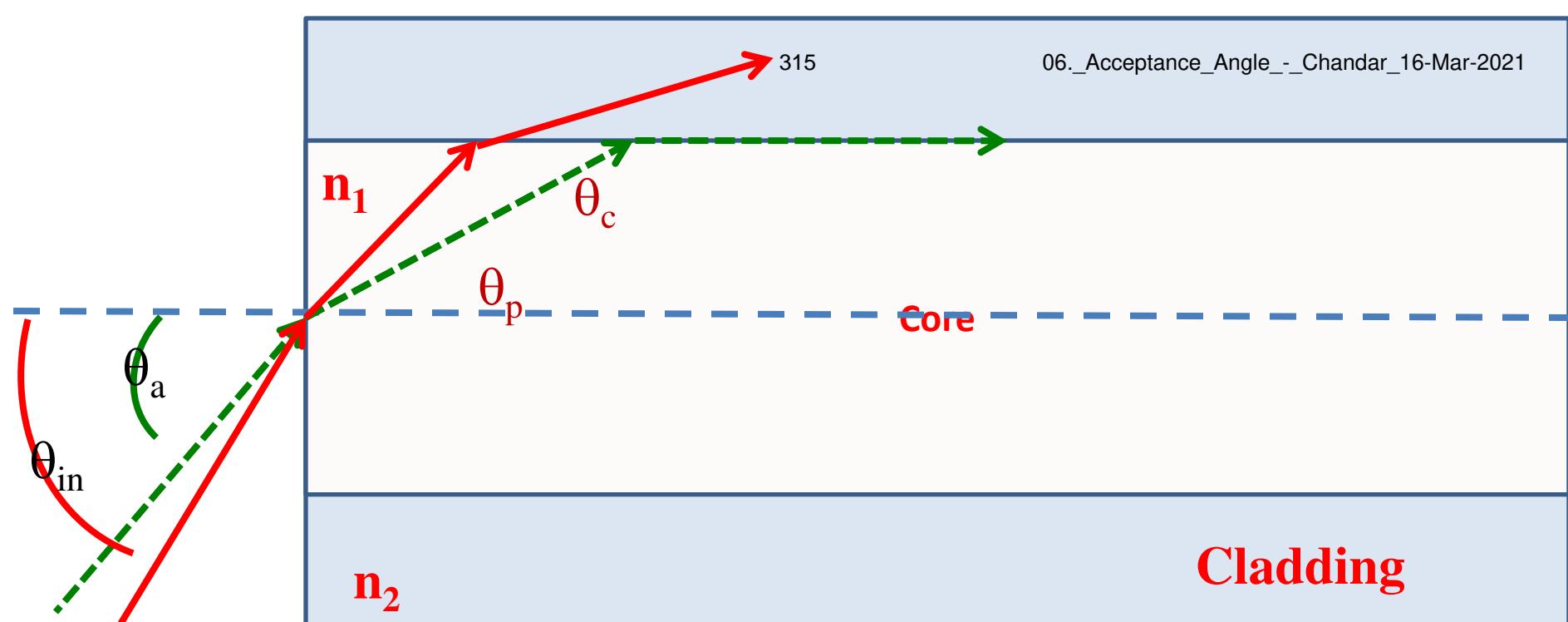
$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{v_1}{v_2} = \frac{n_2}{n_1}$$



Acceptance Angle

An Optical Fiber will only propagate light that enters the fiber within a certain cone, known as the acceptance cone of the fiber. The half angle of this cone is called Acceptance angle θ_a





To propagate the light beam down the optical fiber, the light beam at the core and cladding interface must take an angle less than the critical θ_c , From Snell's law,

$$n \sin \theta_a = n_1 \sin \alpha_c$$

$$\sin \theta_a = n_1 \sin(90 - \theta_c)$$

$$\sin \theta_a = n_1 \cos \theta_c$$

$$\sin \theta_a = n_1 \sqrt{1 - \sin^2 \theta_c}$$

from core to cladding

$$n_1 \sin \theta_c = n_2 \sin 90$$

$$\sin \theta_a = n_1 \sqrt{1 - \frac{n_2^2}{n_1^2}}$$

$$\sin \theta_a = \sqrt{n_1^2 - n_2^2}$$

θ_a – Acceptance angle

Numerical Aperture

Numerical Aperture

$$NA = \sin \theta_a$$

$$NA = \sqrt{n_1^2 - n_2^2}$$

NA describes the ability of an optical fiber to gather light signals from the sources and to preserve them within the fiber

Relative index , Δ

$$\Delta = \frac{n_1 - n_2}{n} = \frac{(n_1 - n_2)(n_1 + n_2)}{n(n_1 + n_2)} = \frac{(n_1^2 - n_2^2)}{n(n_1 + n_2)}$$

Where 'n' average index

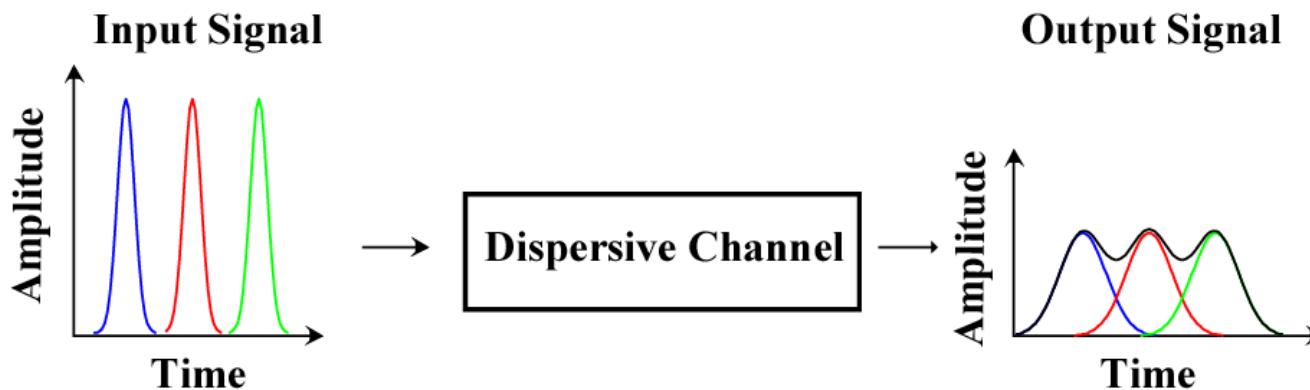
$$\Delta = \frac{(n_1^2 - n_2^2)}{2n_1^2} \longrightarrow (n_1^2 - n_2^2) = 2n_1^2 \Delta$$

$$\sqrt{n_1^2 - n_2^2} = n_1 \sqrt{2\Delta} \longrightarrow NA = n_1 \sqrt{2\Delta}$$

- Fiber optic communication system suffers from the following three major impediments
 - Dispersion
 - Attenuation
 - Nonlinear effects
- Dispersion
 - Spreading of light pulse as it travels down the length of an optical fiber

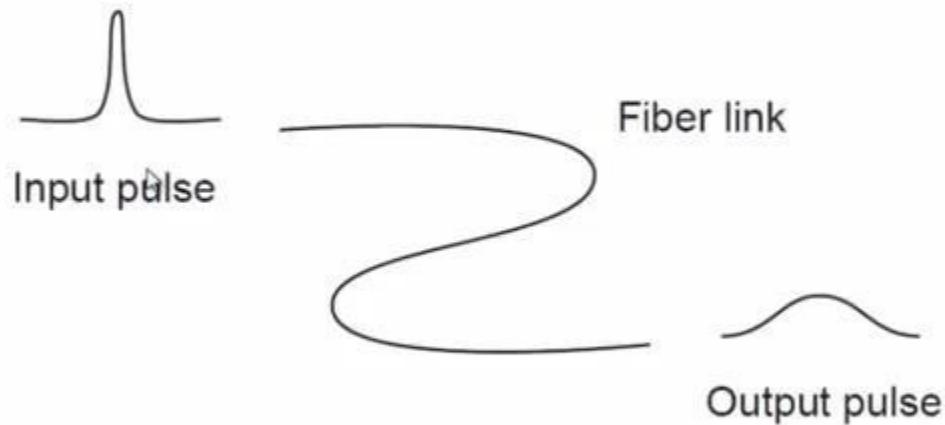
317

06_Dispersion_18-Mar-2021



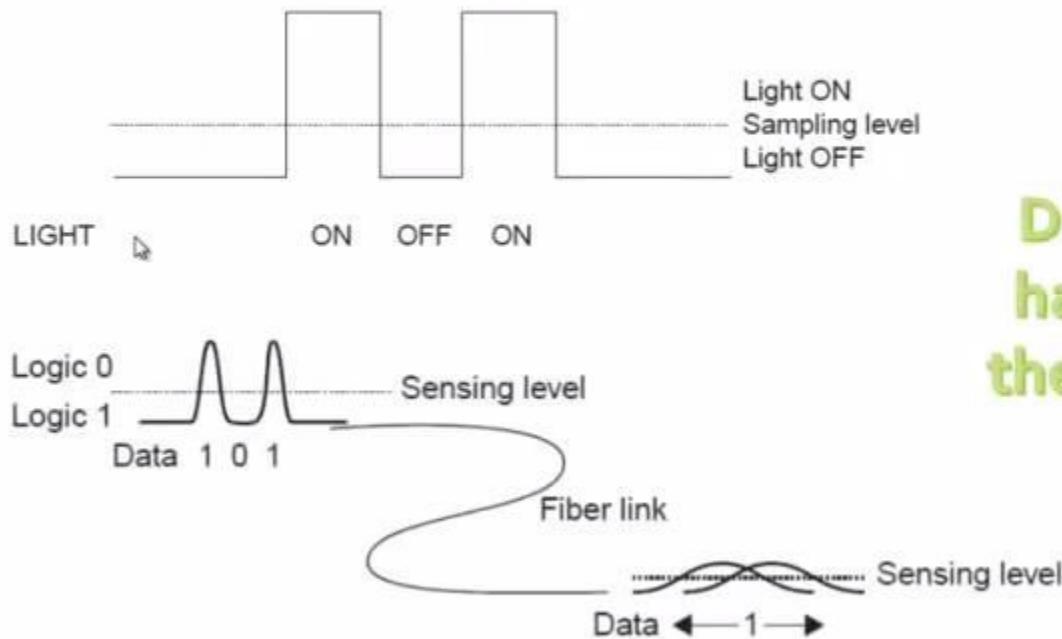
- Dispersion causes temporal pulse spreading
 - Pulse overlap results in indistinguishable data
 - Inter symbol interference (ISI)
- Dispersion is related to the velocity of the pulse

Dispersion



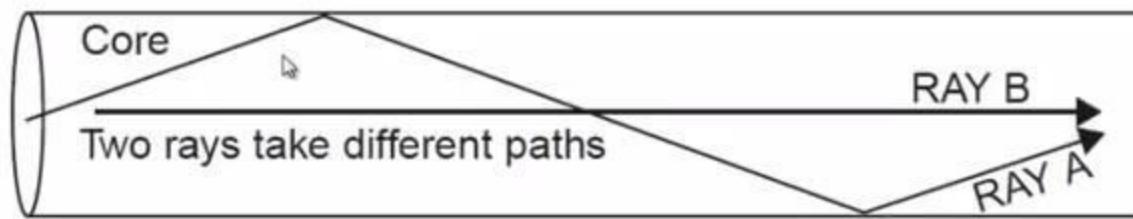
The pulse
Spreads out

The effect on the data



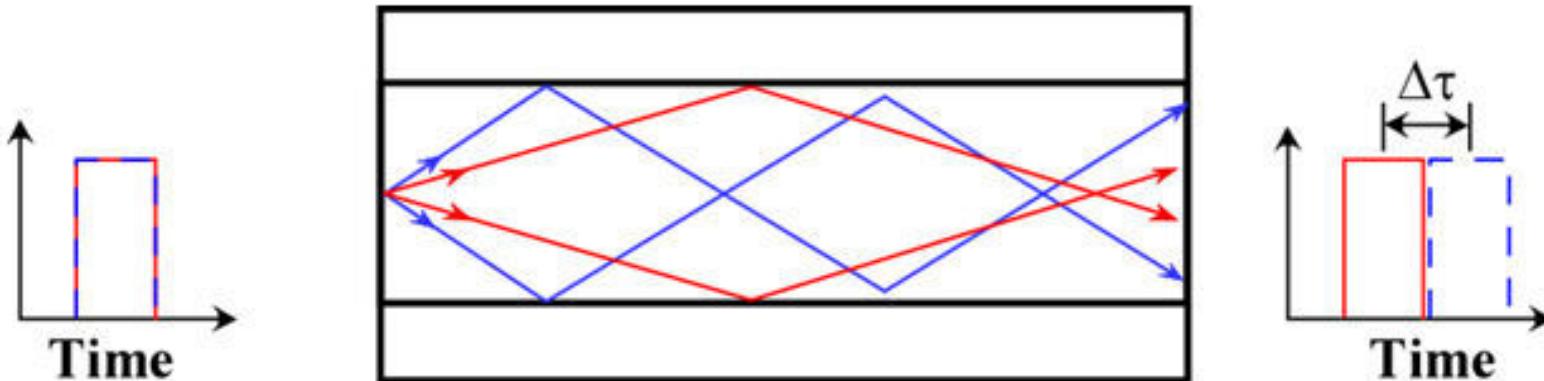
Dispersion
has caused
the pulses to
merge

Intermodal dispersion



Ray B will
arrive first

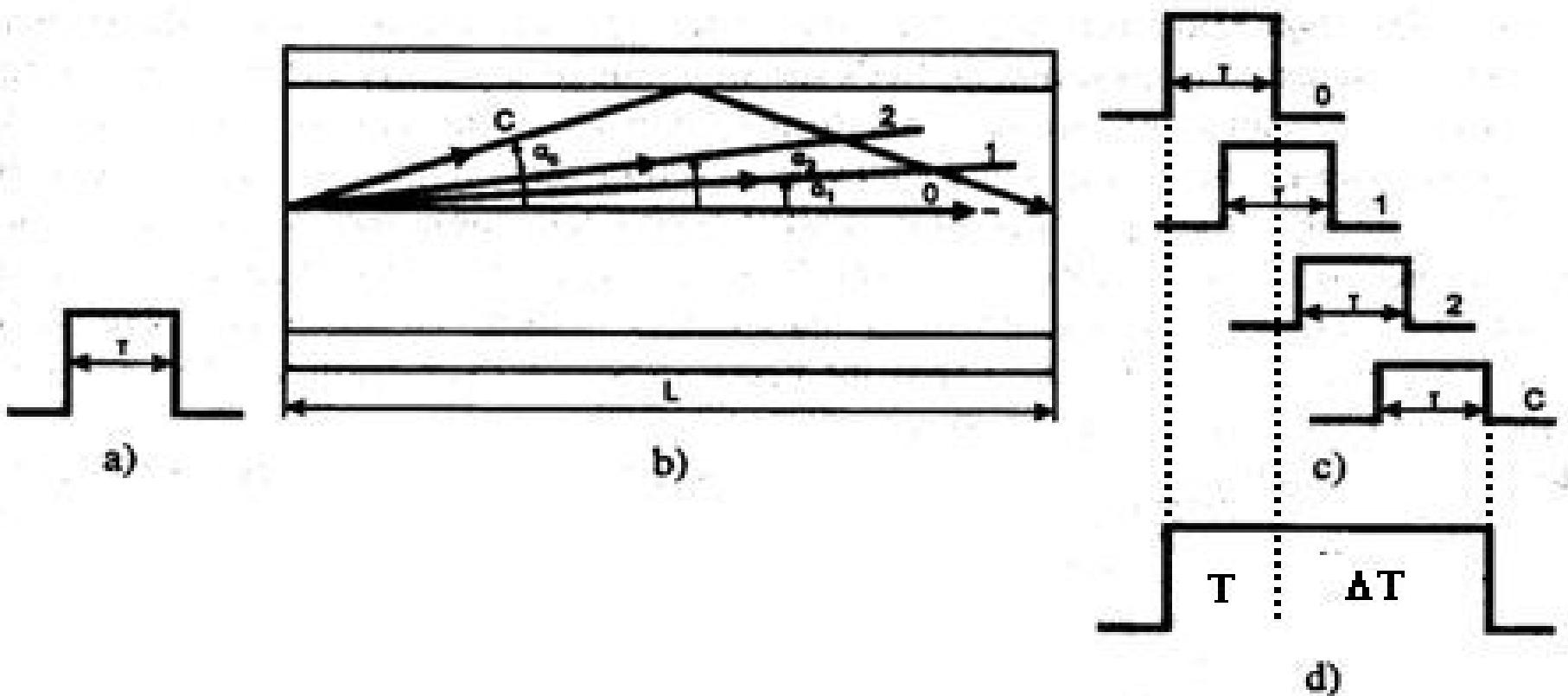
- There are two major types of dispersion in fiber-optics
 - Intermodal
 - Intramodal
- Intermodal
 - When an optical pulse is launched into the fiber, the optical pulse is distributed over all modes of fiber
 - Here we consider the propagation of light within the fiber in terms of guided electromagnetic waves called “modes”.
 - Different modes will travel with different propagation angles, hence these modes takes different routes but travel with the same velocity, but at the end of fiber they come at different timings.
 - This causes pulse widening
 - This is called intermodal dispersion or modal dispersion.



