Notes for

Modern and nonlinear Optics

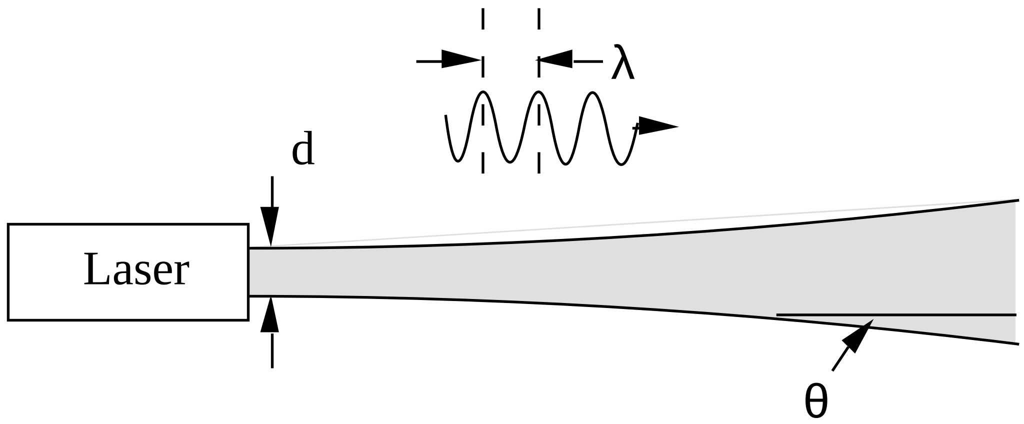
Miles Padgett

m.padgett@physics.gla.ac.uk

Properties of laser light

Collimation

As one can see from a laboratory He-Ne, the output from a laser is well collimated

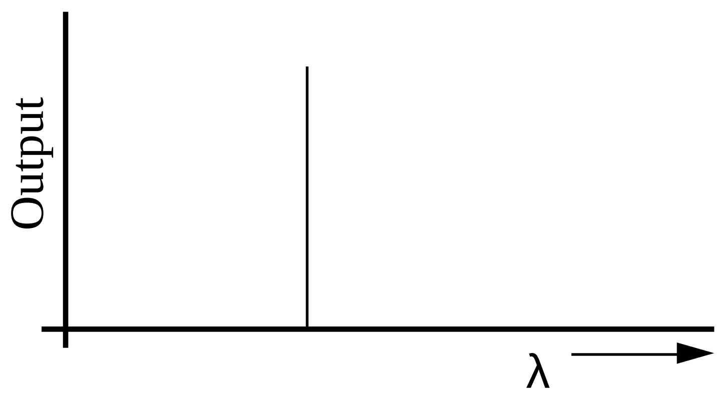


The far-field diffraction angle is $\theta \approx \frac{\lambda}{d}$ [1]

where
d = diameter of output beam
λ = wavelength of laser light

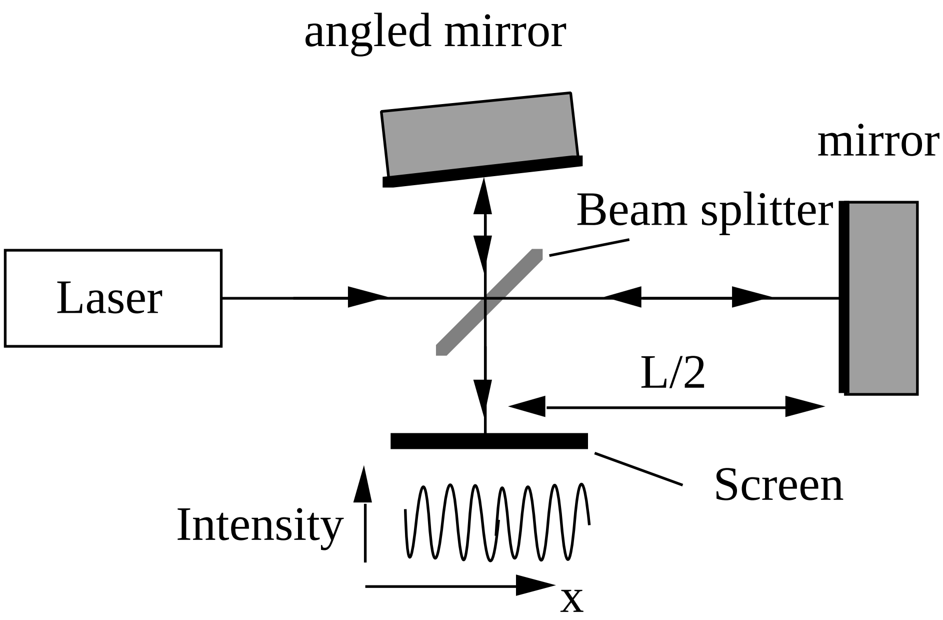
Monochromaticity

If a spectrally pure beam of laser light is examined using a monochromator it is found to comprise narrow range of optical frequencies.



Temporal Coherence

By splitting and subsequently recombining a laser beam after varying path differences, the coherence can be assessed.



Even when L is large, interference fringes are still observed. This tells us the present phase of the light is strongly related to the past phase.

We define length over which fringes are still visible as the **coherence length.** We also define a **coherence time,** where:

Coherence length = Coherence time x c [2]

Coherence time - Optical bandwidth

We know that over the coherence time, the phase of the electric field vector must not 'slip' by more than 2π

By writing the electric field vector as:

$$E = E_0 \sin((\omega_0 + \Delta\omega)t)$$

we can state that the spread of optical frequency (i.e. the bandwidth) $\Delta\omega$

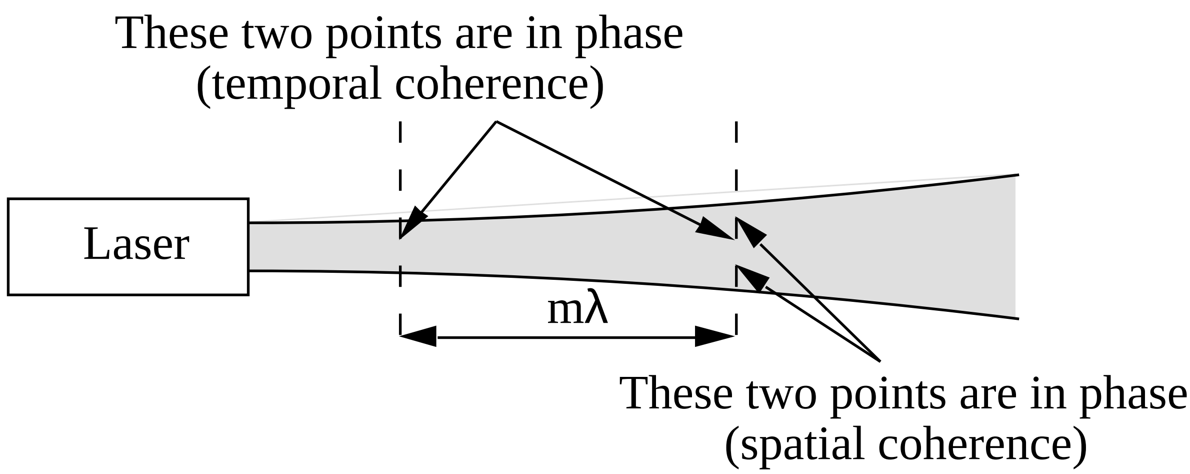
$$\Delta\omega \, t_{coherence} \approx 2\pi$$

i.e. the spread of optical frequencies in the laser beam ($\Delta\nu$) is related to the coherence time by:

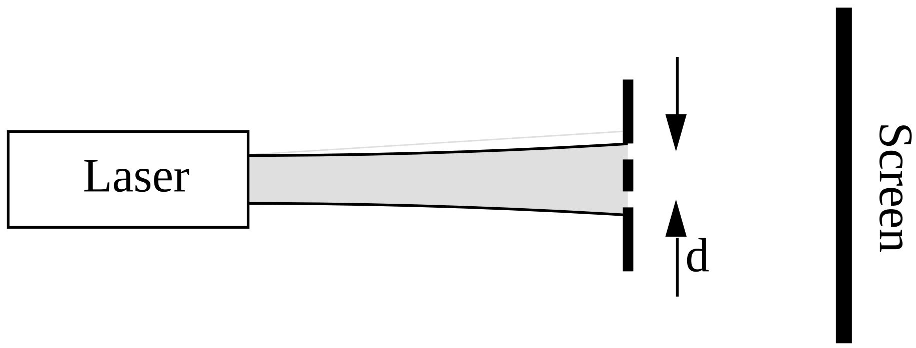
$$\Delta\nu = \frac{1}{t_{coherence}} \qquad [3]$$

Spatial coherence

As well as being temporally coherent, a typical laser beam can be spatially coherent as well.



Spatial coherence can be measured using Young's slits



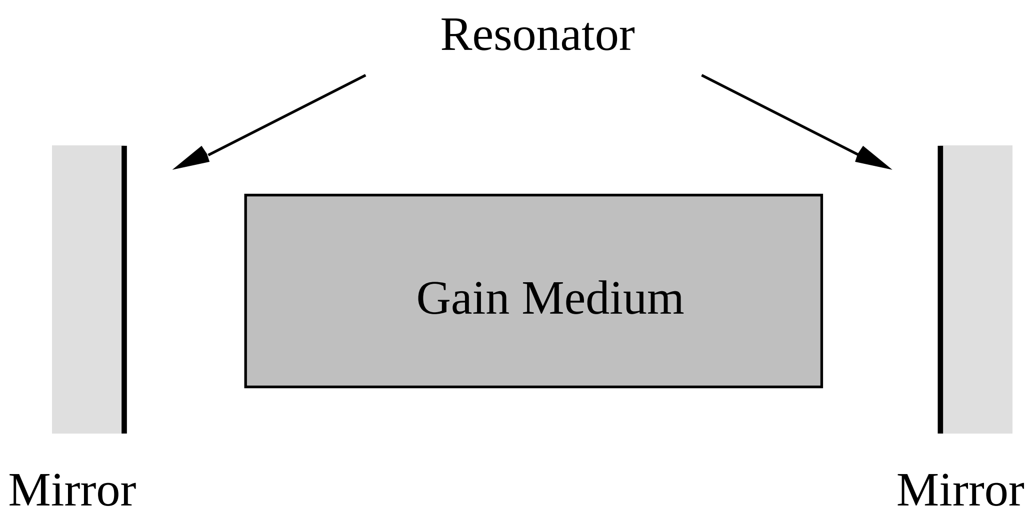
In order to see a high contrast interference pattern on the screen the phase difference between the light at the two slits must be constant. Typically, laser light is found to be spatially coherent across the whole of the beam.

What makes up a laser?

There are two components to a laser

Gain Medium (something to amplify the light)

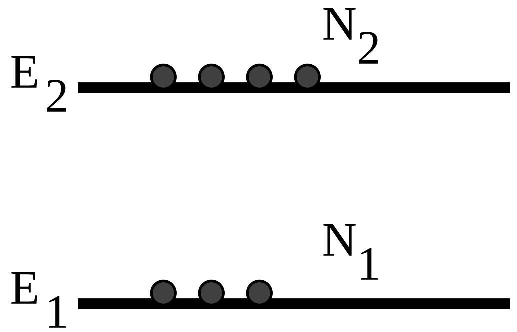
Resonator (something to provide suitable optical feedback)



Gain Medium

Consider two energy levels, E_1 and E_2, with populations N_1 and N_2.

We will use an atomic system in which the electrons occupy well defined energy states.



There are three ways in which this system can interact with light.

Spontaneous emission

An electron in the upper state can spontaneously relax to the lower state and in doing so will emit a photon.

Conservation of energy means the photon energy will equal the change in electron energy, i.e.

$$h\nu = E_2 - E_1$$

where h is Planck's constant

The average time an electron will remain in one state before relaxing is called the lifetime (τ)

The exact time at which the electron "chooses" to relax is totally random and therefore the radiation emitted will be **incoherent.**

The spontaneous emission 'flow' of electrons from the state 2 to state 1 is given by:

$$\frac{dN_2}{dt} = - \frac{N_2}{\tau_2}$$

where

τ_2 = lifetime of state 2

N_2 = population of state 2, (no. m^{-3})

Alternatively, we write this as:

$$\frac{dN_2}{dt} = - A_2 N_2$$

where

$$A_2 = \text{Einstein 'A' coefficient} = \frac{1}{\tau_2} \qquad [4]$$

Stimulated absorption

Similarly, an incoming photon can be absorbed by the system and an electron will be excited from the lower to the upper state. The frequency of the photon must be such that:

$$h\nu = E_2 - E_1$$

The spontaneous absorption 'flow' of electrons from state 1 to state 2 is given by:

$$\frac{dN_2}{dt} = B_{12} \, \rho_\nu \, N_1$$

$$\rho_\nu = D_\nu \, h\nu$$

where

B_{12} = Einstein 'B' coefficient
ρ_ν = Photon energy density at frequency ν
D_ν = Photon number density at frequency ν

Stimulated emission

An incoming photon can **cause** an electron in the upper state to relax and an additional photon will be emitted. The two photons not only have the same frequency but also the same phase, i.e. they are **coherent**.

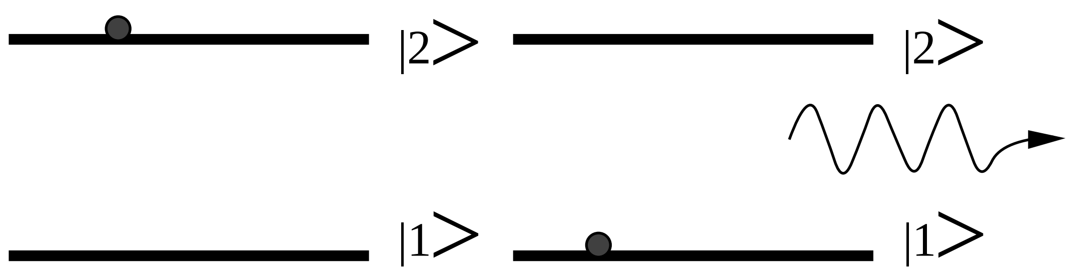
The stimulated emission 'flow' of electrons from state 2 to state 1 is given by:

$$\frac{dN_2}{dt} = -B_{21}\,\rho_\nu\,N_2$$

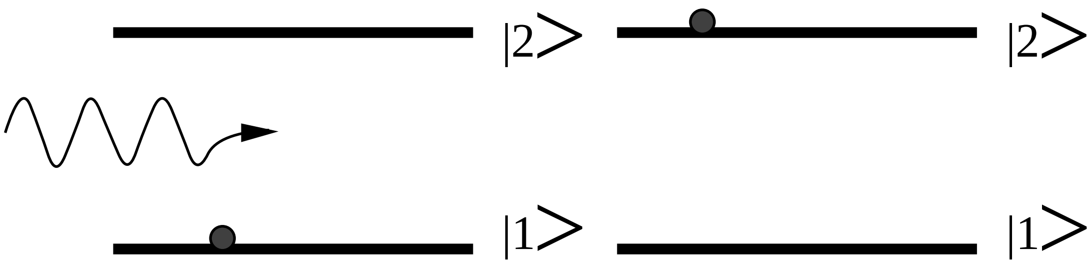
where
B_{21} = Einstein 'B' coefficient

It is stimulated emission that's the key to laser action.

