OPTOELECTRONICS

Handout 2

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

Contents

2 Incoherent sources (the LED)

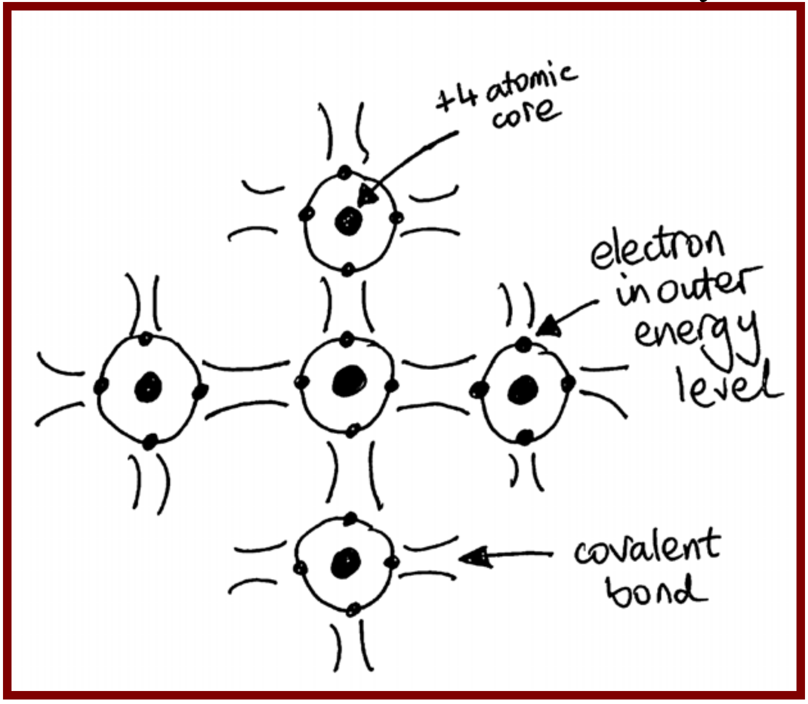
2 An incoherent source: the light emitting diode (LED)

An LED is one of a range of devices that is based on a ***semiconductor junction***. In order to understand how the device works we shall need to know more about semiconductors.

2.1 What is a semiconductor?

2.1.1 Semiconductor material structure

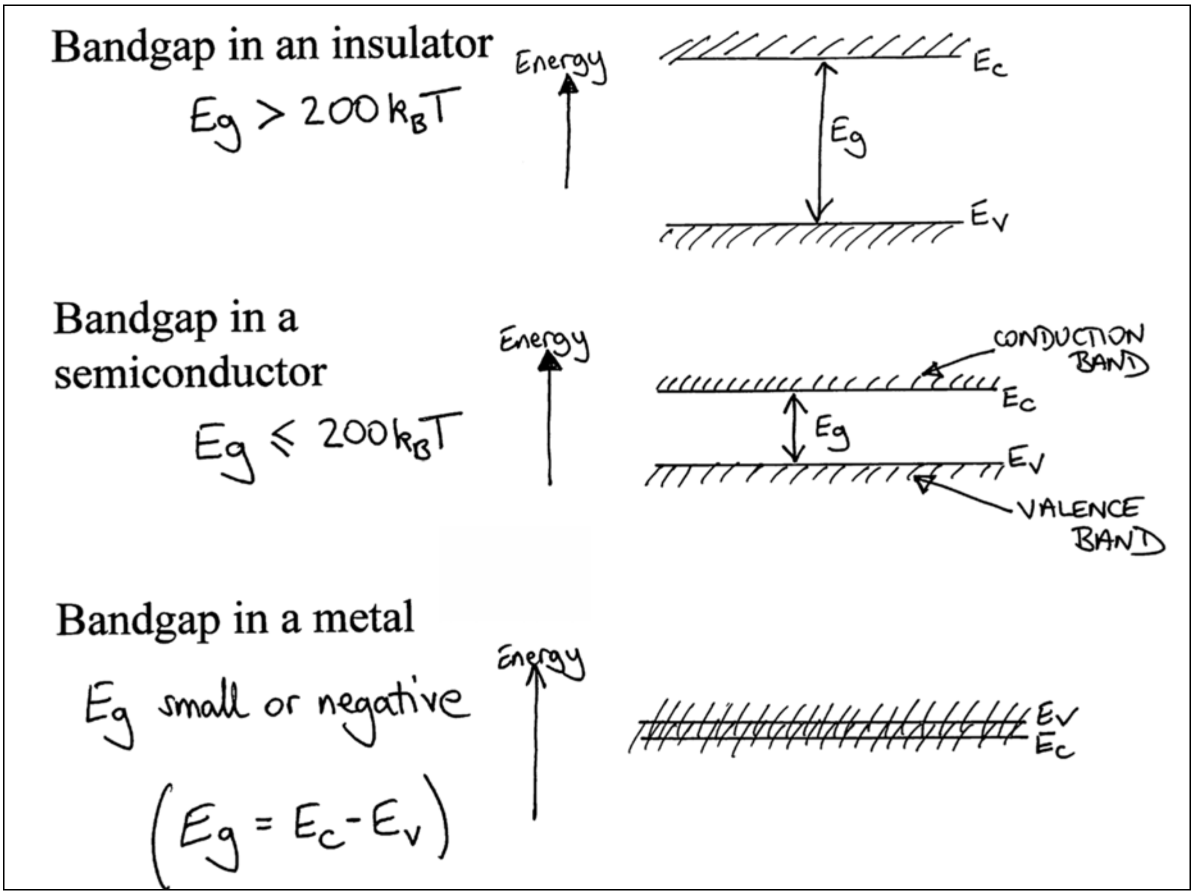
If we think of electrical conduction then at one end of the conduction scale we have metals and at the other we have insulators. Semiconductors are placed somewhere in the middle and are materials that sometimes behave as conductors and sometimes behave as insulators. Materials that are naturally semiconducting are the elements called ***intrinsic semiconductors*** and come from Group IV of the periodic table. Silicon (Si) and germanium (Ge) are natural semiconductors. Consider the structure of a crystal of silicon:



Schematic of the structure of silicon

Silicon has four electrons in its highest occupied energy level. These four electrons are each used in a covalent bond with another silicon atom, forming a lattice with a diamond structure. It is relatively easy to remove one of these electrons, enabling it to take part in conduction. This leaves a hole where the electron was. If we regard electrons as ***negative charge carriers*** then we may regard the hole as a ***positive charge carrier*** which may also take part in the conduction process (it will simply move in the opposite direction to the electrons). We have created an ***electron-hole pair***.

The minimum amount of energy required to create an electron-hole pair is called the ***bandgap energy, E_g***, of the material. This is represented in the energy level diagrams shown below.



Insulators, semiconductors, metals: energy bandgaps

Several points may be made about the diagram:

- the ***valence band*** is the highest occupied energy band involved in bonding the material together
- the ***conduction band*** is the energy band relating to free electrons, i.e. electrons that are able to conduct electricity
- we label the lowest energy in the conduction as E_C and the highest energy in the valence band as E_V , whereby the bandgap energy is $E_g = E_C - E_V$
- for a semiconductor,

$$\boxed{E_g \leq 200 k_B T \qquad (13)}$$

where k_B is Boltzmann's constant ($= 1.381\times10^{-23}$ JK^{-1}) and T is the temperature in kelvin.

From the diagram we can also see that:

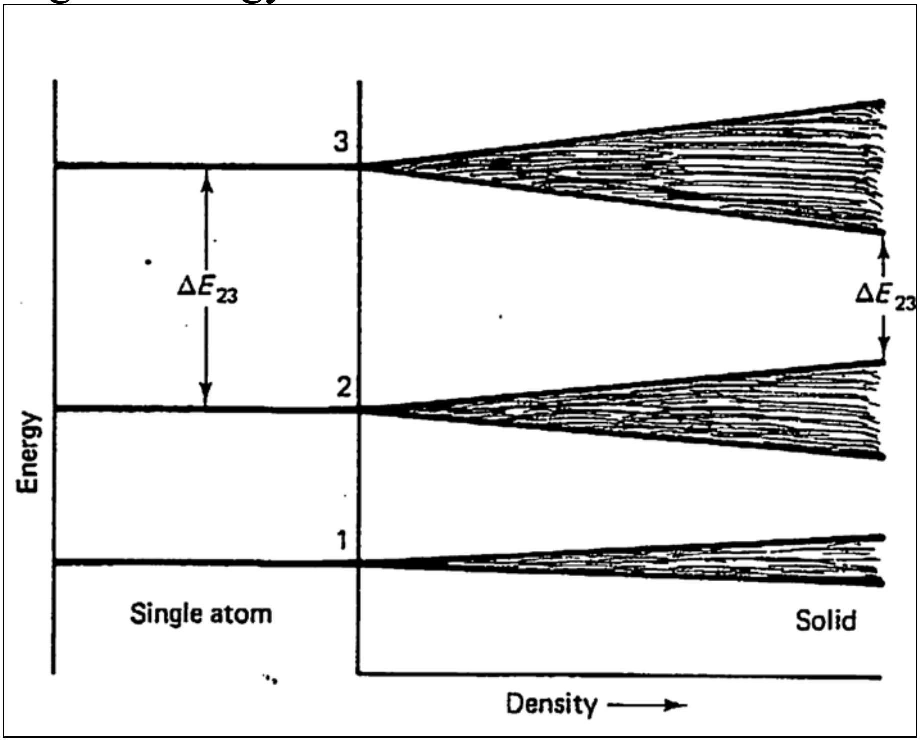
(i) for a metal the valence band and the conduction band overlap, therefore there are always electrons that have energies corresponding to a free electron

(ii) for an insulator the gap between the bands is so large that providing enough energy to move an electron from the valence to the conduction band would destroy the material

(iii) for a semiconductor the gap is no greater than about $200k_BT$, so enough thermal energy can be provided to move some electrons up to the conduction band without damaging the material.

It is worth noting here that we are no longer talking about energy *levels* but energy *bands*.

In a solid the atoms are so close together that their orbitals overlap and, where there used to be a single energy level when the atoms were far apart (e.g. as in a gas), there is now a whole range of energy levels that form bands.



Changes in available energy states of an atom as a function of the proximity of this atom to other atoms in a material.

As the density of the material increases, the progressive coupling of atoms can be thought in terms of: perturbing the isolated atom states (removing degeneracy), along with introducing new coupled-oscillator states. The Pauli exclusion principle dictates that new states need to be distinct, so more and more energy levels are introduced. The uncertainty principle also means a statistical spread of energy associated with each new state. This results in effective continuums ('bands') of possible electron states in solids.

Example

Estimate the maximum size of E_g in an intrinsic semiconductor.

$$\text{Maximum } E_g \approx 200 k_B T \approx 200 \times 1.38 \times 10^{-23} \text{ JK}^{-1} \times 293 \text{ K}$$

$$\approx 8.1 \times 10^{-19} \text{ J} .$$

where we have taken room temperature as $20\,^{\circ}\mathrm{C}$
implying T (Kelvins) = 273 + 20 = 293 K.

Alternative energy units (electron-volts, eV):

Size of electronic charge, $e \approx 1.6 \times 10^{-19}$ C (in coulombs)

and 1 CV = 1 J , therefore $1 \text{ eV} = 1.6 \times 10^{-19} \text{ CV}$

$$= 1.6 \times 10^{-19} \text{ J}$$

$$\text{and } \frac{1}{1.6 \times 10^{-19}} \text{ eV} = 1 \text{ J.}$$

$$\text{So, maximum } E_g \approx \frac{8.1 \times 10^{-19}}{1.6 \times 10^{-19}} \approx 5 \text{ eV} .$$

2.1.2 Semiconductor doping

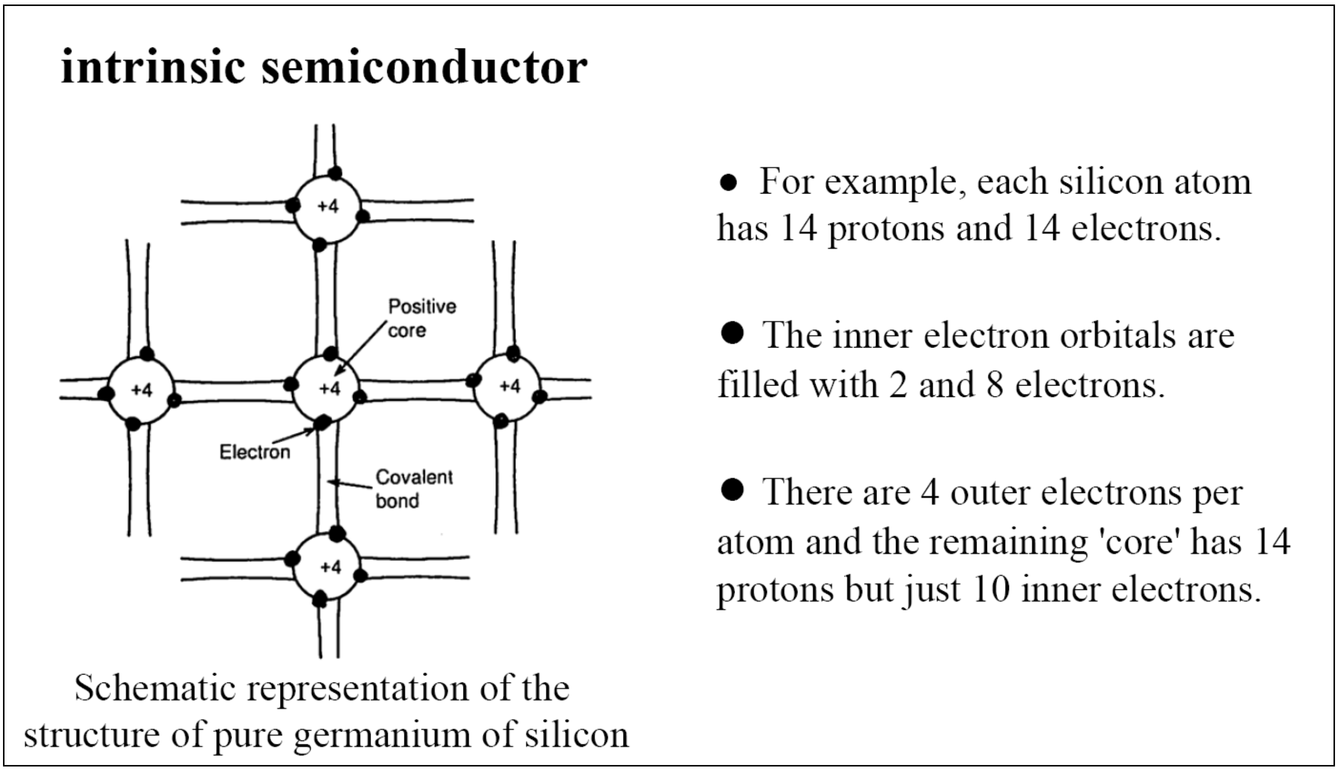
Unfortunately, intrinsic semiconductors do not make very good conductors as there are only about 10^{13} carriers per cm^3 (compared to 10^{28} carriers per cm^3 in a metal).

Normally, semiconductors are ***doped***, i.e. impurities (‘dopants’) are added, which changes the concentration of electrons or holes. these are ***extrinsic semiconductors***.

Extrinsic semiconductors come in two forms: ***n-type*** or ***p-type***.

n-type semiconductors
Here, some of the Si atoms are replaced by atoms with *five* electrons in their highest occupied energy level. Only four of these are required for bonding so the fifth is mostly free. Thus n-type semiconductors have an ***excess of electrons*** (relative to bonding requirements).

p-type semiconductors
Here, some of the Si atoms are replaced by atoms with *three* electrons in their highest occupied energy level. There is therefore one electron missing when it comes to bonding, i.e. there is a hole. Thus p-type semiconductors have an ***excess of holes*** (when compared to bonding requirements).



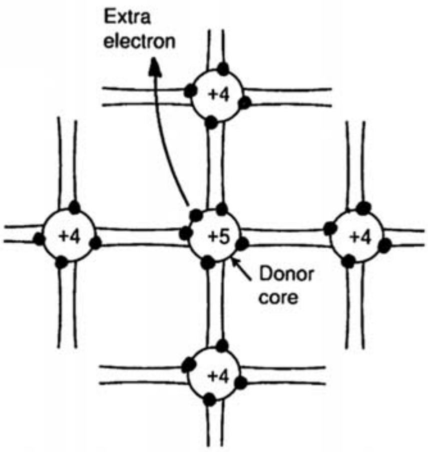
intrinsic semiconductor

- For example, each silicon atom has 14 protons and 14 electrons.
- The inner electron orbitals are filled with 2 and 8 electrons.
- There are 4 outer electrons per atom and the remaining 'core' has 14 protons but just 10 inner electrons.

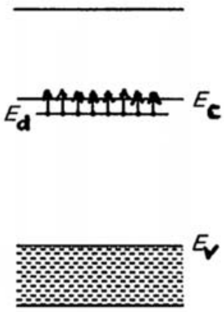
Schematic representation of the structure of pure germanium of silicon

Bonding structure of intrinsic group IV semiconductors

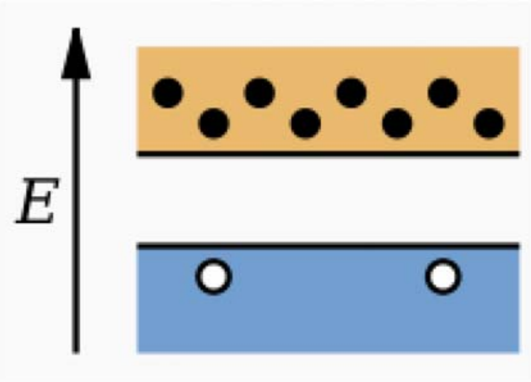
n-type semiconductor



(a) Schematic representation of a n-type semiconductor

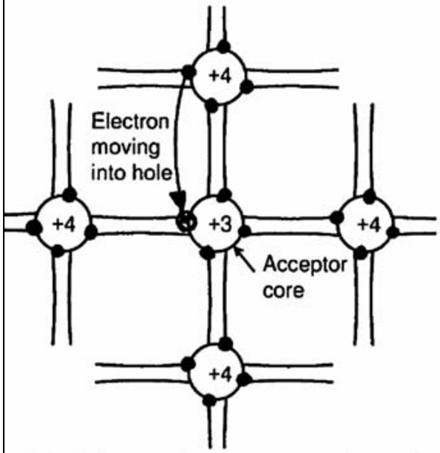


(b) Excess-to-bonding electrons effectively lie in 'donor states' very close to the conduction band

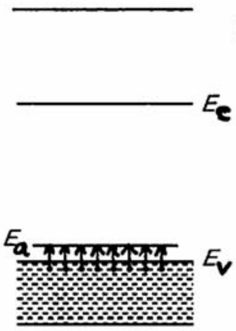


(c) Thermal excitation allows donor electrons, from dopant elements, easily to jump the small gap (about 1% of E_g). While the material is overall electrically neutral, free charge carriers (electrons) are donated to the conduction band.