- Measuring intermodal Dispersion

322

06._Dispersion_18-Mar-2021



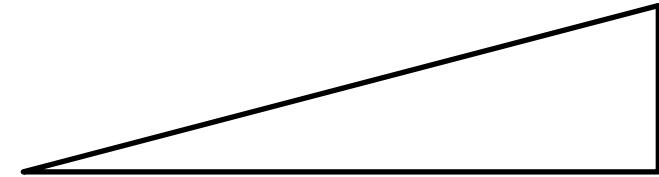
- To ascertain this let us go for some mathematical calculations

323

06_Dispersion_18-Mar-2021

- A zero order mode travelling along the fiber axis needs some time to reach the receiver it is given by

$$t_0 = L / v$$



L – length of the link

$v = c / n_1$ – velocity of light within the core

- The highest order mode propagating at critical angle needs time of

$$t_c = L / (v \cos \alpha_c)$$

Therefore, pulse widening due to intermodal dispersion is

$$\begin{aligned} \Delta t_{SI} &= t_c - t_0 \\ &= [L / (v \cos \alpha_c)] - [L / v] \\ &= L / v [(1/\cos \alpha_c) - 1] \\ &= L / v [(n_1 / n_2) - 1] \quad [\text{since } \cos \alpha_c = n_2 / n_1] \\ &= L / v [(n_1 - n_2) / n_2] \\ &= L n_1 / c [(n_1 - n_2) / n_2] \quad \text{since } v = c / n_1 \end{aligned}$$

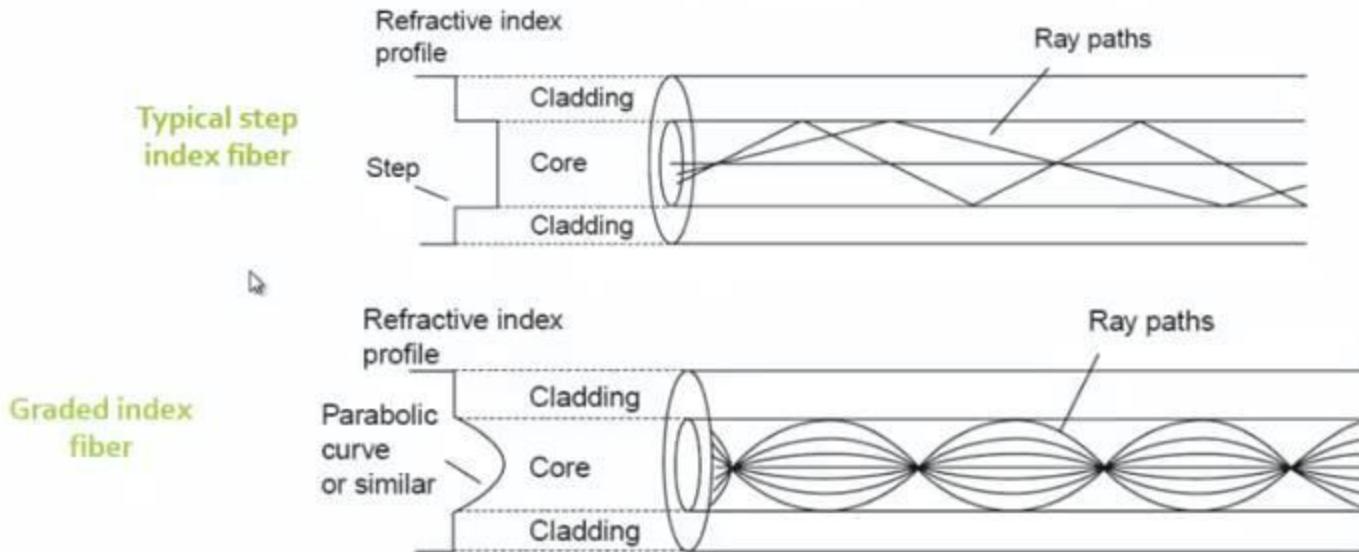
Since $n_2 \approx n$ we can write the equation

$$\Delta t_{SI} = L n_1 / c [(n_1 - n_2) / n]$$

$$\Delta t_{SI} = [L n_1 / c] (\Delta)$$

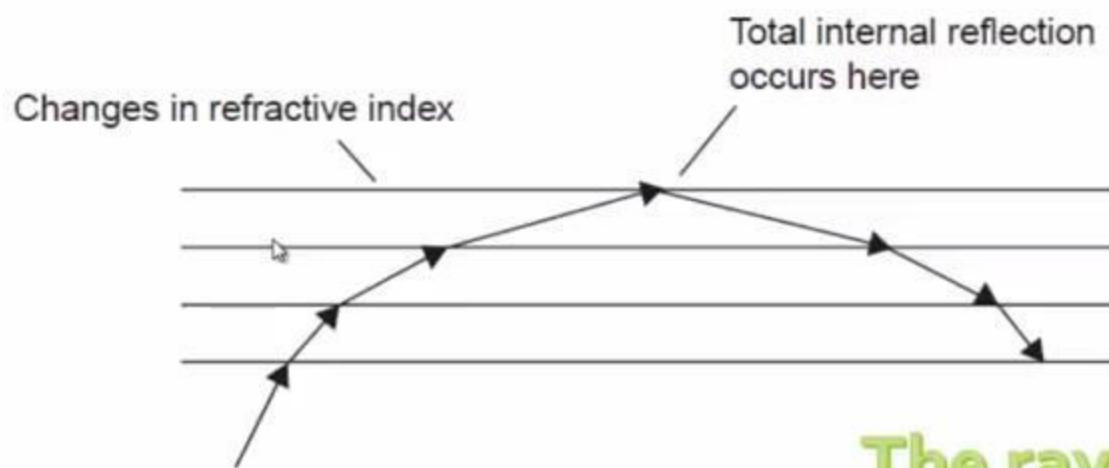
Where Δ is the relative refractive index

How to overcome intermodal dispersion



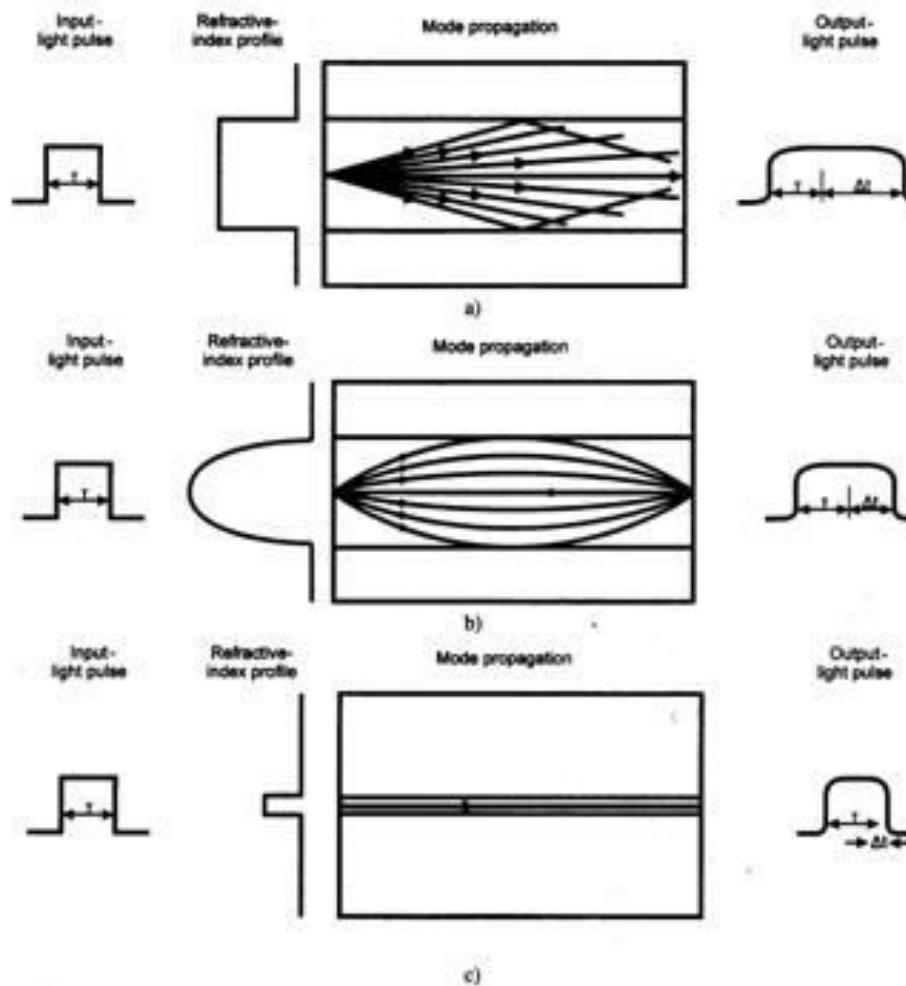
speed of light in the material = speed of light in free space / refractive index

How to overcome intermodal dispersion



**The ray is refracted
slightly at each
boundary**

- Solution to this intermodal dispersion was done with graded index fiber.
326 06_Dispersion_18-Mar-2021
 - A graded index fiber has the center of the core having highest refractive index and gradually decreasing towards the end of the core.



- It is estimated that the modal dispersion of graded index fiber is $\Delta/8$ times less than in the case of step index fiber

Transmission Characteristics

328 06. Attenuation and Dispersion 18-Mar-2021

- ❖ The transmission through an optical fiber is limited by **attenuation** (or loss) and dispersion.
- ❖ In 1970s, it was realized that the attenuation was **largely due to absorption** in the glass caused by impurities such as iron, copper, manganese etc.

Hence, research was stimulated towards a new generation of “pure” glasses for use in optical fiber communication. It lead to **silica based glass fibers** with losses less than 0.2 dB/km.

- ❖ The other characteristic is **bandwidth** which is mostly limited by signal dispersion within the fiber. It determines the number of bits of transmission transmitted in a given time period.

- ❖ Attenuation determines the **maximum transmission distance** prior to signal restoration. OFC became especially attractive when the transmission losses of fibers were reduced below those of the competing metallic conductors.
(< 5 db/km)

- ❖ Signal attenuation in optical fibers (or that of metallic cable) is usually expressed in the units of decibel. Decibel is used for comparing two power levels.

$$dB = 10 \log_{10} \frac{P_i}{P_o}$$

For a particular optical wavelength,
 $P_i \rightarrow$ input (transmitted) optical power
 $P_o \rightarrow$ output (received) optical power

In OFC, attenuation is usually expressed in dB per unit length (dB/km)

$$\alpha_{dB} L = 10 \log_{10} \frac{P_i}{P_o}$$

$\alpha_{dB} \rightarrow$ signal attenuation/length
 $L \rightarrow$ Fiber length

❖ This loss mechanism is related to material composition and the fabrication process for the fiber. Absorption of light may be :

Intrinsic: caused by the interaction with one or more of the major components of the glass

Extrinsic: caused by impurities within the glass

Intrinsic absorption

- Pure silica-based glass has *two* major intrinsic absorption mechanisms at optical wavelengths:
 - (1) a *fundamental UV absorption edge*, the peaks are centered in the *ultraviolet wavelength region*. This is due to the *electron transitions* within the glass molecules. The tail of this peak may extend into the the shorter wavelengths of the fiber transmission spectral window.
 - (2) A fundamental *infrared and far-infrared absorption edge*, due to *molecular vibrations* (such as Si-O). The tail of these absorption peaks may extend into the longer wavelengths of the fiber transmission spectral window.

Electronic and molecular absorption

- ***Electronic absorption:*** the bandgap of fused silica is about 8.9 eV (~ 140 nm). This causes strong absorption of light in the UV spectral region due to electronic transitions across the bandgap.

In practice, the bandgap of a material is not sharply defined but usually has ***bandtails*** extending from the conduction and valence bands into the bandgap due to a variety of reasons, such as *thermal vibrations of the lattice ions* and *microscopic imperfections of the material structure*.

An *amorphous* material like fused silica generally has very long bandtails. These bandtails lead to an absorption tail extending into the visible and infrared regions. Empirically, the absorption tail at photon energies below the bandgap falls off exponentially with photon energy.

- ***Molecular absorption:*** in the infrared region, the absorption of photons is accompanied by transitions between different *vibrational modes* of silica molecules.
- The *fundamental vibrational transition* of fused silica causes a very strong absorption peak at about 9 μm wavelength.
- *Nonlinear effects* contribute to important harmonics and combination frequencies corresponding to minor absorption peaks at 4.4, 3.8 and 3.2 μm wavelengths.
=> a long absorption tail extending into the near infrared, causing a sharp rise in absorption at optical wavelengths longer than 1.6 μm .

Extrinsic absorption³³⁴

06._Attenuation_and_Dispersion_18-Mar-2021

- Major extrinsic loss mechanism is caused by absorption due to water (*as the hydroxyl or OH⁻ ions*) introduced in the glass fiber during *fiber pulling by means of oxyhydrogen flame*.
- These OH⁻ ions are bonded into the glass structure and have absorption peaks (due to *molecular vibrations*) at **1.39 μm**. The fundamental vibration of the OH⁻ ions appear at 2.73 μm.
- Since these OH⁻ absorption peaks are sharply peaked, narrow spectral windows exist **around 1.3 μm and 1.55 μm which are essentially unaffected by OH⁻ absorption**.
- The lowest attenuation for typical silica-based fibers occur at **wavelength 1.55 μm at about 0.2 dB/km**, approaching the *minimum possible attenuation* at this wavelength.