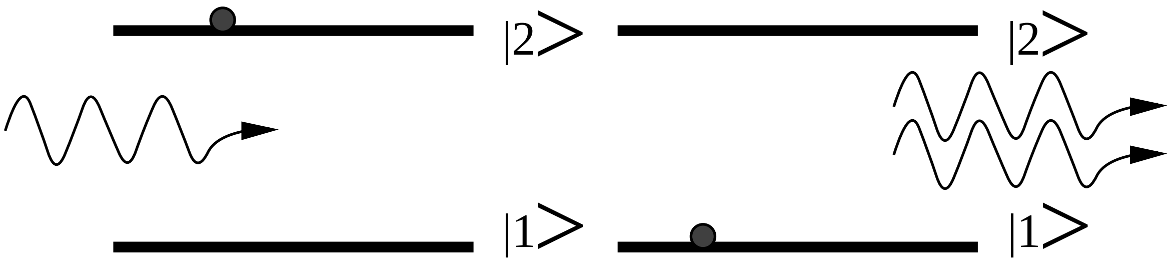
The three processes.



Spontaneous emission



Stimulated absorption



Stimulated emission

The Einstein relations

Einstein showed that the A and B coefficients are related.

In equilibrium, the rate of change of upper and lower populations must be 0, i.e.

$$\frac{dN_2}{dt} = \frac{dN_1}{dt} = 0$$

$\frac{dN_2}{dt}$ = spontaneous emission 'flow'
+ stimulated emission 'flow'
- stimulated absorption 'flow'

$$\frac{dN_2}{dt} = A_2 N_2 + B_{21} \rho_\nu N_2 - B_{12} \rho_\nu N_1 = 0 \qquad [5]$$

likewise,

$$\frac{dN_1}{dt} = - A_2 N_2 - B_{21} \rho_\nu N_2 + B_{12} \rho_\nu N_1 = 0 \qquad [6]$$

From [5]:

$$B_{12} \rho_\nu N_1 = A_2 N_2 + B_{21} \rho_\nu N_2$$

Re-arrange for ρ_ν

$$\rho_\nu = \frac{A_2 N_2}{B_{12} N_1 - B_{21} N_2}$$

or

$$\rho_\nu = \frac{\frac{A_2}{B_{21}}}{\frac{B_{12}}{B_{21}} \frac{N_1}{N_2} - 1} \qquad [7]$$

However, in thermal equilibrium, Boltzmann statistics will tell us the relative population of state 1 and state 2

$$\frac{N_1}{N_2} = \frac{g_1}{g_2} \exp((E_2 - E_1)/kT)$$

where

g_1 = degeneracy of state 1
g_2 = degeneracy of state 2
k = Boltzmann's constant
T = temperature in Kelvin

In our case

$$h\nu = E_2 - E_1$$

Therefore

$$\frac{N_1}{N_2} = \frac{g_1}{g_2} \exp(-h\nu/kT) \qquad [8]$$

sub into equ. [7]

$$\rho_\nu = \frac{\dfrac{A_2}{B_{21}}}{\left(\dfrac{g_1 B_{12}}{g_2 B_{21}} \exp(h\nu/kT)\right) - 1} \qquad [9]$$

Since our system is in thermal equilibrium, ρ_ν must be identical the black body emission , i.e.

$$\rho_\nu = \frac{8\pi n^3 h\nu^3}{c^3} \left(\frac{1}{\exp(h\nu/kT) - 1}\right) \qquad [10]$$

Equating equs. [9] and [10] we get the Einstein relations:

$$g_1 B_{12} = g_2 B_{21} \qquad [11]$$

and

$$\frac{A_2}{B_{21}} = \frac{8\pi n^3 h\nu^3}{c^3} \qquad [12]$$

The ratio of the spontaneous to stimulated emission is given by:

$$\text{Ratio} = \frac{\text{spont.}}{\text{stim.}} = \frac{A_2}{\rho_\nu B_{21}} \qquad [13]$$

Re-arrange equ. [9], to get:

$$\text{Ratio} = \left(\frac{g_1 B_{12}}{g_2 B_{21}} \exp(h\nu/kT)\right) - 1)$$

but $g_1 B_{12} = g_2 B_{21}$, therefore:

$$\text{Ratio} = \exp(h\nu/kT) - 1$$

e.g. electric light bulb, T= 2000K, $\nu = 5 \times 10^{14}$ Hz

$$\text{Ratio} = 1.5 \times 10^{5}$$

Steps to achieve laser action

Stimulated emission will enable us to amplify an incoming stream of photons (One photon in, two photons out).

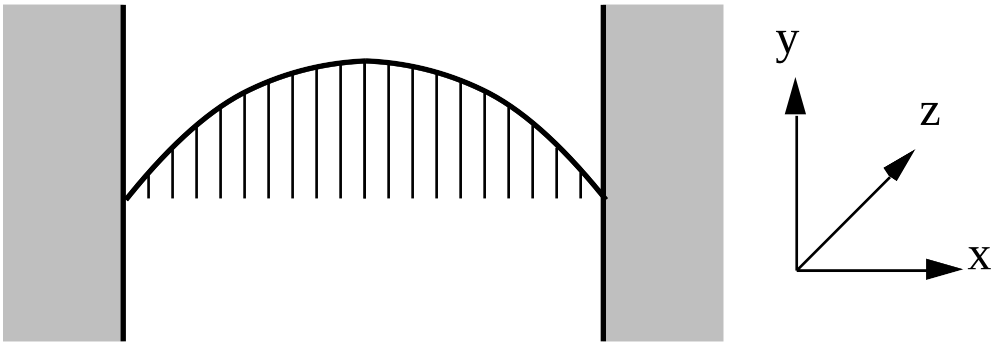
However, under normal conditions, the ratio between stimulated
n
d spontaneous emission is
very poor.

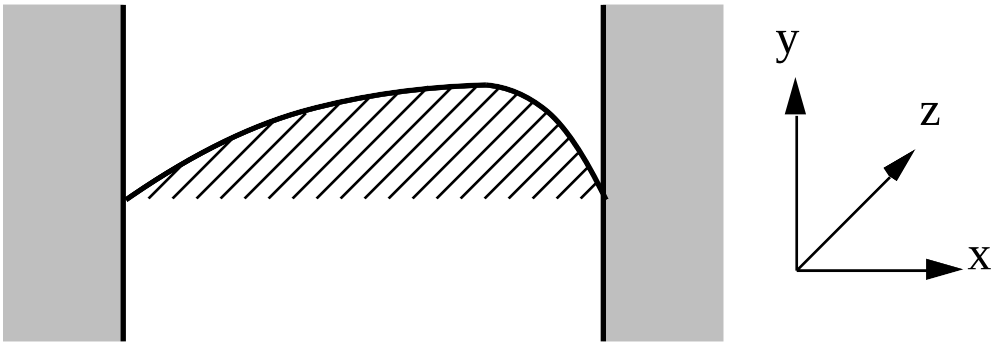
To boost the amount of stimulated emission, we need

- High ρ_ν
- High N_2

Explanation of 'g' degeneracy

If two states have different quantum numbers but identical energies, they are said to be degenerate, e.g.





$g_1 N_1$ is the total number of occupied states with the energy E_1.

Note that if the degeneracies are equal (e.g. both =1) then we have:

$$B_{21} = B_{12} \text{ (for } g_1 = g_2) \quad [14]$$

Absorption of light

It is normal to define an absorption coefficient (α), such that:

$$I_\nu(x) = I_\nu(x=0) \exp(-\alpha x) \quad [15]$$

i.e.

$$\frac{dI_\nu}{dx} = -\alpha I_\nu(x)$$

where

I_ν = light intensity at frequency ν

If α is +ve then incoming light is absorbed

If α is -ve then incoming light is amplified

How can we relate I_ν to ρ_ν ?

Units of I, energy per unit area per unit time

Units of ρ, energy per unit volume

Therefore,

$$I_\nu = \rho_\nu \times \text{speed of light}$$

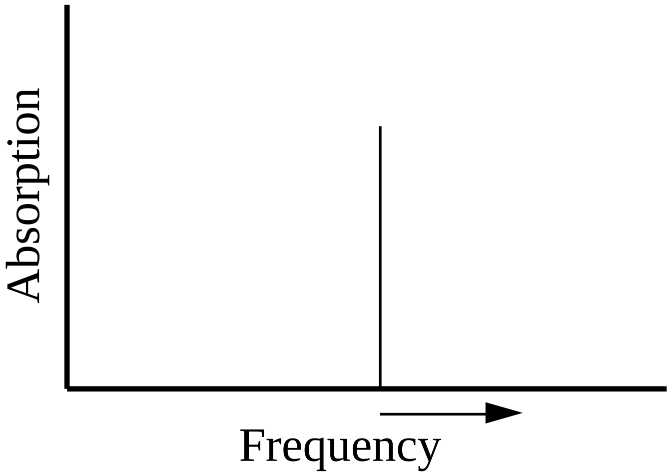
$$I_\nu = \rho_\nu \frac{c}{n}$$

where

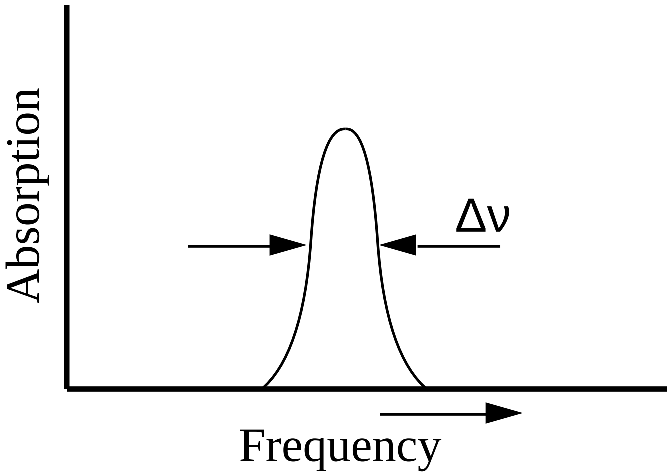
n = refractive index

Emission Linewidth

So far we have spoken of emission and absorption as if it only happened at one specific frequency and all the light emitted was monochromatic, i.e.



This is not the case, atoms may absorb or emit photons over a narrow range of neighbouring frequencies Δν, i.e.



The shape is given by the **lineshape function =** g(ν).

What causes the broadening of the line?

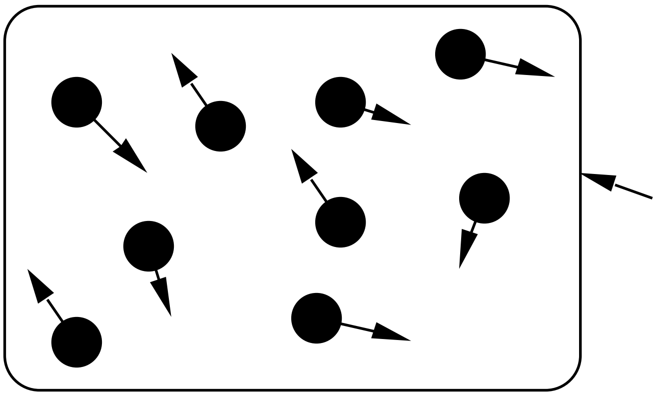
In gases:
Doppler broadening, pressure broadening, natural lifetime, observation time.

In solids:
Lattice variations, phonon interactions, natural lifetime

The various broadening mechanisms fall into two groups:

Inhomogeneous Broadening

e.g. Doppler broadening in gases



Atoms/molecules moving with thermal velocity

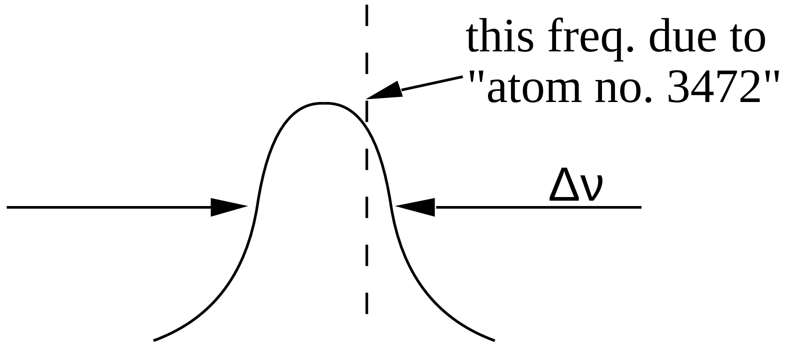
When a photon is emitted from a moving atom, its frequency will be Doppler shifted by an amount $\delta\nu$ where:

$$\delta\nu = \nu_0 \frac{v_x}{c}$$

where
ν_0 = transition centre frequency, = $(E_2 - E_1)/h$
v_x = velocity of atom/molecule along line of sight

Each individual atom/molecule has a specific velocity and therefore contributes to a specific value of $\delta\nu$.

Each individual atom/molecule has a specific velocity and therefore contributes to a specific value of $\delta\nu$.



The values of $\delta\nu$ correspond exactly to the spread of v_x. Therefore, the lineshape function for Doppler broadening has the same form as the velocity distribution

$$P[v_x] = \sqrt{\frac{m}{2\pi kT}} \exp\left(\frac{-m v_x^2}{2kT}\right)$$

i.e. Gaussian lineshape

value of v_x when $P[V_x] = 0.5\, P[V_x=0]$

$$v_x = \sqrt{\frac{2 \ln(2)\, kT}{m}}$$

Therefore defining $\Delta\nu$ to be the Full Width of the transition at Half Maximum (FWHM), we get:

$$\Delta\nu_{Doppler} = \frac{\nu_0}{c} \sqrt{\frac{8 \ln(2)\, kT}{m}} \qquad [35]$$

For **Inhomogeneous** broadening, each frequency within the linewidth **always** corresponds to an **individual** atom/molecule.

Even in solids Inhomogeneous broadening mechanisms also exist.

Different atoms are sited in slightly different Lattice positions and therefore 'see' different perturbing fields, again we get a Gaussian lineshape.

Homogeneous Broadening

e.g. Natural Lifetime

All Homogeneous broadening mechanisms can be understood in terms of the **Uncertainty Principle**.

$$\Delta E.\ \Delta t \approx h$$

For our atoms /molecules, the lifetime of state 2 is τ_2. Therefore we can write that the uncertainty in the energy of state 2 is:

$$\Delta E_{state\ 2} \approx \frac{h}{\tau_2}$$

Relating the uncertainty in energy to frequency we get:

$$\Delta E_{state\ 2} = h\ \Delta\nu \approx \frac{h}{\tau_2}$$

$$\Delta\nu \approx \frac{1}{\tau_2} \qquad [36]$$

Within the Homogeneous Linewidth, **no distinction** can be made between different atoms/molecules. "All the individual atoms are broadened by the same amount"

The lineshape can also be derived from the uncertainty principle. We tend to write:

$$\Delta E.\ \Delta t \approx h$$

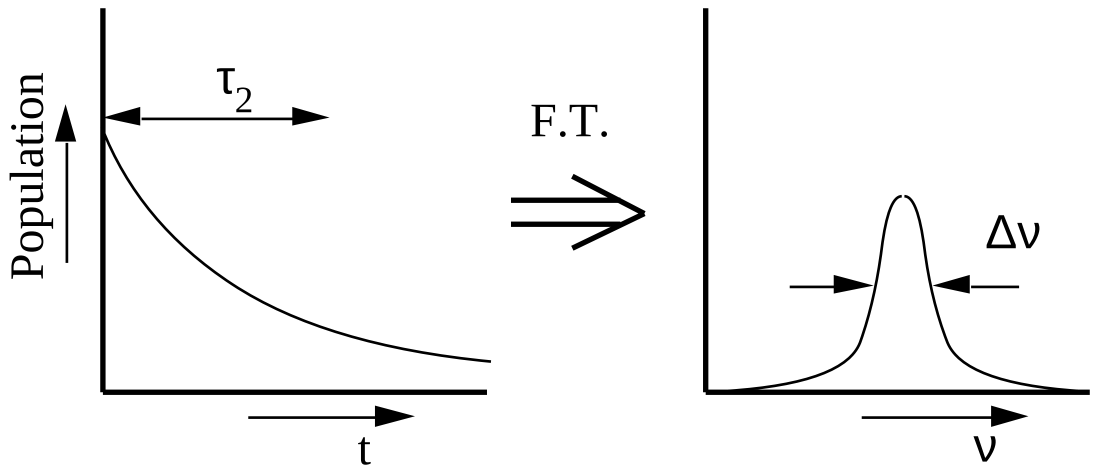
However, ΔE & Δt are related as a Fourier Transform Pair (as are Δp & Δx)

Fourier Transform

$$\Delta E \Longrightarrow \Delta t$$

h = 'scaling factor'

In our case,



So far we have only considered Δt to be limited by the natural spontaneous emission lifetime of the transition, i.e. the **natural linewidth**.