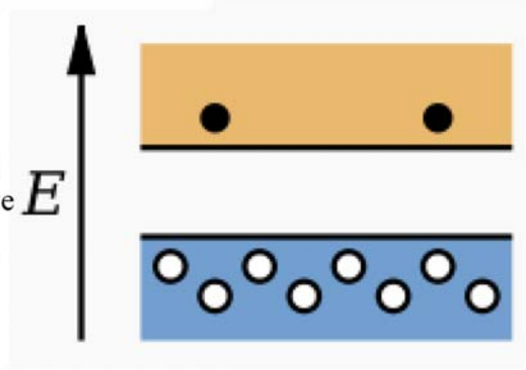
p-type semiconductor



(a) Schematic representation of a p-type semiconductor

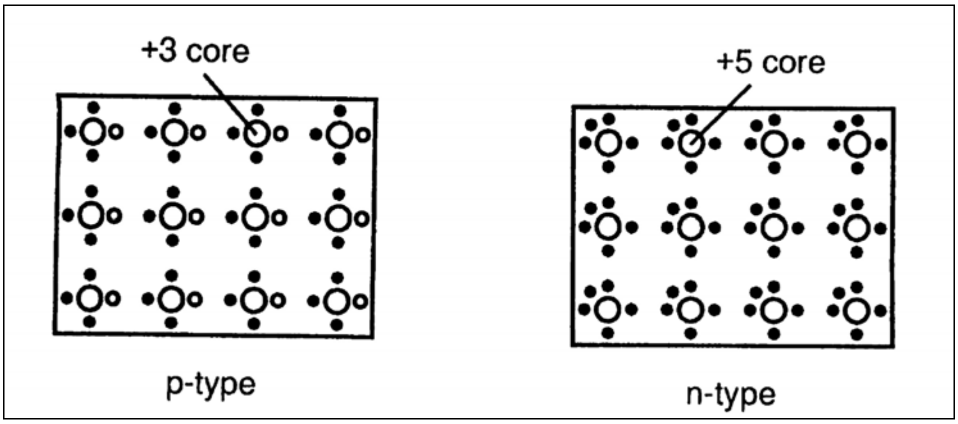


(b) A tiny amount of energy is needed to allow bonding electrons to move into 'acceptor states' that are very close to the valence band



(c) Thermal excitation allows electrons to move from one bonding/valence role to another. The effect is that (effectively positive) holes move in other direction. While the material is overall electrically neutral, the impurity atoms accept electrons - which creates mobile holes in the valence band.

If we now consider the fabrication of these two types of semiconductors alongside each other, then there is a greater concentration of conduction electrons in the n-type material. Conversely, there is a higher density of holes in the p-type material.

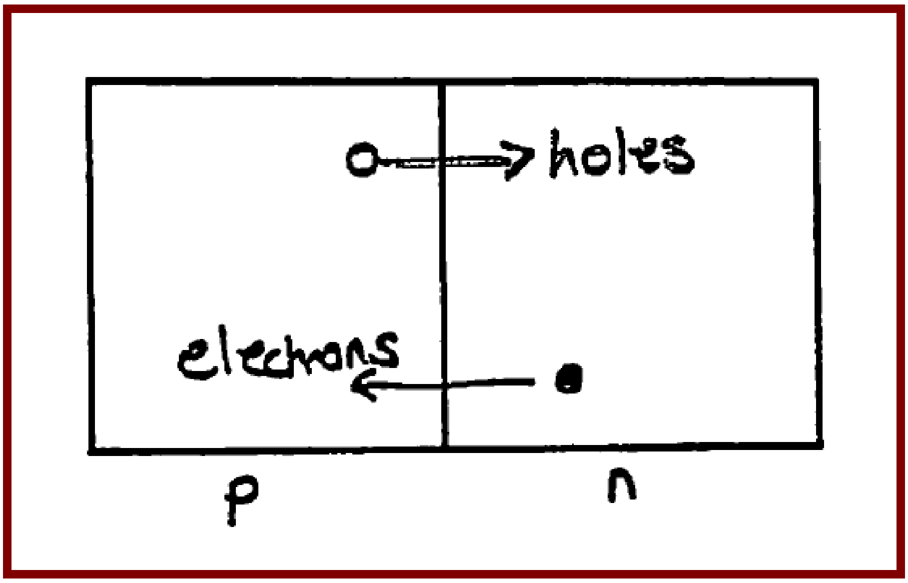


Two separated doped-semiconductors (only outer electrons and cores of some of the impurity atoms are shown)

2.1.3 The pn junction

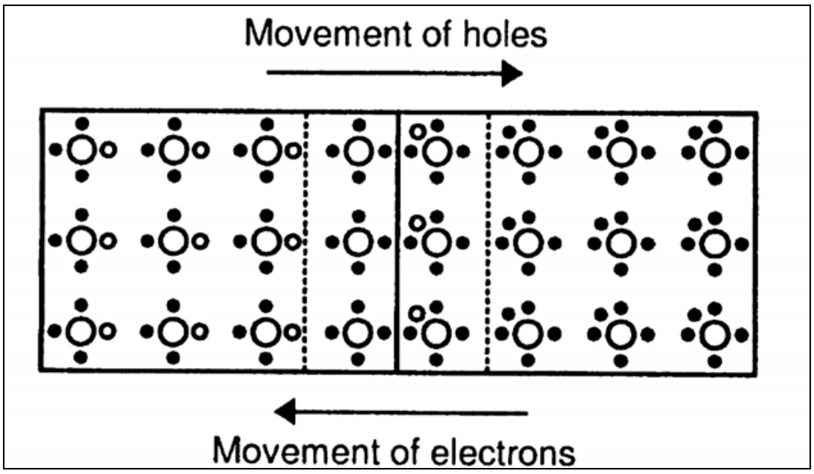
Since these carriers can move, *diffusion creates net currents of*:

free holes from the p-side to the n-side, and;
free electrons from the n-side to the p-side.



Initial flow of carriers across a pn junction

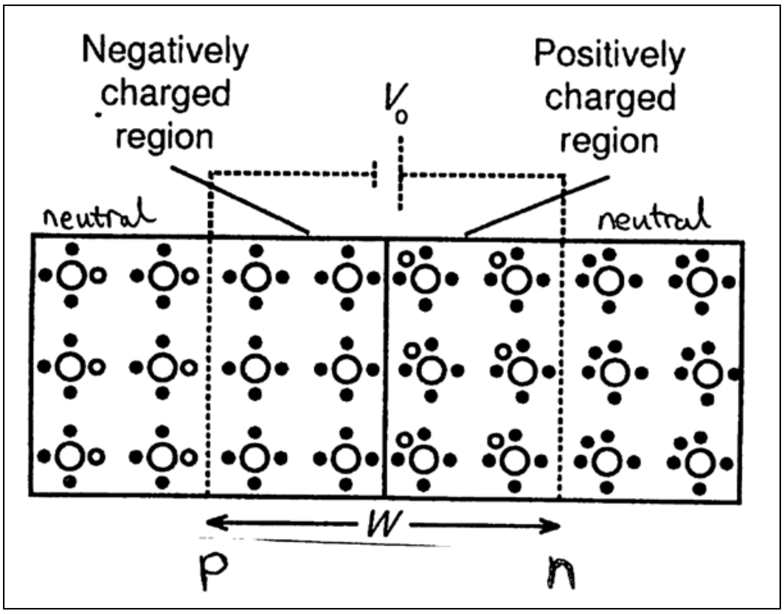
When the two types of carriers meet we get ***electron-hole recombination***. For this reason, the above currents are sometimes called ***recombination currents***.



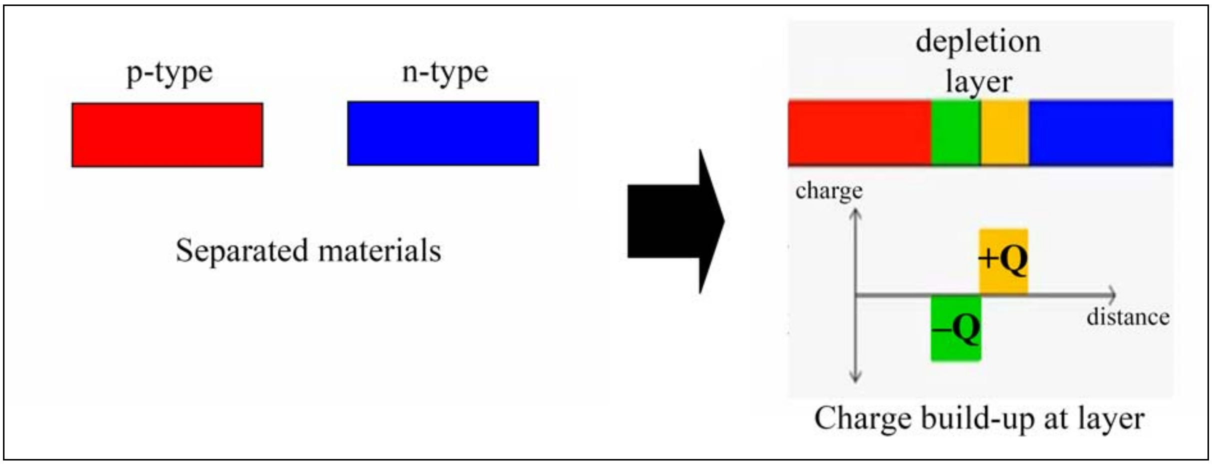
Recombination currents, due to diffusion, result in excess electrons falling into excess holes (becoming trapped) in the region of the pn junction

In an isolated piece of semiconductor this process does not continue indefinitely: an electric field is set up between the electrons that have moved into the p-type material and the holes that have moved into the n-type material. This field opposes the movement of further electrons and holes and eventually it is large enough to stop movement altogether.

We now have a layer that extends partly into the p-type and partly into the n-type, where all the excess holes have found an electron and all the excess electrons have found a hole. Thus, this region has been depleted of carriers and is known as the ***depletion layer***. The width of the depletion layer (e.g. μm's thick) depends on the original density of excess carriers, i.e. the dopant density on each side. A potential difference exists between the two sides of the depletion layer called the ***contact potential,*** $\boldsymbol{V_o}$. Typically, V_o is of the order of 1 V.

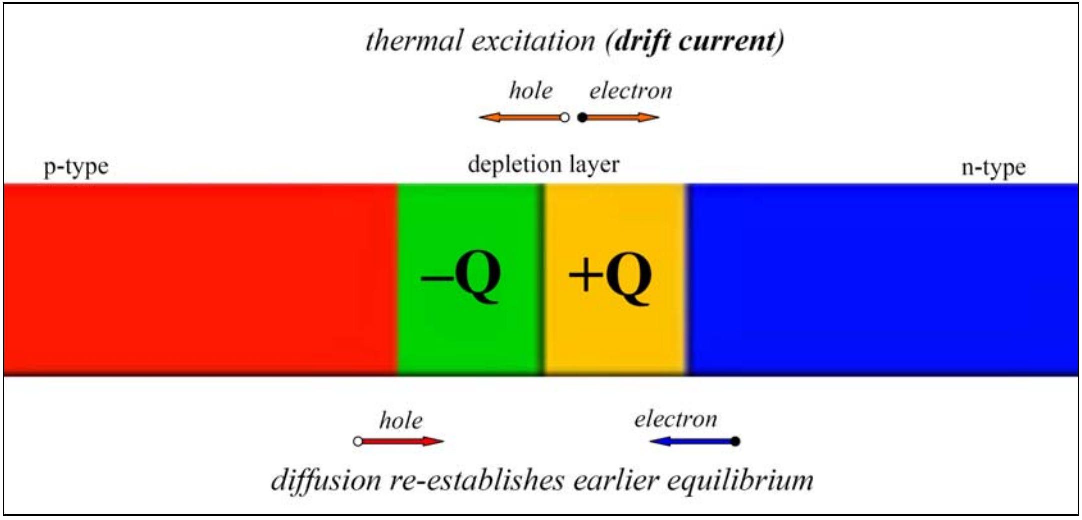


Depletion layer, of width W, forms and a contact potential V_0 forms due to charge build-up (halting further net currents)



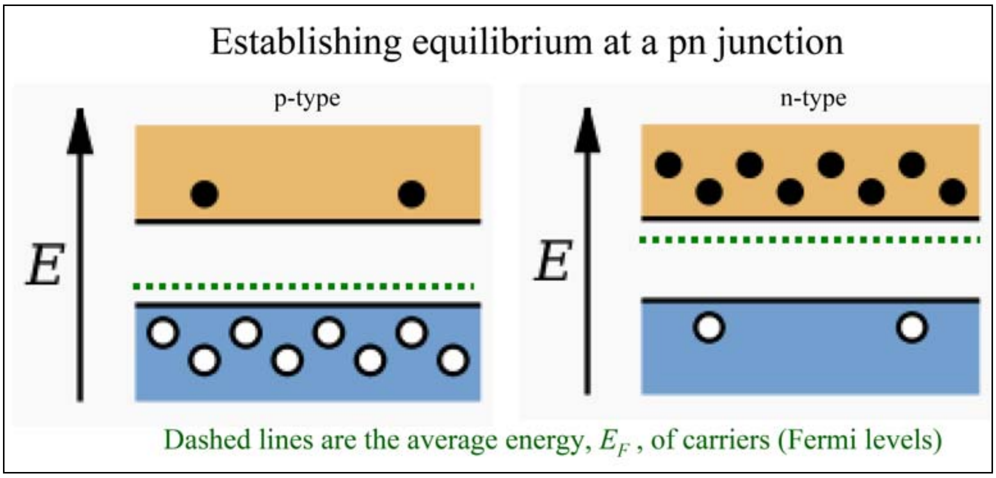
Electron captures result in charged impurity atoms (ions) that have fixed positions in the semiconductor material lattices

While there is no longer any net current, ***thermal excitation*** may still create electrons and holes in the depletion layer. This depends on the energy required to create a carrier pair, E_g , and the thermal energy factor k_BT. The size of the resultant (so-called ***drift current***) is proportional to the Boltzmann term $\exp(-E_g/k_BT)$, and reflects the rate of thermal generation of carriers. When these new carriers appear in the depletion layer, they experience the electric field arising from the contact potential and thus move in the *opposite* direction to the earlier recombination currents. Since the drift current acts to slightly reduce the contact potential, then diffusion recommences to allow establishment of the full contact potential and the earlier equilibrium. If the total diffusion current is I_0 then this balances the drift current, $-I_0$.



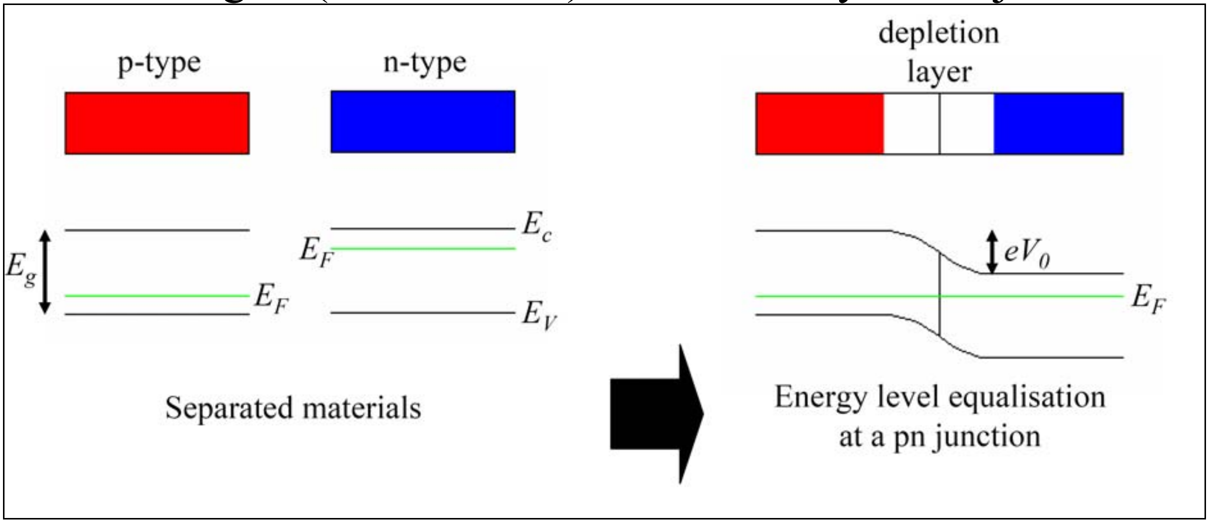
Equal and opposite diffusion and drift currents (open circuit)

Associated with the formation of the contact potential is the establishment of the equilibrium of the average *energy* of the excess carriers on the two sides of the pn junction. These are shown in the diagram below as dashed lines. In semiconductor theory, these average energy levels are called the ***Fermi levels***.



Fermi levels of two separated doped-semiconductors

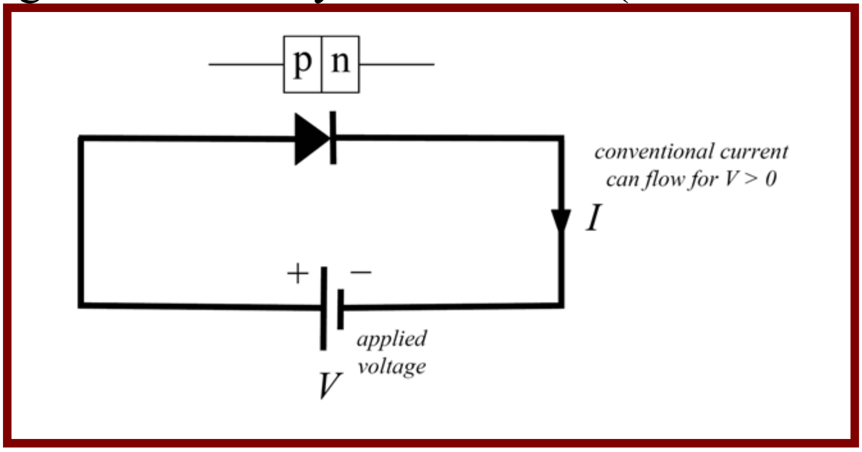
Away from the junction, the two materials retain their original individual energy configurations. The depletion layer presents an energy jump of size eV_0 for electrons trying to move from right to left. The conduction and valence bands of the materials bend to accommodate equalisation of the average carrier energies (Fermi levels) in the vicinity of the junction.