Impurity absorption

- **Impurity absorption:** most impurity ions such as OH^- , Fe^{2+} and Cu^{2+} form absorption bands in the *near infrared* region where both electronic and molecular absorption losses of the host silica glass are very low.
- Near the peaks of the impurity absorption bands, an impurity concentration as low as *one part per billion* can contribute to an absorption loss as high as 1 dB km^{-1} .
- In fact, fiber-optic communications were not considered possible until it was realized in 1966 (Kao) that most losses in earlier fibers were caused by impurity absorption and then ultra-pure fibers were produced in the early 1970s (Corning).
- Today, impurities in fibers have been reduced to levels where losses associated with their absorption are negligible, with the exception of the OH^- radical.

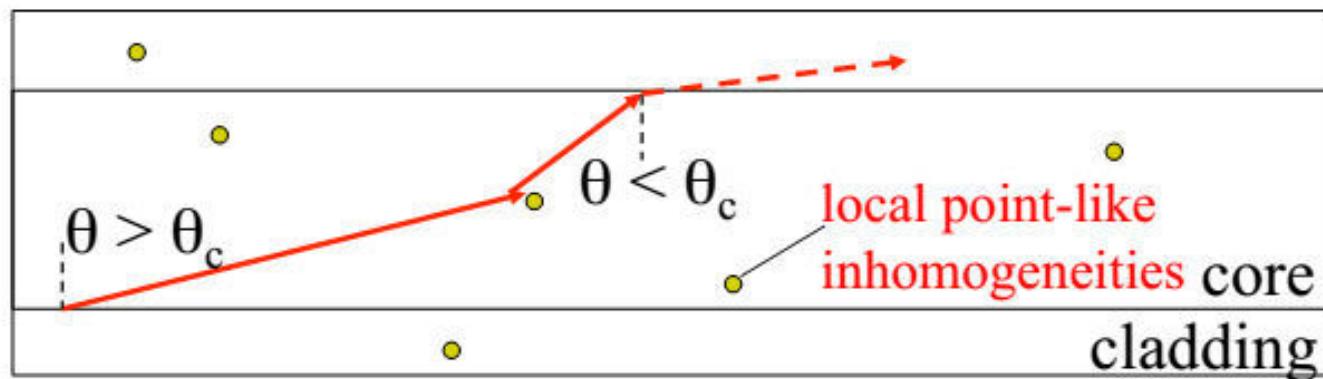
Scattering loss

336

06._Attenuation_and_Dispersion_18-Mar-2021

Scattering results in attenuation (*in the form of radiation*) as the scattered light may not continue to satisfy the total internal reflection in the fiber core.

One major type of scattering is known as *Rayleigh scattering*.



The scattered ray can escape by refraction according to Snell's Law.

Rayleigh scattering

- *Rayleigh scattering* results from **random *inhomogeneities*** that are **small in size** compared with the wavelength.
 - $\bullet \quad << \lambda$
- These inhomogeneities exist in the form of *refractive index fluctuations* which are frozen into the *amorphous* glass fiber upon fiber pulling. Such fluctuations *always exist and cannot be avoided* !

Rayleigh scattering results in an attenuation (dB/km) $\propto 1/\lambda^4$

Waveguide scattering (**Mie Scattering**)

- ***Imperfections in the waveguide structure*** of a fiber, such as nonuniformity in the size and shape of the core, perturbations in the core-cladding boundary, and defects in the core or cladding, can be generated in the manufacturing process.
- Environmentally induced effects, such as stress and temperature variations, also cause imperfections.
- The imperfections in a fiber waveguide result in additional scattering losses. They can also induce coupling between different guided modes.

Bending loss

339

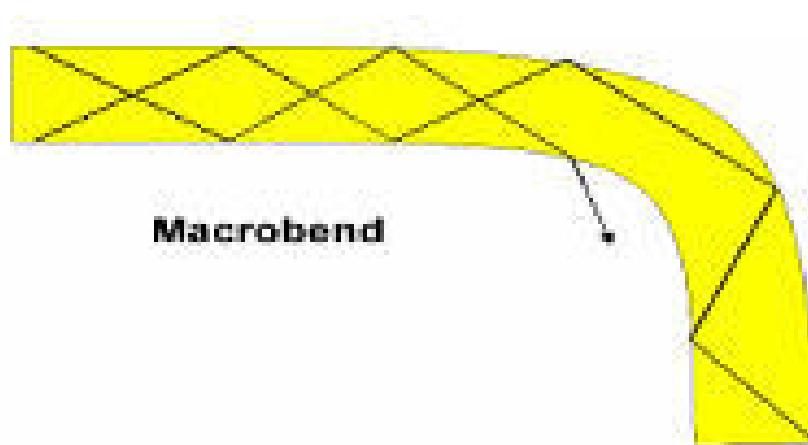
06._Attenuation_and_Dispersion_18-Mar-2021

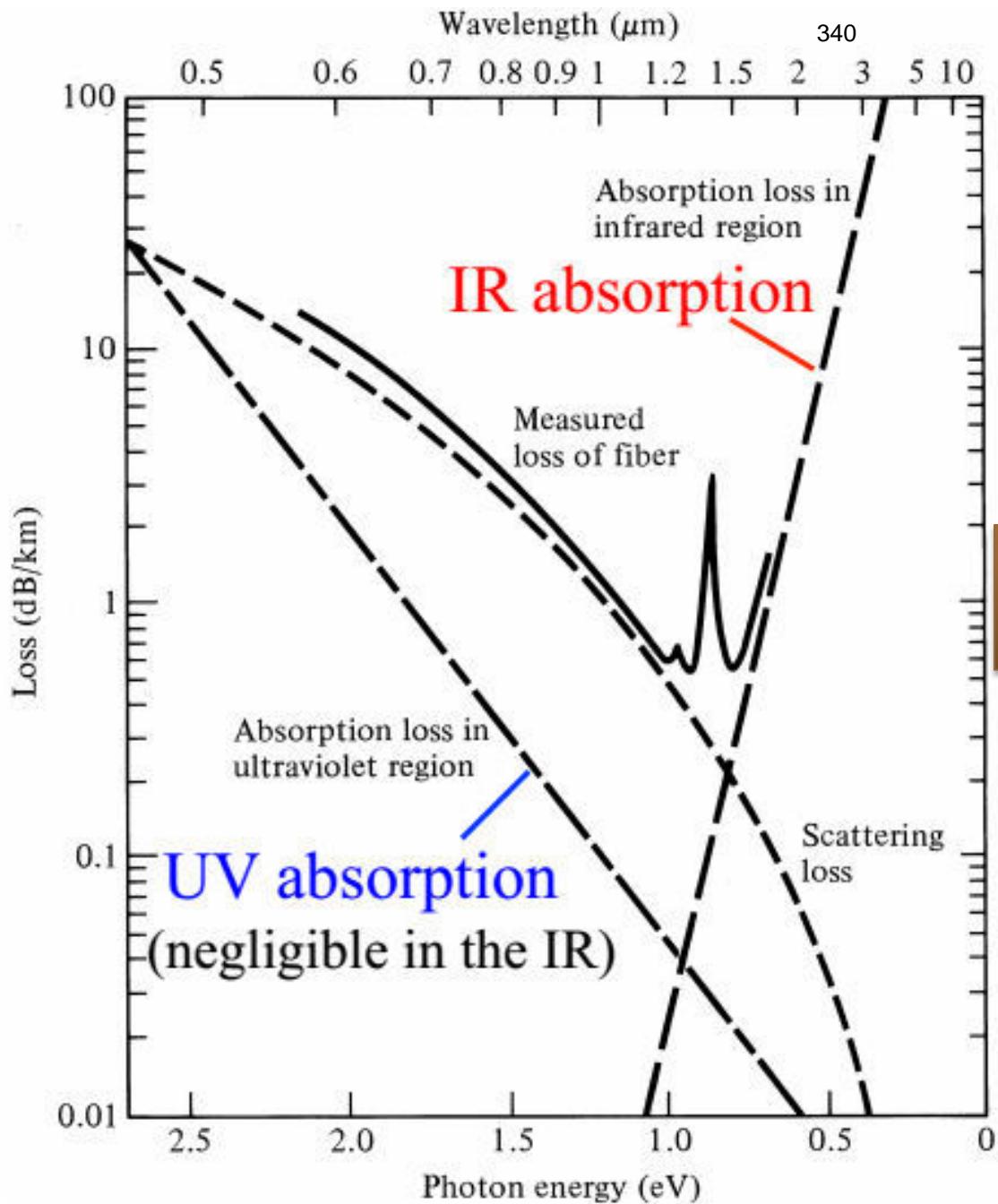
- At a bend the propagation conditions alter and light rays which would propagate in a straight fibre are lost in the cladding.
- Macrobending, for example due to tight bends
- Microbending, due to microscopic fibre deformation, commonly caused by poor cable design

Microbending is commonly caused by poor cable design



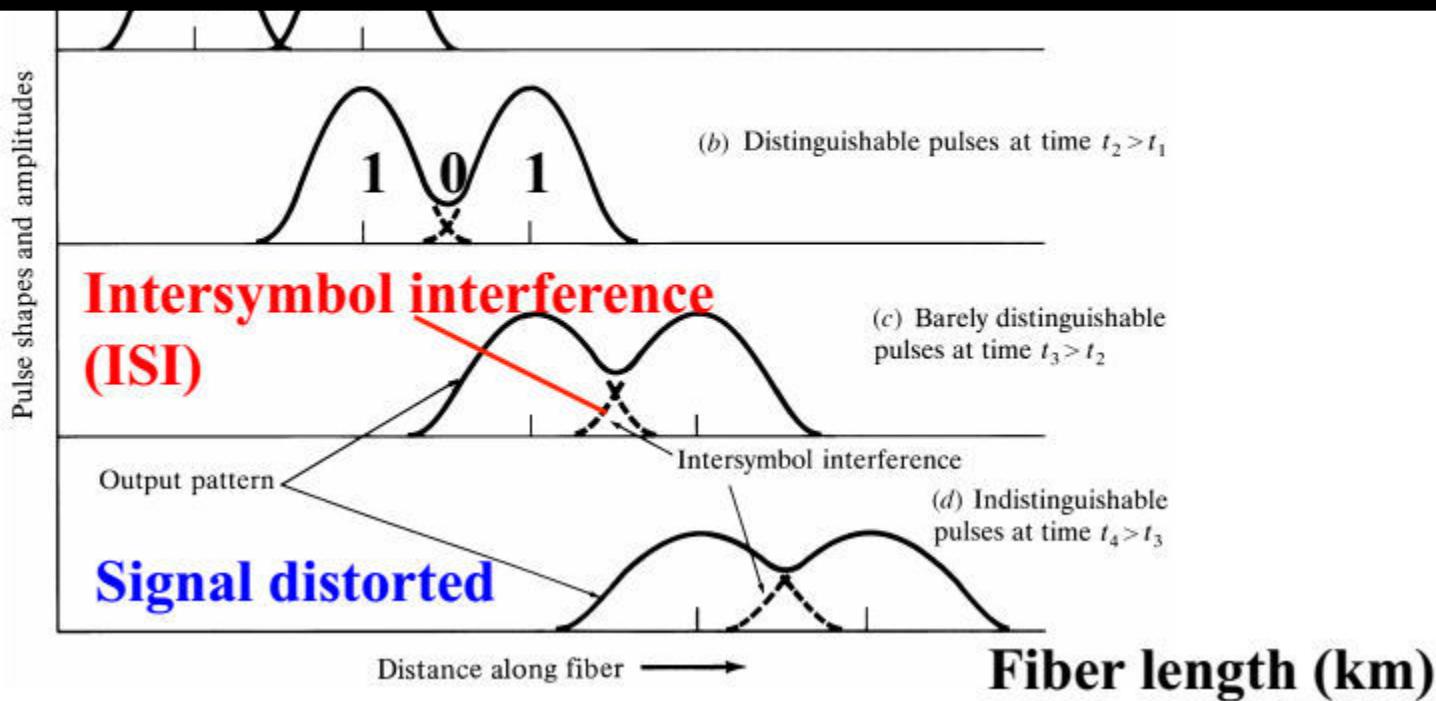
Macrobending is commonly caused by poor installation or handling





Attenuation spectra for
fused silica based glass

Hence, the number of optical signal pulses which may be transmitted in a given period and therefore the information carrying capacity of the fiber, is restricted by the amount of pulse dispersion per unit length. The pulse broadening increases linearly with fiber length and thus the bandwidth is inversely proportional to distance.