However, Δt may be further reduced by:

Collisions between atoms, i.e. **pressure broadening**

Rapid stimulated emission, i.e. **power broadening**

Inhomogeneous and Homogeneous Broadening

All transitions are homogeneously broadened (e.g. the spontaneous emission lifetime results in a natural linewidth)

All transitions are inhomogeneously broadened (e.g. Brownian motion giving a Doppler shift)

In most cases, one mechanism is far larger than the other and the transition is said to be homogeneously or inhomogeneously broadened.

Linewidth as a function of frequency

Consider a gas which has a number of different transitions throughout the spectrum.

For Doppler broadening, we have:

$$\Delta\nu_{Doppler} = \frac{\nu_0}{c}\sqrt{\frac{8\ln(2)\,kT}{m}} \qquad [35]$$

Therefore:

$$\Delta\nu_{inhom} \; \alpha \; \nu_0$$

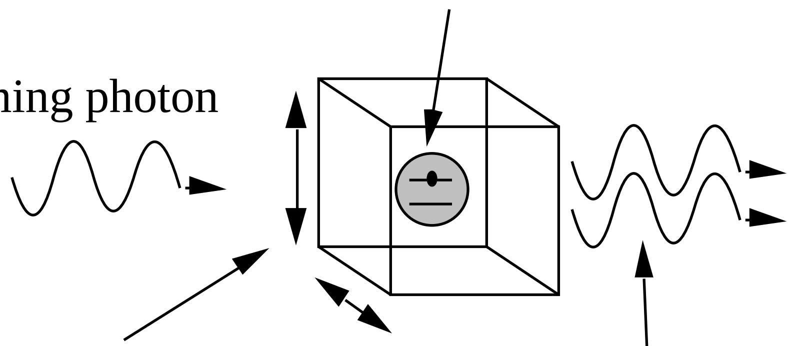
For the natural linewidth, it can be shown that:

$$\Delta\nu_{homo} \; \alpha \; \nu_0^2$$

Therefore, infra red transitions are long lived and Doppler broadening is very significant. UV transitions are short lived and the natural linewidth is more important.

Absorption / Emission Cross Section

Rather than using the A and B coefficients with g(ν), it is often more convenient to think in terms of an effective cross section.



The probability of a single atom undergoing stimulated emission is:

$$Prob_{stim} = \text{incident photon flux} \times \text{x-section}$$

$$Prob_{stim} = D_\nu \frac{c}{n} \,.\, \sigma_\nu$$

Where:

$$D_\nu = \text{photon number density} = \frac{\rho_\nu}{h\nu} = I_\nu \frac{n}{h\nu c}$$

$$\sigma_\nu = \text{x-section for stimulated emission}$$

i.e.

$$Prob_{stim} = \frac{I_\nu}{h\nu} \,.\, \sigma_\nu$$

For N atoms the number of emissions will be:

$$\text{Number} = \frac{I_\nu}{h\nu} \,.\, N.\, \sigma_\nu$$

Rewrite the rate equations we get:

$$\frac{dN_1}{dt} = -\frac{N_1}{\tau_1} + \frac{N_2}{\tau_2} + \sigma_\nu (N_2 - N_1) \frac{I_\nu}{h\nu} \quad [45]$$

$$\frac{dN_2}{dt} = -\frac{N_2}{\tau_2} - \sigma_\nu (N_2 - N_1) \frac{I_\nu}{h\nu} \quad [46]$$

Note by comparison to [5]:

$$\frac{dN_2}{dt} = A_2 N_2 + B_{21} \rho_\nu (N_2 - N_1) = 0 \quad [5]$$

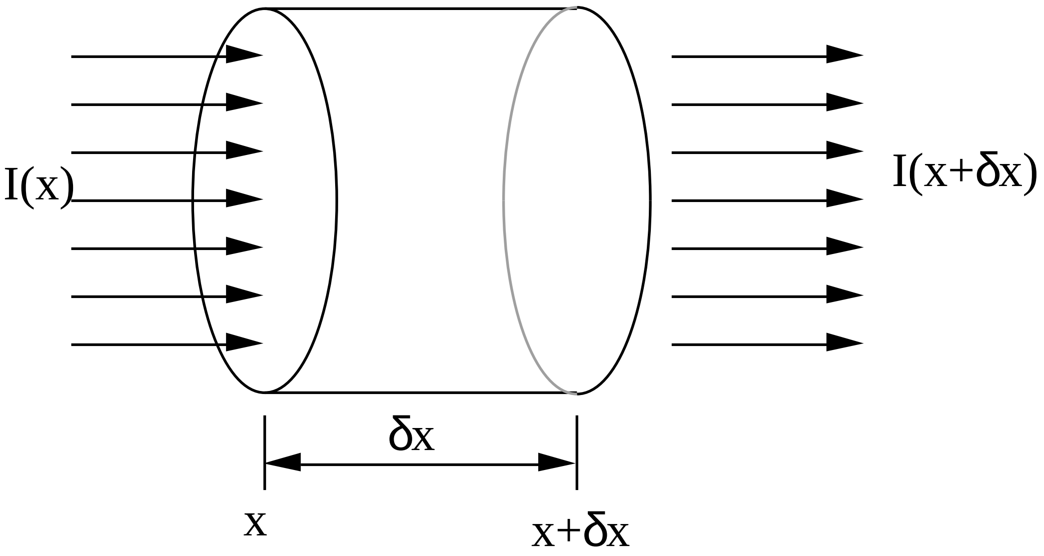
we can write:

$$\sigma_\nu = B_{21} h\nu \frac{n}{c}$$

or, using [12]:

$$\sigma_\nu = \frac{A_2 c^3}{8\pi n^3 h\nu^3} h\nu \frac{n}{c} = \frac{c^2}{8\pi n^2 \nu^2 \tau_2} \quad [47]$$

Let us consider a volume of the laser media



The number of photons per unit volume is D_ν:

where $I_\nu = \rho_\nu \frac{c}{n} = D_\nu h\nu \frac{c}{n}$

$$\frac{dD_\nu}{dt} = \text{stimulated emission / unit vol.} - \text{stimulated absorption/unit vol.}$$

At this point we can ignore spontaneous emission since the emitted photons will not be coherent with the incoming light and therefore do not give a useful output:

$$\frac{dD_\nu}{dt} = N_2 \sigma_\nu D_\nu \frac{c}{n} - N_1 \sigma_\nu D_\nu \frac{c}{n}$$

$$\frac{dD_\nu}{dt} = (N_2 - N_1) \sigma_\nu D_\nu \frac{c}{n} \qquad [16]$$

The change in D_ν can be related to the change in intensity by:

$$\Delta D_\nu = \frac{dI_\nu}{dx} \frac{\Delta x\, n}{h\nu\, c}$$

or simply we can say D_ν is proportional to I_ν

$$\frac{dI_\nu}{dt} = (N_2 - N_1)\,\sigma_\nu\, I_\nu \frac{c}{n} \qquad [17]$$

We want to express things with respect to x not t, hence

$$\frac{dI_\nu}{dx}\frac{dx}{dt} = (N_2 - N_1)\,\sigma_\nu\, I_\nu \frac{c}{n}$$

but we know that $\frac{dx}{dt} = \frac{c}{n}$

and therefore

$$\frac{dI_\nu}{dx} = (N_2 - N_1)\,\sigma_\nu\, I_\nu \qquad [18]$$

but we've already defined $\frac{dI_\nu}{dx} = -\alpha\, I_\nu(x)$

and therefore:

$$\alpha = (N_1 - N_2)\,\sigma_\nu \qquad [19]$$

α is called the **absorption coefficient.**

Under thermal equilibrium, we have a Boltzmann distribution and because $E_2 > E_1$, we have:

$$N_1 > N_2$$

therefore α is always +ve and the medium always absorbs.

However, if we can create a situation where:

$N_1 < N_2$ (known as a population inversion)

then α is -ve and the medium will amplify

When α is -ve, it is usually called k, where $k = -\alpha$, i.e.

$$k = (N_2 - N_1)\,\sigma_\nu \qquad [20]$$

k is called the **small signal gain coefficient.**

Gain Threshold

Once a population inversion is established, the gain medium will amplify an incident beam of light.

By reflecting the light backwards and forwards through the gain medium an intense beam of radiation can be established. By analogy to the electrical case we call this process **optical feedback** and the laser becomes a **laser oscillator.**

In reality, a gain of 1.000001 would not be enough to achieve laser action since we need enough gain not only to overcome the stimulated absorption in the gain medium but also to overcome the losses due to poor mirrors etc. i.e.

Gain ≥ Loss

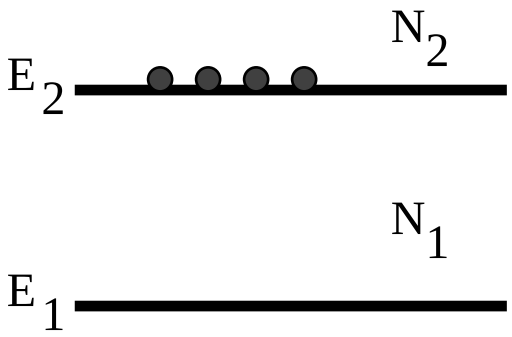
In the limit

Minimum Gain (or Threshold Gain) = Loss

The size of the population to achieve the threshold inversion gain is called the **Threshold Inversion**

Population inversion

In order to get amplification we need to create a population inversion i.e.



$N_2 > N_1$

We create the population inversion by **pumping.**

We need to find a way in which we can overcome the thermal distribution of states.

One way in which we could try to increase the population of state 2 is to illuminate the medium with light of frequency ν, where:

$$\nu = \frac{E_2 - E_1}{h}$$

However, since $B_{12} = B_{21}$ this light will just as likely cause stimulated absorption as stimulated emission. In the steady state we see from [5]:

$$B_{12}\, \rho_\nu\, N_1 = A_2\, N_2 \; + \; B_{21}\, \rho_\nu\, N_2$$

that for any value of ρ_ν:

$$N_2 \leq \frac{g_2}{g_1} N_1$$

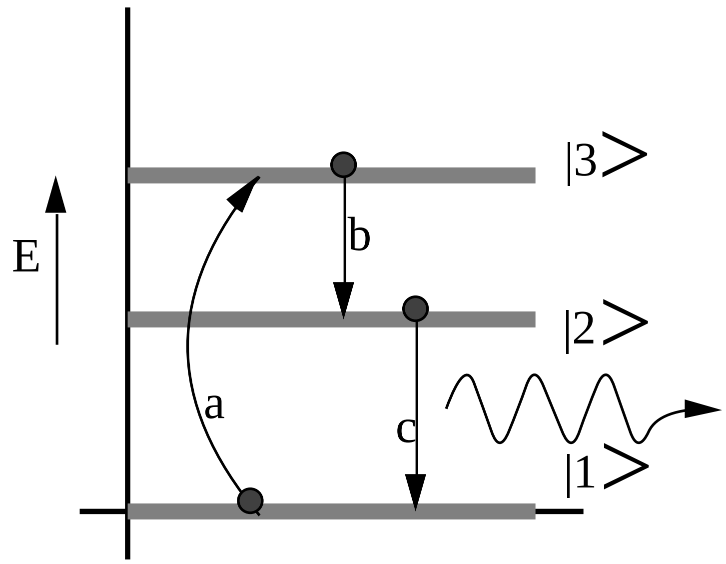
and hence, at best the medium will be transparent

Making the medium transparent by an intense optical pumping is called **bleaching.**

To achieve a population inversion we need to use media with more than two energy states. All practical laser systems can be simplified to three or four energy levels.

These are called
Three level laser systems
or **Four level laser systems**

Three level Laser



a) Pump from state 1 to state 3

The pump transfers population from the ground state to higher energy levels. (State 3 may be a collection of different levels)

b) Non radiative decay from state 3 to state 2

In a good laser medium, the lifetime of state 3 is short and all the population in state 3 rapidly decays to state 2

c) Stimulated emission from state 2 to state 1

In a good laser medium the lifetime of state 2 is long so that the population will grow and an inversion can be created with respect to state 1. Once an inversion is obtained, stimulated emission will give optical gain.

Three level system, getting an inversion

Under thermal equilibrium, nearly all the population resides in state 1. To get an inversion we need to pump at least half the total population via state 3 into state 2.

If the total population is N_{Tot}, then for an inversion we need:

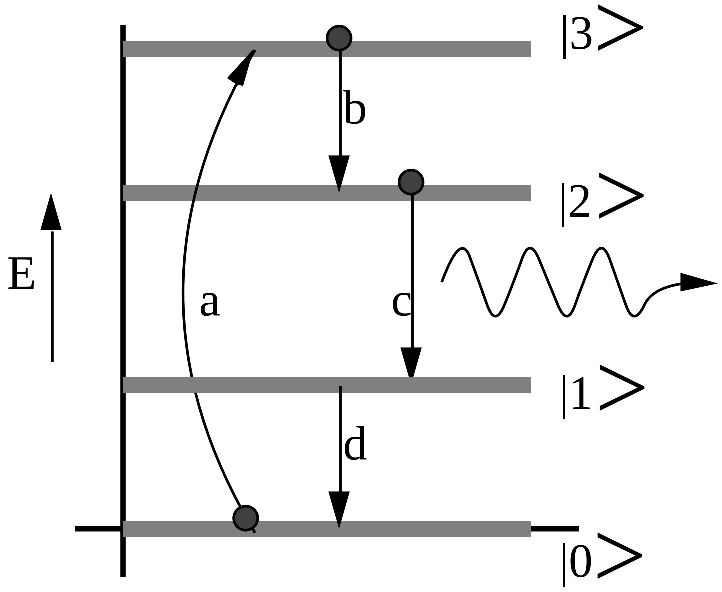
$$N_2 \geq \frac{N_{Tot}}{2}$$

However, the population in state 2 will decay due to spontaneous emission. To maintain the inversion we need to pump at:

$$R_{th\text{-}3\ level} = \frac{N_{Tot}}{2} \frac{1}{\tau_2} \qquad [24]$$

where
R_{th} = threshold pump level, excitations per sec, per unit vol.
τ_2 = lifetime of state 2

Four level Laser



a) Pump from state 0 to state 3

The pump transfers population from the ground state to higher energy levels. (State 3 may be a collection of different levels)

b) Non radiative decay from state 3 to state 2

In a good laser medium, the lifetime of state 3 is short and all the population in state 3 rapidly decays to state 2

c) Stimulated emission from state 2 to state 1

In a good laser medium the lifetime of state 2 is long so that the population will grow and an inversion can be created with respect to state 1. Once an inversion is obtained, stimulated emission will give optical gain.

d) Non radiative decay from state 1 to state 0

In a four level laser, terminal state of the laser transition is **not** the ground state and therefore a population inversion is easier to maintain.