Energy equilibrium in pn junction formation (lectures movie)

2.2 Semiconductor diode in a circuit

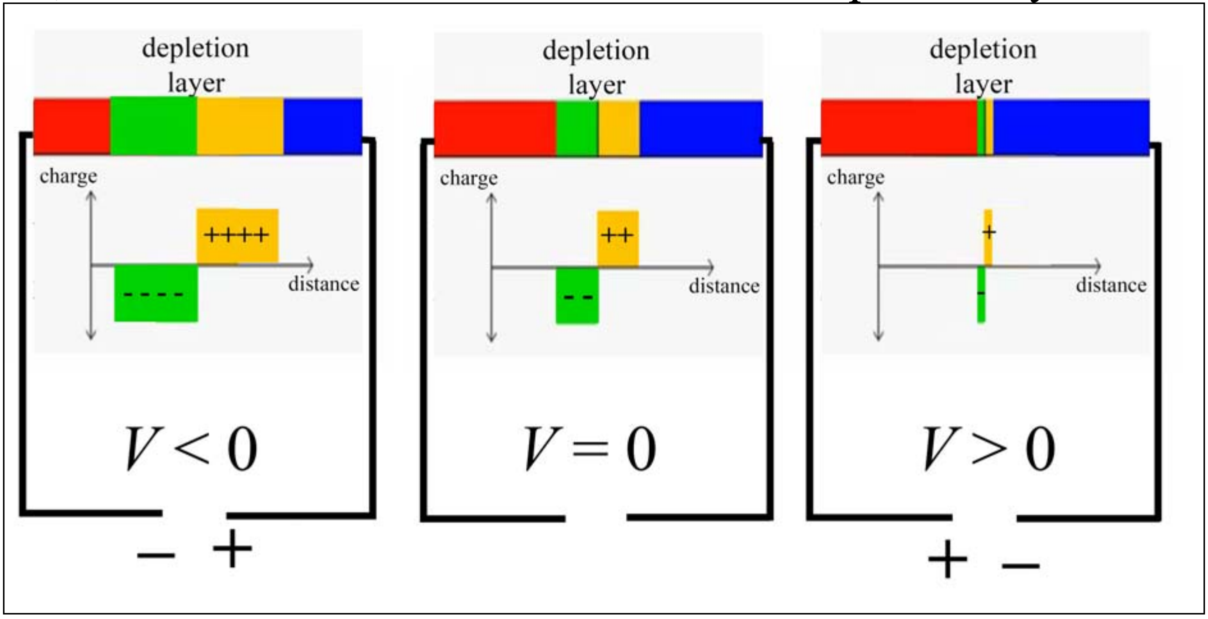
We can now consider the consequences of applying a potential difference across the pn junction. This creates a ***semiconductor diode***, as shown schematically below. As with earlier valve-based diodes, the component acts as a ***rectifier*** – conducting current in only one direction (for moderate V).



Semiconductor diode with forward bias ($V > 0$)

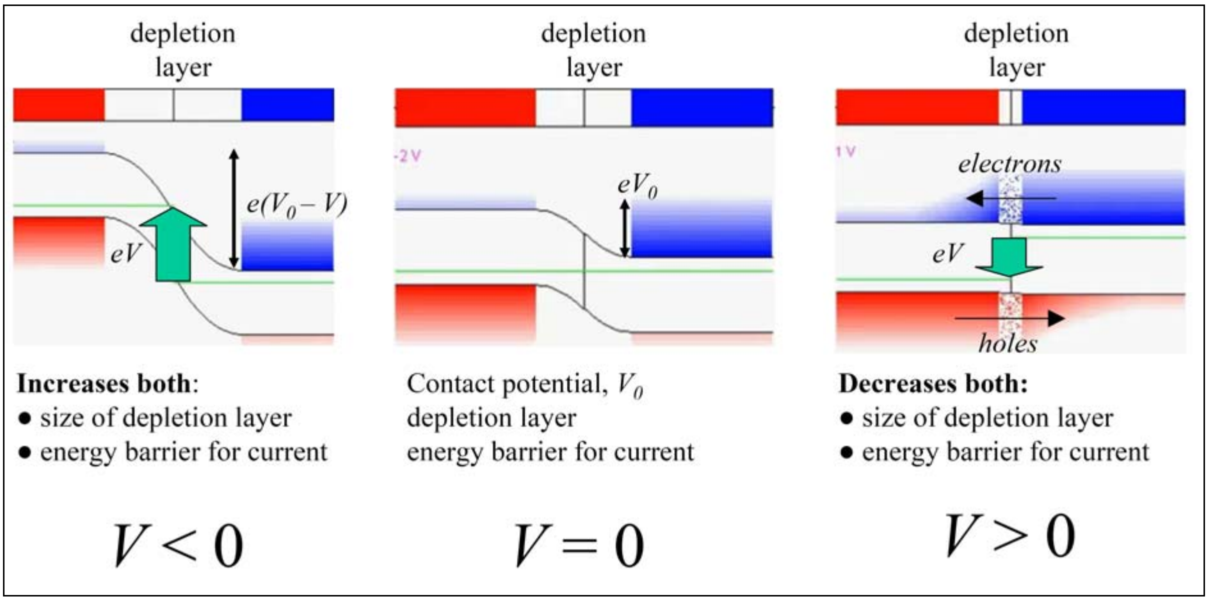
Application of the positive voltage terminal to the p-side and the negative terminal (or ground) to the n-side is called ***forward bias*** ($V > 0$). The reverse connection, where $V < 0$ and current does initially flow, is called ***reverse bias***.

In terms of the *charge build-up* that gives rise to the contact potential, reverse bias pushes electrons towards the p-side and draws electrons away from the n-side. This creates a wider layer that is depleted of the carriers supplied by the impurity atoms. Forward bias has the opposite effect: electrons are stripped away from the negatively-charged p-side and pushed toward the positively-charged n-side (filling the holes there). So, forward bias reduces the width of the depletion layer.



Variation of depletion layer width with applied voltage V

In terms of the *energy of the carriers*, and the shape of the conduction and valence bands in the vicinity of the pn junction, open circuit ($V = 0$) gives rise to an overall contact potential V_0 and an energy barrier of size eV_0 (i.e. charge times voltage, where e is the electronic charge). An additional applied voltage creates non-equilibrium between the average energy levels of the carriers on each side of the junction. With *reverse bias*, the depletion layer is extended and an energy imbalance of size $|eV|$ is imposed between the Fermi levels (see diagram below). Overall, the *energy barrier is increased* to $e(V_0 - V)$, recalling that $V < 0$ here. On the other hand, application of a *forward bias* reduces the depletion layer width and *reduces the energy barrier* to current flow. There will thus be a threshold amount of applied voltage, required to overcome the contact potential and to drive carriers across the junction.



Colour online: The variation of energy bands (black) and Fermi levels (green) with applied voltage V. Blue/red shading illustrates the density of electrons/holes, respectively

A forward bias of size V changes the energy barrier to current flow by an amount $\Delta E = -eV$. The Boltzmann factor of $\exp(-\Delta E/k_BT) = \exp(+eV/k_BT)$ estimates the fractional change in carrier population densities across the junction, and reflects the factor by which the diffusion current I_0 (which is driven by carrier concentration imbalance) is increased. The total current also includes the drift current $-I_0$ (such as from thermal excitations). This latter current depends on the number of background excitations per second and does not originate from processes associated with the applied voltage bias. The net current in a biased pn junction is thus

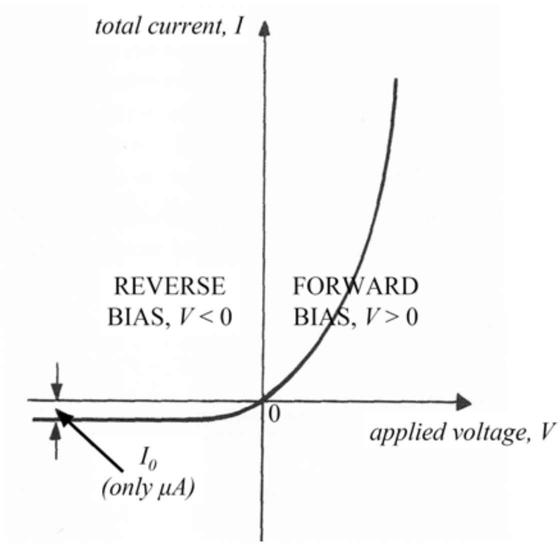
$$I = I_0 \exp\left(eV / k_B T\right) + \left(-I_0\right),$$

which yields the ***ideal diode equation***:

$$\boxed{I = I_0\left[\exp\left(eV / k_B T\right) - 1\right] \qquad (14)}$$

where

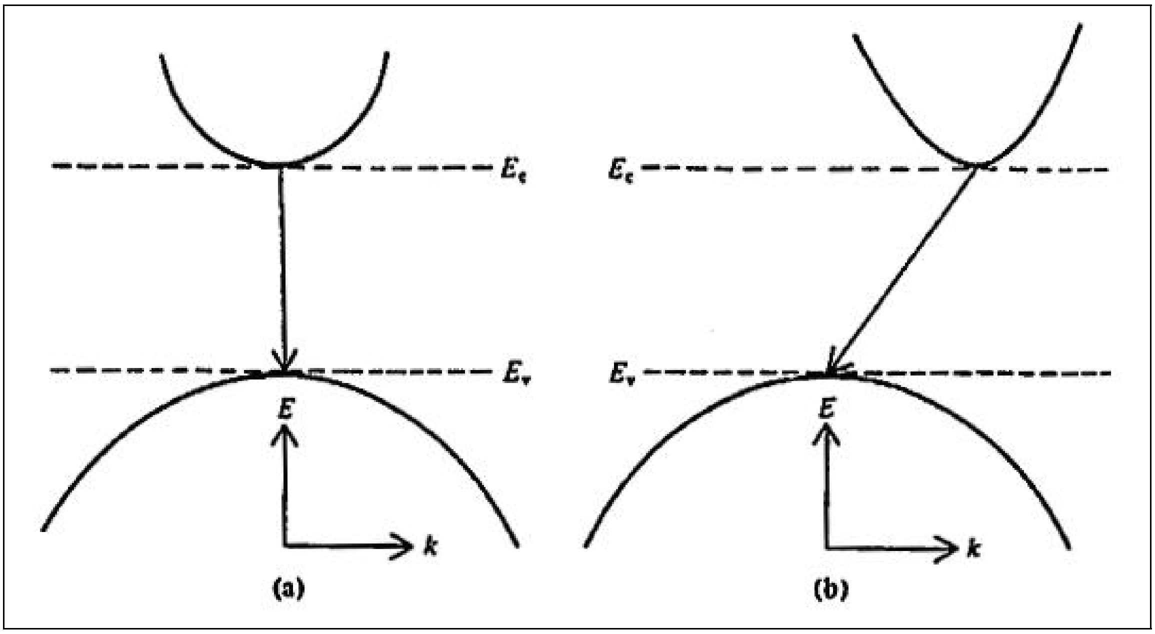
I_0 is often referred to as the ***reverse saturation current***
V is the applied voltage
e is the electron charge
k_B is Boltzmann’s constant
T is absolute temperature



When I_0 is due to thermal excitations, it involves the material energy bandgap parameter, through $I_0 \propto \exp(-E_g/k_BT)$.
It needs to be stressed that the above diode equation only models an ‘ideal diode’. Particular current-voltage characteristics of actual diodes do vary according to the details of the manufacturing process involved, such as the assumption of an abrupt transition between the p-type and n-type material layers.

2.3 Generation of light at a semiconductor junction

If the conditions are correct then light will be created when electrons and holes recombine across a semiconductor junction. Silicon based semiconductors do not produce light because they are what are known as ***indirect bandgap*** materials.



Electron transitions: (a) Direct bandgap; (b) Indirect bandgap

Note: total electron energy = potential energy + kinetic energy

$$= E_g + \tfrac{1}{2}mv^2 = E_g + \frac{p^2}{2m} = E_g + \frac{h^2}{2m\lambda^2} = E_g + \frac{h^2}{2m}\left(\frac{k}{2\pi}\right)^2$$

i.e. total electron energy = $E_g + \dfrac{h^2}{8\pi^2 m}k^2$

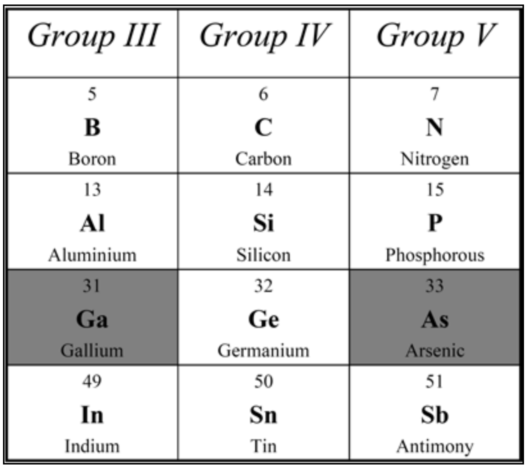
where momentum $p = \dfrac{h}{\lambda} = \left(\dfrac{h}{2\pi}\right)\left(\dfrac{2\pi}{\lambda}\right) = \dfrac{h}{2\pi}k$ has been used.

In ***indirect bandgap*** materials, when electron-hole recombination occurs, momentum must be lost by the electron as the transition involves not only a change in energy but also a change in *k*. If a photon were to generated from the electron's energy change, then that photon would also need to be the entity that preserves the momentum conservation law.

However, a photon cannot carry away the relatively large momentum associated with an indirect bandgap electron transition. Instead of this, the recombination process involves the absorption of *both* the electron's energy *and* its momentum by the material lattice. This type of material excitation is essentially the generation of heat at the junction.

Light can be produced by ***direct bandgap*** materials. In this case no electron momentum needs to be lost in the transition. So, the energy can result in producing a photon. This process of producing light is known as ***injection luminescence***.

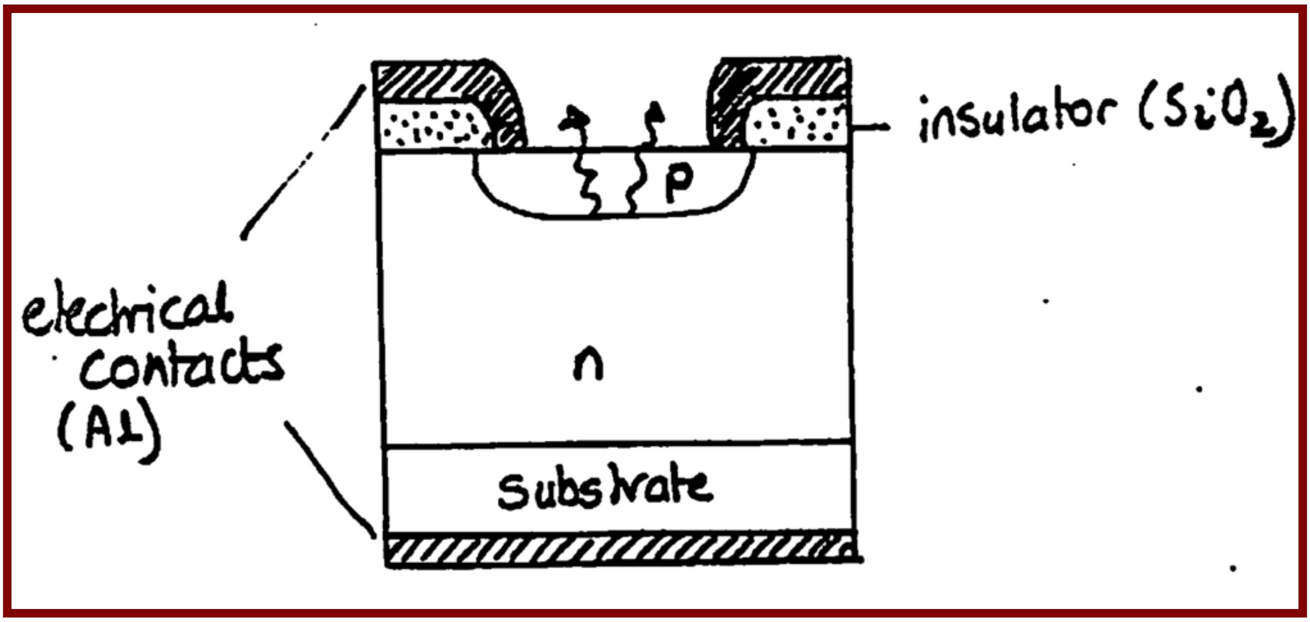
An example of a direct bandgap material is gallium arsenide (GaAs), which is commonly used for making LEDs. GaAs is an alloy semiconductor, often known as a III-IV semiconductor as its elements come from groups III and IV of the periodic table.



Group III	*Group IV*	*Group V*
5 **B** Boron	6 **C** Carbon	7 **N** Nitrogen
13 **Al** Aluminium	14 **Si** Silicon	15 **P** Phosphorous
31 **Ga** Gallium	32 **Ge** Germanium	33 **As** Arsenic
49 **In** Indium	50 **Sn** Tin	51 **Sb** Antimony

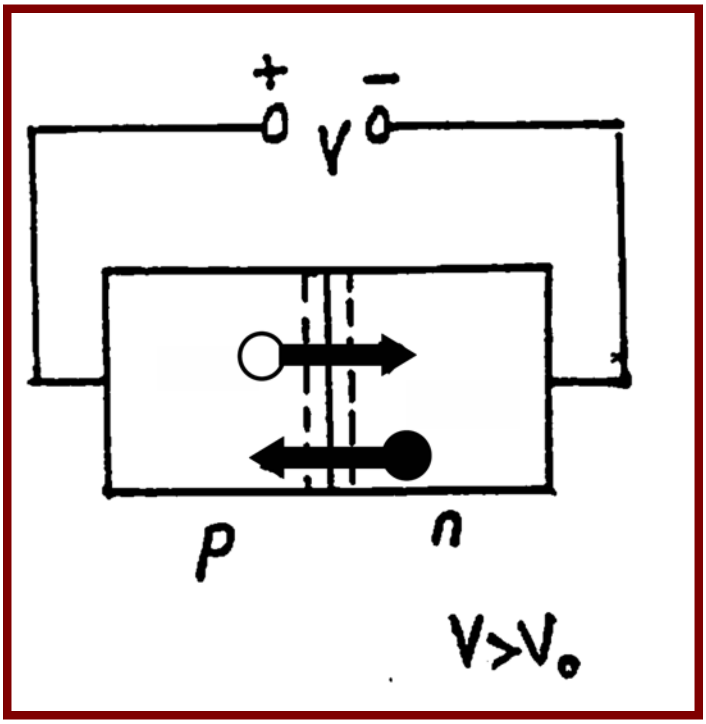
2.4 Construction and operation of an LED

A typical LED construction is shown below.