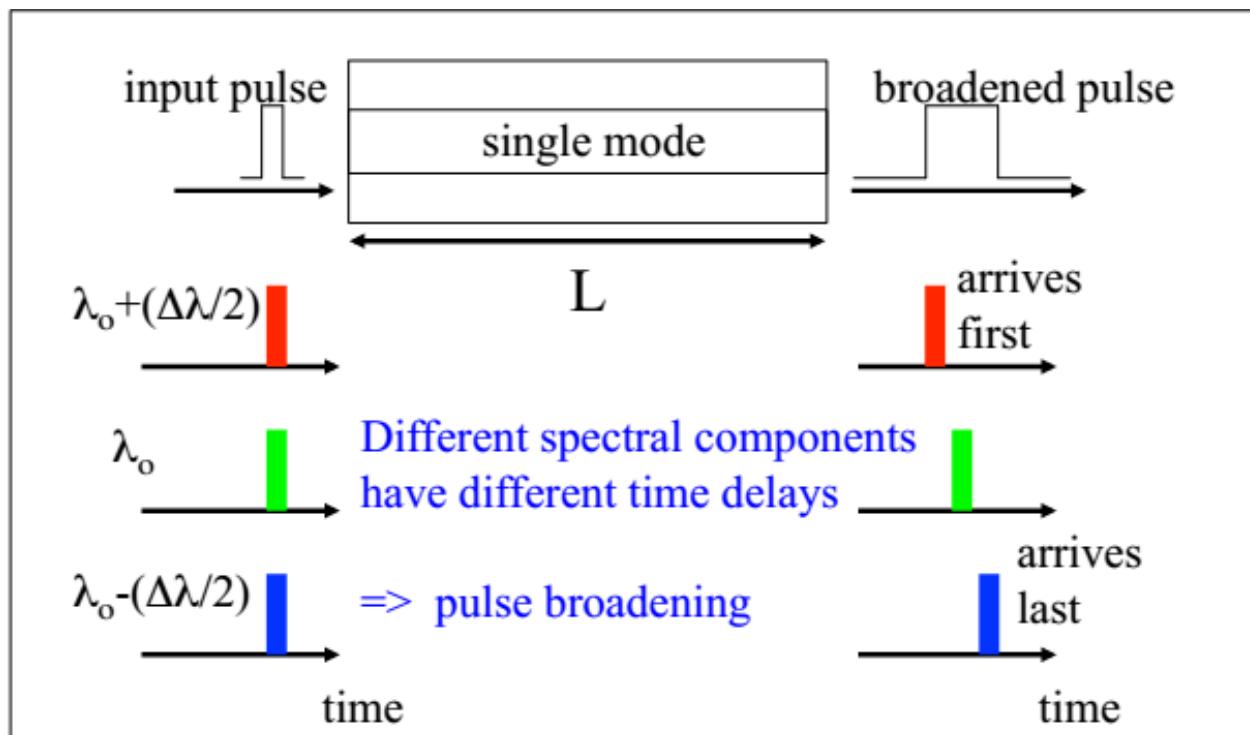


Intramodal (Chromatic) dispersion

342

06. Attenuation_and_Dispersion_18-Mar-2021

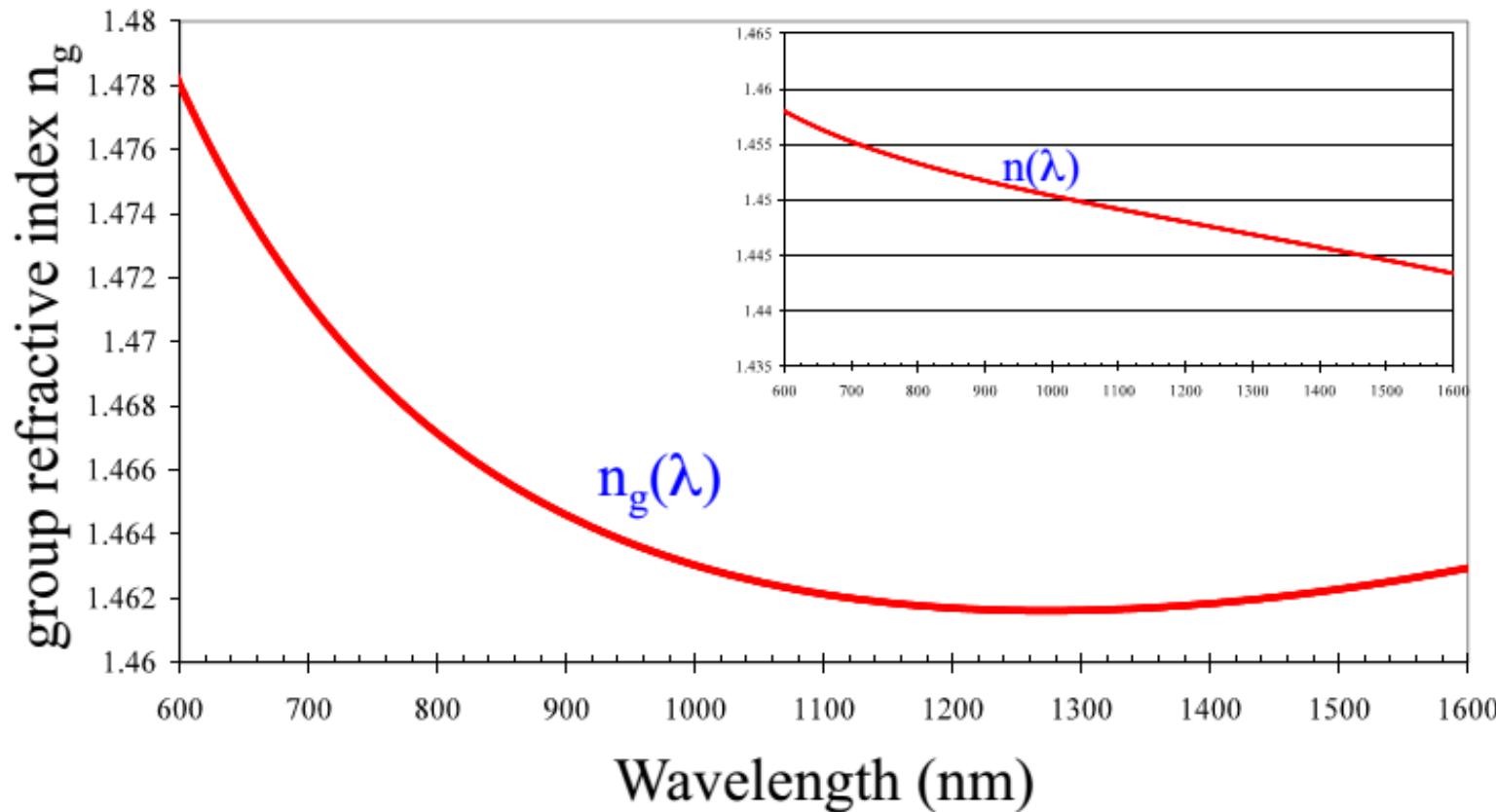
- ❖ Results from the finite spectral linewidth of the optical source.
- ❖ Optical light sources do not emit just a single frequency but a band of frequencies. Hence, there may be propagation delay differences between the different spectral components of the transmitted signal. This causes broadening of each transmitted mode and hence intramodal dispersion.



- ❖ The delay differences may be caused by:
 - ❖ Dispersive properties of the waveguide material (**material dispersion**)
 - ❖ Guidance factors within the fiber structure (**waveguide dispersion**)

- ❖ Results when different spectral components of a pulse travel at different group velocities.

Group refractive index n_g vs. λ for fused silica

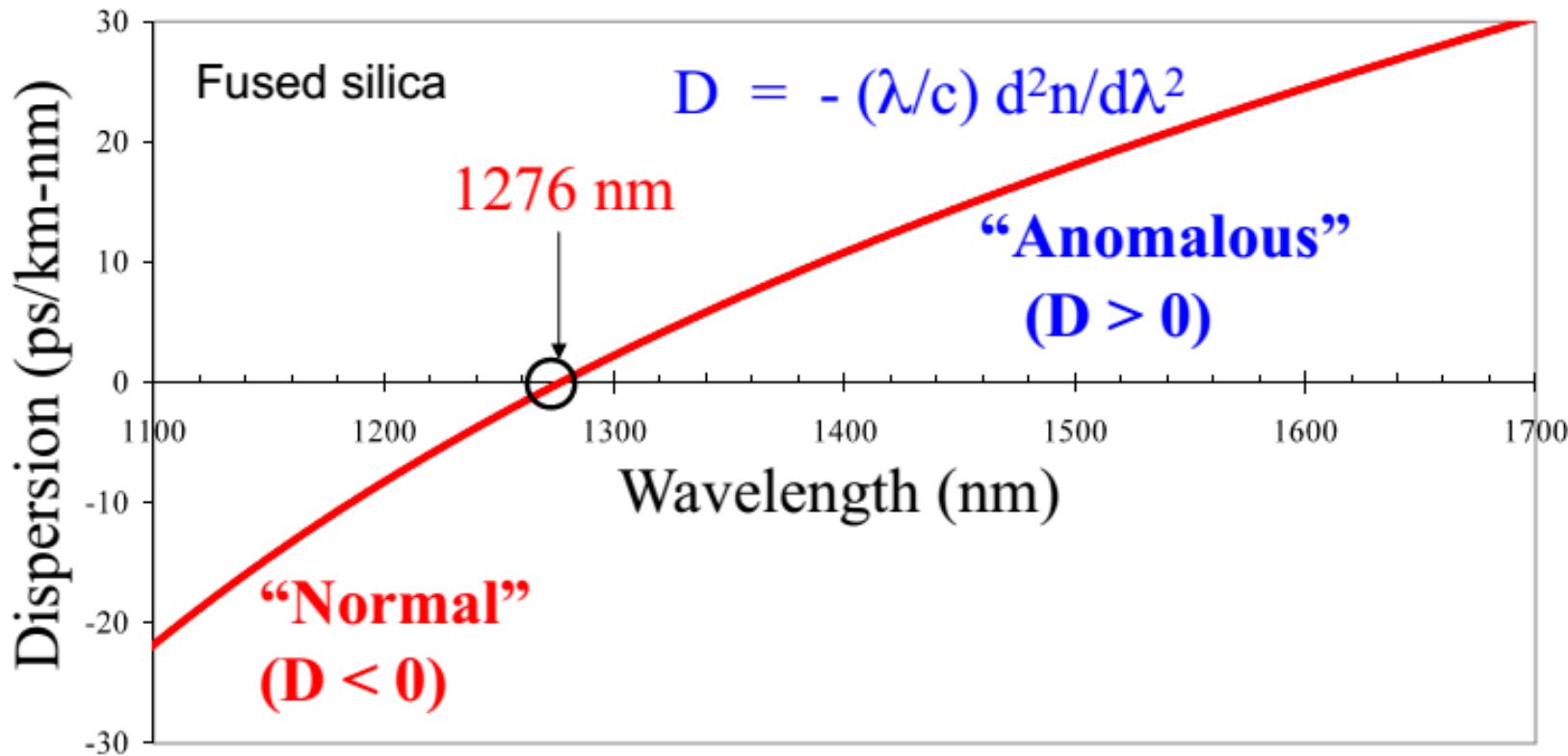


- ❖ A material is said to exhibit material dispersion when the 2nd order refractive index of core with respect to wavelength is not equal to zero.

- ❖ Material dispersion $D(\lambda)$ is given by:

$$D(\lambda) = -\frac{\lambda}{c} \frac{d^2n}{d\lambda^2}$$

$$\frac{d^2n}{d\lambda^2} \neq 0$$



Material dispersion $D_{\text{mat}} = 0$ at $\lambda \sim 1276$ nm for fused silica.

This λ is referred to as the *zero-dispersion wavelength* λ_{ZD} .

Chromatic (or *material*) dispersion $D(\lambda)$ can be zero;

or

negative \Rightarrow longer wavelengths travel *faster* than shorter wavelengths;

or

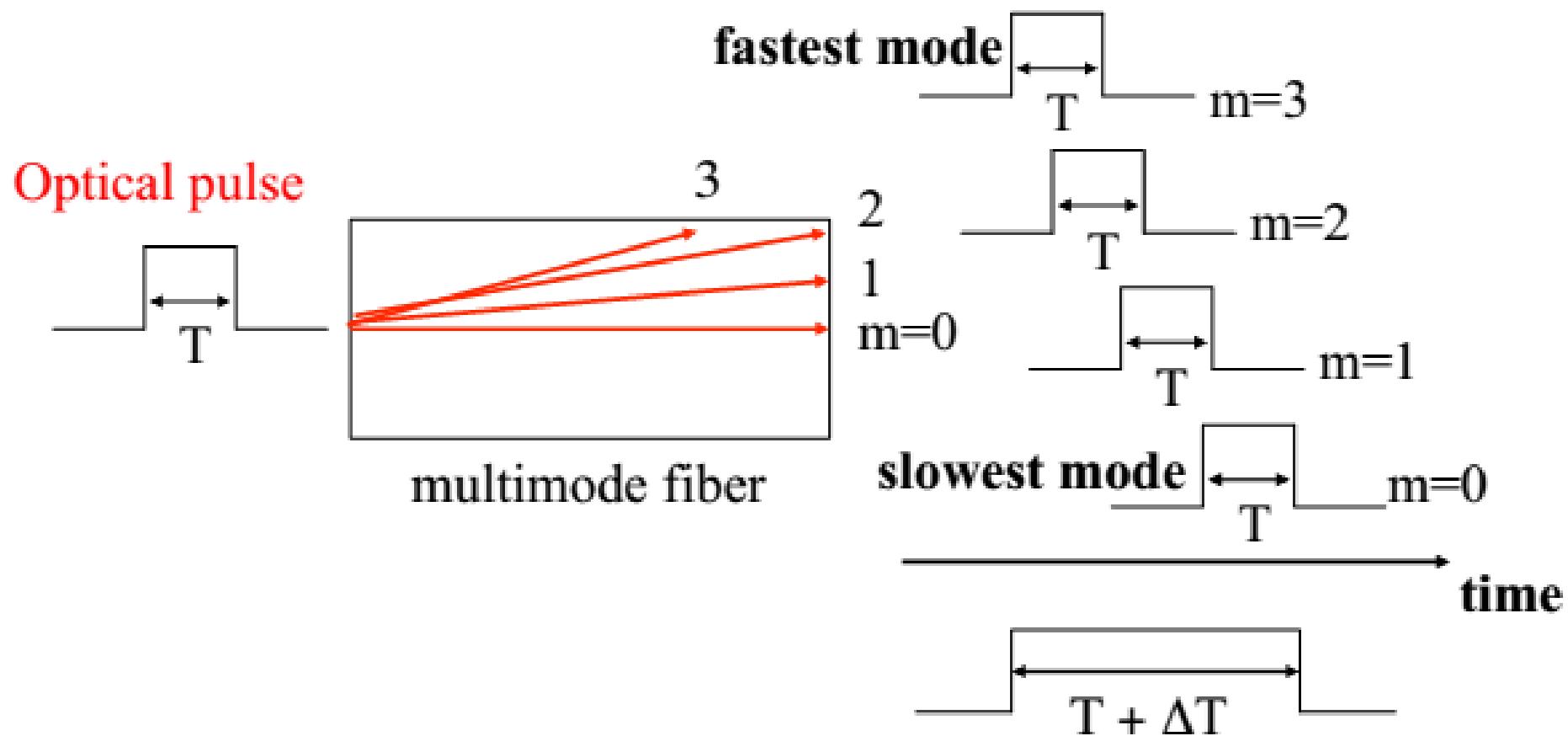
positive \Rightarrow shorter wavelengths travel *faster* than longer wavelengths.

- ❖ Results from the variation in group velocity with wavelength for a particular mode.
- ❖ Angle between the ray and the fiber axis varies with wavelength which subsequently leads to a variation in the transmission times for the rays, hence dispersion.
- ❖ More prominent in case of single mode fibers than in the multimode fibers.

$$\Rightarrow D(\lambda) = D_{\text{mat}}(\lambda) + D_{\text{wg}}(\lambda)$$

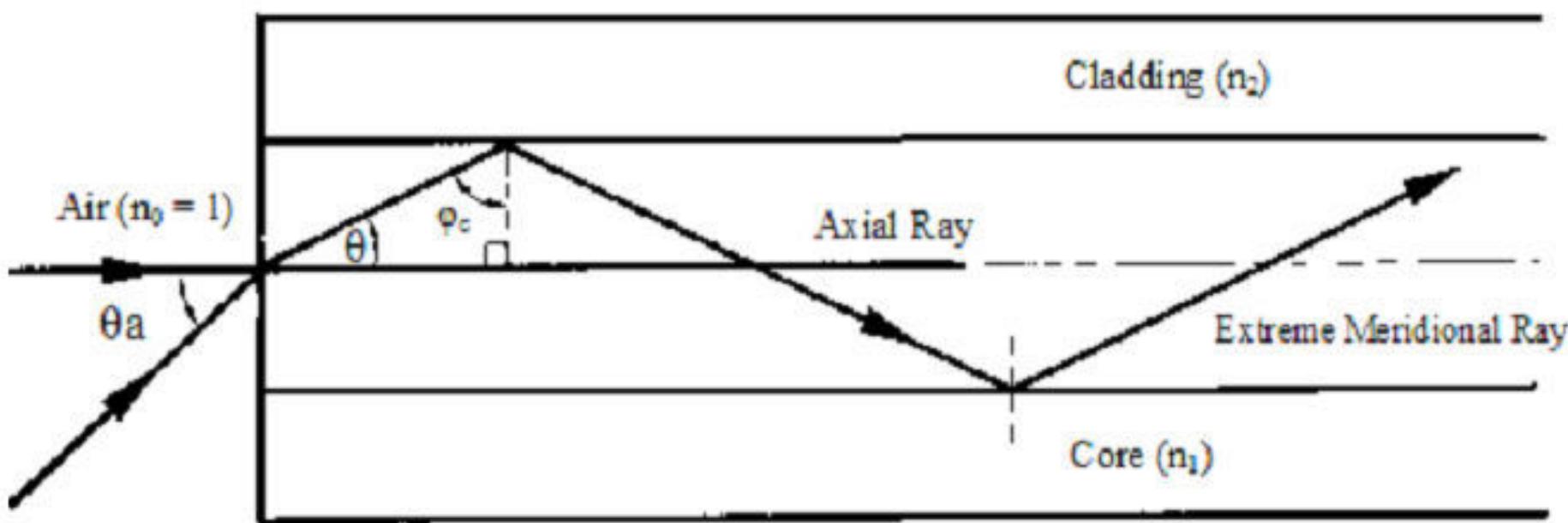
- ❖ Sometimes referred to as **Modal** (or mode) dispersion.
- ❖ When numerous waveguide modes are propagating, they all travel with different group velocities.
- ❖ Parts of the wave arrive at the output before other parts, spreading out the waveform. Hence, it is also known as multimode dispersion.
- ❖ It is **independent** of the source linewidth.
- ❖ It **does not occur** in a single mode fiber.