Four level system, getting an inversion

Under thermal equilibrium, nearly all the population resides in state 0, i.e. state 1 is empty. A population inversion between states 1 and 2 and be obtained even for small populations in state 2.

However, to overcome losses in the resonator etc., we need to achieve a gain somewhat greater than unity. As mentioned before, this corresponds to a threshold inversion N_{th}. (but typically $N_{th} << N_{Tot}/2$).

The required level of pumping is:

$$R_{th\text{-}4\ level} = N_{th} \frac{1}{\tau_2} \qquad [25]$$

What is the threshold inversion?

To achieve laser operation we need to create a large enough inversion so that the gain exceeds the loss.

Source of loss:

1) Transmission of mirrors (we need a partially transmitting mirror to get some output light!).

2) Absorption and scattering by the mirrors.

3) Unwanted absorption by the laser gain media (these would be by transitions other than those we have considered).

4) Scattering by optical imperfections (e.g. contamination within the laser crystal).

5) Diffraction losses.

All the losses other than mirror transmission are grouped into an overall loss, giving a loss coefficient γ,

We can characterise a cavity by the time it takes the oscillating light to escape. The intensity in the cavity will decay exponentially with a time constant t_c.

We can work out the minimum gain required by equating the exponential gain to the exponential loss

Gain: $I = I_o \exp(kx)$

Loss: $I = I_o \exp\left(\frac{-t}{t_c}\right)$

Time (t) is related to distance (x) by the speed of light $\left(\frac{c}{n}\right)$

Therefore the cavity loss time t_c, equates to a cavity loss length of $t_c \left(\frac{c}{n}\right)$.

At threshold by equating the "gain length" to the "loss length" we get:

$$\frac{1}{k_{th}} = t_c . \left(\frac{c}{n}\right)$$

which can be rearranged to give [30]:

$$k_{th} = \frac{n}{c\, t_c} \qquad [30]$$

Then sub [30] into [20] to get the population inversion required to achieve threshold:

$$k = (N_2 - N_1)\, \sigma_\nu \qquad [20]$$

$$N_{th} = (N_2 - N_1) = k_{th} / \sigma_\nu$$

$$N_{th} = \frac{n}{\sigma_\nu\, c\, t_c} \qquad [31]$$

N_{th} is called the **threshold inversion.**

Also sub in for σ_ν in terms of τ_2: from [47]

$$\sigma_\nu = \frac{c^2}{8\pi n^2 \nu^2 \tau_2}$$

Our expression for the threshold inversion becomes:

$$N_{th} = \frac{8\pi n^3 \nu^2\, \tau_2}{c^3\, t_c} \qquad [32]$$

For some lasers, the pump power required to attain a threshold can only be achieved for a short time. These lasers are **pulsed lasers.**

If the inversion can be maintained indefinitely the output is continuous. Such a laser is a continuous-wave, i.e. a **CW laser.**

Examples of laser systems

Helium Neon Laser

Active Medium: 90% He, 10% Ne, ≈10 torr gas

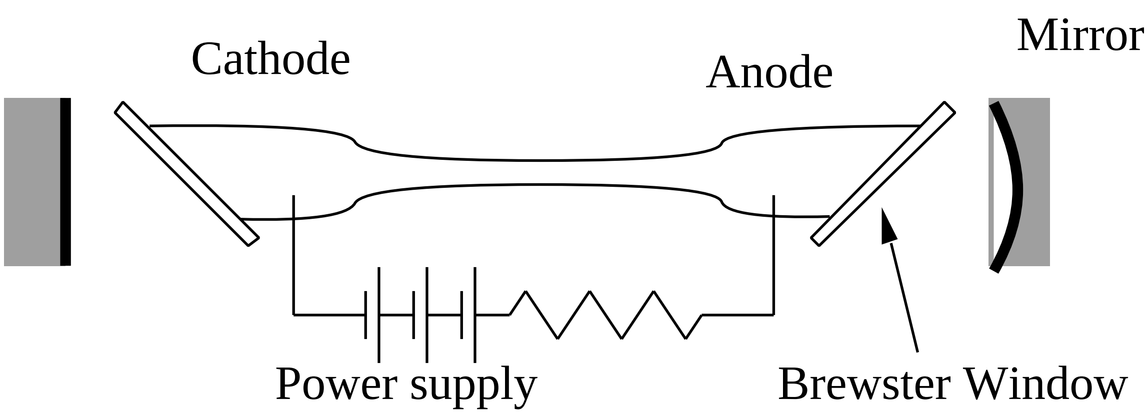
Pumping: Electrical discharge

Output Wavelength: 632nm, 1.15µm & 3.39µm (select by mirror choice)

Typical power levels: 1-10mW

Cost: £100's

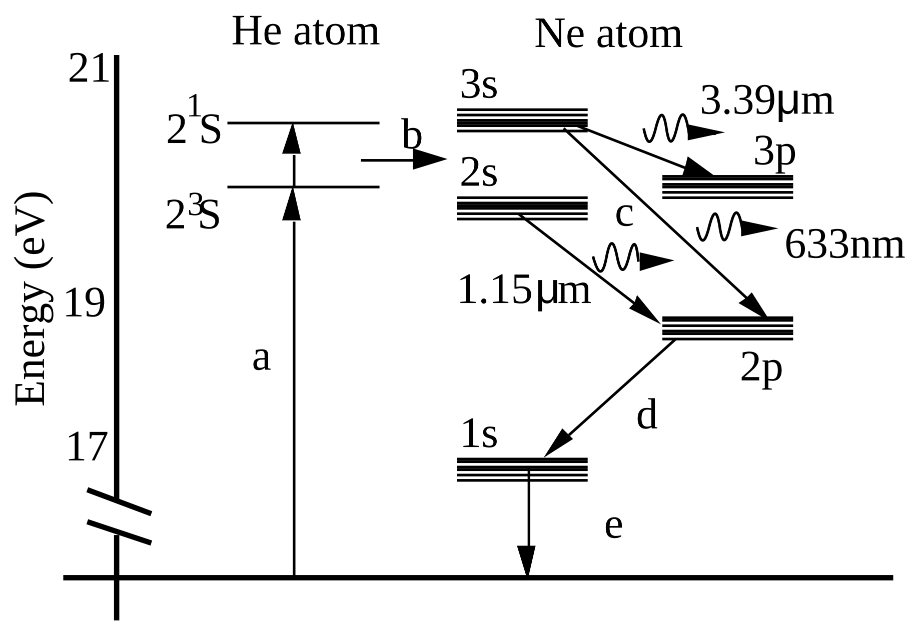
Laser Type: 4 level, inhomogeneous



In the He-Ne laser, the population inversion is created between two energy levels in the neon atoms.

Similar to many other systems pumped by electrical discharge, the neon is excited indirectly.

Energy level diagram for He-Ne laser



a Electron impact excites the helium atoms into the long-lived 2^3S and 2^1S states

b Collision between He and Ne atoms excites the neon into the 2s and 3s states.

c Population inversion created between the 3s/3p and 2s/2p states in neon. Stimulated emission gives gain.

d The lifetime of the 2p and 3p states is short and they rapidly decay to the 1s state.

e Collisions between the neon atoms and the tube walls returns the neon atom to the ground state.

With the He-Ne laser the tube walls play an important part in maintaining the population inversion (step e).

Therefore cannot operate at large tube diameter or high gas pressure.

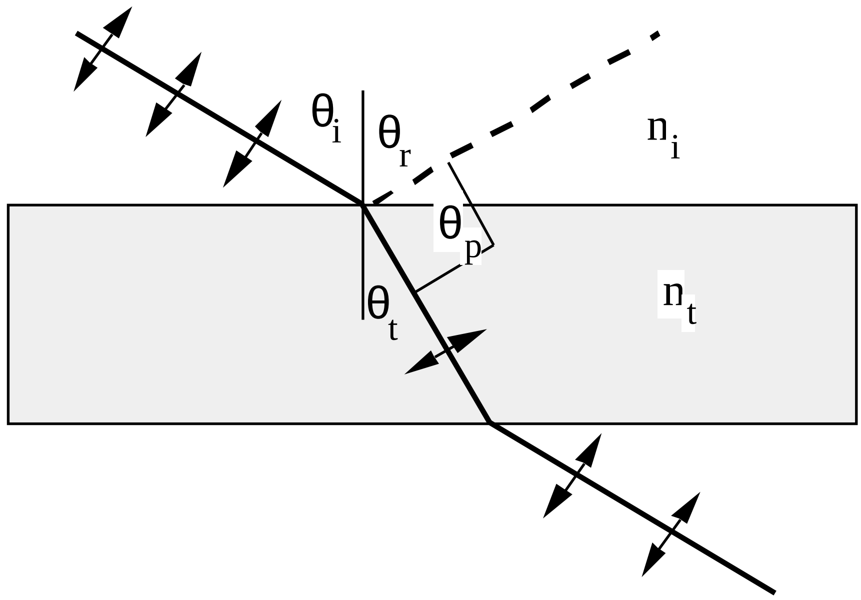
Brewster Windows

It is often convenient to contain the gain medium inside the resonator in a separate container/vessel. This protects the mirrors from the medium (which may be a high temperature gas) and allows them to be easily adjusted.

Often this is done using Brewster windows. Any transparent material aligned so the angle of incidence for the incoming light ray is

$$\theta = \tan^{-1}\frac{n_t}{n_i} \qquad \text{(n = refractive index)}$$

exhibits no reflection for one polarisation.

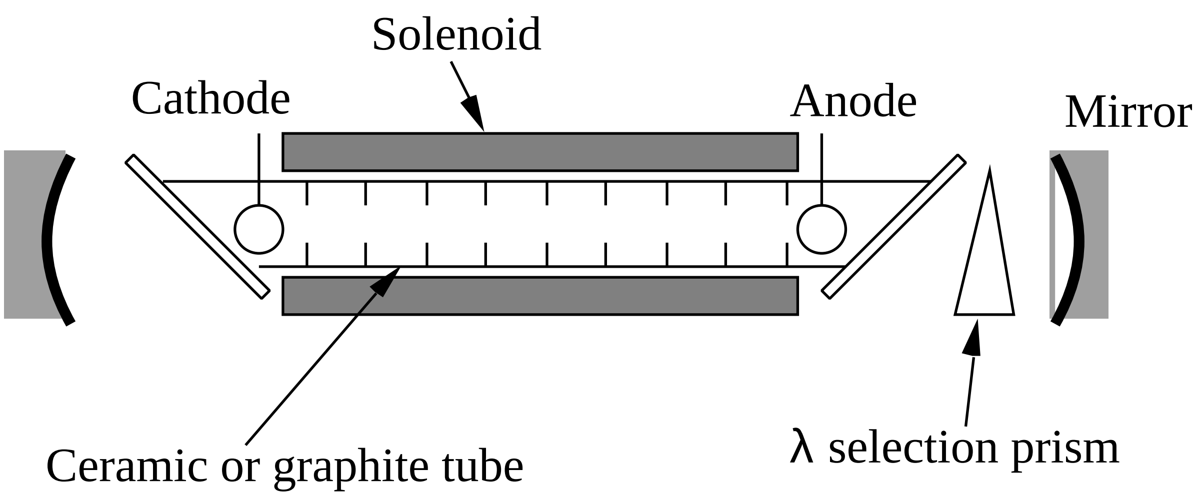


When θ_p = 90° there is no reflection of the || polarisation state. The ⊥ polarisation state is reflected and therefore suffers increased loss.

Lasers with Brewster windows give linearly polarised outputs.

Argon Ion Laser

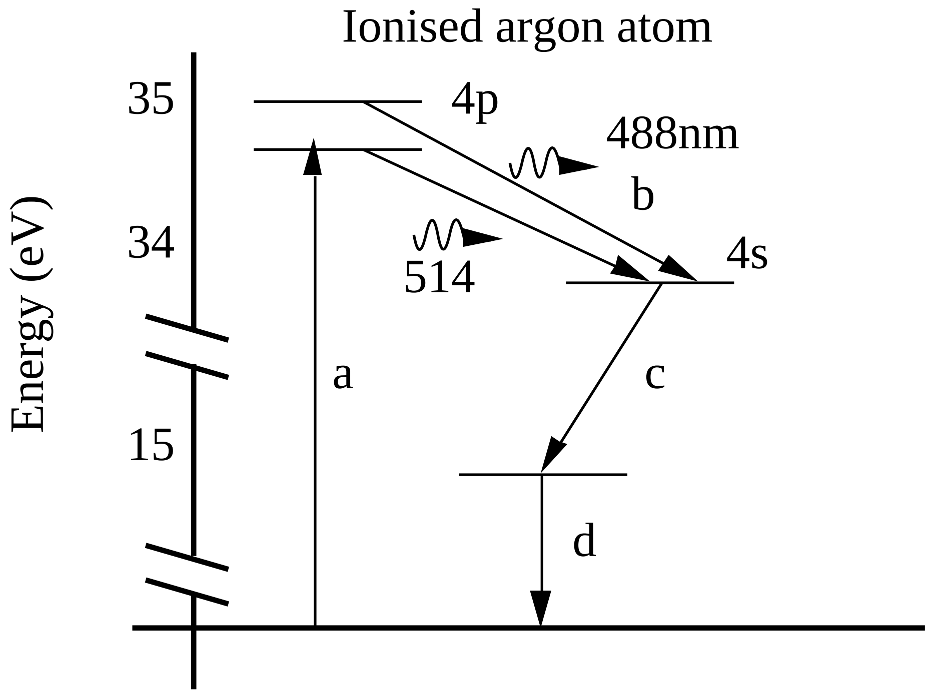
Active Medium:	Ionised argon atoms (gaseous)
Pumping:	Electrical discharge
Output Wavelength:	Mainly 514nm and 488nm
Typical power levels:	1-10W
Cost:	£10k -100k
Laser Type:	4 level, inhomogeneous



In the Ar^+ laser, the population inversion is created between two energy levels in ionised argon

Because of the high currents, a solenoid is often used to contain the plasma in the centre of the tube, away from the walls. Discs within the tube acts as heat exchangers and the whole tube is cooled with a water jacket (low power lasers can be air cooled)

Energy level diagram for Ar^+ laser



a The argon atoms are ionised by electron collision within the discharge (up to 50amps!). Further multiple collisions excite the ions to the 4p states.

b A population inversion is created between the 4p/4s states. Stimulated emission gives gain.

c The 4s state has a short lifetime and decays to the ground state of the argon ion (giving off UV in the process).

d The argon ion recaptures an electron

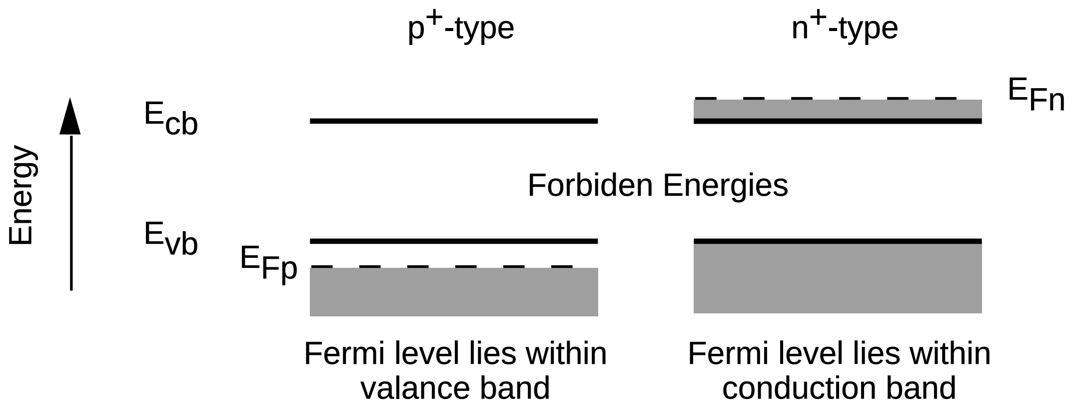
Ar^+ lasers have a number of laser transitions in the blue/green region of the spectrum. An internal prism can be used to select the desired output line. Or for maximum power, the prism can be removed and simultaneous output on "all lines" can be obtained.