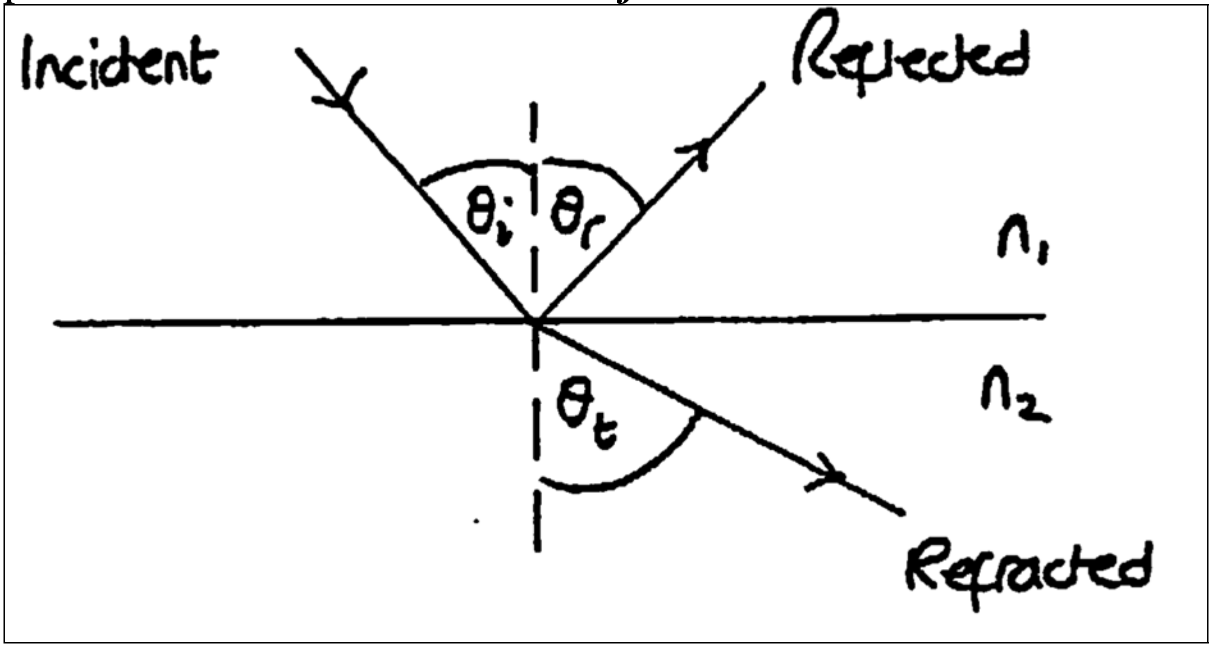
Construction of a typical LED

The semiconductor on the output side of the junction is made as thin as possible to allow the maximum amount of light to escape.



The electrodes allow an external voltage to be applied to the junction which overcomes the contact potential and drives the holes and electrons across the junction so that they can recombine and produce light. Only a few volts are required.

The internal ***quantum efficiency*** of the LED may be 100% (i.e. 1 photon per electron-hole pair recombination) but the external efficiency is much lower. This is because much of the light is trapped inside the material, mostly by the phenomenon of ***total internal reflection***.



Reflection and refraction at an interface

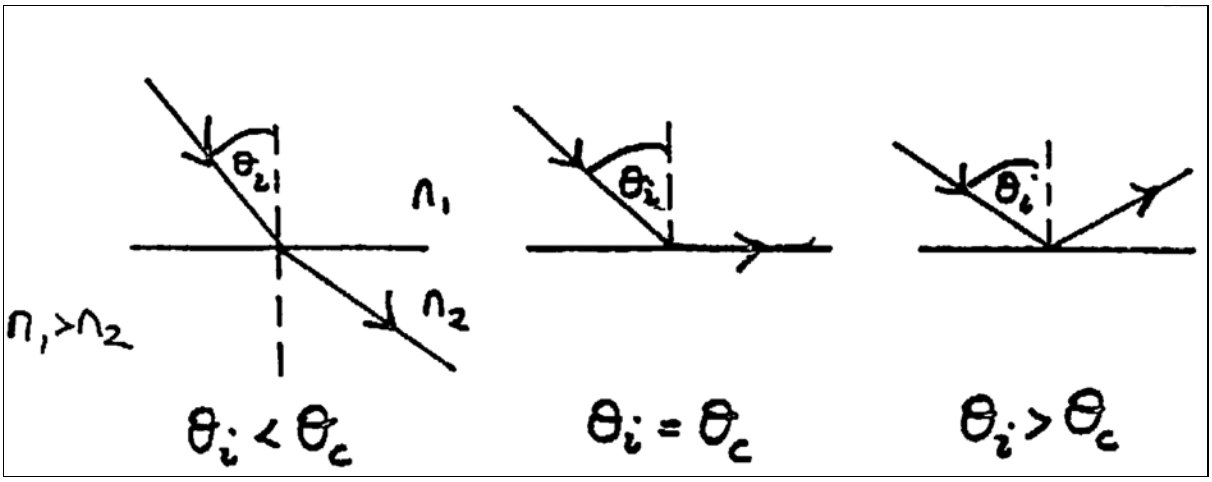
Snell's law states that

$$\boxed{n_1 \sin\theta_i = n_2 \sin\theta_t \qquad (15)}$$

If $n_1>n_2$ then we may reach a situation where $\theta_t = 90°$. The incident angle for which $\theta_t = 90°$ is called the ***critical angle,*** θ_C, which is given by:

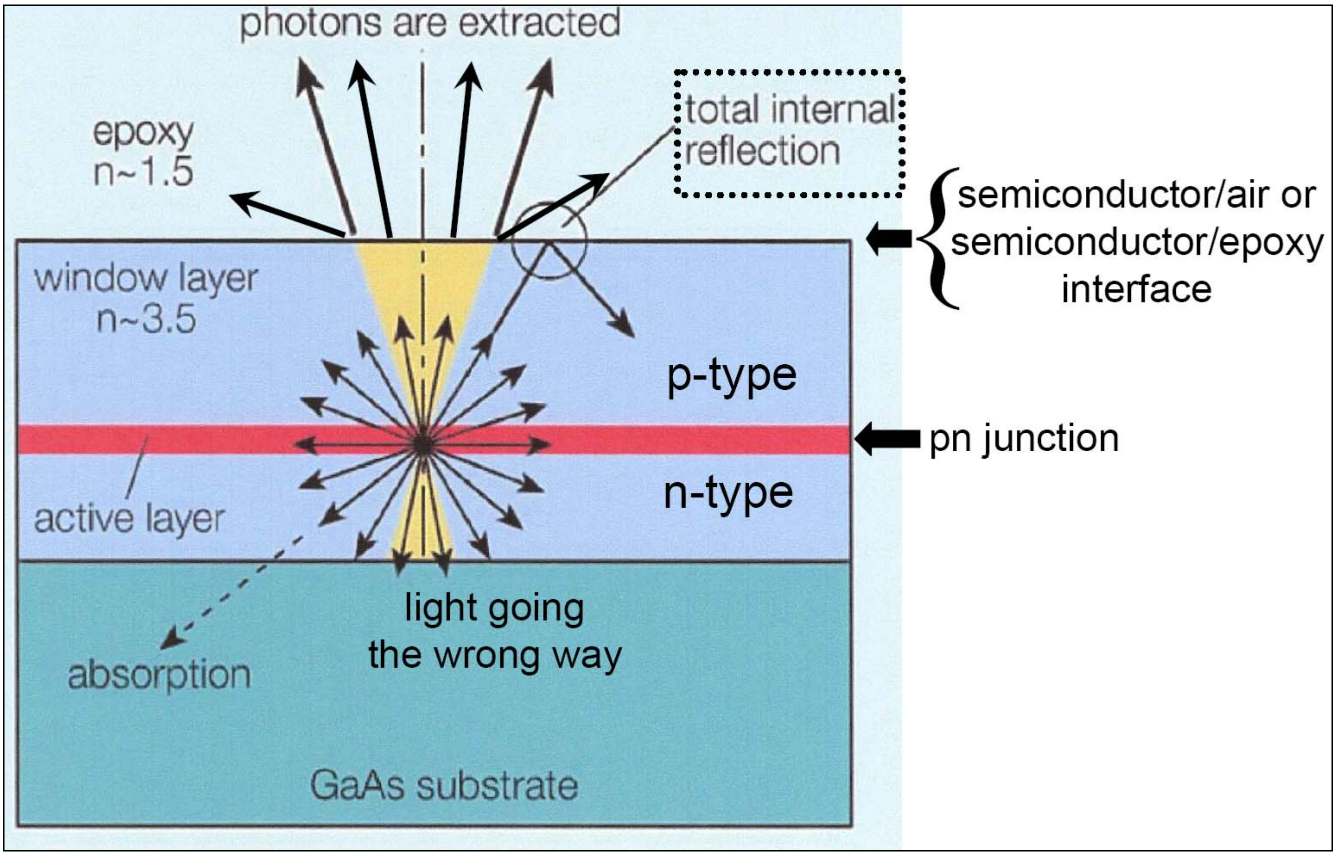
$$\boxed{\theta_C = \sin^{-1}\left(\frac{n_2}{n_1}\right) \qquad (16)}$$

If we increase θ_i such that $\theta_i > \theta_C$ then we achieve total internal reflection, i.e. all of the light incident at the boundary is reflected back into the first material.



The achievement of total internal reflection

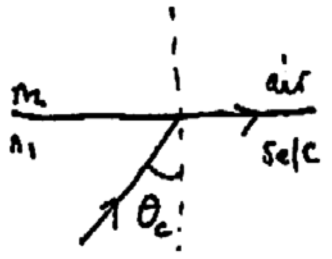
Semiconductors have relatively high refractive indices, so θ_C is quite small. Therefore, if we assume that light is given off in all directions, only a small proportion is incident on the boundary with $\theta_i < \theta_C$ and therefore can escape.



Loss of light at the exit of an LED

Example

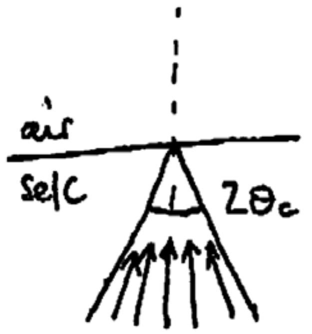
Calculate θ_C for a GaAs/air interface, where $n_{GaAs} = 3.6$ and $n_{air} = 1$. Hence, estimate the percentage of light leaving the device.



$$\theta_c = \sin^{-1}\left(\frac{n_2}{n_1}\right) = \sin^{-1}\left(\frac{n_{air}}{n_{GaAs}}\right)$$

$$\simeq \sin^{-1}\left(\frac{1}{3.6}\right) \simeq 16.1^\circ.$$

Light generated at all angles within device (i.e. 360°)

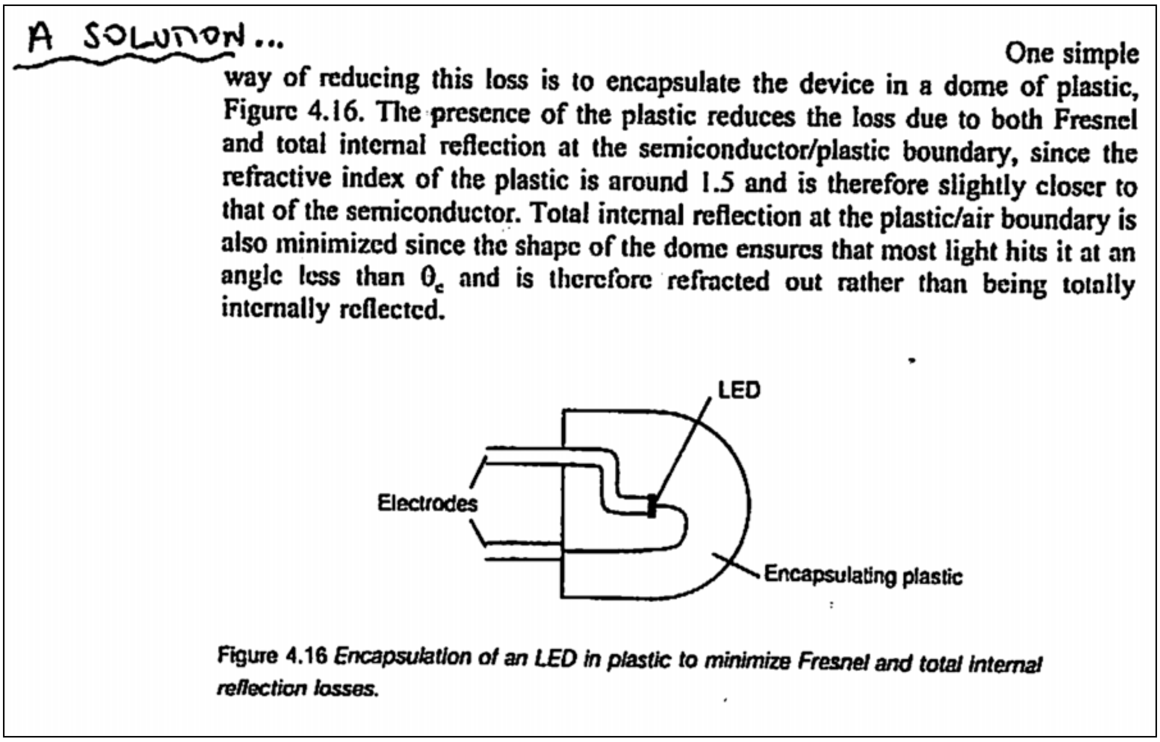


Cone of escaping light has angle $2\theta_C \simeq 32.2^\circ$

$$\therefore \text{\% escaping light} \simeq \frac{32.2^\circ}{360^\circ} \times 100\%$$

$$\simeq 9\%.$$

… and this does not include ‘Fresnel’ reflection losses, which tend to be at a level of 4% to 5% .



A SOLUTION...

One simple way of reducing this loss is to encapsulate the device in a dome of plastic, Figure 4.16. The presence of the plastic reduces the loss due to both Fresnel and total internal reflection at the semiconductor/plastic boundary, since the refractive index of the plastic is around 1.5 and is therefore slightly closer to that of the semiconductor. Total internal reflection at the plastic/air boundary is also minimized since the shape of the dome ensures that most light hits it at an angle less than θ_c and is therefore refracted out rather than being totally internally reflected.

Figure 4.16 *Encapsulation of an LED in plastic to minimize Fresnel and total internal reflection losses.*

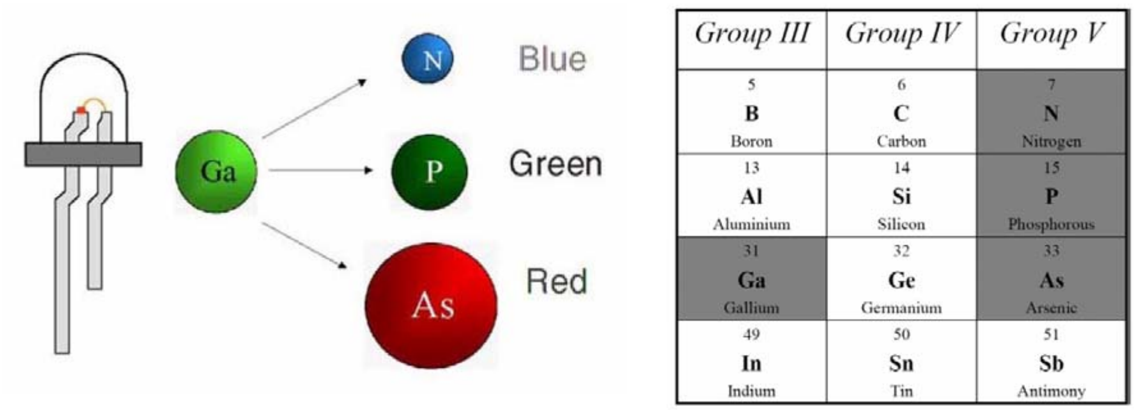
2.5 LED Materials

There are a number of requirements when considering materials that are suitable for LED construction. Clearly they must have bandgap energies that relate to visible photons, but they must also have low resistivities (low power consumption) and high quantum efficiencies. This means that single element semiconductors are not suitable.

Some common commercial LED materials are shown in the table below.

Material	*Wavelength*	*Colour*
GaAs	9110-1020nm	Infrared
GaAlAs	650, 870-890	Red, infrared
GaP	570, 590, 700	Green, yellow, red
GaAsP	589, 632, 650	Yellow, orange, red
(In)GaN	550-360nm	Green, blue, ultraviolet

Some common LED materials



Group III	Group IV	Group V
5 **B** Boron	6 **C** Carbon	7 **N** Nitrogen
13 **Al** Aluminium	14 **Si** Silicon	15 **P** Phosphorous
31 **Ga** Gallium	32 **Ge** Germanium	33 **As** Arsenic
49 **In** Indium	50 **Sn** Tin	51 **Sb** Antimony

As atoms get larger, outer electron energy states get closer, transitions between these levels are thus of lower frequency

Trends of LED output with respect to the periodic table

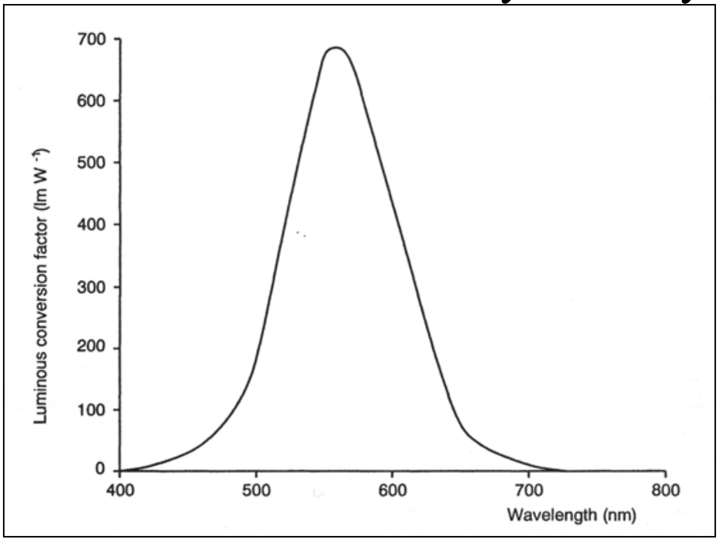
2.6 Further LED updates (non-examinable)

LED technology has advanced tremendously since the first demonstration of a practical visible-light LED in 1962. Not only have they become more powerful but breakthroughs in materials technologies have led to many new developments. For example, in 1976 their materials were specifically refined for optical fibre communications.

While it is estimated that around 80% of LEDs are still of the classic ‘3mm’ and ‘5mm’ dome-shaped variety, detailed earlier, much recent attention has centred on many developments of blue-emitting indium-gallium-nitride (InGaN) LED systems. It was not until the mid 1990s that the first efficient blue LEDs were demonstrated. This was quickly followed by deployment of this blue light in the design of white light LEDs. It was near the end of the 1990s when usefully high power LEDs were made commercially available.