Modal dispersion results in pulse broadening



modal dispersion: different modes arrive at the receiver with different delays \Rightarrow pulse broadening

Paths taken by the axial and an extreme meridional ray in a perfect multimode step index fiber is shown here.



$T_{Min} \rightarrow$ Minimum delay time (time taken for the axial ray to travel along a fiber of length L)

$T_{Max} \rightarrow$ Maximum delay time (time taken for the meridional ray to travel along a fiber of length L)

$$T_{Min} = \frac{distance}{velocity} = \frac{L}{(C/n_1)} = \frac{Ln_1}{C}$$

$$T_{Max} = \frac{L/\cos\theta}{(C/n_1)} = \frac{Ln_1}{C\cos\theta} \quad \sin\phi_c = \frac{n_2}{n_1} = \cos\theta$$

$$T_{Max} = \frac{Ln_1^2}{Cn_2}$$

$$\delta T_s = T_{Max} - T_{Min} = \frac{Ln_1^2 \Delta}{Cn_2} \cong \frac{Ln_1 \Delta}{C} \cong \frac{L(NA)^2}{2n_1 C}$$

Delay difference for
 $\Delta \ll 1$

• Optical Sources

- Optical source is often considered to be the active component in an optical fiber communication system
- Fundamental function is to convert electrical energy into optical energy (light)
- Three main types of optical sources
 - Wide band **continuous spectra source** (incandescent lamp)
 - Monochromatic **incoherent** sources (**Light Emitting Diodes LED**)
 - Monochromatic **coherent** sources (**Laser**)

Characteristics of optical sources for OFC

- ❖ Light output should be highly **directional**.
- ❖ Most accurately track the electrical input signal to minimize distortion and noise. Ideally, the **source should be linear**.
- ❖ Should **emit light at wavelengths where the fiber has low losses** and low dispersion and where the detectors are efficient.
- ❖ Should have a **very narrow spectral linewidth** in order to minimize dispersion in the fiber.

Types of Fibers

354 06._Types_of_Fibers_22-Mar-2021

CLASSIFICATION BASED ON MATERIALS

1. Glass fiber: Made by fusing mixtures of metal oxides and silica glasses.

Ex: $\text{GeO}_2\text{-SiO}_2$ core, SiO_2 cladding

SiO_2 core, $\text{P}_2\text{O}_5\text{-SiO}_2$ cladding

2. Plastic fiber: Made up of plastic polymers and is of low cost and flexible. Can be handled without any special care due to its toughness and durability.

Ex: Polysterene core, Methyl methacrylate cladding

Polymethyl methacrylate core and co-polymer cladding

Number of modes

- ❖ Optical fiber is a dielectric waveguide.
- ❖ Energy in the fiber is propagated by electric and magnetic field vectors of electromagnetic wave; which can be analysed by Maxwell's field equations.
- ❖ Maxwell's equations have discrete sets of solutions called the modes.
- ❖ Number of modes propagating in an optical fiber can be determined by a factor known as “horizontal wave number” (V).

$$V = \frac{2\pi a}{\lambda} NA$$

$a \rightarrow$ radius of the core

- ❖ Maximum number of modes supported by a **step index fiber** is:

$$N_m = \frac{1}{2} V^2$$

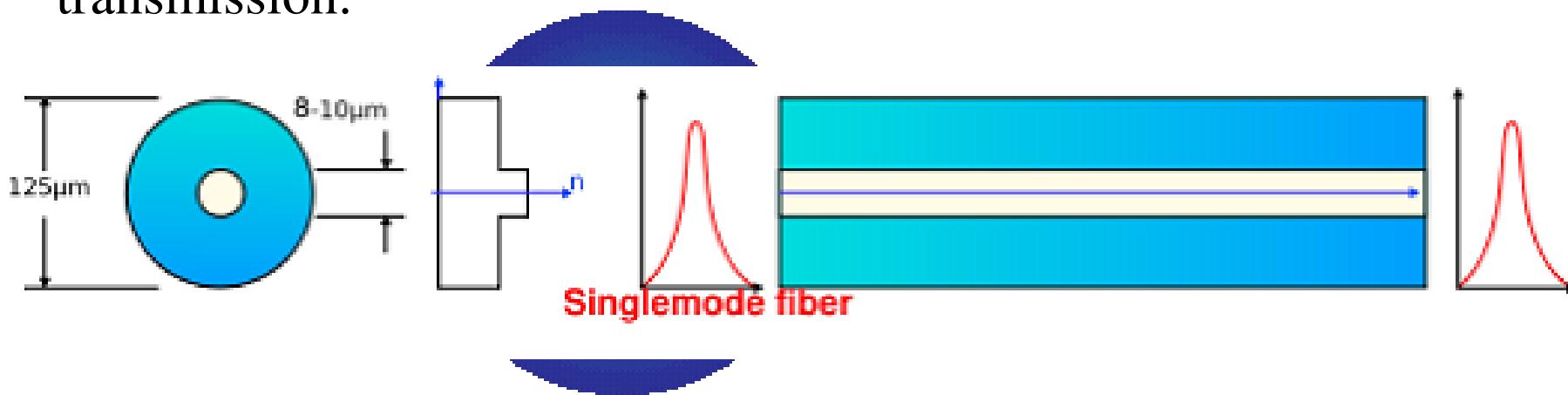
- ❖ Maximum number of modes supported by a **graded index fiber** is:

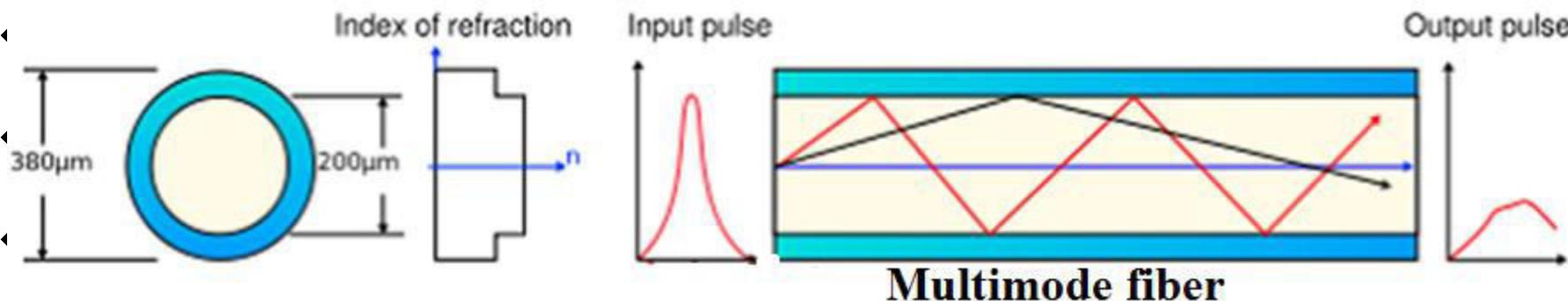
$$N_m = \frac{1}{4} V^2$$

- ❖ For $V < 2.405$: Only one mode is supported (single mode fiber)
- ❖ For $V > 2.405$: Can support more than one mode (Multimode fiber)
- ❖ The wavelength corresponding to $V = 2.405$ is known as the cut-off wavelength of the fiber.

$$\lambda_c = \frac{\lambda V}{2.405}$$

- ❖ These fibers have very narrow core ($\sim 10 \mu\text{m}$ in diameter).
- ❖ Hence allow only one mode (TE, TM or TEM) to pass through it.
- ❖ NA and acceptance angles are small for these fibers which allows only the transmission of fundamental modes.
- ❖ Amount of dispersion is very less.
- ❖ Used for very high speed, large bandwidth and long distance transmission.

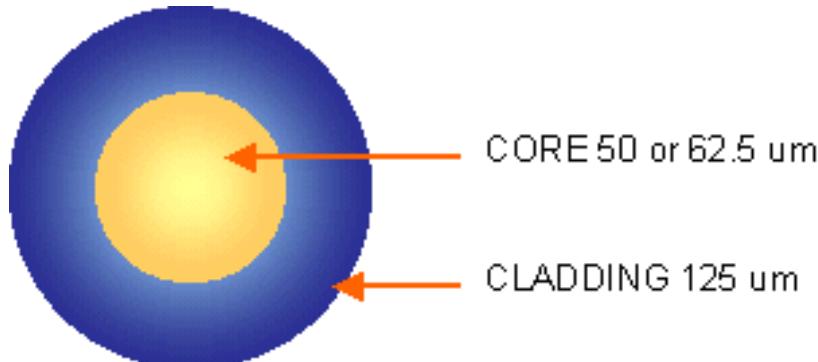


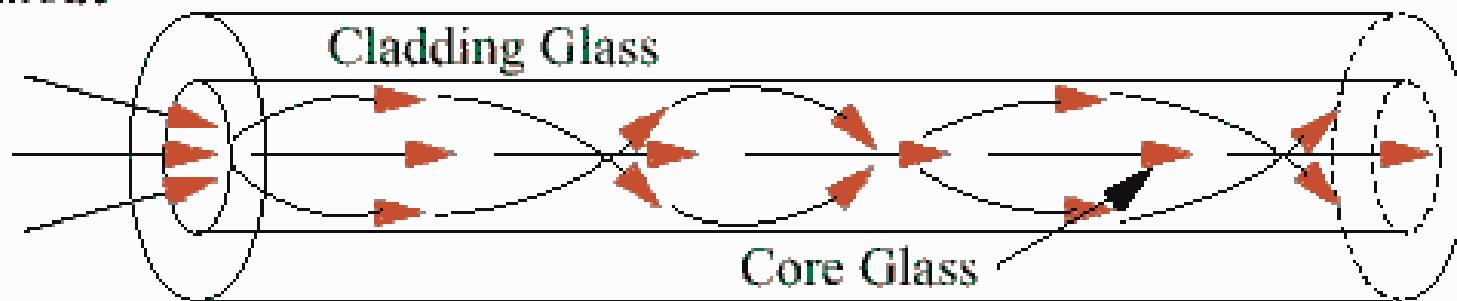
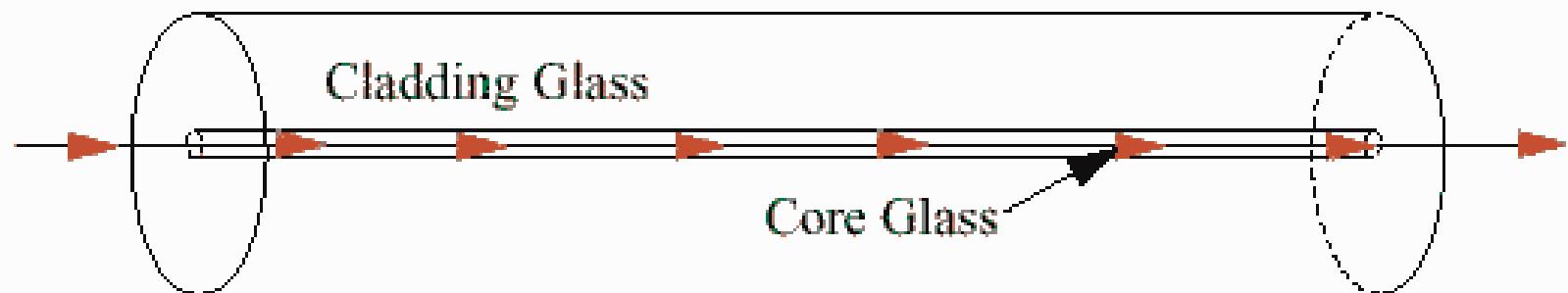


and **marginal ray** (near the fiber surface).

- ❖ Marginal ray travels longer distance than the axial ray.
- ❖ This time delay causes distortion in the pulse leading to dispersion.
- ❖ Results in broadening of light pulses reducing the transmission speed and transmission bandwidth.

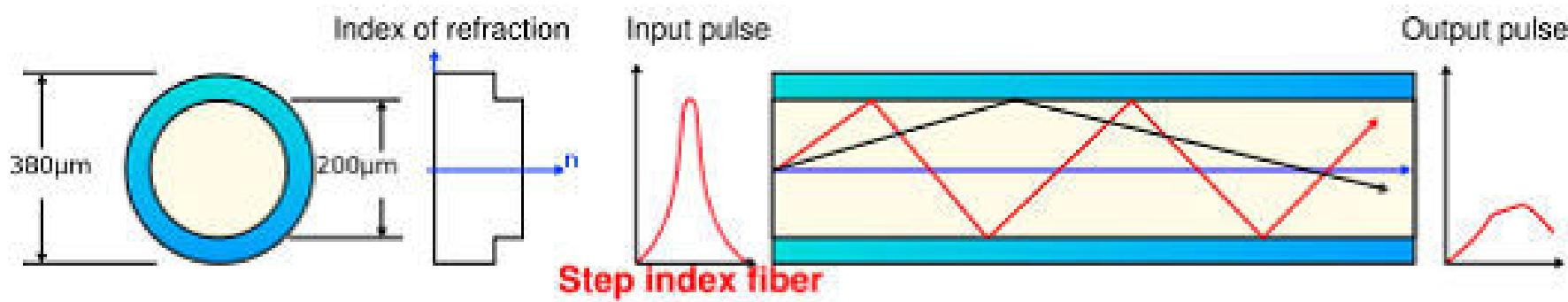
Best designed for short transmission distances and is suited for use in LAN systems and video surveillance



Multimode*Single-Mode*

STEP-INDEX FIBER

- ❖ Refractive index of the core is uniform throughout and undergoes an abrupt change (step) at the cladding boundary.



Multimode step-index fiber

- ❖ Core refractive index is made to vary as a function of the radial distance from the center of the fiber. Also known as inhomogeneous core fibers.