Argon Ion lasers use lots of power and breakdown often!

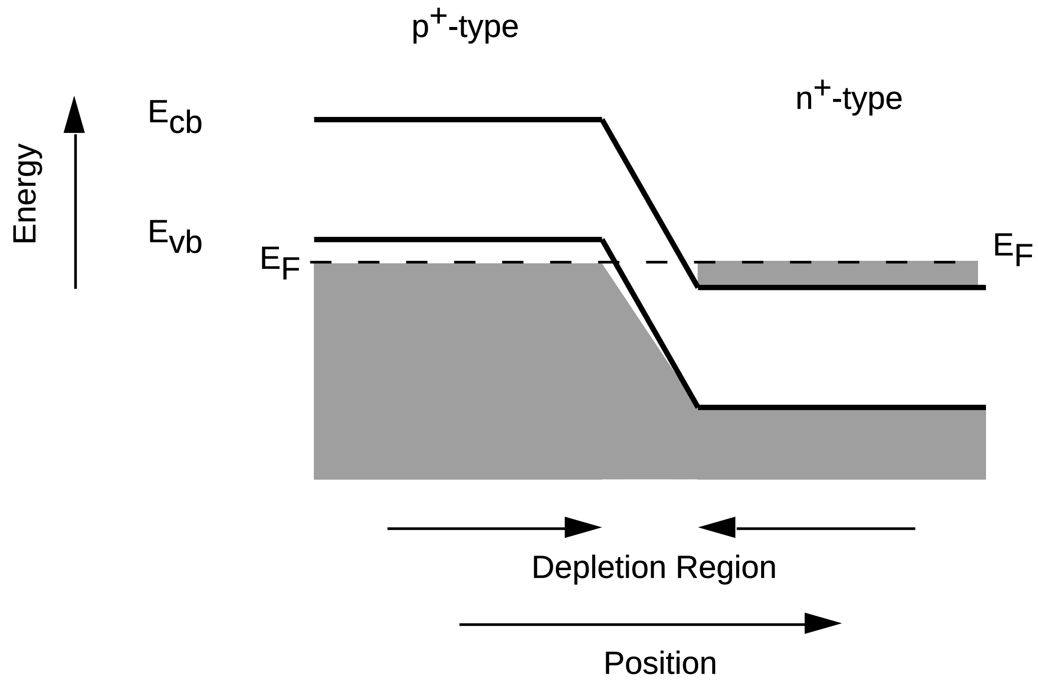
Laser diodes

Active Medium:	Direct bandgap semiconductor e.g. Gallium Arsenide
Pumping:	Electrical
Output Wavelength:	360nm - 7μm (different diodes and extending)
Typical power levels:	1mW- 1W
Cost:	£1- £20k
Laser Type:	4 level

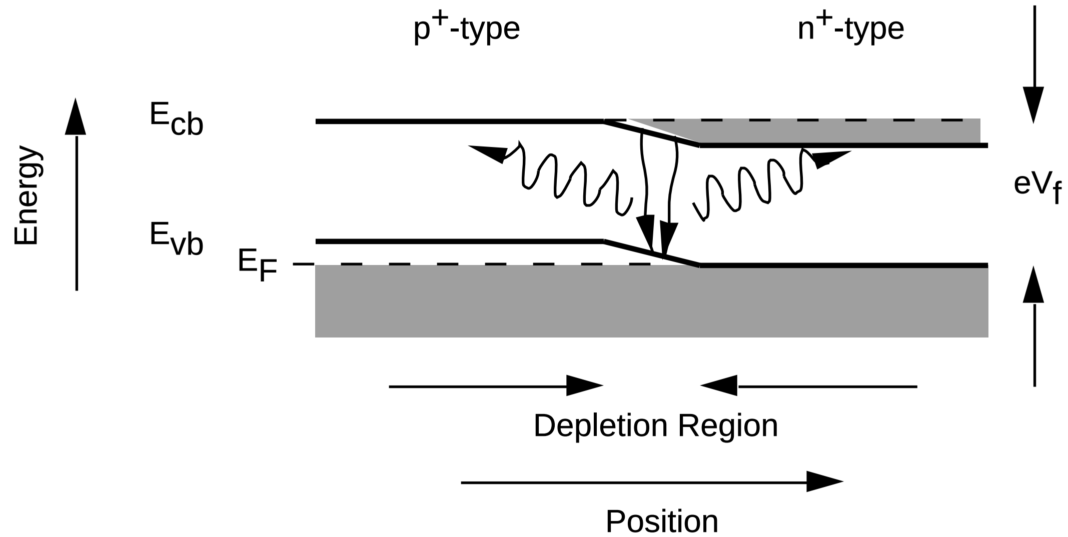
Laser diodes are based on very heavily doped p-n junctions (heavy doping is denoted p^+ or n^+). In heavily doped p^+-type and n^+-type materials, the Fermi energy lies within the valence and conduction bands respectively.



In a p^+-n^+ junction the contact potential is nearly equal to the energy gap.



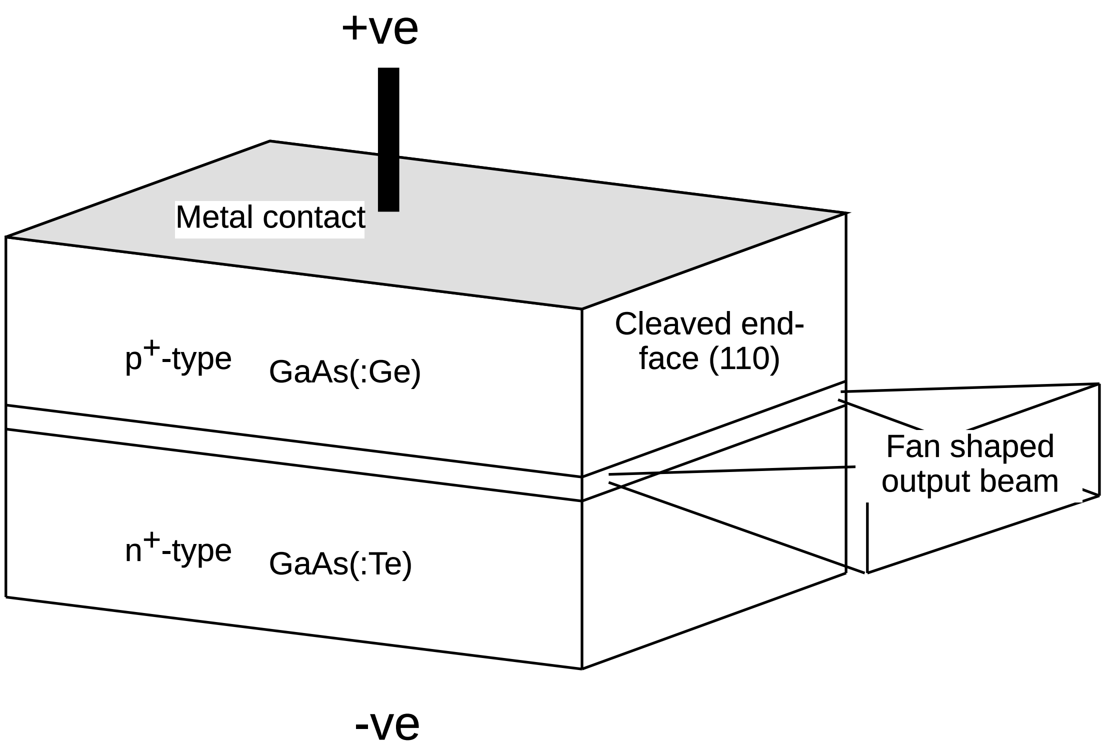
Under forward bias, the electrons and holes recombine extremely close (in) to the depletion region. This is called the active region



The size of the active region is approximately equal to the diffusion length, which for heavily doped material is 1-3µm.

For materials with a direct band-gap the recombination will be radiative, although the heavy doping also leads to strong reabsorption of the emitted light.

If the injected carrier concentration is large enough then the stimulated emission dominates over the absorption and optical gain is observed. A homojunction laser is shown below.



Points to note:

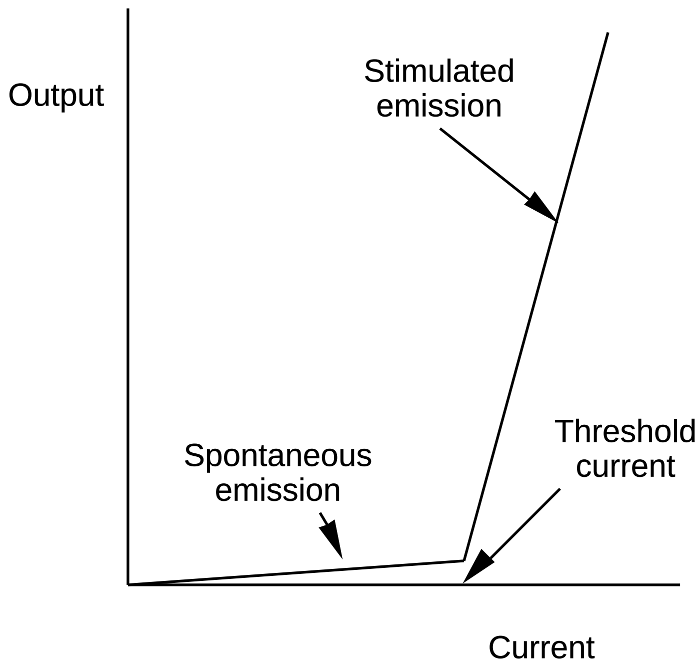
- cavity mirrors formed by Fresnel reflection from uncoated end faces of the crystal

$$R = \left(\frac{n_2 - n_1}{n_2 + n_1}\right)^2$$

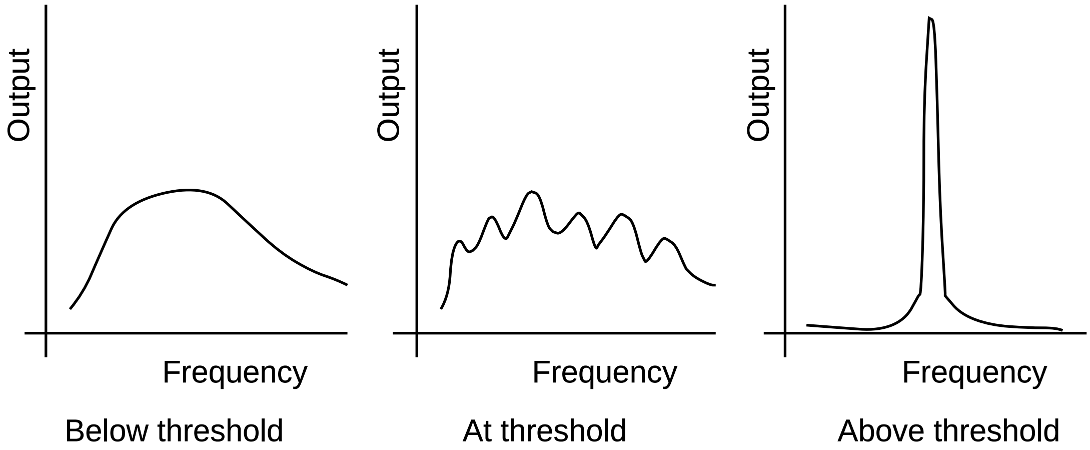
- the µm width of the active region results in light spreading outside the active region. Therefore mode volume > active region
- the excess of carriers in the active region modifies the refractive index (+0.1) to give slight waveguiding
- principal losses are due to scattering from crystal defects and free-carrier absorption

Homojunction (p and n materials the same) lasers have high losses due to free-carrier absorption. Combined with the large active area this results in high threshold currents (≈400A mm^{-2}). Therefore, operation is only possible in a pulsed mode or at low temperatures.

The power output of a laser diode increases rapidly as threshold is reached.



Additionally the spectral profile of the laser diode changes as a function of forward bias

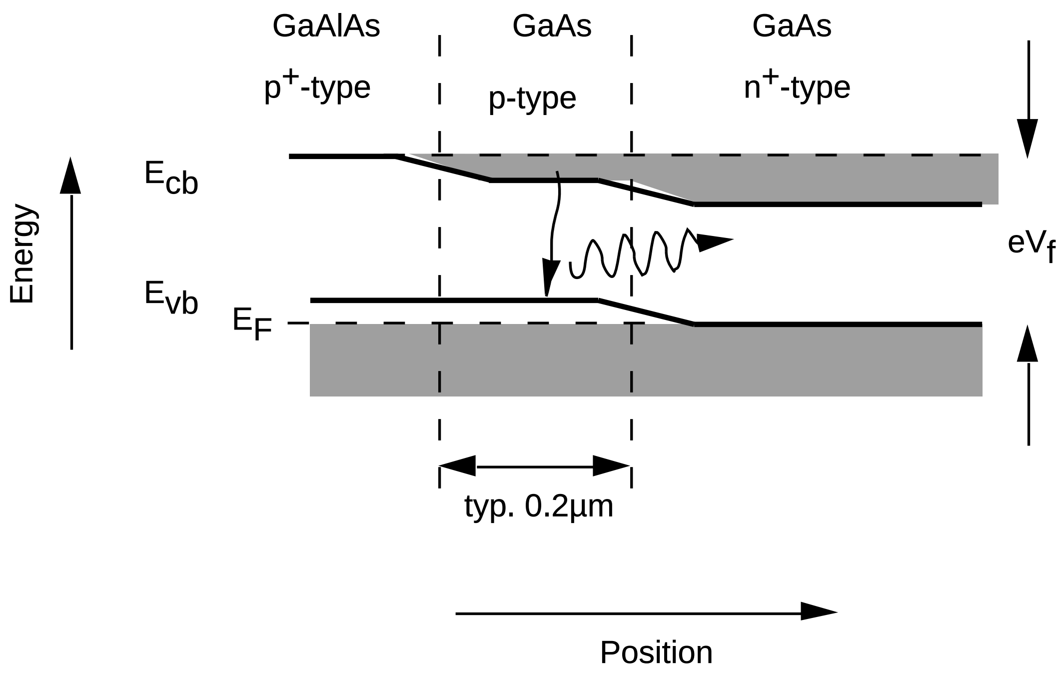


The key to reducing the losses in a diode laser is to improve the light guiding in the active region.