2.6.1 Lumens versus watts

Before reviewing developments, some units quantifying light output need to be outlined. *Photometric units* are matched to the response of the human eye, and how well we perceive brightness of different colours – e.g. the ***lumen*** (lm). *Radiometric units* give light strength regardless of the wavelength involved e.g. the ***luminous flux*** (watts). For describing LEDs, one often quotes the ***luminous performance/efficiency/efficacy***, which gives the lumens output per electrical watt (lmW^{-1}). The relationship between lumens and watts reflects the sensitivity of the eye:

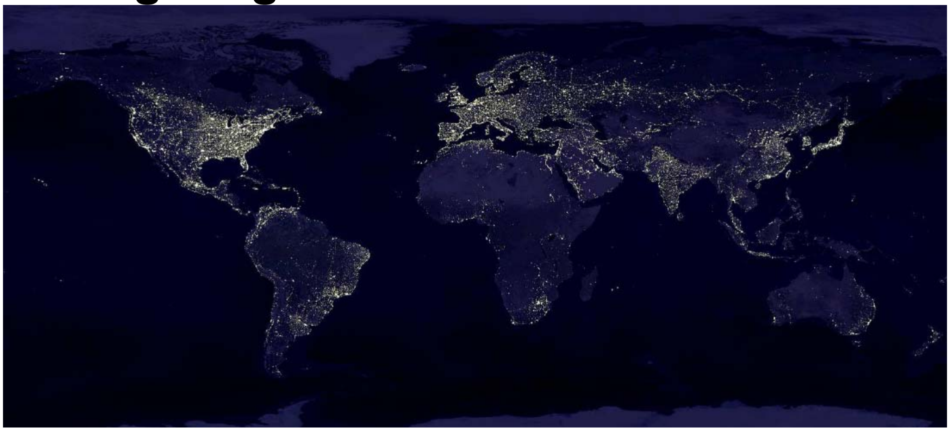


Lumens versus watts as a function of wavelength

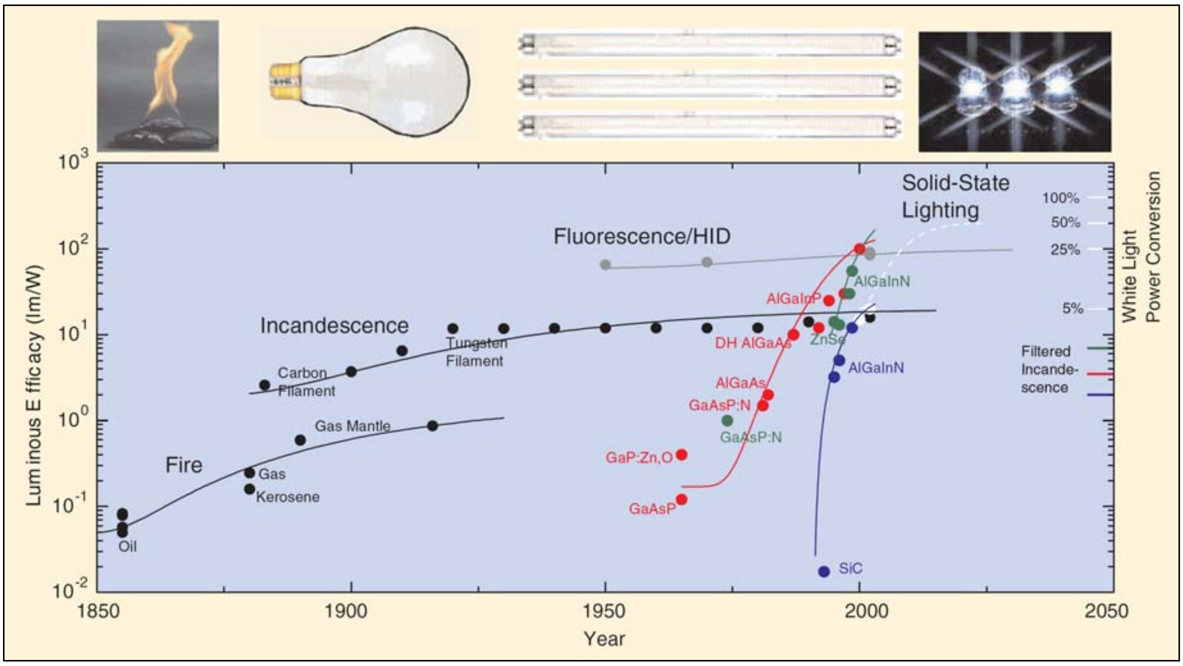
For example, at $\lambda = 550$ nm 1 W $\Rightarrow$ 680 lm
while at $\lambda = 600$ nm 1 W $\Rightarrow$ 430 lm.

2.6.2 Solid-state lighting trends

Lighting accounts for a fifth of the world's energy consumption.



LEDs are twice as energy efficient as energy-saving light bulbs (which work on fluorescence). With their high energy efficiency and other strong environmental attributes (no lead, no mercury, long operating lifetime implying reduced waste), LEDs are certain to be dominant in the future of lighting.

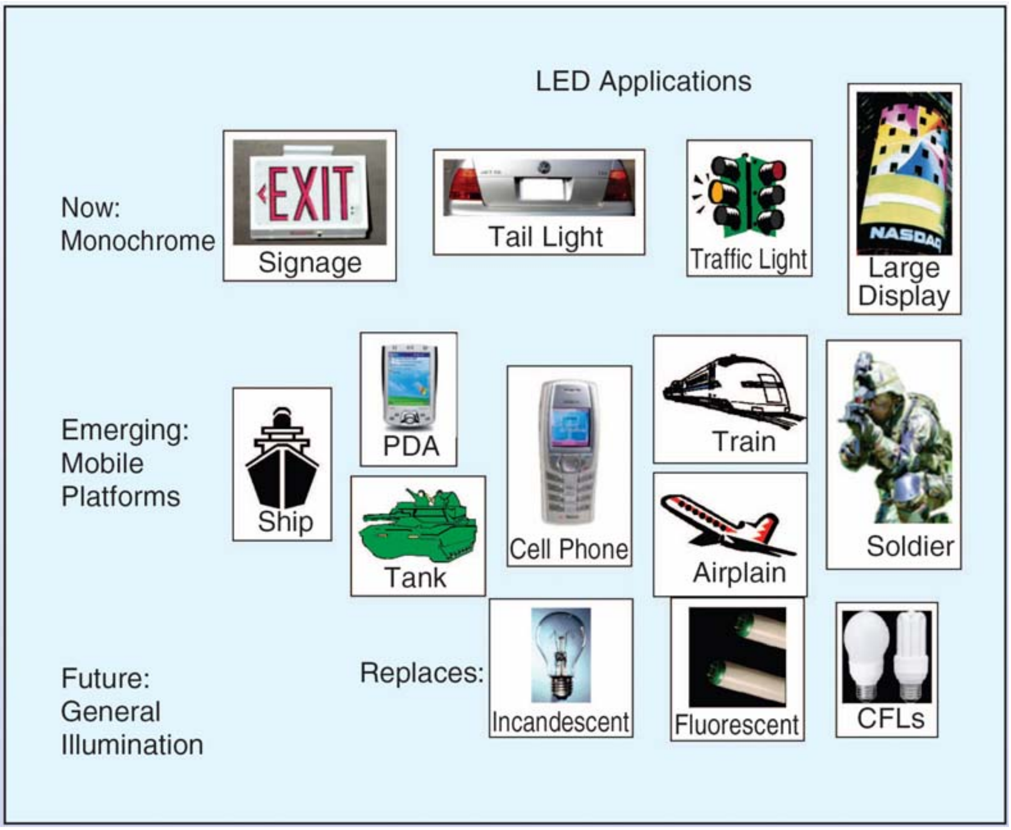


History of luminous efficiency of various light sources

Fast LED switching times, efficient high power, and potential dynamic control of the specific spectral content of generated white light offer many application opportunities, such as more responsive and ‘mood-orientated’ lighting applications.

The main current barrier to widespread deployment of LEDs in lighting systems is that their cost needs driven down by smarter designs and more larger-scale production techniques.

But world-wide research in this area is very intense and most experts expect continuation of Moore’s Law-like growth in developments and improvements of performance parameters.

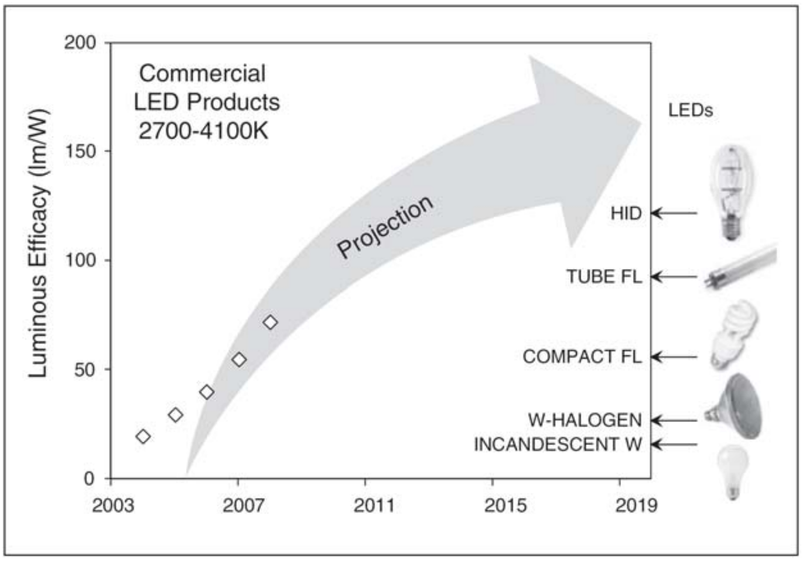


Selection of existing and emerging LED application areas

LEDs are advantageous in many ways, when compared to other light sources. Their compact size, robustness, lifetimes and fast switching times lend themselves to deployment in many areas such as large scale and local communications technologies (e.g. current usage in device remote controls). A key area is in general background lighting applications.

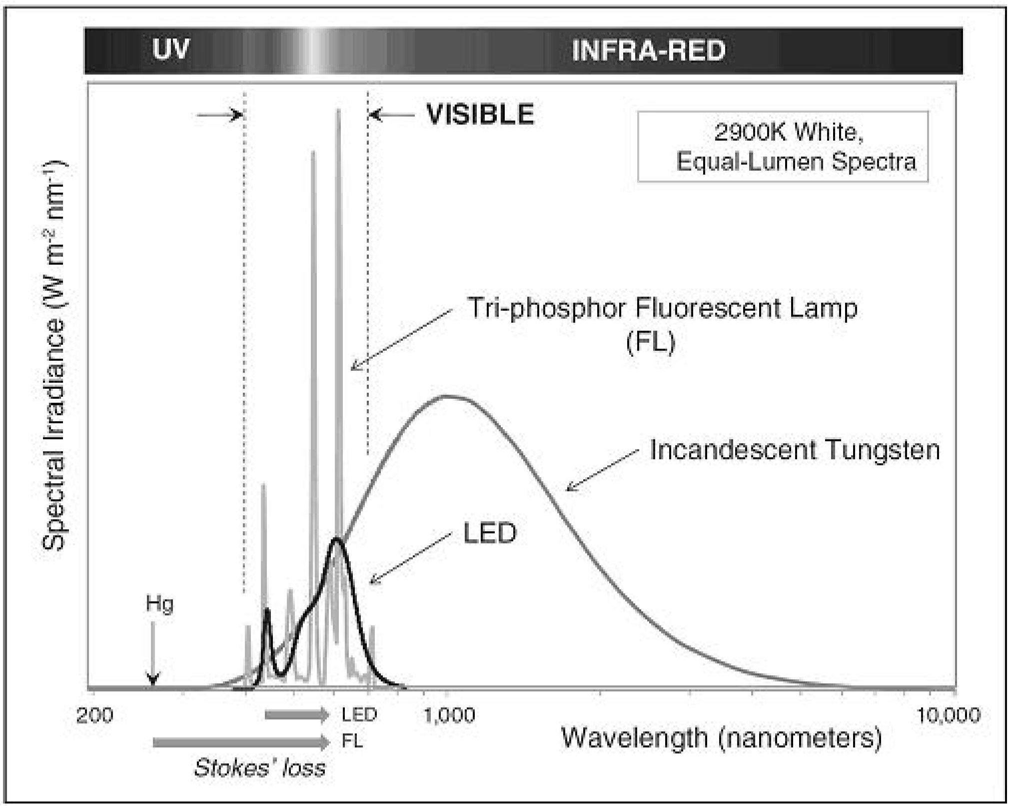
Bruning, Philips, 2009

Classic bulb vs energy-saving bulb vs LED bulb



Predicted evolution of the luminous efficacy (lumens per electrical watt) for commercial LED products. At the right, typical luminous efficacies are indicated for conventional lighting technologies.

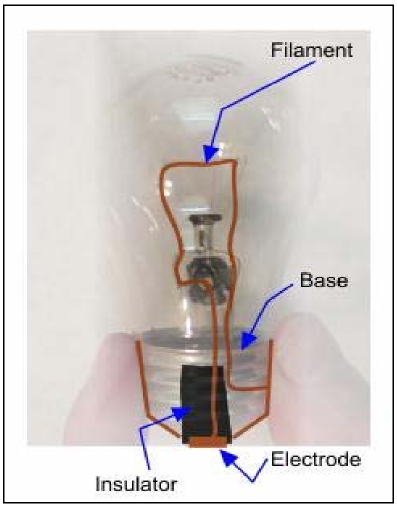
2.6.3 Efficiency comparison of lighting sources



Comparison of spectral compositions of ‘white light’

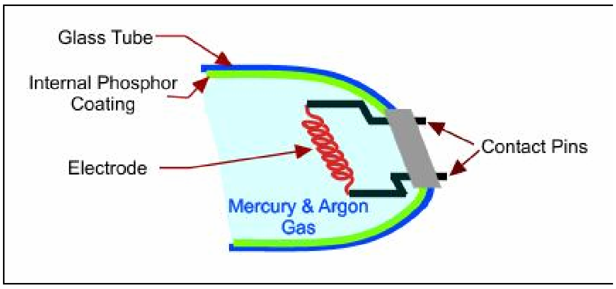
The tungsten filament (household bulb) radiates most of its energy in the infrared range (i.e. radiation outside of the visible spectrum).

As little as 5% of the energy output is light, the rest is effectively heat.



Tungsten filament bulb

Also shown is the spectrum from a typical Hg (mercury) vapour fluorescent lamp that employs three types of phosphors. Electrical discharge in the Hg vapour gives off ultraviolet radiation that (frequency) down-converts to visible light in the fluorescent phosphors of the glass tube coating.



Fluorescent Bulb

The heat generated through converting the light to lower frequency is labelled as the Stokes' loss in the diagram. This is still quite a lot of energy loss to heat.

Finally, blue LEDs are shown to be able to down-convert to a range of visible wavelengths with much less energy loss to heat (note the much smaller Stokes' shift). They can thus be a fundamentally far more energy efficient than the other two means of generating visible light.

2.6.4 Predictions for LED operational parameters

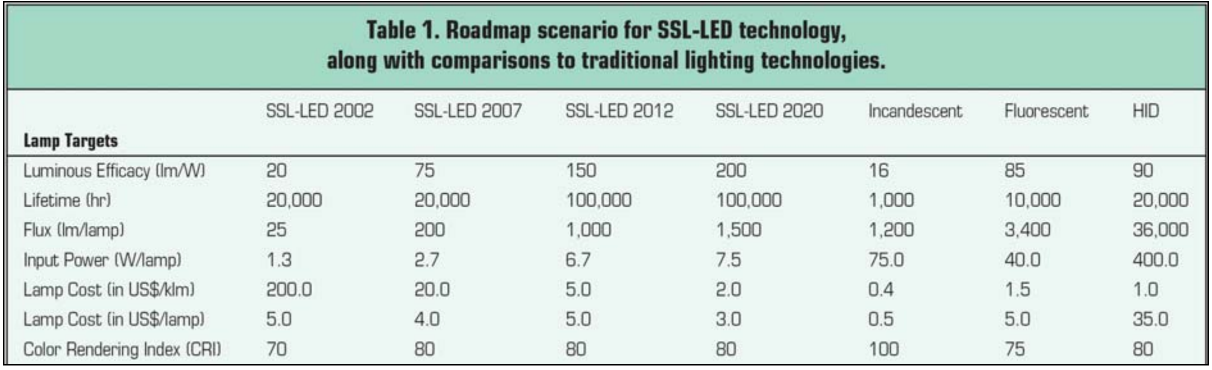
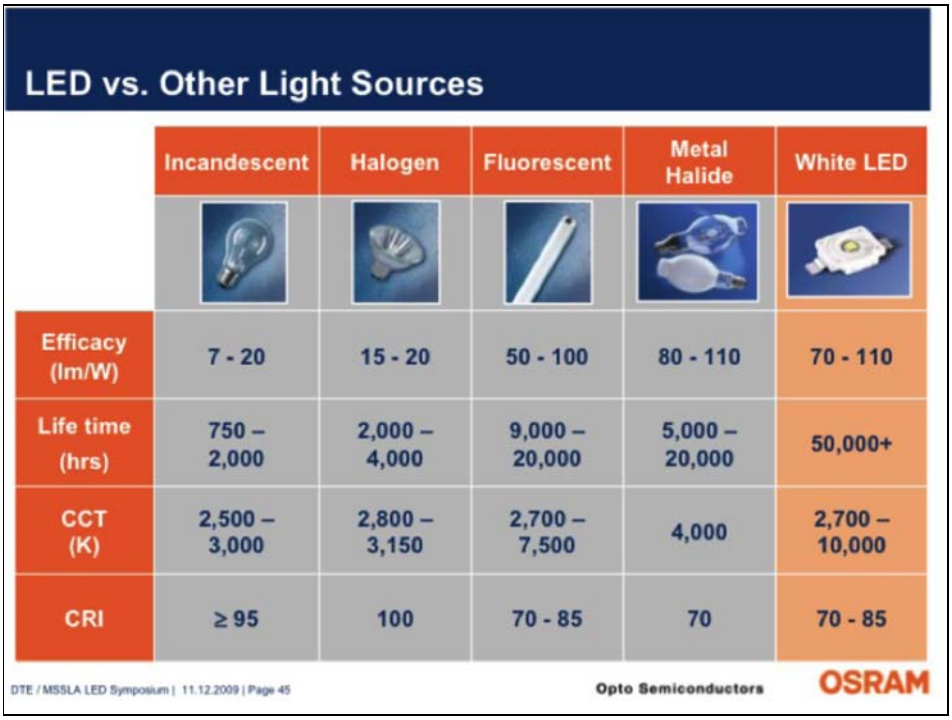


Table 1. Roadmap scenario for SSL-LED technology, along with comparisons to traditional lighting technologies.

Lamp Targets	SSL-LED 2002	SSL-LED 2007	SSL-LED 2012	SSL-LED 2020	Incandescent	Fluorescent	HID
Luminous Efficacy (lm/W)	20	75	150	200	16	85	90
Lifetime (hr)	20,000	20,000	100,000	100,000	1,000	10,000	20,000
Flux (lm/lamp)	25	200	1,000	1,500	1,200	3,400	36,000
Input Power (W/lamp)	1.3	2.7	6.7	7.5	75.0	40.0	400.0
Lamp Cost (in US$/klm)	200.0	20.0	5.0	2.0	0.4	1.5	1.0
Lamp Cost (in US$/lamp)	5.0	4.0	5.0	3.0	0.5	5.0	35.0
Color Rendering Index (CRI)	70	80	80	80	100	75	80

Projections for solid-state lighting (SSL) LED technology



LED vs. Other Light Sources

	Incandescent	Halogen	Fluorescent	Metal Halide	White LED
Efficacy (lm/W)	7 - 20	15 - 20	50 - 100	80 - 110	70 - 110
Life time (hrs)	750 – 2,000	2,000 – 4,000	9,000 – 20,000	5,000 – 20,000	50,000+
CCT (K)	2,500 – 3,000	2,800 – 3,150	2,700 – 7,500	4,000	2,700 – 10,000
CRI	≥ 95	100	70 - 85	70	70 - 85

DTE / MSSLA LED Symposium | 11.12.2009 | Page 45

Opto Semiconductors

OSRAM

Actual Conference Data from November 2009

In the above table there are two characteristics (CCT, CRI) that are defined below.