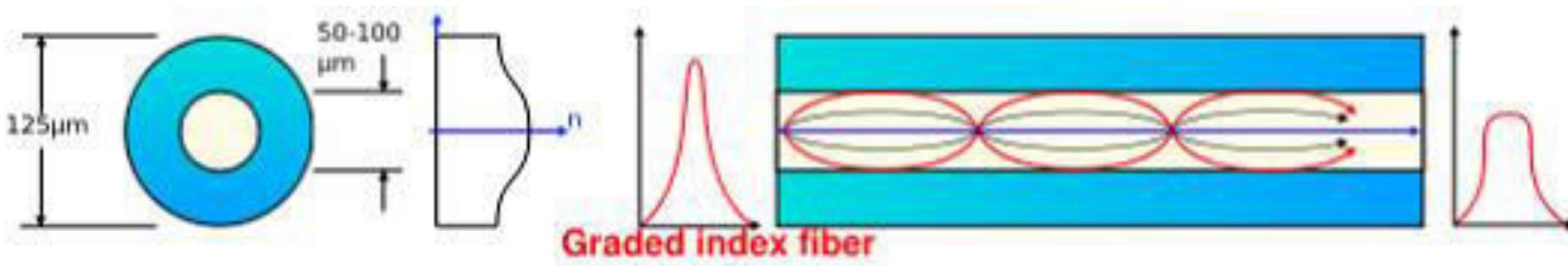
$$n(r) = n_1 \left[\left(1 - 2\Delta \left(\frac{r}{a} \right)^\alpha \right) \right]^{1/2} \quad \text{For } r < a, \text{ Core}$$

$$n(r) = n_1 (1 - 2\Delta)^{1/2} = n_2 \quad \text{For } r \geq a, \text{ Cladding}$$

$\Delta \rightarrow$ Relative refractive index difference

$\alpha \rightarrow$ Profile parameter (Gives the characteristic refractive index of core)

= 1 (triangular profile), 2 (parabolic), ∞ (step-index)



- ❖ Graded index profiles, giving best result for multimode optical propagation have **nearly parabolic refractive index profile**.
- ❖ In this case, the **pulse dispersion is less than that in step-index fiber**.

The Periodic Table

1 H																		2 He
3 Li	4 Be																	10 Ne
11 Na	12 Mg																	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Al	14 Si	15 P	16 S	17 Cl		36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I		54 Xe
55 Cs	56 Ba	57-71	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At		86 Rn
87 Fr	88 Ra	89-103	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112 Cn	113 Nh	114 Fl	115 Mc	116 Lv	117 Ts		118 Og
	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb	71 Lu			
	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No	103 Lr			

Conductors

Semiconductors

Insulators

Semiconductors

Intrinsic

Extrinsic

Pure

Doped

p-type

n-type

Electron-hole pair

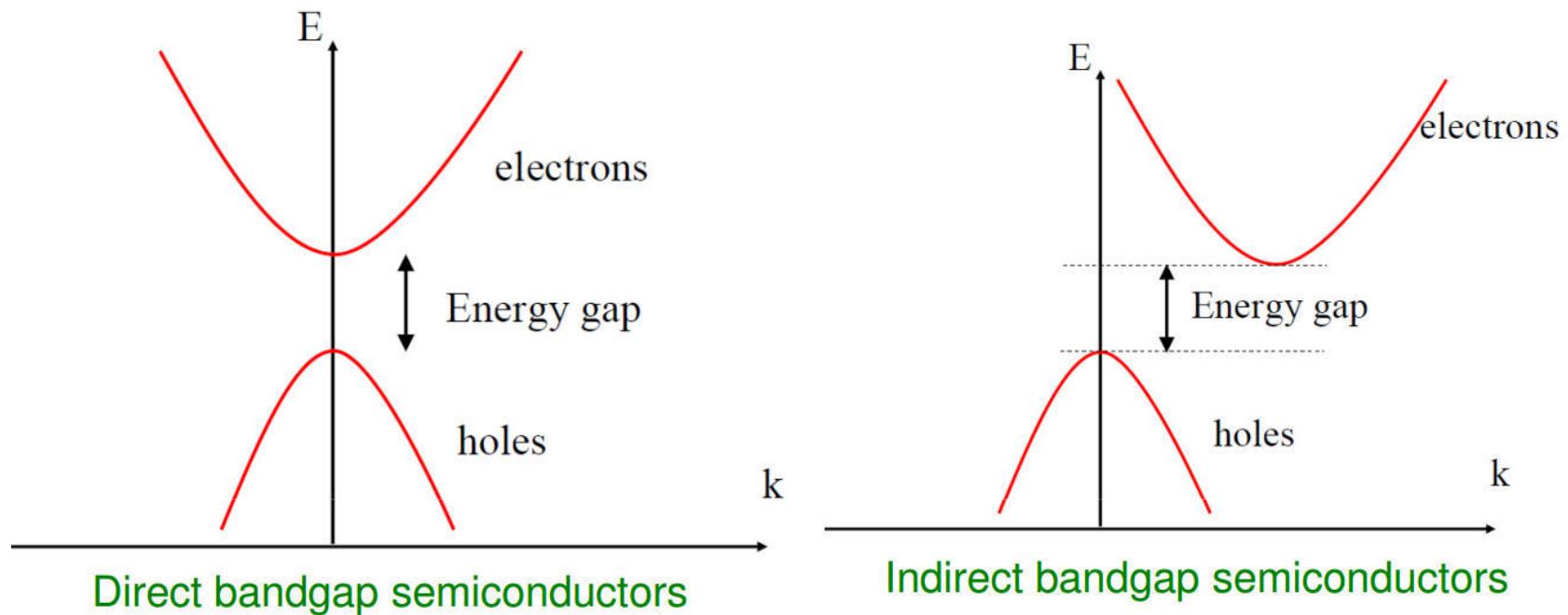
holes

electrons

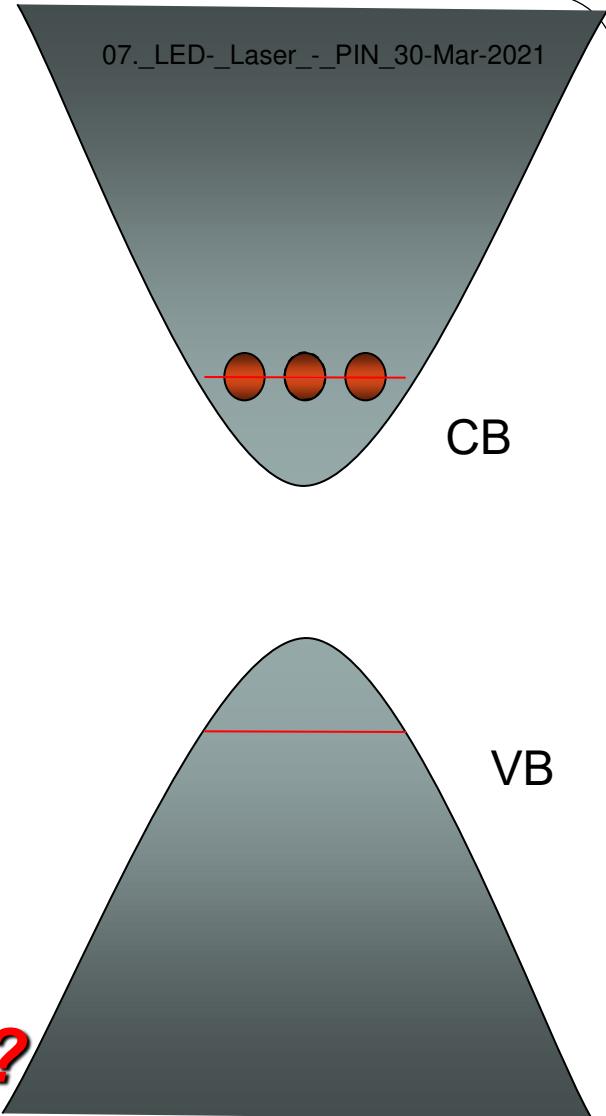
Indirect Bandgap

Direct Bandgap

Direct and Indirect Bandgap



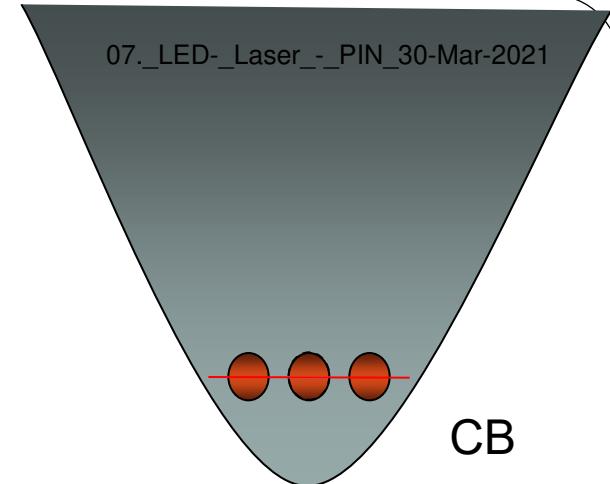
When the electron falls down from conduction band and fills in a hole in valence band, there is an obvious loss of energy



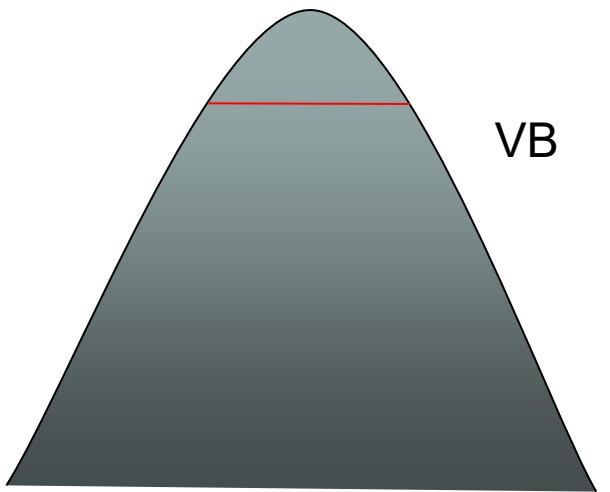
***The question is;
where does that energy go?***

In order to achieve a reasonable efficiency for photon emission, the semiconductor must have a direct band gap

368
07._LED-_Laser_-_PIN_30-Mar-2021



**The question is;
what is the mechanism
behind photon emission in LEDs?**



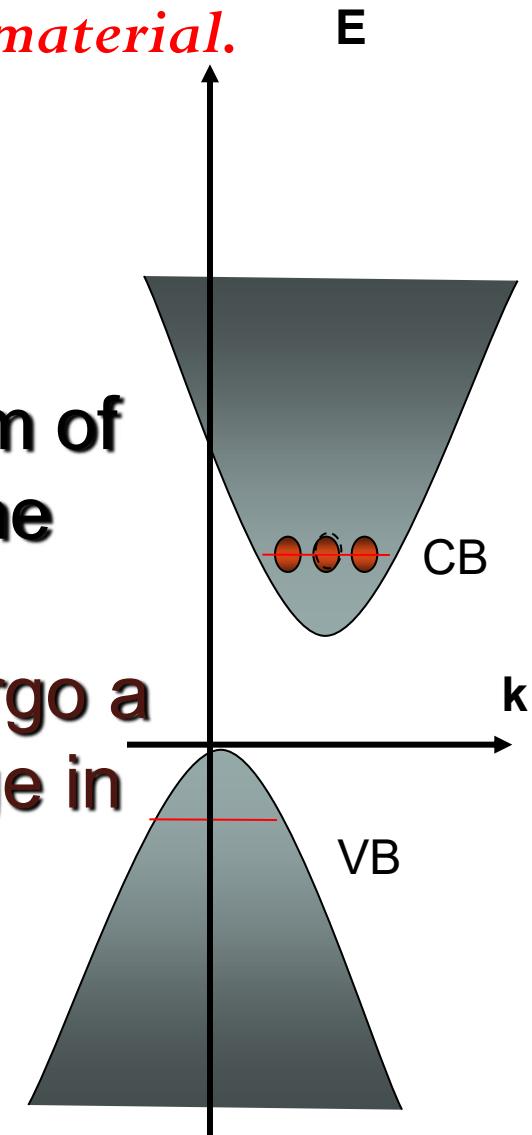
For example;

Silicon is known as an *indirect band-gap material.*

What this means is that

**as an electron goes from the bottom of
the conduction band to the top of the
valence band;**

**it must also undergo a
significant change in
momentum.**



- As we all know, whenever something changes³⁷⁰ state, one must conserve not only energy, but also momentum.
- In the case of an electron going from conduction band to the valence band in silicon, both of these things can only be conserved:

The transition also creates a quantized set of lattice vibrations, called phonons, or "heat".

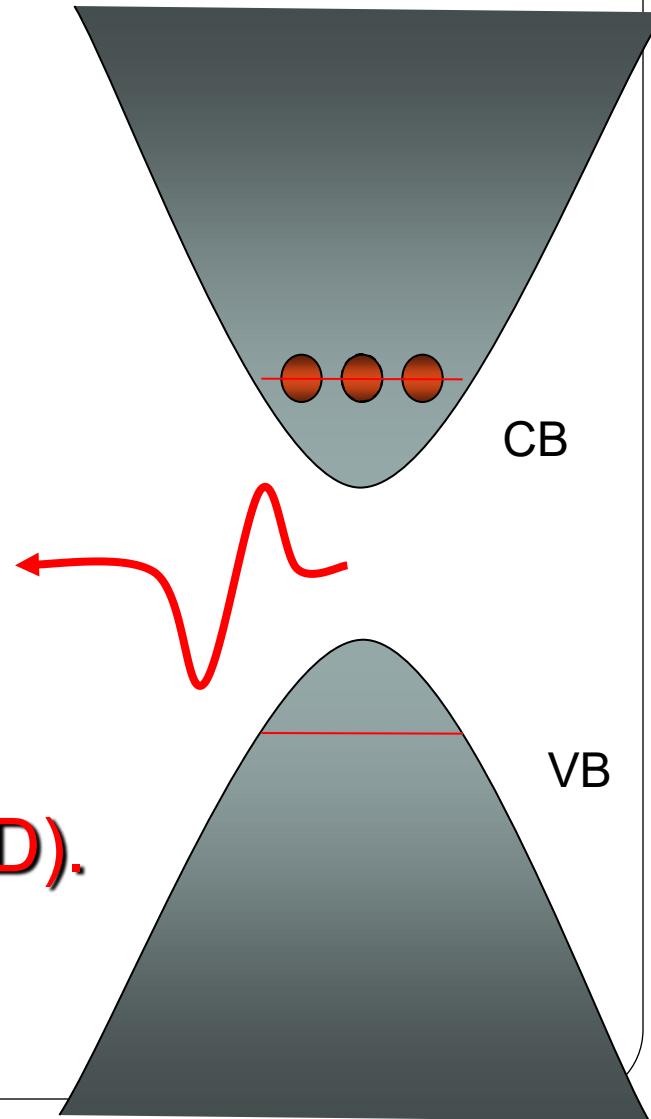
- Phonons possess both energy and momentum.
- Their creation upon the recombination of an electron and hole allows for complete conservation of both energy and momentum.
- All of the energy which the electron gives up in going from the conduction band to the valence band (1.1 eV) ends up in phonons, which is another way of saying that the electron heats up the crystal.

In a class of materials called ***direct band-gap semiconductors***;

- the transition from conduction band to valence band involves essentially **no change in momentum**.
- Photons, it turns out, possess a fair amount of energy (several eV/photon in some cases) but they have very little momentum associated with them.

- Thus, for a direct band gap₃₇₃ material, the excess energy of the electron-hole recombination can either be taken away as heat, or more likely, as a photon of light.
- This radiative transition then conserves energy and momentum by giving off light whenever an electron and hole recombine.