Heterojunction Lasers

A heterojunction is a junction between materials of different types.

In a heterojunction laser an extra thin layer of material creates a region of higher refractive index as the active region which also acts as a waveguide for the photons

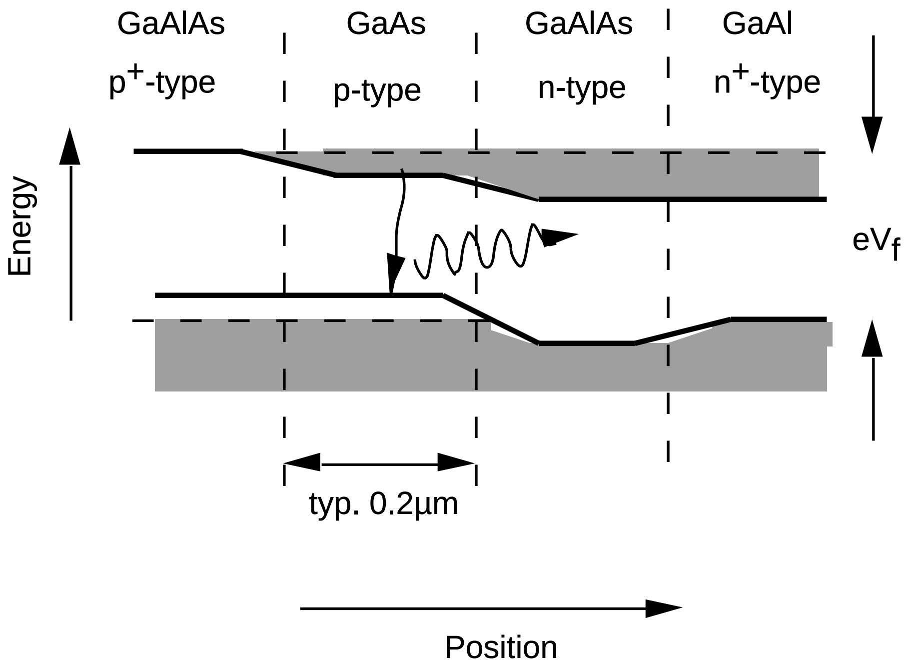


Points to note:

- GaAs acts as waveguide for emitted photons
- GaAlAs has larger band-gap than GaAs, therefore design reduces photon absorption in p^+-type material
- threshold reduces to 10A mm^{-2}

Double Heterojunction laser

The heterojunction principle can be taken a further step to a double heterojunction.



Points to note:

- GaAs acts as waveguide for emitted photons
- GaAlAs has larger band-gap than GaAs, therefore design reduces photon absorption on BOTH sides of active region
- threshold reduces to 10A mm^{-2}

Quantum Well Lasers

If the addition layer within a heterojunction is reduced from 100's nm to 10nm or less it becomes comparable to the wavelength of the electrons within the crystal. The electron is effectively trapped within a one-dimensional quantum well. Consequently, the energy levels and corresponding wavelengths of emission are modified. To increase the overall power output, many adjacent quantum wells can be formed in the region of the junction. These are called multiple quantum well lasers.

Stripe Geometry Lasers

Heterojunction designs allow the threshold to be reduced to a low current density. To reduce the current to a low value it is necessary to reduce the area of the laser diode.

If the length is reduced the round trip gain may fall below threshold, therefore need to reduce the width.

The current flow into the active region can be restricted to a single stripe along the length of the laser which may only be a few microns wide.

The stripe can be formed using a patterned electrode or by selective processing of the semiconductor material.

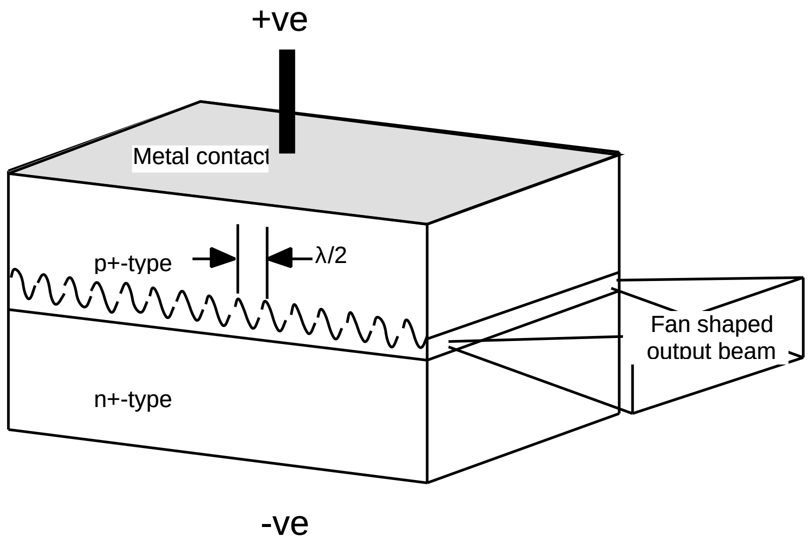
A stripe geometry gives threshold currents as low as 50mA with power outputs of 10mW.

Other advantages of stripe geometry are
- power emitted from small area therefore collimation and focusing are simplified
- power output can be more stable since lateral stability of active region is improved.

Distributed feedback lasers

As an alternative to mirrors or cleaved crystal ends a diode laser cavity can be formed using a Bragg Grating. A Bragg Grating is a similar idea to a multilayer dielectric coating which relies upon stacked reflectors separated by half wavelengths such that constructive interference between reflection from each layer results in a strong reflection for one particular wavelength.

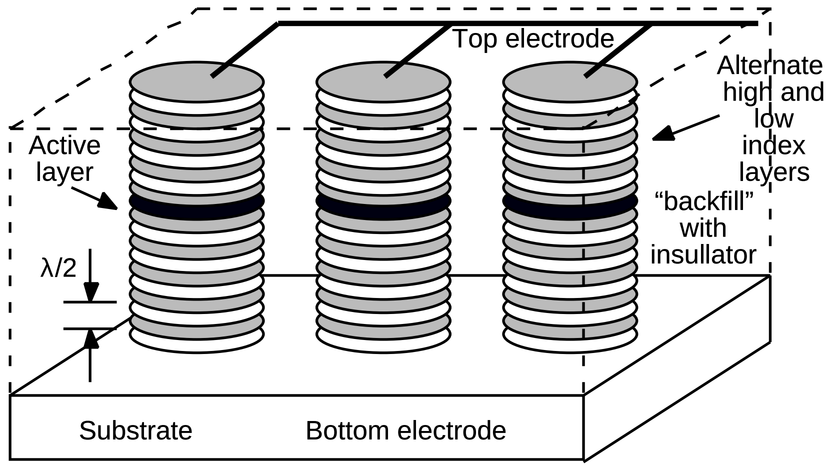
In this case, each “reflector” is created by changing the thickness of the active region. By placing a corrugation on the surface this introduces a periodic change mode index (see later) resulting in a strong distributed reflection at a particular wavelength. Hence the term distributed feedback laser (DFB)



DFB lasers are extremely useful since the exact wavelength of the output is determined by the period of the structure.

Vertical cavity lasers

One particular interesting form of DFB laser is the vertical cavity surface emitting laser (VCSEL). In which the light is emitted vertically, perpendicular to the plane of the substrate.



These VCSEL devices can be made many to the chip and require no coating consequently they now represent nearly half the world market for laser diodes

Tuning diode lasers

There are two methods by which laser diodes can be easily tuned

- temperature tuning, the wavelength of a GaAs laser diode tunes by approximately 0.3nm per °C
- current tuning, the frequency typically tunes over several 10's GHz as the drive current is altered.

Failure modes of laser diodes

Failure may be gradual or catastrophic

Catastrophic failure occurs when one of the end facets is damaged. The small active volume and cross-section results in high intensities even at low powers. The maximum optical flux is typically 10^9W m^{-2}.

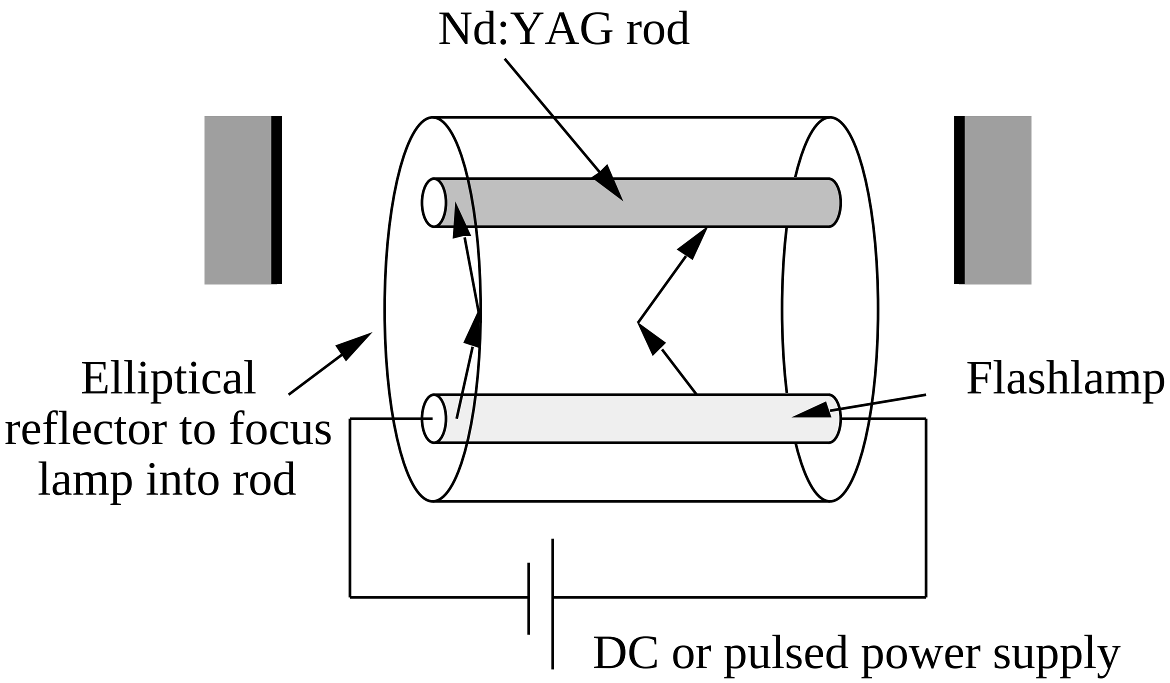
Gradual failure gives rise to an increase in threshold current with time. Dark stripes (regions of low gain) appear in the active region. These occur at high current density and are associated with the migration and generation dislocations in the lattice. Consequently, fabrication process and control are of critical importance. The dark stripes act as centres for non-radiative recombination.

Nd:YAG

Active Medium:	Neodymium impurity in Yttrium Aluminium Garnet (solid state)
Pumping:	Optical
Output Wavelength:	1064nm
Typical power levels:	up to 50Wcw

Cost: £10k - 50k

Laser Type: 4 level, inhomogeneous



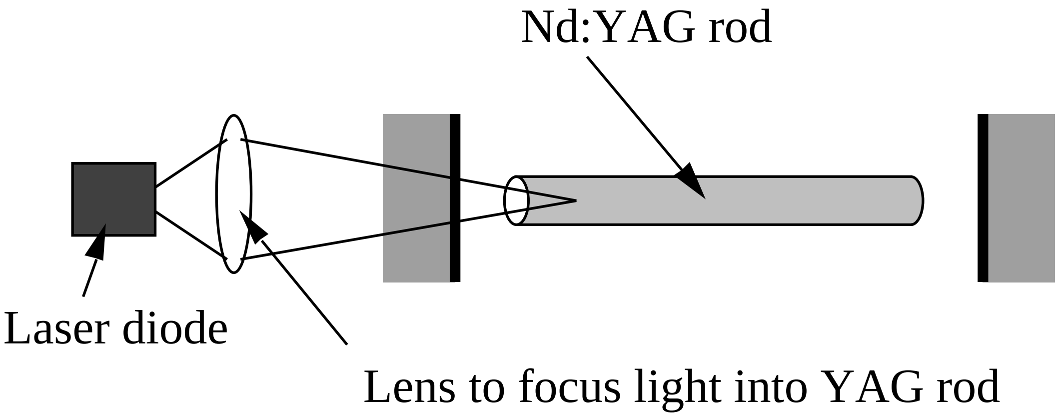
The YAG is doped with Nd^{3+} ions replacing some of the Yttrium.

The Nd^{3+} ions absorb light from the pump (particularly at ≈810nm) and are excited into a number of high energy states from which they decay to the $^4F_{3/2}$ state.

The inversion is formed between the $^4F_{3/2}$/$^4I_{11/2}$ states.

The $^4I_{11/2}$ is short lived and it rapidly decays to the ground state, thereby maintaining the inversion.

Within the last decade, Nd:YAG lasers have also been pumped using laser diodes, temperature tuned to match the absorption of the Nd:YAG at ≈810nm. Frequency doubling is frequently employed to obtain a 532nm output in the green.



Nd:YAG lasers are found in both pulsed and CW forms.

Copper Vapour Laser

Active Medium:	Copper vapour at ≈1500°C in a neon buffer gas
Pumping:	Electric discharge
Output Wavelength:	578nm and 510nm
Typical power levels:	up 50W average power
Cost:	£10'sk
Laser Type:	4 level, but metastable lower laser level, inhomogeneous

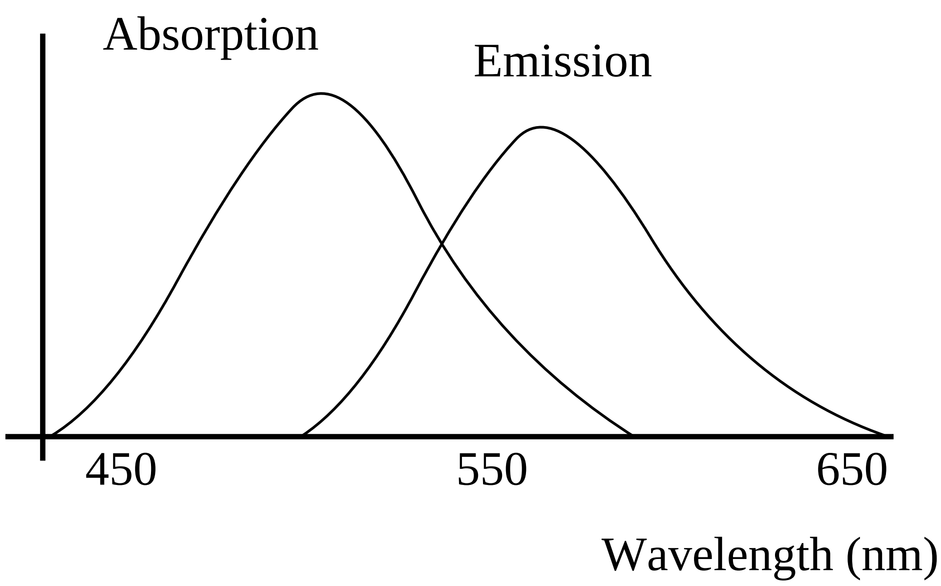
Solid copper is heated in an oven to ≈1500°C, the vapour pressure is ≈0.1 torr.

An electric current passed down the tube generates a population inversion in the copper. However, the lower laser level is long lived and therefore the population inversion cannot be sustained. Therefore, copper vapour lasers are always pulsed. Rep rates as high as several kHz may be obtained.