Correlated colour temperature (CCT)

Correlated colour temperature (CCT) is usually stated in units of absolute temperature, K, and is derived from the theory of black body radiation. For example, high values of over 5,000K are called *cool colours* (and appear as a blueish white), while lower colour temperatures (2,700 to 3,000 K) are called *warm colours* (and appear as a yellowish white through red).

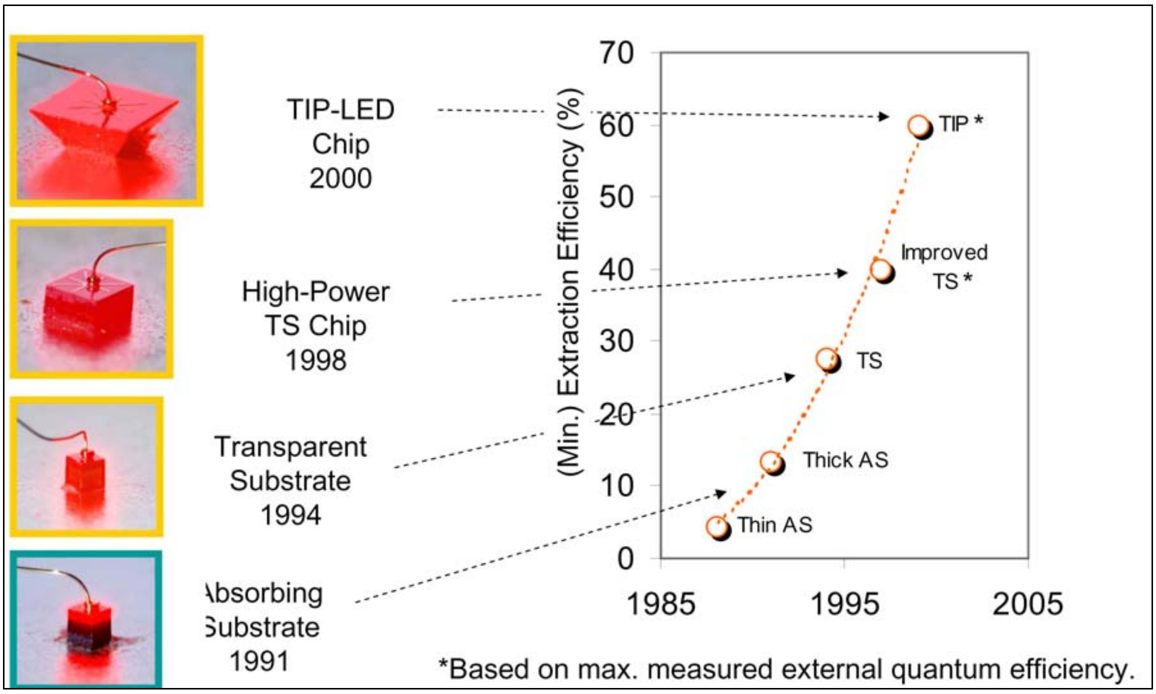
Colour Rendering Index (CRI)

The colour rendering index (CRI) is a measure of the ability of a light source to reproduce the colours of various objects being lit by the source (100 is the best CRI).

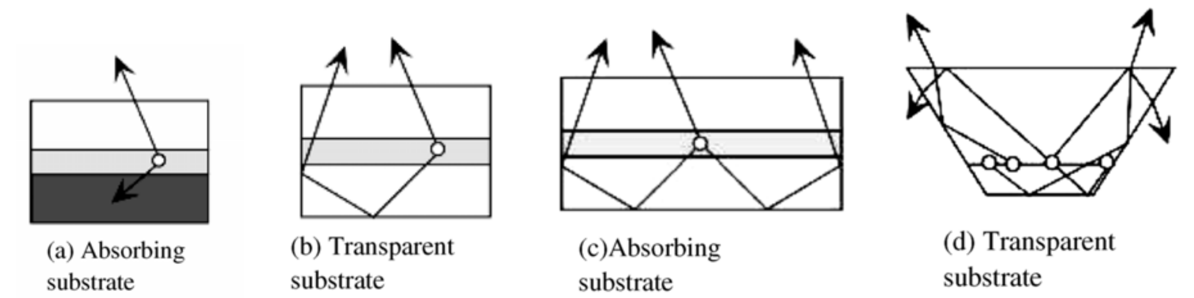
Examples of the meaning of CRI values

Improving this for economically-viable white light LED bulbs is still a subject of on-going research.

2.6.5 Light extraction developments: shape & substrates



Historical developments of variation of the substrate materials and shape of the LED chip ('die').



Schematic cross-section of historical developments

Chip geometry has been found to have a large impact upon extraction efficiency:

(a) Earlier absorbing substrates introduced too much loss before the light could escape;

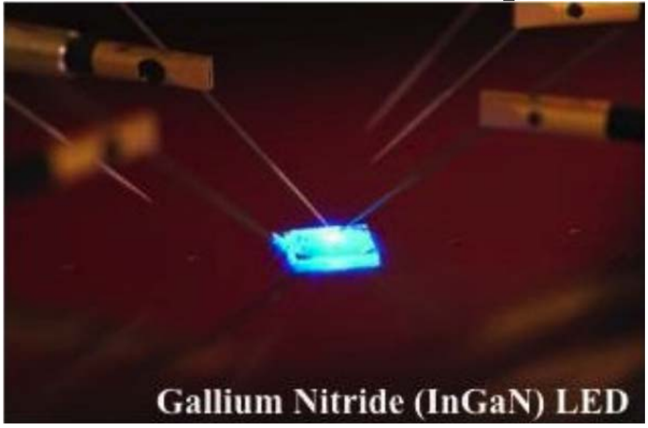
(b) these were replaced by transparent substrates, such as sapphire, and reflecting bottom face;

(c) wider area layers gave higher output power;

(d) an efficient chip geometry, for all-angle emission, is the Truncated Inverted Pyramid (TIP) structure, which minimises the number of internal reflections. A very thin emission layer (see MQW later) is also indicated here.

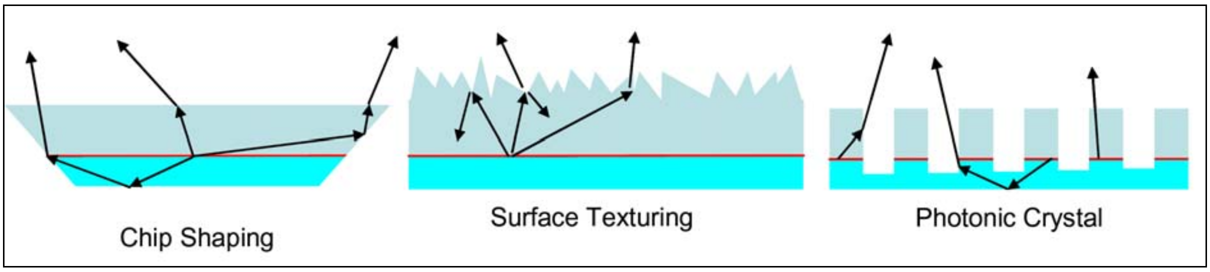
2.6.6 Light extraction developments: Blue LEDs and the output face

Beyond substrate and chip shape considerations, variations of the details of the output surface have also been investigated. Much of this work has concerned developments of blue LEDs.



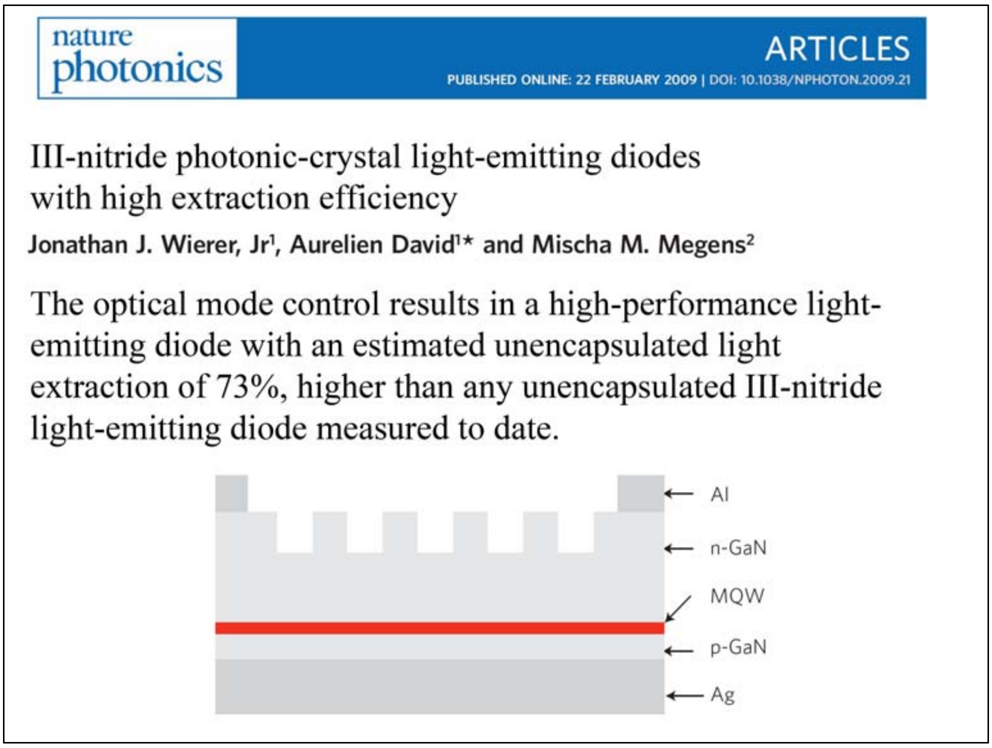
Blue LEDs based on GaN alloys

This has included enhancements gained through roughening the output surface to create a wider range of light incidence angles. One can think of this as a huge number of tiny microscopic encapsulations sitting on the output surface.



Developments from chip shape to output surface refinements

Also, instead of randomising the output surface, recent research has focused on creating highly-regular patterning to guide the light out of the chip. This patterning is on μm or nm scales, i.e. comparable to the optical wavelength. Such periodic material patterning results in what are called ***photonic crystals***, which have very strong light-guiding properties.


nature photonics

ARTICLES

PUBLISHED ONLINE: 22 FEBRUARY 2009 | DOI: 10.1038/NPHOTON.2009.21

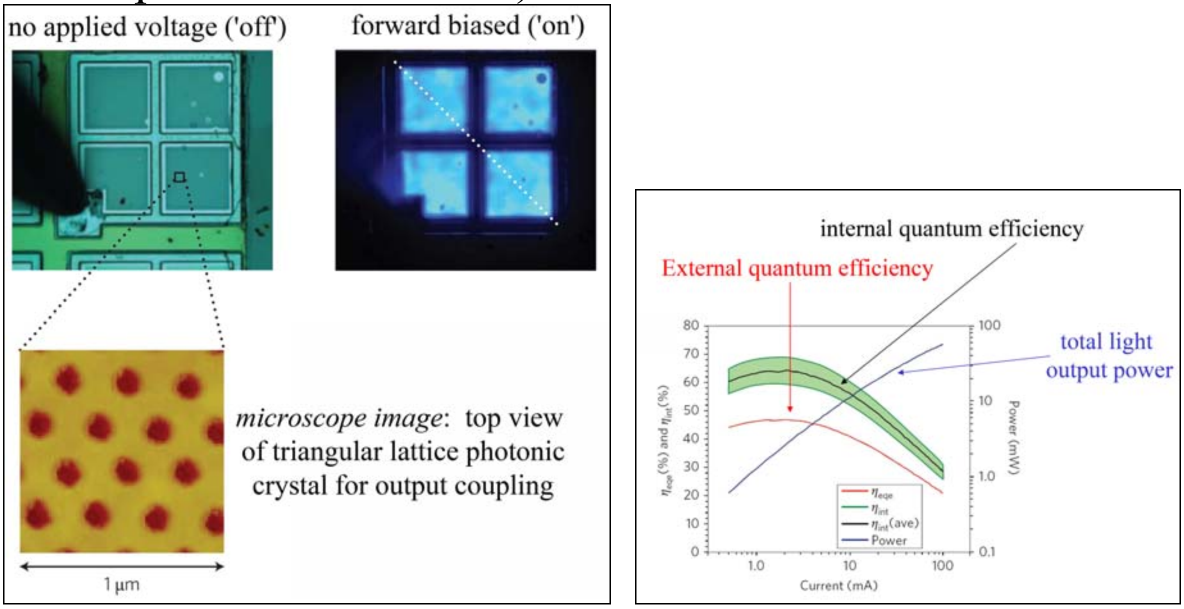
III-nitride photonic-crystal light-emitting diodes with high extraction efficiency

Jonathan J. Wierer, Jr[1], Aurelien David[1]* and Mischa M. Megens[2]

The optical mode control results in a high-performance light-emitting diode with an estimated unencapsulated light extraction of 73%, higher than any unencapsulated III-nitride light-emitting diode measured to date.

Photonic crystal patterning geometry for light extraction

The patterning involved is actually two-dimensional (across the output area of the LED).

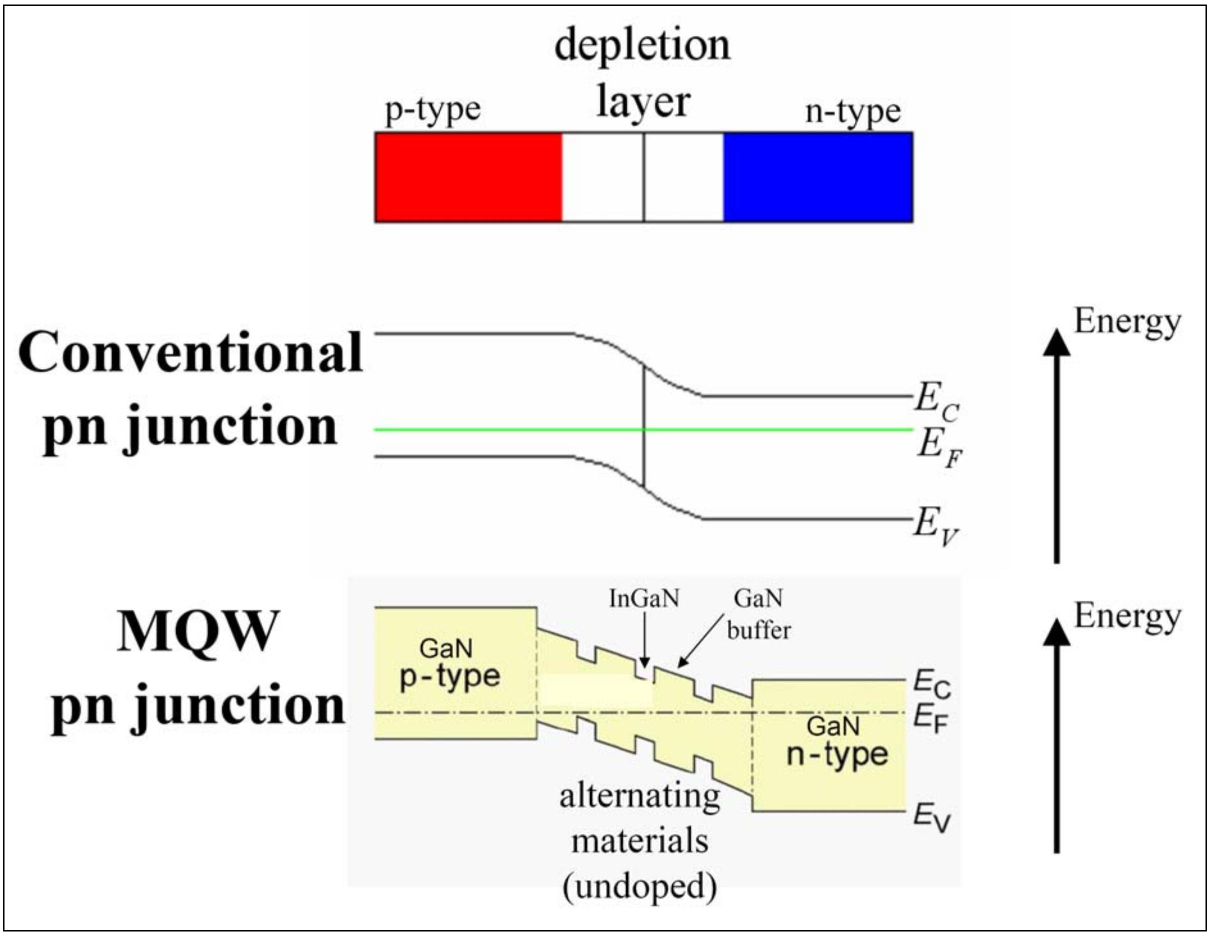


Left: top view of the output face (with and without bias voltage) and microscope image. Right: output efficiencies and total output power as a function of the device current.

2.6.7 Multiple Quantum Well (MQW) devices

For many LED devices, several layers of semiconductor are deposited between the p-type and n-type sections to form what are called ***multiple quantum wells (MQW).*** For GaN diodes, layers of indium gallium nitride (InGaN) that are just a few atoms thick are sandwiched between thicker (buffer) layers of GaN.