**This gives rise to
(for us) a new type
of device;
the light emitting diode (LED).**



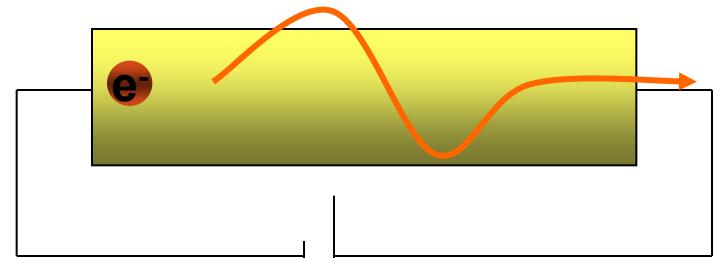
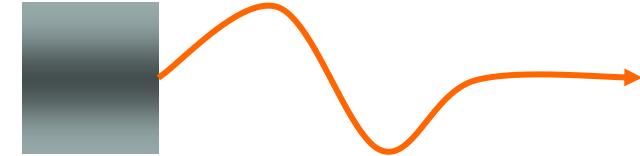
Mechanism behind photon emission in LEDs?

Mechanism is “**injection Electroluminescence**”.

Luminescence part tells us that we are producing photons.

Electro part tells us that the photons are being produced by an electric potential.

Injection tells us that photon production is by the injection of charge carriers.

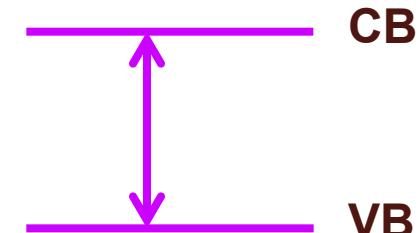


Producing photon

Electrons recombine with holes.

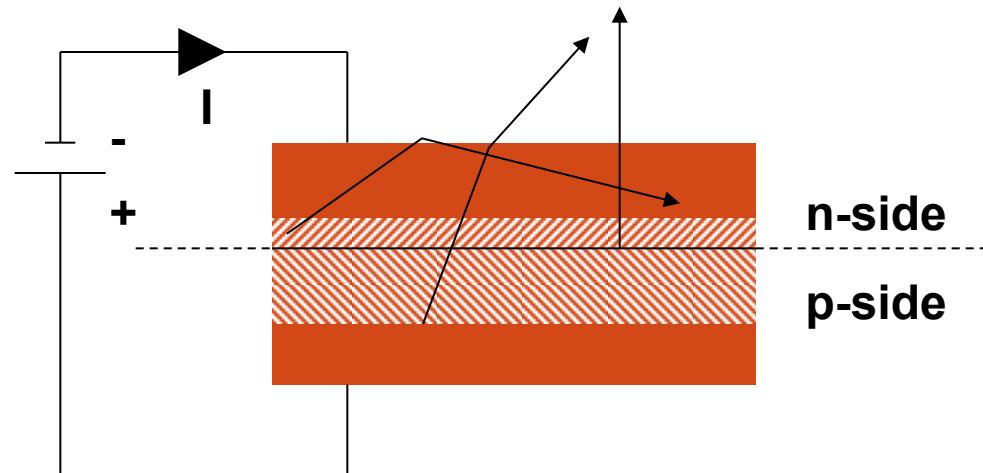


Energy of photon is the energy of band gap.



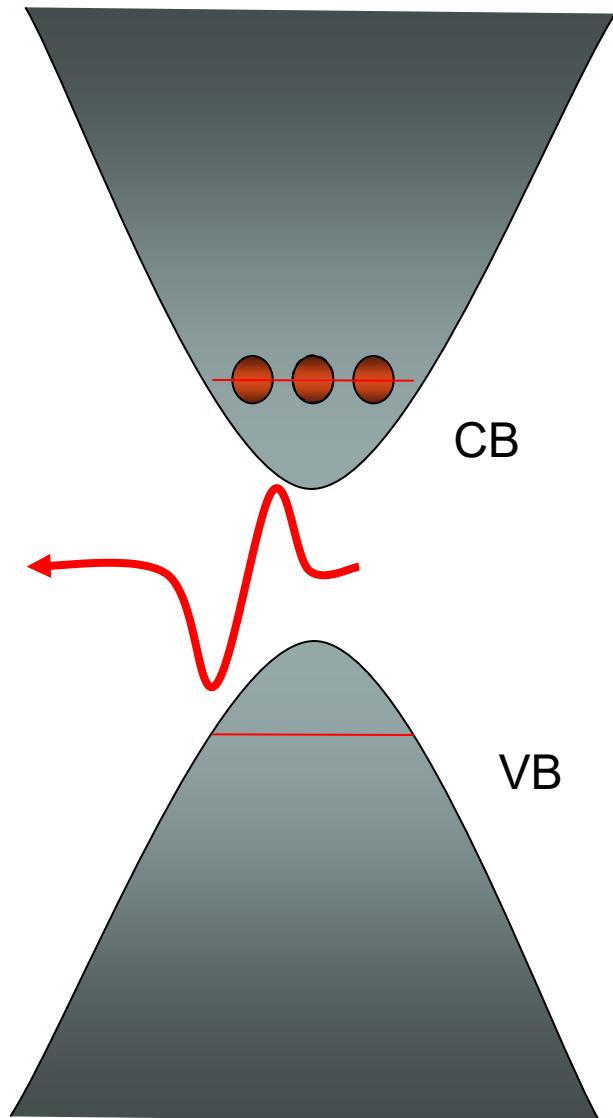
Method of injection

- We need putting a lot of e^- 's where there are lots of holes.
- So electron-hole recombination can occur.
- **Forward biasing** a p-n junction will inject lots of e^- 's from n-side, across the depletion region into the p-side where they will be combine with the high density of majority carriers.



MATERIALS FOR LEDs

- The semiconductor bandgap energy defines the energy of the emitted photons in a LED.
- To fabricate LEDs that can emit photons from the infrared to the ultraviolet parts of the e.m. spectrum, then we must consider several different material systems.
- No single system can span this energy band at present, although the 3-5 nitrides come close.



- Unfortunately, many of potentially useful II-VI group of direct band-gap semiconductors (ZnSe, ZnTe, etc.) come naturally doped either p-type, or n-type, but they don't like to be type-converted by overdoping.
- The material reasons behind this are complicated and not entirely well-known.
- The same problem is encountered in the III-V nitrides and their alloys InN, GaN, AlN, InGaN, AlGaN, and InAlGaN. The amazing thing about III-V nitride alloy systems is that appear to be direct gap throughout.

- When we talk about light, it is conventional to specify its wavelength, λ , instead of frequency.
- Visible light has a wavelength of the order of nanometers.

$$\lambda(nm) = \frac{hc}{E(eV)}$$

$$\lambda(nm) = \frac{1240}{E(eV)}$$

- Thus, a semiconductor with a 2 eV band-gap should give a light at about 620 nm (in the red). A 3 eV band-gap material would emit at 414 nm, in the violet.
- The human eye, of course, is not equally responsive to all colors.

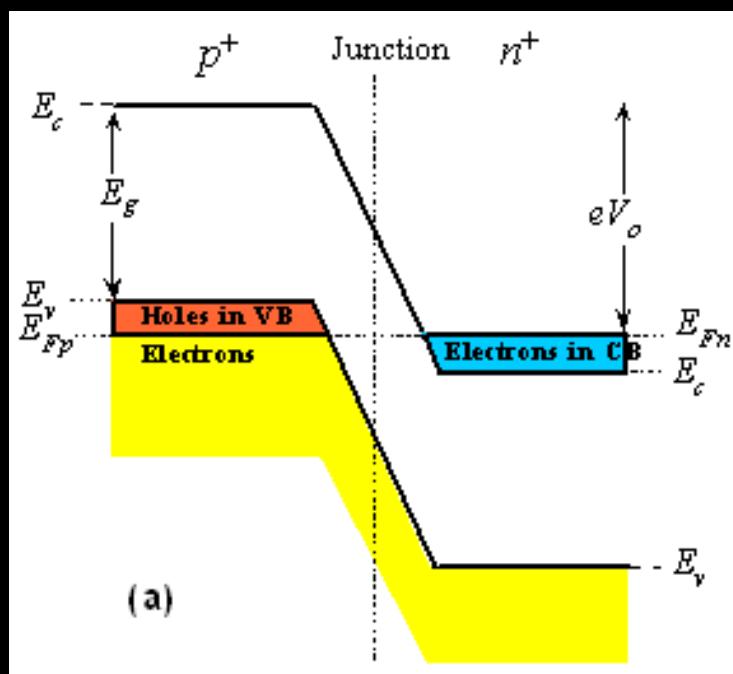
Color Name	Wavelength (Nanometers)	Semiconductor Composition
Infrared	880	GaAlAs/GaAs
Ultra Red	660	GaAlAs/GaAlAs
Super Red	633	AlGaN/P
Super Orange	612	AlGaN/P
Orange	605	GaAsP/GaP
Yellow	585	GaAsP/GaP
Incandescent White	4500K (CT)	InGaN/SiC
Pale White	6500K (CT)	InGaN/SiC
Cool White	8000K (CT)	InGaN/SiC
Pure Green	555	GaP/GaP
Super Blue	470	GaN/SiC
Blue Violet	430	GaN/SiC
Ultraviolet	395	InGaN/SiC

Material	Wavelength (μm)	Material	Wavelength (μm)
ZnS	0.33	GaAs	0.84-0.95
ZnO	0.37	InP	0.91
GaN	0.40	GaSb	1.55
ZnSe	0.46	InAs	3.1
CdS	0.49	Te	3.72
ZnTe	0.53	PbS	4.3
GaSe	0.59	InSb	5.2
CdSe	0.675	PbTe	6.5
CdTe	0.785	PbSe	8.5

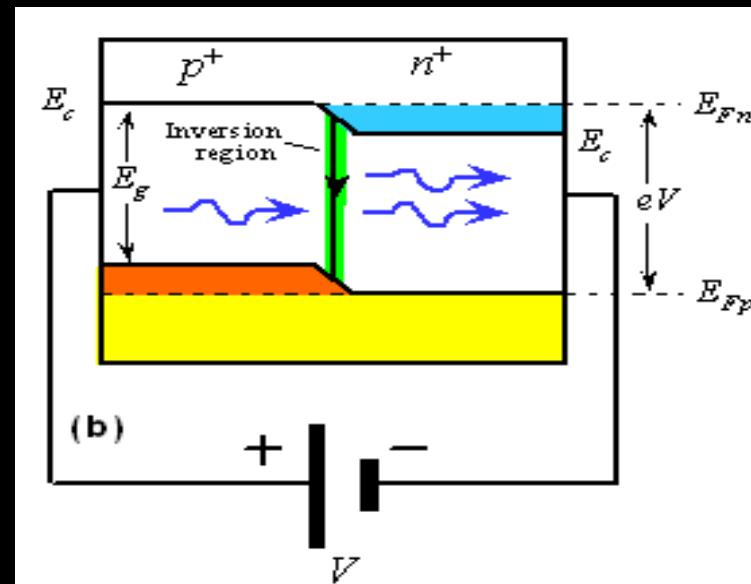
• **LASER DIODE**

- Basic mechanism for light emission is the **recombination of electrons and holes at p-n junction when current is passed through the diode**
- Any laser system has 3 possible interaction process
 - a) **excitation from VB to CB** by absorbing energy from incident radiation
 - b) **spontaneous transition from CB to VB** with the emission of radiation
 - c) **stimulated emission from CB to VB** by the stimulation of incident radiation
- If we can have large number of electrons in the bottom of **CB** and large number of **holes in the upper part of the valance band**, then semiconductor can amplify optical radiation at frequency which corresponds to energy slightly greater than the band gap energy.

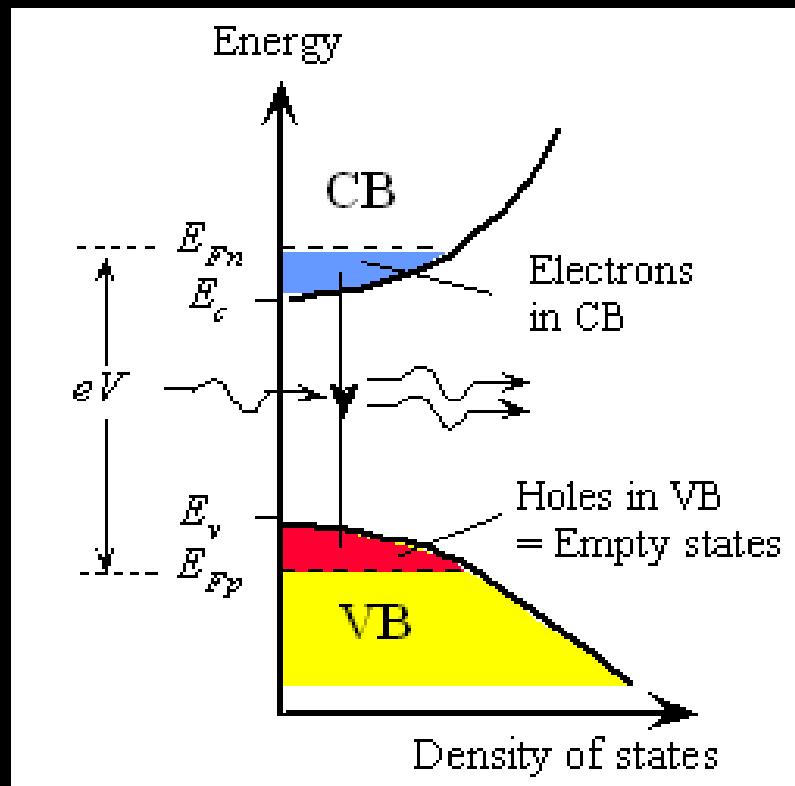
- When current is passed through the p-n junction it will increase the carrier concentration in the CB and the VB close to the junction.
- For some value of current stimulated emission will exceed the absorption rate and amplification will begin.
- On further increase of current (threshold value of current) the amplification will overcome the losses in the cavity and laser will begin to emit coherent radiation.
- Consider a degenerately doped direct band gap semiconductor p-n junction



- **Degenerate doping** – Fermi level of p side is in its VB and Fermi level of n side is in its CB, all the energy levels can be considered to be occupied in the absence of applied voltage
- Fermi level is continuous across the diode
- Depletion region or the space charge region in **such p-n junction is very narrow**.
- There is a **built in voltage V_0** which give rise to a potential energy barrier eV_0 that prevents the electron in CB of n side to diffuse into CB of p and the hole diffusion too.
- If this junction is forward biased by a voltage V greater than the band gap voltage

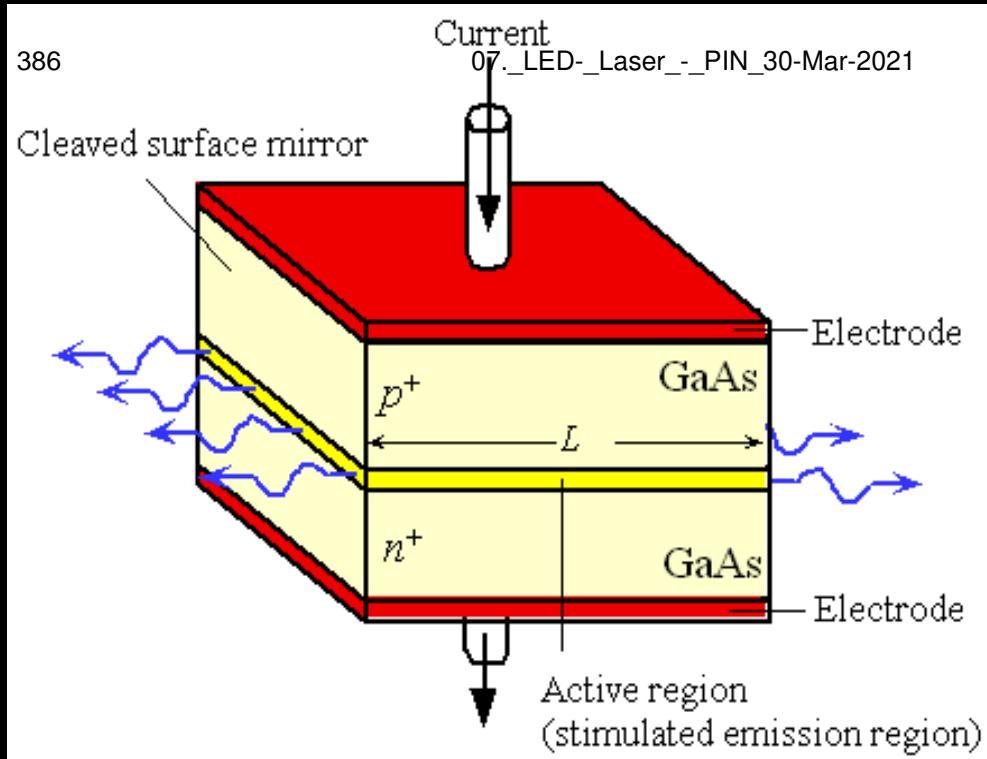


- Now the separation between E_{Fn} and E_{Fp} is the applied voltage.
- This applied voltage reduces the built-in potential barrier almost to zero- which ensues flow of carriers.
- There are more number of electrons in CB when compared to VB which could be assumed as population inversion stage
- The region where the population inversion occurs develops a layer along the junction called an ***inversion layer or active region***



- An incoming photon with energy of $(E_c - E_v)$ will not see electrons to excite from E_v due to the absence of electrons at E_v .
- The photon can cause an electron to fall down from E_c to E_v .
- The incoming photon is stimulating ***direct recombination***.