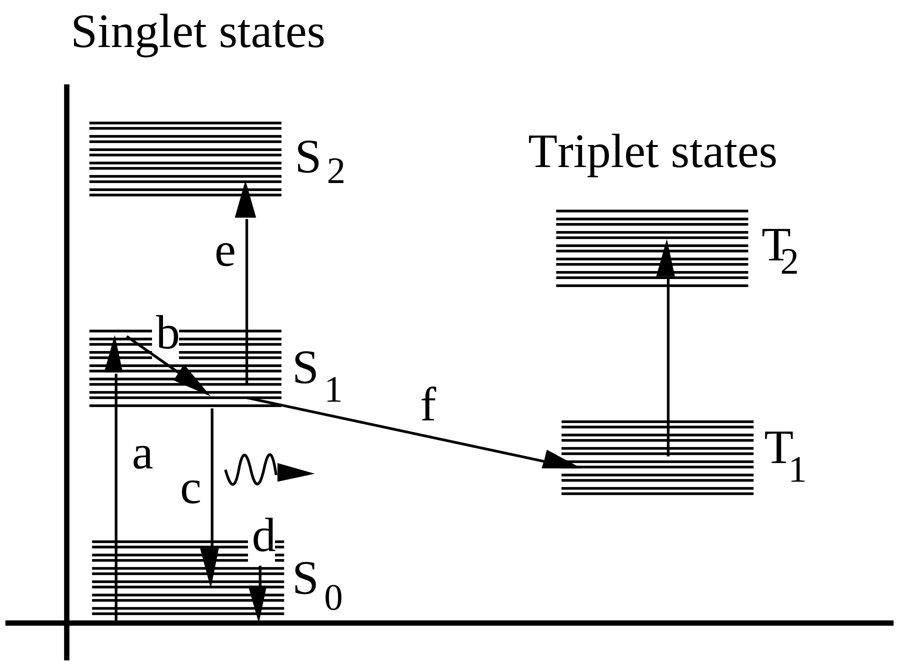
Dye Laser

Active Medium:	Organic dyes in liquid solvent (e.g. Rhodamine 6G)
Pumping:	Optical
Output Wavelength:	350nm - 900nm (with 10 different dyes)
Typical power levels:	up to 1W CW
Cost:	£10'sk + pump source
Laser Type:	4 level, homogeneous

Organic dyes have very broad absorption (spontaneous absorption) and fluorescence spectra (spontaneous emission) if used as the laser gain medium we would expect them to give us broad tunability.



Dye lasers are pumped optically. This can be via flashlamps (similar to flashlamp pumped Nd:YAG) , or for CW output by another laser.



Each electronic energy level is significantly broadened by the vibration and rotation of the dye molecule.

The terms singlet and triplet refer to the spin state of the excited electrons.

a Optical pumping excites a dye molecule into the first singlet state.

b The molecule quickly relaxes to the bottom of the excited state.

c An inversion is created between the bottom of the excited singlet state and the upper levels within the ground state.

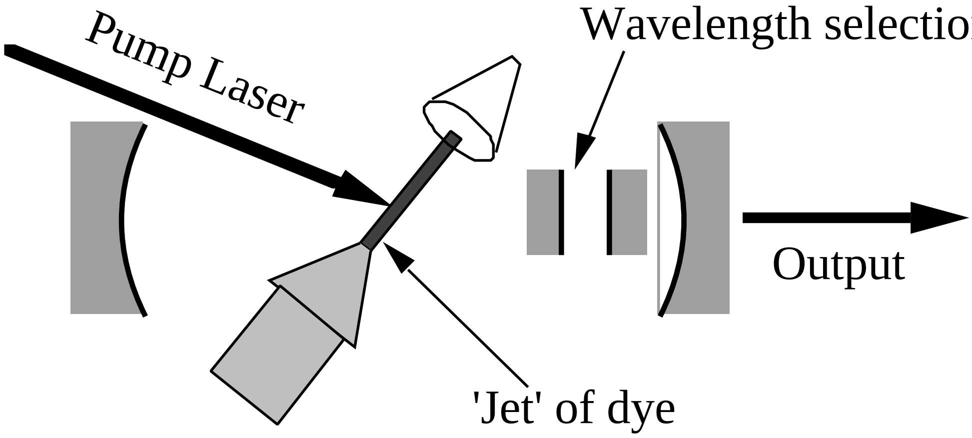
d The molecule quickly relaxes to the bottom of the ground state, thereby maintaining the inversion.

e Unwanted re-absorption of the laser light may cause further excitation of the dye, leading to reduced gain.
This is called excited state absorption (ESA)

f In particular, CW systems are limited by the gradual build up of the triplet population.

In a CW system a rapid flow of dye is maintained so that the triplet state population is removed from the active region.

Argon ion are the lasers most frequently used to pump a dye laser. The rapid flow of dye is maintained using a jet of liquid dye formed by a nozzle and recycled through the system. A particular dye will give laser output over ≈100nm, wavelength selective optics (e.g. prisms and etalons) are used within the cavity to select a particular wavelength and longitudinal mode.



Titanium Sapphire Laser

Active Medium:	Ti (3+) ion in sapphire host
Pumping:	Optical
Output Wavelength:	700nm - 900nm
Typical power levels:	up to 1W CW
Cost:	£10'sk + pump source

Laser Type: 4 level, homogeneous

Ti-sapph lasers have largely replaced dye lasers as tunable laser sources within the research laboratory. Despite having a restricted tuning range, the vastly improved "ease of use" (i.e. lack of leaking liquid dyes!) has persuaded people of their scientific merit.

They look and behave very similar to the dye lasers they replace, and are also pumped with an additional laser source (either argon ion or more recently frequency doubled YAG).

As with Dye lasers the broad tuning range allows the generation of ultrashort pulses (see modelocking).

More recently similar laser materials such as Cr Li saF which absorb in the red can be directly pumped using laser diodes. This will potentially yield "shoe boxed" sized tunable laser sources for reduced cost.

Gain saturation

What happens to the gain (k), the population inversion (N) and the radiation density (I_ν) as we pump harder and harder (increasing R)?

We have shown that there is a threshold value for the population inversion, below which the total gain (i.e. gain + loss) is less than unity and no laser output results.

Once above threshold, the total gain is greater than unity and the energy density within the cavity (I_ν) will grow.

Obviously, for a steady state or CW (continuous wave) output, I_ν must be constant and therefore the gain must be equal to unity.

As I_ν increases, so does the rate of stimulated emission, which acts to remove population from state 2 and hence reduce the population inversion.

The steady state condition corresponds to the case when the level of stimulated emission is just sufficient to offset the pumping and maintain the inversion at the threshold value. Thereby giving a total gain of one.

Consider the populations of states 1 and 2, but this time include the effect of pumping:

$$\frac{dN_1}{dt} = -\frac{N_1}{\tau_1} + \frac{N_2}{\tau_2} + \sigma_\nu (N_2 - N_1)\frac{I_\nu}{h\nu} \qquad [45]$$

$$\frac{dN_2}{dt} = R - \frac{N_2}{\tau_2} - \sigma_\nu (N_2 - N_1)\frac{I_\nu}{h\nu} \qquad [50]$$

Where
R = pumping rate excitations/sec per unit vol.

For the steady state condition:

$$\frac{dN_1}{dt} = \frac{dN_2}{dt} = 0$$

Adding [45] and [50] we get:

$$N_1 = R\,\tau_1 \qquad [51]$$

Subtracting [45] and [50] we get:

$$N_2 = \frac{N_1/2\tau_1 + N_1\sigma_\nu \dfrac{I_\nu}{h\nu} + R/2}{1/\tau_2 + \sigma_\nu \dfrac{I_\nu}{h\nu}} \qquad [52]$$

sub [51] into [52]:

$$N_2 = \frac{R\left(1 + \sigma_\nu \dfrac{I_\nu}{h\nu}\tau_1\right)}{1/\tau_2 + \sigma_\nu \dfrac{I_\nu}{h\nu}}$$

Therefore, the population inversion is given by:

$$N = N_2 - N_1 = \frac{R\,(1 - \tau_1/\tau_2)}{1/\tau_2 + \sigma_\nu \dfrac{I_\nu}{h\nu}}$$

Note, if $\tau_1 > \tau_2$ then N is always negative, i.e. there can be no population inversion

But in an ideal four level laser system, $\tau_2 >> \tau_1$, therefore:

$$N = N_2 - N_1 = \frac{R}{1/\tau_2 + \sigma_\nu \dfrac{I_\nu}{h\nu}} \qquad [53]$$

At threshold, $I_\nu = 0$, therefore: (see also [25])

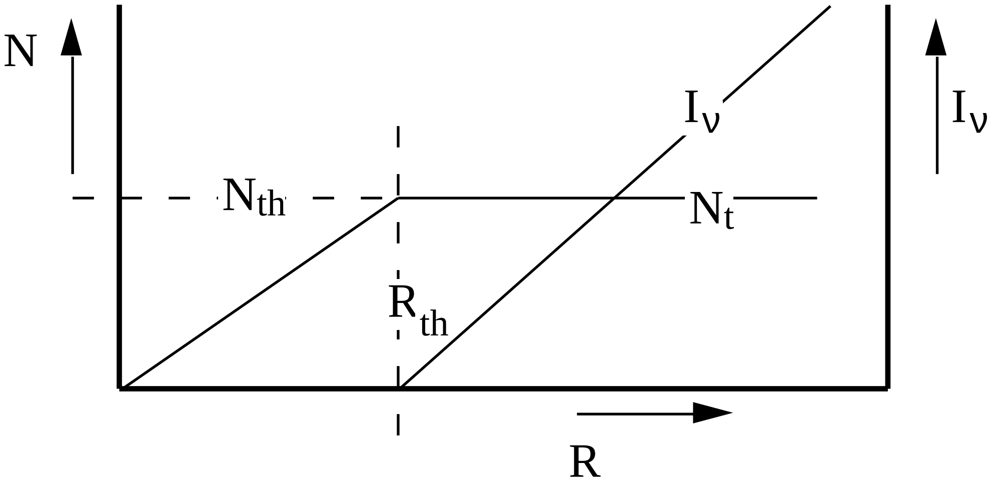
$$N_{th} = R_{th}\,\tau_2 \qquad [54]$$

Even when $R > R_{th}$, $N = N_{th}$. Sub [54] into [53]:

$$R_{th}\,\tau_2 = \frac{R}{1/\tau_2 + \sigma_\nu \dfrac{I_\nu}{h\nu}}$$

$$I_\nu = \frac{R\,h\nu}{\sigma_\nu R_{th}\,\tau_2} - \frac{h\nu}{\sigma_\nu\,\tau_2}$$

$$I_\nu = \frac{h\nu}{\sigma_\nu\,\tau_2}\left(\frac{R}{R_{th}} - 1\right) \qquad [55]$$



How much light is required to saturate the gain?

$$N = N_2 - N_1 = \frac{R}{1/\tau_2 + \sigma_\nu \dfrac{I_\nu}{h\nu}}$$

Remember:

N = Population inversion no. m^{-3}

R = Pump rate no. m^{-3} per sec

σ_ν = x-section for stimulated emission

I_ν = Light intensity = $\rho_\nu \dfrac{c}{n} = D_\nu\, h\nu \dfrac{c}{n}$

τ_2 = Lifetime of upper laser level

$$N = N_2 - N_1 = \frac{R\tau_2}{1 + \tau_2 \sigma_\nu \frac{I_\nu}{h\nu}}$$

Prior to lasing when $I_\nu = 0$, $N = R\tau_2$
We define this to be N_i , the initial population inversion:

$$N = \frac{N_i}{1 + \tau_2 \sigma_\nu \frac{I_\nu}{h\nu}} \qquad [58]$$

We define a **saturation intensity** I_{sat}:

$$I_{sat} = \frac{h\nu}{\tau_2 \sigma_\nu} \qquad [59]$$

Sub [59] into [58], to get an expression for the population inversion as a function of I_ν:

$$N = \frac{N_i}{1 + \frac{I_\nu}{I_{sat}}} \quad \text{(homogeneous only)} \qquad [60]$$

When $I_\nu = I_{sat}$, $N = N_i/2$.

Since the gain (k) is proportional to the inversion (N), we can write:

$$k = \frac{k_i}{1 + \frac{I_\nu}{I_{sat}}} \quad \text{(homogeneous only)} \qquad [61]$$

Where k_i is the gain when $I_\nu \approx 0$

Likewise, when $I_\nu = I_{sat}$, $k = k_i/2$

Note that when $I_\nu = I_{sat}$, from [59]:

$$\frac{I_\nu \sigma_\nu}{h\nu} = \frac{1}{\tau_2}$$

Remember,

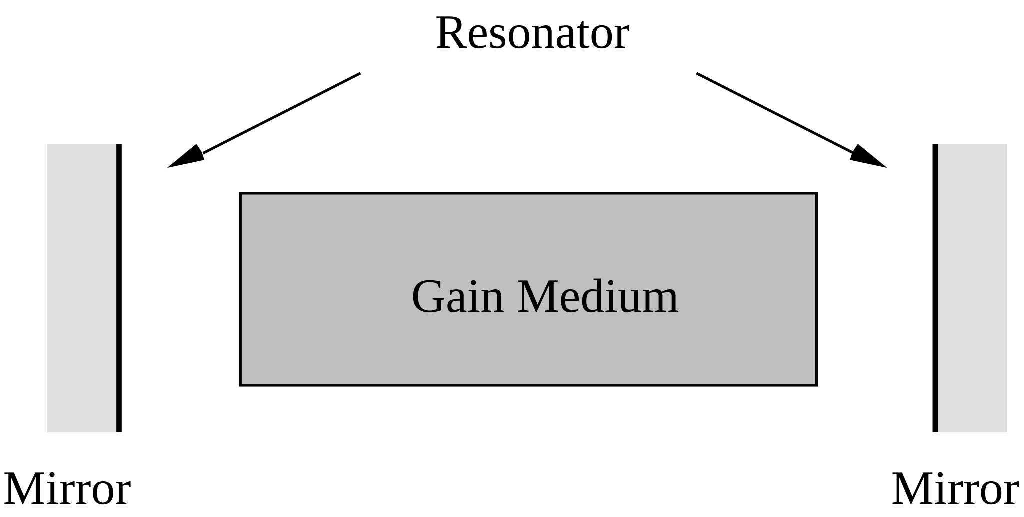
Rate of stimulated emission = $\frac{I_\nu \sigma_\nu}{h\nu} N_2$

Rate of spontaneous emission = $\frac{1}{\tau_2} N_2$

Therefore when $I_\nu = I_{sat}$

stimulated emission rate = spontaneous emission rate

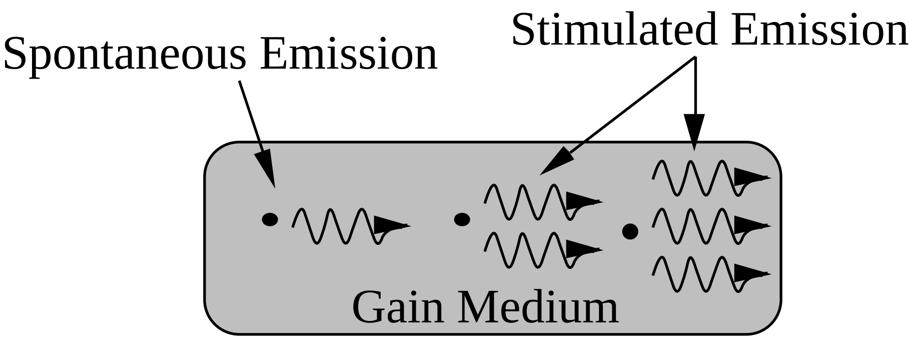
The Laser Resonator



So far we have only discussed the gain medium.