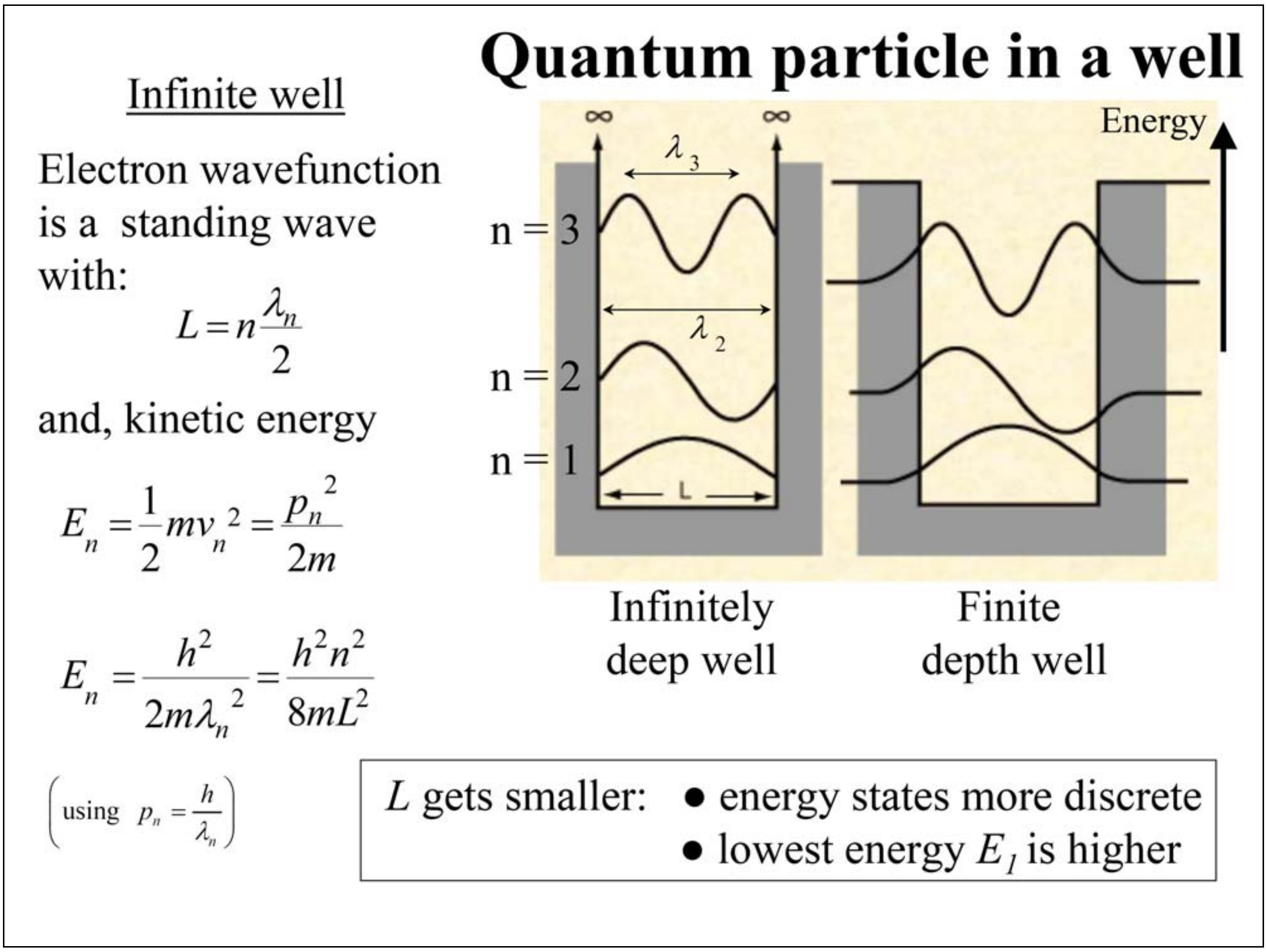
As shown below, the *different materials* have *different energy bandgaps* so that conduction electrons fall into and accumulate in the InGaN wells. In the valence band, the picture is reversed. Since electrons fall into lower energy states the holes accumulate in the same thin InGaN layers (effectively pushing to the highest available energy states).



Extra material layers to make multiple quantum wells

The extremely narrow width of the InGaN layers (of the order of the electron de Broglie wavelength) means that the available states in the wells are strongly quantum mechanical in character. To understand what electron quantisation features arise, and how these depend on the narrow width of the well, one can consider the possible confinement states of an electron in an *energy well of infinite depth*. In this case, the electron cannot escape the well and its wave representation has zero amplitude at the walls of the well.

The total electron energy is the sum of its kinetic energy and the potential energy (due to position in the material). Since the electron is entirely confined to an infinite well, its *potential energy is constant* within the trap. The kinetic energy can be written in terms of the electron momentum, *p*, which in turn is expressible in terms of the de Broglie wavelength λ.



Quantum particle in a well

Infinite well

Electron wavefunction is a standing wave with:

$$L = n\frac{\lambda_n}{2}$$

and, kinetic energy

$$E_n = \frac{1}{2}mv_n^{\,2} = \frac{p_n^{\,2}}{2m}$$

$$E_n = \frac{h^2}{2m\lambda_n^{\,2}} = \frac{h^2n^2}{8mL^2}$$

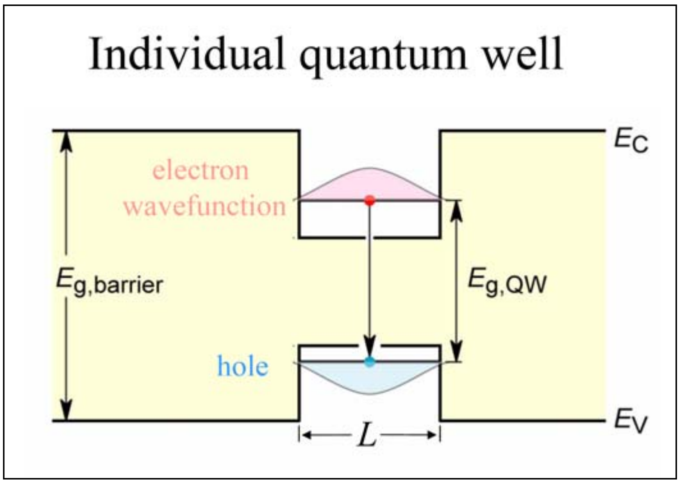
$$\left(\text{using } p_n = \frac{h}{\lambda_n}\right)$$

L gets smaller:
- energy states more discrete
- lowest energy E_1 is higher

Available electron states in a potential energy well

For *potential wells of finite depth*, a similar picture emerges except that there is now a probability that the electron can ‘tunnel out’ of the well. The distributions given by the electron wavefunctions then extend into, and beyond, the walls of the well (see the right side of the above figure).

We can now focus on the role of quantum wells in LEDs and what benefits multiple quantum wells have to offer.



Electron-hole recombination in a quantum well

A crucial point to note is that the lowest energy level of the electron is not at the bottom of the well. Instead, this depends on the width of the well. The energy of optical emissions due to electron-hole recombination does not match the bandgaps of the materials involved and is instead determined by the physical width of the material layer deposited during fabrication.

In effect, from an optical viewpoint …

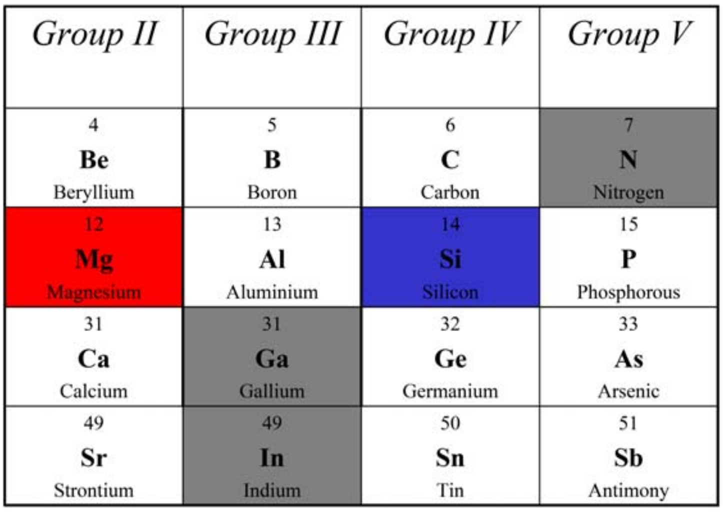
***… the quantum well structures act a new synthetic material for which we can design transition energies*.**

Also, the narrow layers where the electrons and holes accumulate have *low volumes* and thus focus high densities of electrons and holes for ***exceptionally efficient recombination***, and hence photon generation.

2.6.8 GaN doping

We described earlier that, to increase the number of available number of holes or electrons not dedicated to a bonding role, intrinsic semiconductors like Si or Ge in group IV can be doped with group III or group V elements to create (*extrinsic*) p-type or n-type materials.

However, to generate light from electron-hole recombination, *alloy semiconductors* such as GaAs or GaN (where the underlying material is not a group IV element but is a III-V compound) are used. So, how are p- and n-type materials created in this case?



Group II	*Group III*	*Group IV*	*Group V*
4 **Be** Beryllium	5 **B** Boron	6 **C** Carbon	7 **N** Nitrogen
12 **Mg** Magnesium	13 **Al** Aluminium	14 **Si** Silicon	15 **P** Phosphorous
31 **Ca** Calcium	31 **Ga** Gallium	32 **Ge** Germanium	33 **As** Arsenic
49 **Sr** Strontium	49 **In** Indium	50 **Sn** Tin	51 **Sb** Antimony

Alloy semiconductors, such as GaN, may use group II and IV impurities to create n-type and p-type materials

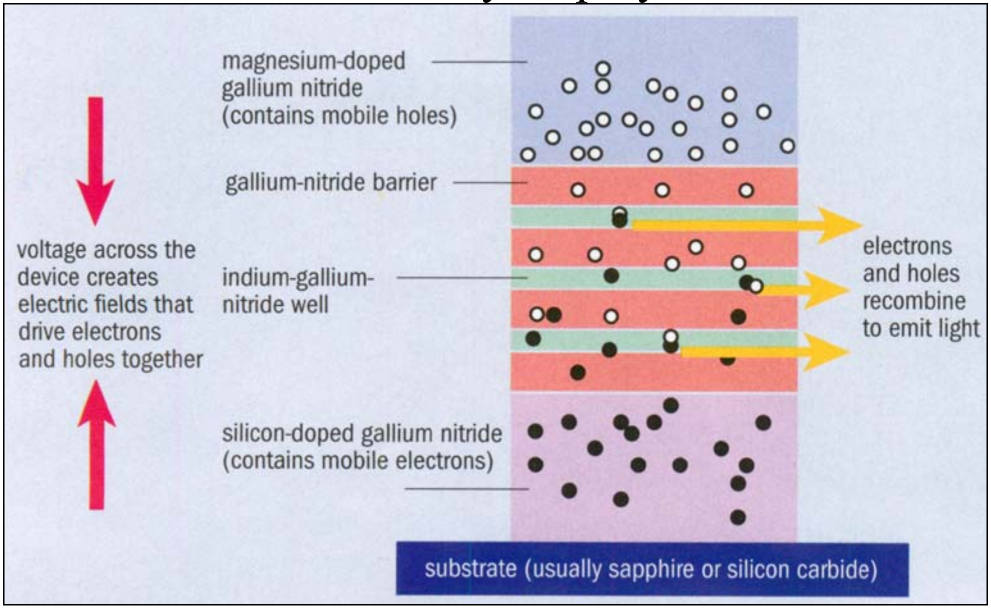
In the case of GaN, p-type material can be made by replacing some of the Ga atoms with Mg (magnesium) from group II. For n-type material, Si from group IV replaces some Ga atoms.

As shown in the above section of the periodic table, the compound InGaN (which is often used in optically active layers) does not involve n- or p-type impurities, it is undoped.

The inclusion of larger size elements, such as In, opens the possibility of closer transition energy levels (and thus a smaller material energy bandgap where electrons and holes can accumulate). On the other hand, GaN alloys that include smaller atoms from group III, such as AlGaN have wider energy transitions and tend to emit at higher frequencies in the ultraviolet region.

2.6.9 Droop – the dark secret of LEDs?

The diagram below, illustrating the core of an operating GaN LED, shows both the doping scheme for GaN and the multiple quantum well structures usually employed.

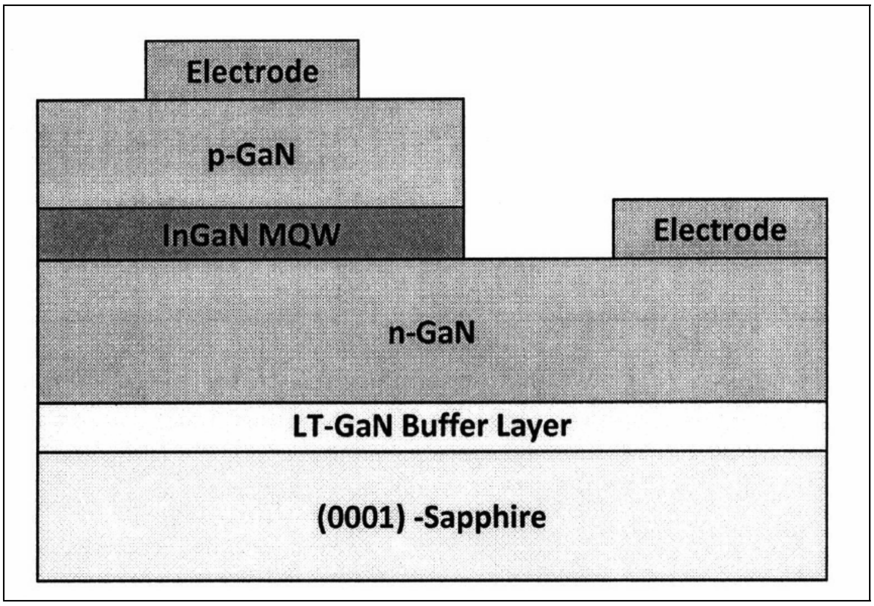


GaN LED operation, from Physics World, Feb 2011

But the Physics World article was actually about an effect that is often called 'droop'. It is a label for the rollover of efficiency with high current densities – an effect that is little understood, but is of significant importance. If efficiency can be sustained at higher driving currents then power LEDs become much efficient and the cost of devices (which otherwise typically employ many LEDs to give high total power) will drop dramatically.

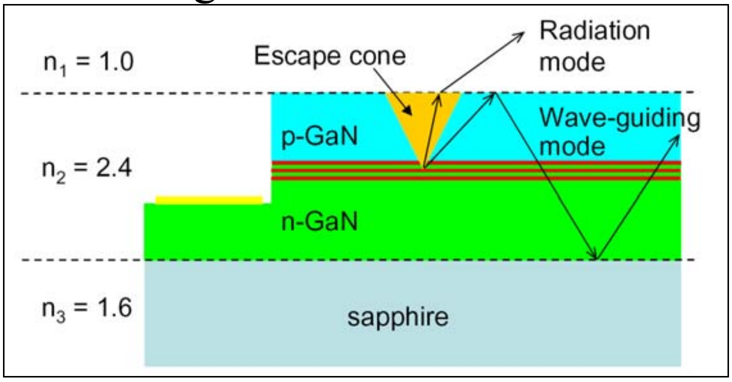
2.6.10 Heterojunctions and edge-emitting LEDs

One can expand upon the above shown construction characteristics of GaN blue LEDs by including the electrode connections:



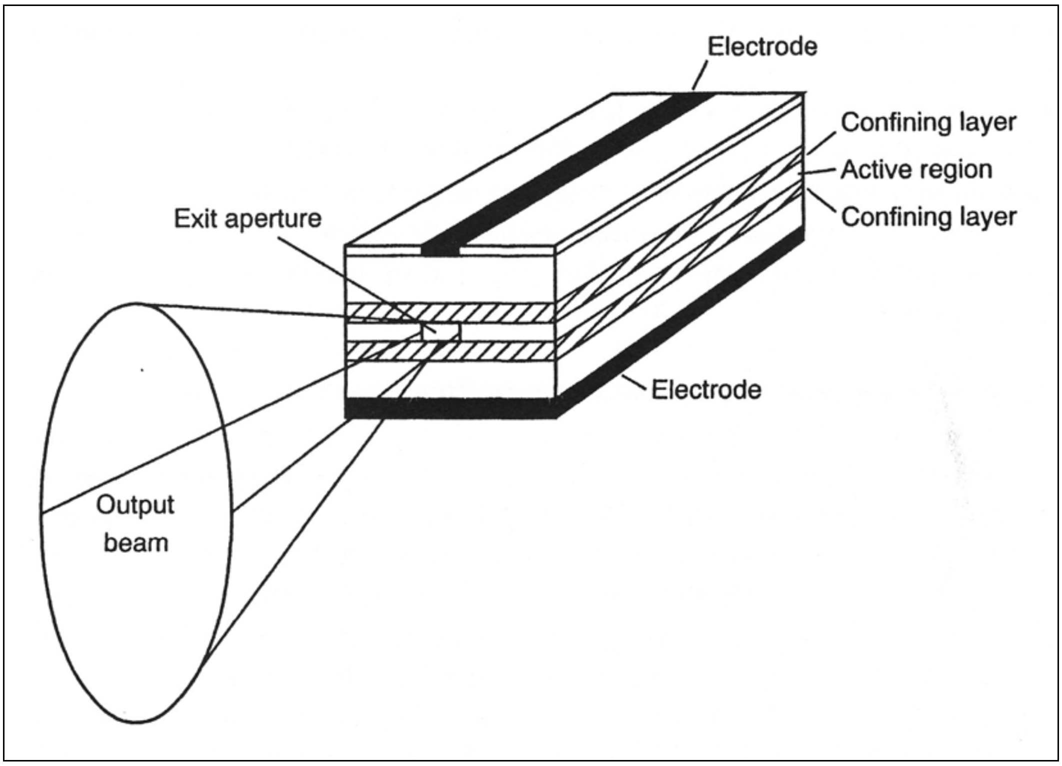
Overview of GaN LED construction (with electodes)

But note that in the previous two figures that the emitted light appears not to be coming from the top of the device. Recall that top-surface light extraction involves overcoming photons being trapped inside the high index semiconductor layer. Multiple reflections can ensue until much of the light is totally absorbed by the host materials. Many additional material layers, such as quantum well and buffer layers, can compound these problems. The figure below indicates an alternative.



Vertical emission versus horizontal waveguiding

The above diagram shows that some of the light trapped inside the LED can be waveguided along the length of the device. The picture is similar to light confinement in optical fibres, where total internal reflection is also the guiding mechanism. One can thus abandon vertical emission strategies and enhance this waveguiding effect by sandwiching the active region (such as MQW) between two layers of semiconductor that have a slightly lower refractive index.



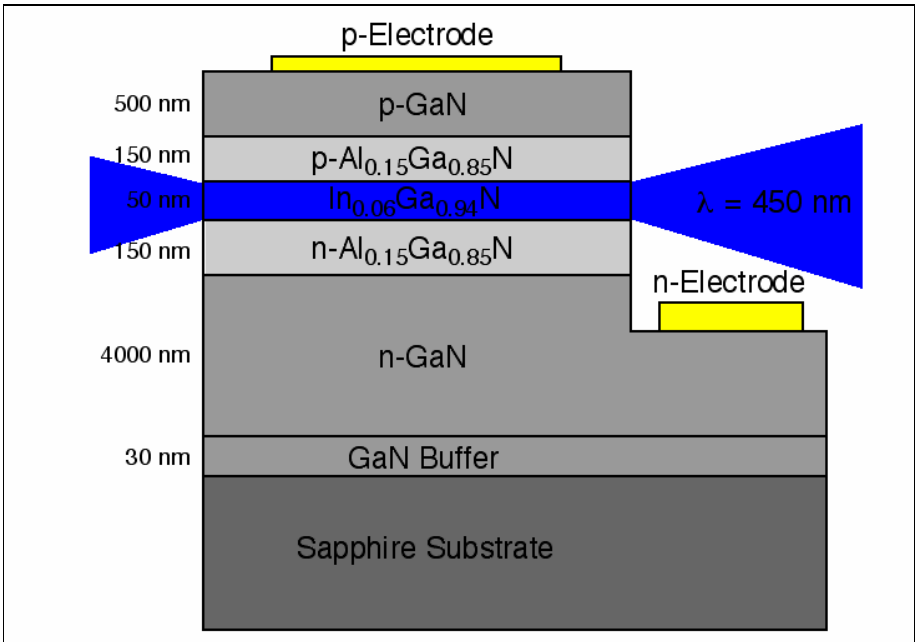
A double heterojunction structure for an edge-emitting LED

Ideally, the energy bandgap of the sandwiching material is greater than that of the active region, so electrons and holes also naturally collect in this waveguide.