- Figure shows a schematic diagram of homo junction laser
- Laser action - only in the active region
- Optical feedback obtained by reflections from cleaved end faces of the semiconductor
- Amount of population inversion and the gain is determined by the current flowing through the junction.
- Beyond a particular value of current (threshold current) lasing commences and radiative output then increases very rapidly with increasing current
- The outrageous difference between semiconductor laser and other lasers is their exceedingly small size.



• **PHOTO-DETECTORS**

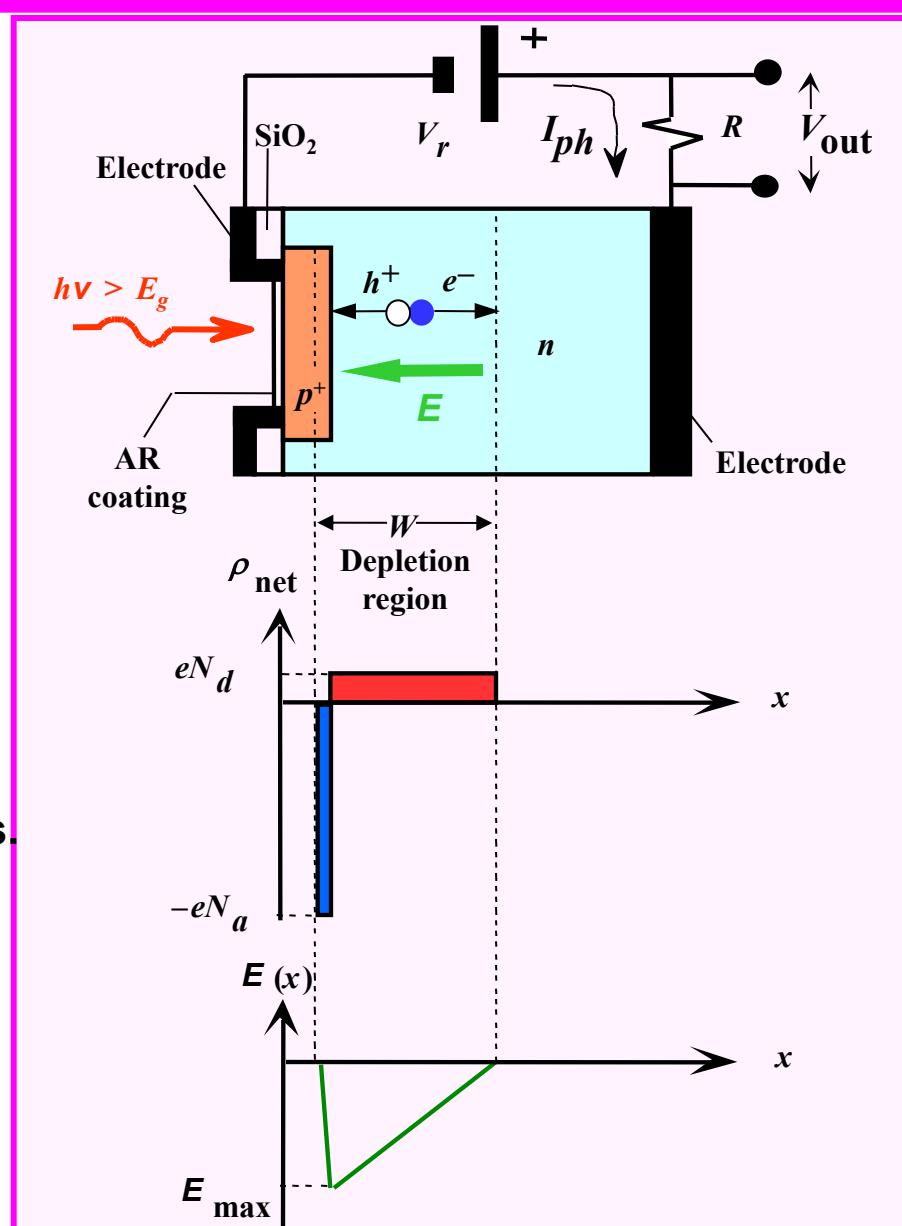
- Convert an **optical signal** into an electrical signal.
- Photodetectors made up of semiconductor materials absorb incident photons and produce electrons/holes
- **Basic requirements** of a photodetector:
 - **Sensitivity** at the required wavelength
 - **Efficient conversion** of photons to electrons/holes
 - **Fast response** to operate at high frequencies
 - **Low noise** for reduced errors
 - **Sufficient area** for efficient coupling to optical fibers
 - **Low cost**

Principle of the p-n junction Photodiode

- Schematic diagram of a reverse biased p-n junction photodiode
 - Photocurrent is depend on number of EHP and drift velocity.
 - The electrode do not inject carriers but allow excess carriers in the sample to leave and become collected by the battery.

- Net space charge across the diode in the depletion region. N_d and N_a are the donor and acceptor concentrations in the p and n sides.

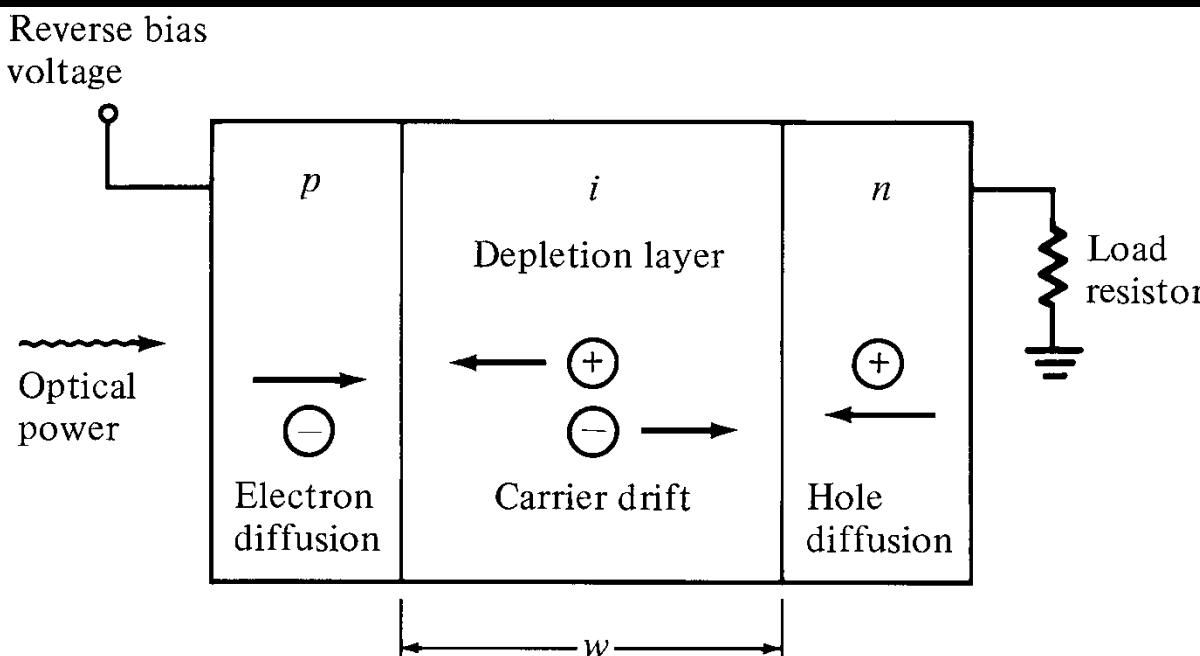
- The field in the depletion region.



• REVERSE BIASING

- Depletion region is devoid of any charge carriers.
- Width of depletion region increases upon reverse biasing the p-n junction leading to higher quantum efficiency.
- The carriers produced due to photons are driven by the potential applied generating photocurrent.
- Reverse biasing also helps in eliminating dark current.

- ✓ It is **positive – intrinsic – negative** photodiode, it consist of a **thick, lightly doped intrinsic layer** sandwiched between thin p and n regions.
- ✓ **Intrinsic layer is the depletion layer** where the absorption of photons occurs.
- ✓ Photons entering these layer produces carrier charges, this action results in **high quantum efficiency** of this device.
- ✓ As it is **reverse biased** the carriers produced are driven by the respective terminals.



RESPONSIVITY (R) AND QUANTUM EFFICIENCY (η)

- Responsivity is a measure of the conversion efficiency of a photodetector.
- Current produced is proportional to the number of incident photon

$$I_p \propto P \text{ (input power)}$$

$$I_p = RP$$

R is a constant called responsivity, measured in Amp/Watt

Photocurrent – number of electrons (N_e) times the electron charge (e) per unit time

$$I_p = (N_e e)/t$$

Light power – light energy per unit time

Light energy – energy of photons E_p times the number of photons (N_p)

$$P = (N_p E_p)/t$$

Substituting, $E_p = hc/\lambda$, $P = (N_p hc)/t\lambda$

But $I_p/P = R$

Substituting for P, and I_p we get

$$R = \frac{N_e e / t}{N_p h c / t \lambda}$$

$$R = \left(\frac{N_e}{N_p} \right) \left(\frac{e\lambda}{hc} \right)$$

The ratio of number of electrons produced (N_e) to the number of photons falling (N_p) shows the efficiency of the semiconductor material to convert light into current.

This ratio is called as quantum efficiency η of a photo diode

$$\eta = N_e / N_p$$

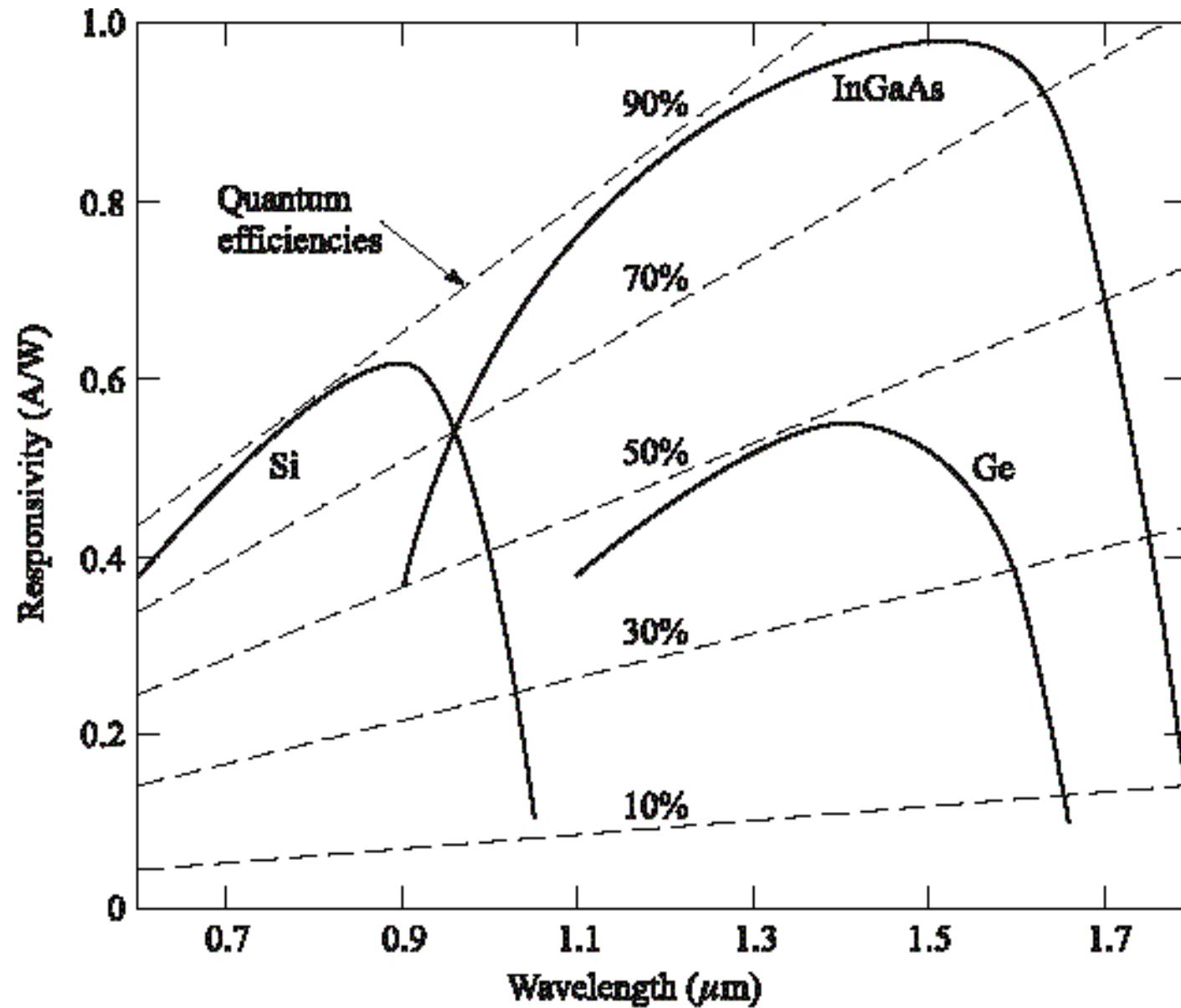
hence

$$R = \eta (e\lambda/hc)$$

The above relation shows the connection between the responsivity and quantum efficiency of a photo diode.

Responsivity of various P-I-N photodiodes

07._LED-_Laser_-_PIN_30-Mar-2021



Quantum efficiency vs. wavelength for various photodetectors