This ***double heterojunction*** structure (see figure above) also provides a more efficient means to couple the light out of the edge of the device by helping to control the divergence of output light. This can be contrasted with the wider divergence that can arise in vertical emission designs.

Manufacture of double heterojunction LEDs involves the layers of semiconductor material being built upon each other.

Each layer are often deposited from a chemical reaction in a vapour that is passed over the existing layers (see movie below).



Journal of Physics: Condensed Matter, **13**, 7089 (2001)

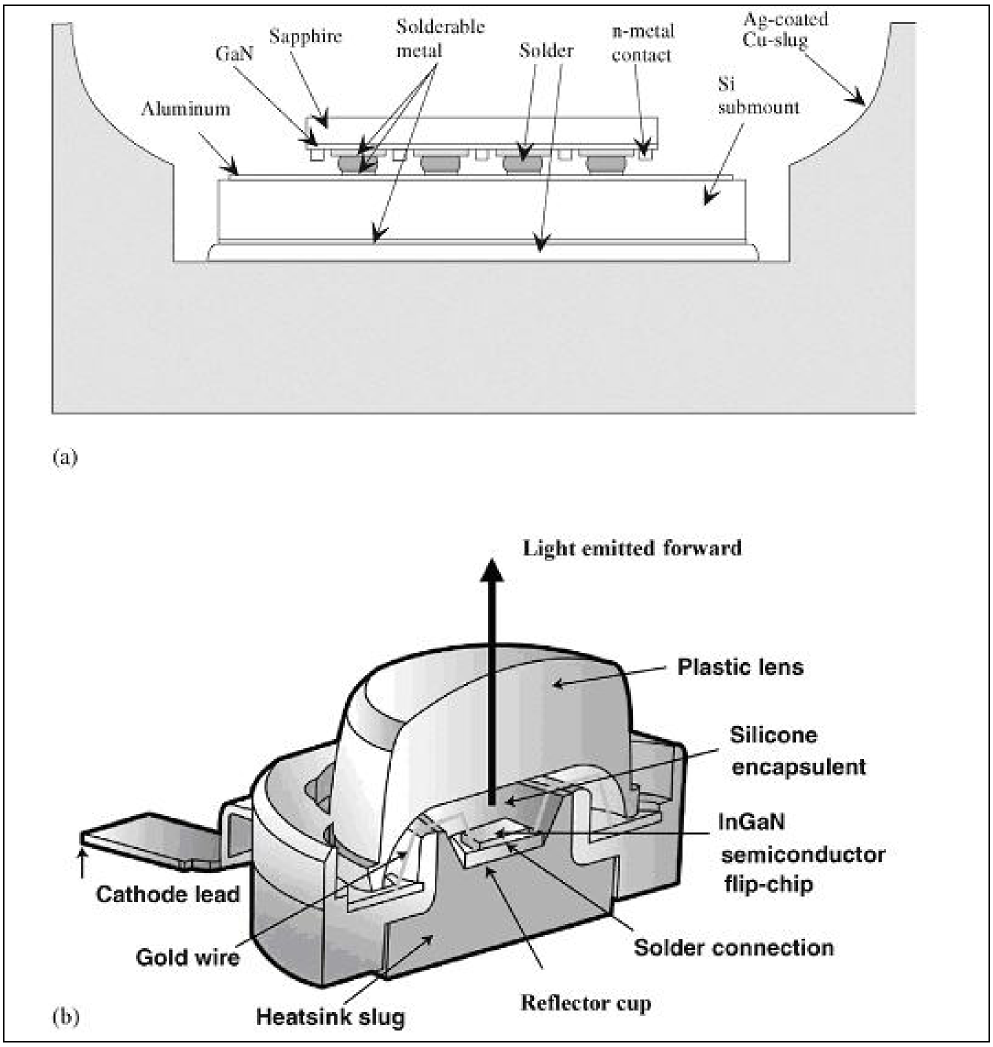
Lecure movie showing manufacture of an edge-emitting LED
http://mrsec.wisc.edu/Edetc/background/LED/#intro

2.6.11 High power LEDs

High power devices require sophisticated packaging with good optical efficiency and means for thermal management.

The figure below illustrates current chip housing and packaging considerations for a high power device.

Individual design and construction aspects are then listed.



Cross-sectional views of a high-power LED. (a) LED chip details; (b) chip with packaging elements

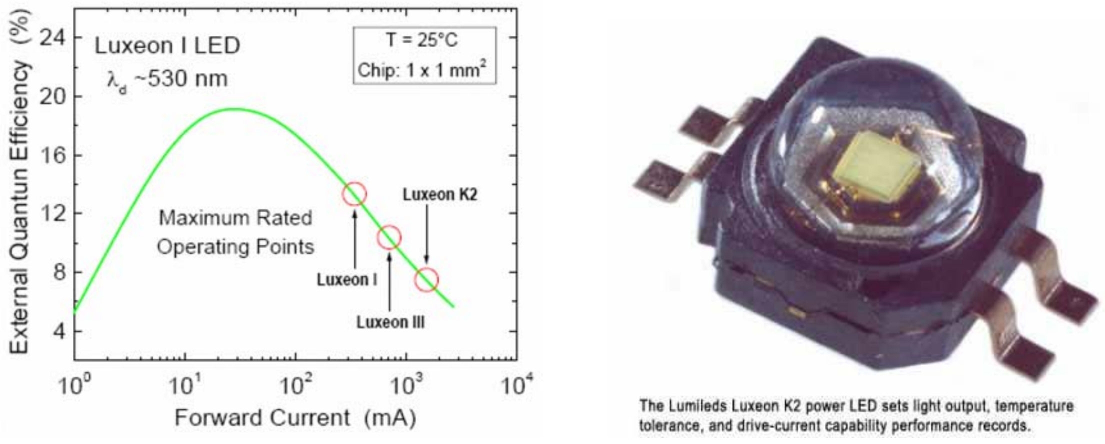
● A relatively large area chip (e.g. 2x2 mm^2) is used. Current, and associated heating, is controlled by division of this area into segmented cells – see the multiple contacts to the bottom of the chip in part (a) of the above figure.

● Many high power designs omit the bottom substrate and have a (e.g. aluminium, Al) conducting and reflecting layer. This is supplemented with a large copper (Cu) heat sink that is coated with highly-reflecting silver (Ag). The substrate-less heat sink structure dramatically improves heat transfer.

- Electrical leads are isolated from the thermal path, supporting greater die size and increased operating current. Transparent electrical contacts (e.g. indium tin oxide) can also be used.

- Silicone encapsulant provides a high refractive index buffer between the chip and the lens. It is thermally stable, moisture sealing and transparent. It is also pliable, easing stresses on the chip and leads.

- A large plastic lens plays a similar role to dome-shaped encapsulation of lower-power LEDs and helps directionality of the output light. The lens is removable, and can be bought separately, offering the option of using a more efficient photonic crystal output coupler.



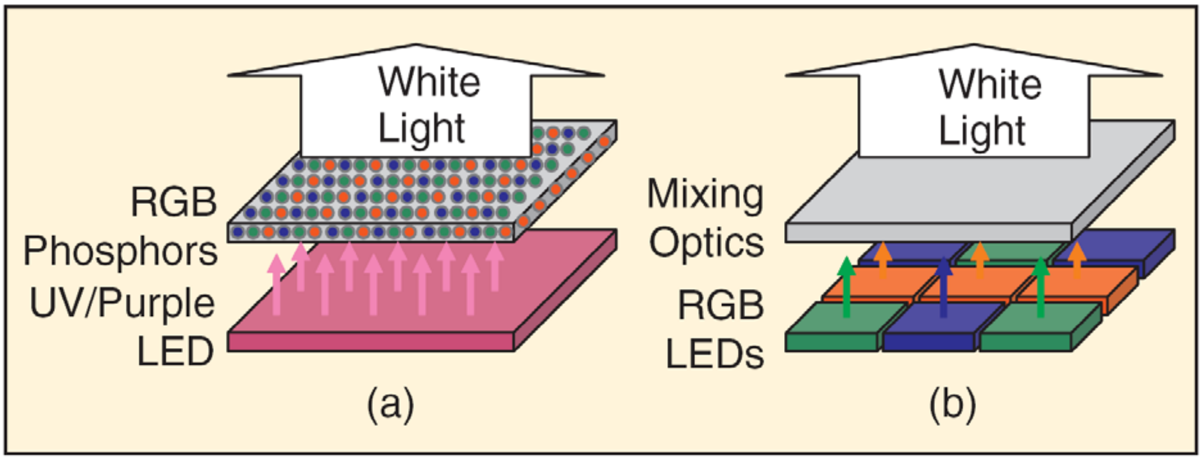
The Lumileds Luxeon K2 power LED sets light output, temperature tolerance, and drive-current capability performance records.

High power operation points (circles) may be a low efficiency. Right: The Luxeon K2 LED is shown at the right.

- Note the pronounced ‘droop’ effect, as the efficiency rolls over as the current increases. The circles on the graph indicate operating points of different versions of a specific make of high power LED. Devices are operating within the high current and lower efficiency regime to deliver maximum output power.

2.6.12 White light LEDS

There are two main strategies for generating white light from LEDs. The first involves high frequency LED light that is partially down-converted to lower frequencies using phosphors. The second entails mixing colours of separate LEDs, e.g. Red, Green, and Blue outputs for RGB LEDs.

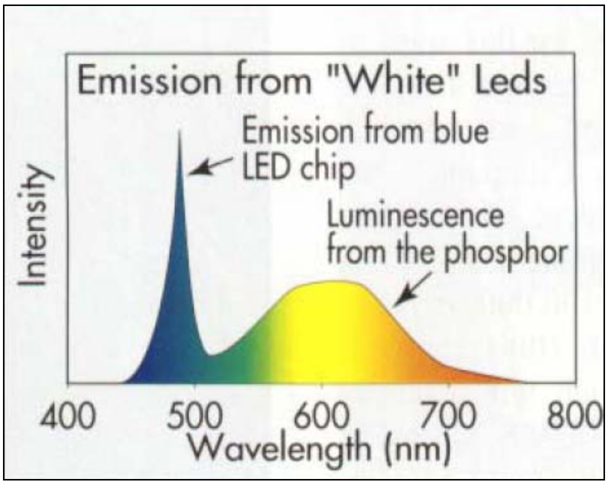


White light LED strategies: (a) down-conversion of blue light using phosphors; and (b) combining separate LEDs of different colours

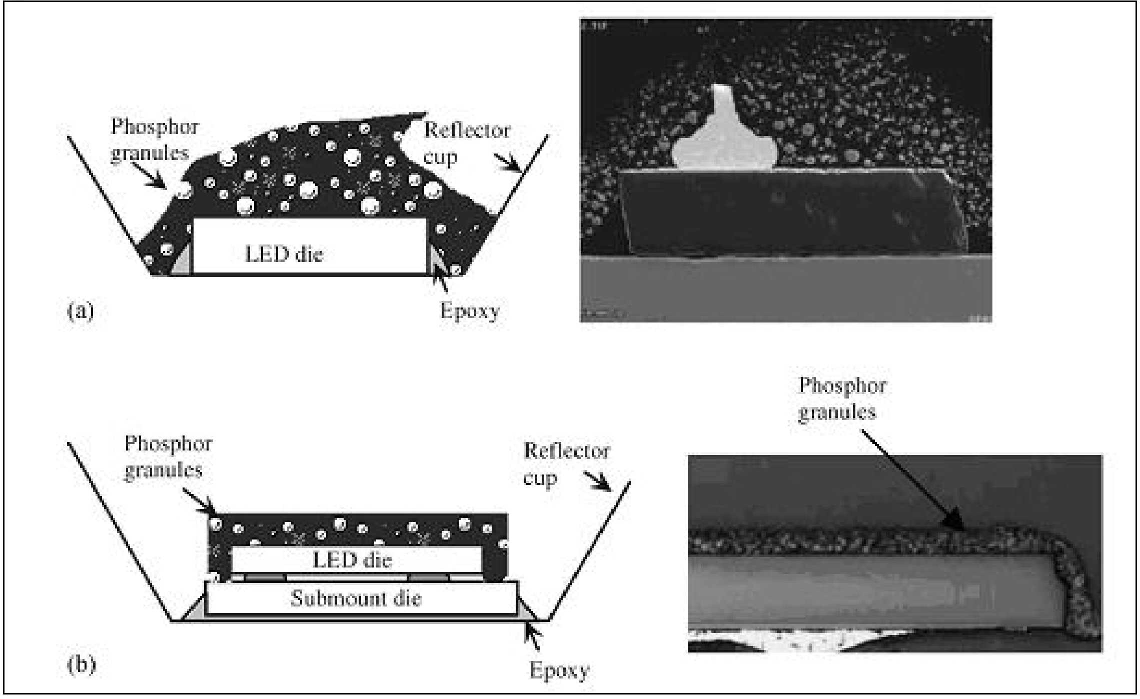
(a) Phosphor-based white light LEDs

The most common approach is to down-convert blue, violet or ultra-violet LED light by using phosphors.

A popular choice is the yellow-emitting 'YAG' phosphor.



Phosphors are applied in various ways around or on top of the LED chip. A uniform coating is preferable to avoid angle-dependent colour variations.

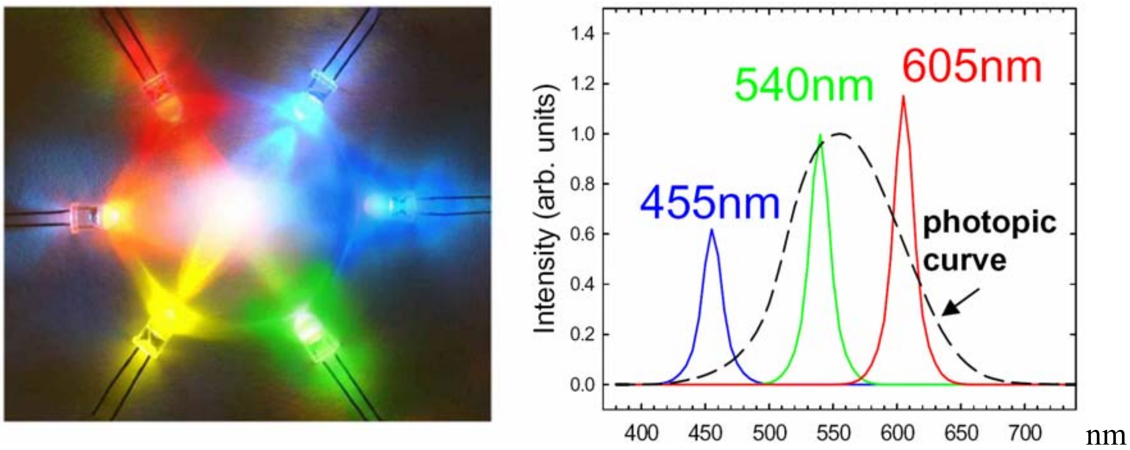


Coating the LED chip with phosphor granules

The phosphor composition lets a precise amount of blue light leak through to combine with the emitted yellow light. Red phosphors are also typically added to mix to generate a 'warmer' white light chromaticity. Adding more layers of different phosphors fills and broadens the final output spectrum, and increases the colour rendering index of the LED. This scheme provides a simple integrated colour mixing strategy, but it is not possible to vary the total spectral output content after manufacture and the process of down-shifting light generates heat that needs to be extracted from the chip.

(b) White light generation with RGB LEDs

This method is more complicated as electronic circuits are required to control the blending of the different LED colours. Optical elements are also needed to combine and to control the contributing light frequencies. It is for these reasons, of complexity and cost, that this method is currently not the preferred strategy.



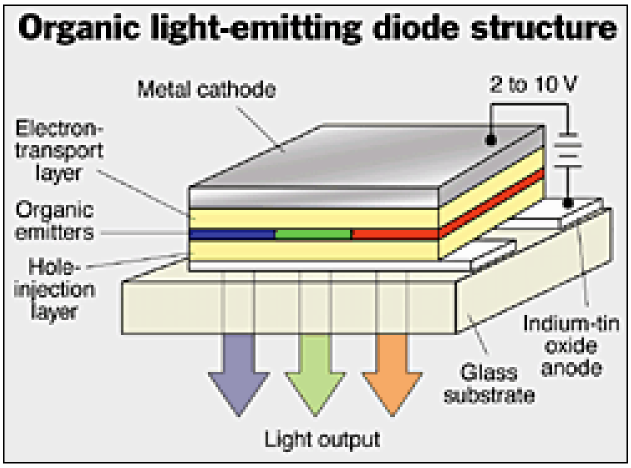
(a) RGB white light LED. (b) Output colours & eye response

Nonetheless, RGB LEDs offer much promise for the future since their precise spectral properties can be dynamically controlled. Two, three or four different LED colours can be mixed. The higher the number of constituent LED colours then the higher the capability for quality colour rendering in the final output image.

2.6.13 Organic Light Emitting Diodes (OLEDS)

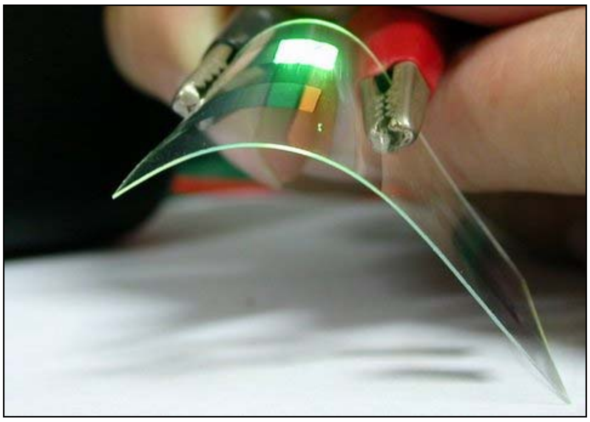
OLEDs have similar working principles to conventional GaAs LEDs.

The difference is that films of organic compounds replace the inorganic semiconductors.



OLED structure

One main advantage of over LEDs is that, currently, they can be produced at a much lower cost. Another advantage is that they can be applied as coatings on thin, flexible and inexpensive substrates. There is potential for their use in low voltage displays in portable electronic devices and benefits such as wide viewing angle and high contrast images.



A flexible organic light emitting diode (OLED)

However, in comparison to inorganic LEDs, they do not match their efficiency and have shorter lifetimes (e.g. greater than 5,000 hrs, but much lower than the current 50,000 to 100,000 hrs for conventional LEDs). Moreover, their scaling to large area, high power devices currently involves loss of efficiency and their cost advantage.