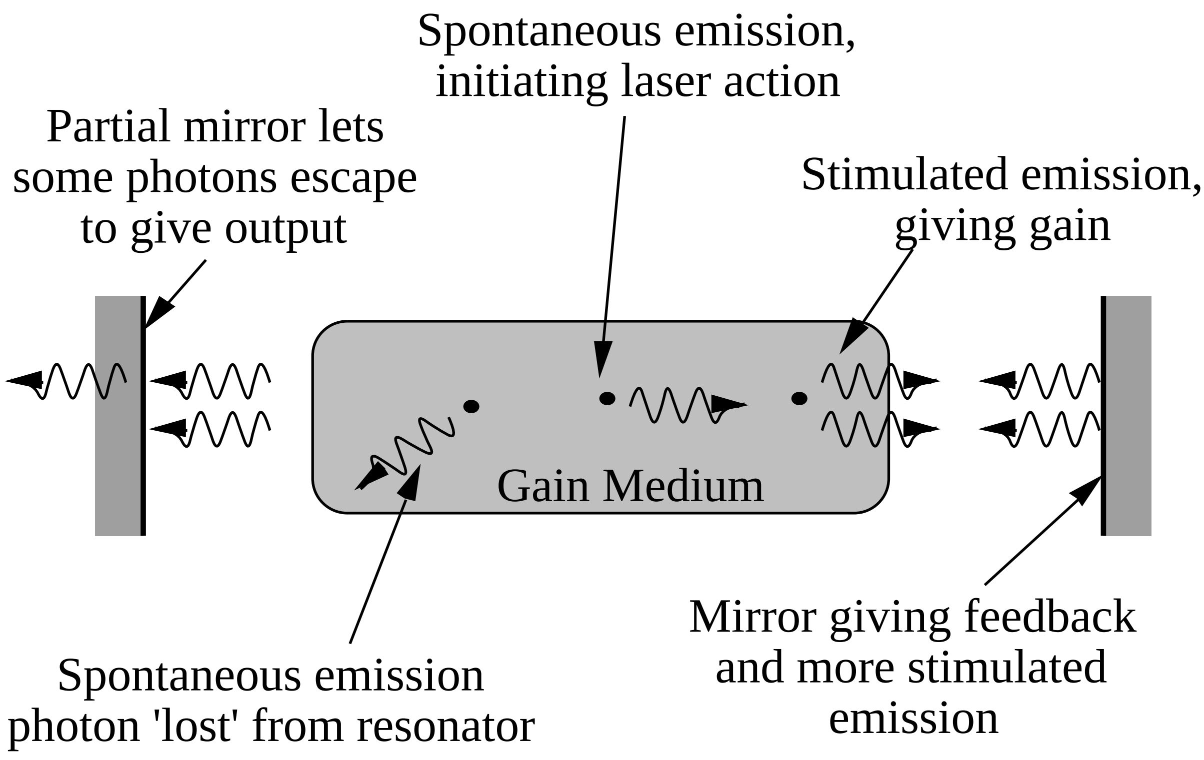
In order to get Stimulated emission, we need to feedback into the gain medium some of the previously emitted light, i.e. increase I_ν.

'Lasers' in which the gain is so high that no feedback is needed are called super-fluorescent.



However, with most laser systems some form of feedback is required.

The simplest system to obtain optical feedback is to place the gain medium between two highly reflective mirrors.



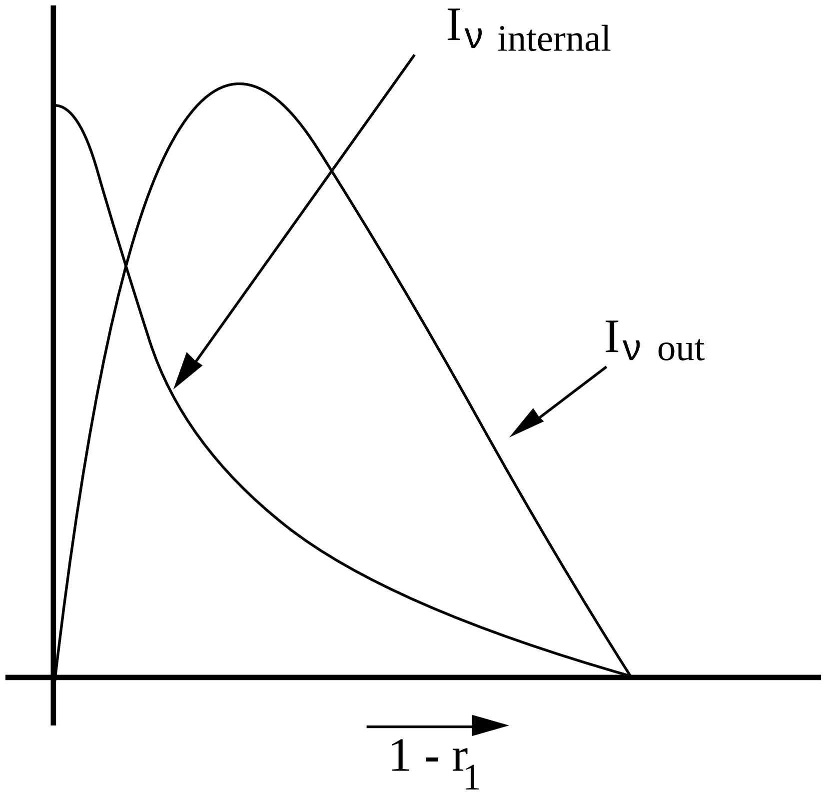
Laser action begins by a 'chance' spontaneous emission along axis of resonator (optical equivalent of electrical noise).

Initial 'spontaneous photon' experiences gain due to stimulated emission.

Mirrors provide optical feedback, and the energy density within resonator increases.

A partially transmitting mirror allows some of resonator field to 'escape', giving an output beam.

If the mirrors are perfect then no light escapes if the mirrors let out too much then the laser will not operate – there is an optimum!

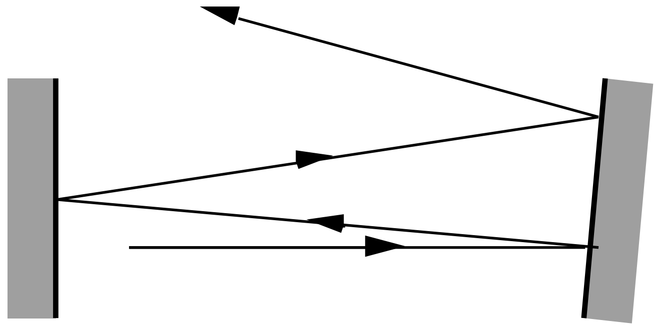


We might assume that light of any frequency can propagate up and down the resonator as a well collimated beam. We would be wrong!

Transverse Modes

Think of two plane mirrors separated by a distance L, if one of them is slightly mis-aligned then successive reflections will cause the ray to escape from the resonator. This is called beam **walk off,** and is a source of cavity loss (γ).

Beam 'walk off'

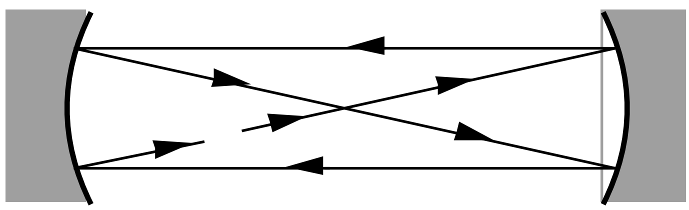


Mis-aligned mirror

Even for perfectly aligned mirrors we have to worry about diffraction of light by the finite aperture of the mirror, also leading to light loss.

We can use curved mirrors to 're-focus' the beam at each reflection. Curved mirror cavities are more stable than plane mirror cavities. The intensity distribution across the width of the cavity is the **Transverse Mode** pattern.

Beam 'held' within resonator



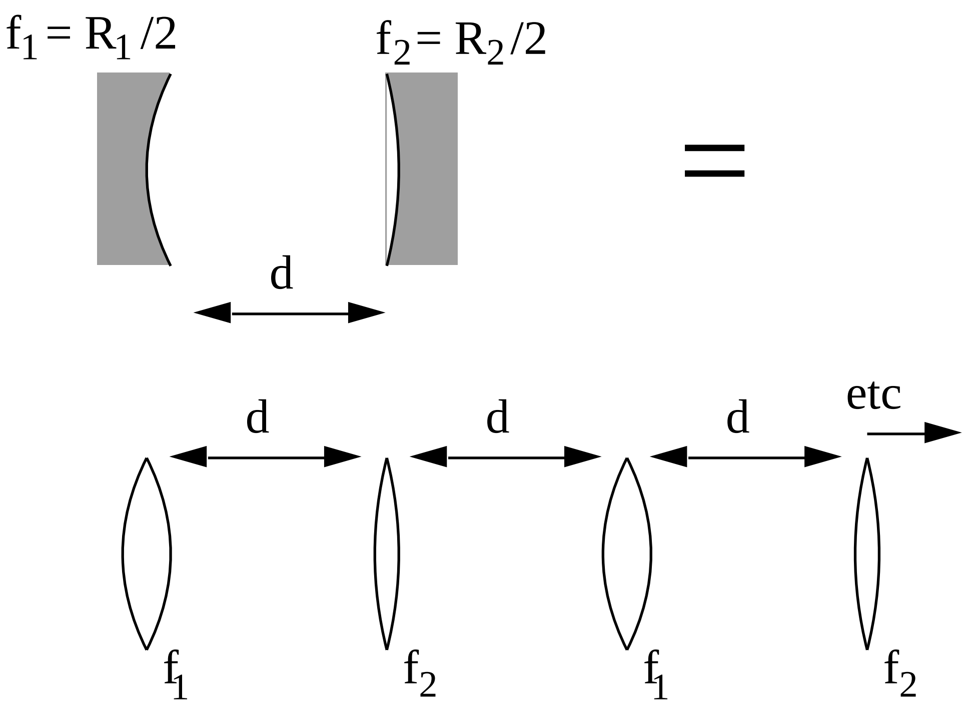
A frequently used mirror geometry is the **confocal** geometry. Mirrors of radius of curvature R (i.e. focal length R/2) are separated also by R (i.e. their focal points are coincident in the centre of the resonator)

For low loss, high reflectivity mirrors are very important. These can be made as high as 99.9%. Such mirrors are made from alternate λ/4 layers of high and low refractive index materials so that the reflections from each of the interfaces add up in phase.

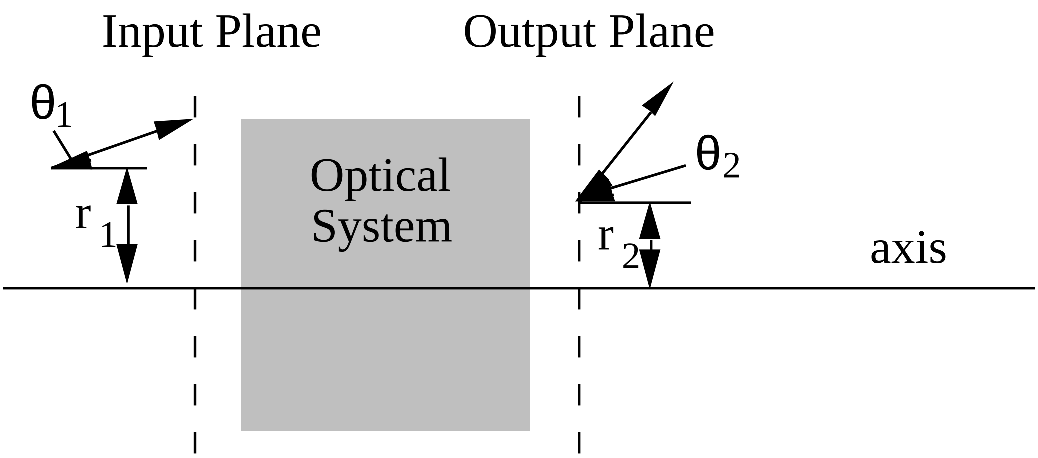
Conditions for Cavity Stability

We have said that curved mirrors can be used to re-focus the light back into the cavity and prevent light 'escaping'. What are the conditions for mirror curvature (i.e. focal length) and mirror separation for a stable cavity?

We can think of our mirror cavity in terms of a series of lenses.



Matrix optics can be used to calculate what happens to a light ray when it passes through an optical system.



In general we can write:

$$\begin{pmatrix} r_2 \\ \theta_2 \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} r_1 \\ \theta_1 \end{pmatrix}$$

For a simple lens (focal length f) the A, B, C, D matrix is:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}_{\text{lens}} = \begin{pmatrix} 1 & 0 \\ \frac{-1}{f} & 1 \end{pmatrix}$$

i.e. $r_2 = r_1$ and $\theta_2 = \frac{-r_1}{f} + \theta_1$

For a free space section (length d) the A, B, C, D matrix is:

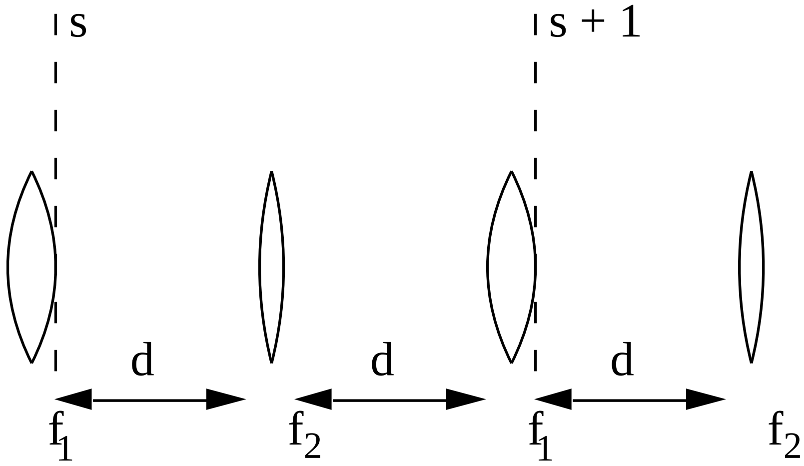
$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}_{\text{free space}} = \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix}$$

i.e. $r_2 = r_1 + d\theta_1$ and $\theta_2 = \theta_1$

For a free space section followed by a lens the A, B, C, D matrix is (remember to multiply the matrix in reverse order):

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{-1}{f} & 1 \end{pmatrix} \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & d \\ \frac{-1}{f} & 1-\frac{d}{f} \end{pmatrix}$$

Our laser cavity is a repeating series of two lenses:



The A, B, C, D matrix to transfer from the s to the s+1 plane is:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & d \\ \frac{-1}{f_1} & 1-\frac{d}{f_1} \end{pmatrix} \begin{pmatrix} 1 & d \\ \frac{-1}{f_2} & 1-\frac{d}{f_2} \end{pmatrix}$$

$$= \begin{pmatrix} 1-\frac{d}{f_2} & d\left(2-\frac{d}{f_2}\right) \\ \frac{-1}{f_1} - \frac{1}{f_2}\left(1-\frac{d}{f_1}\right) & 1 - \frac{2d}{f_1} - \frac{d}{f_2} + \frac{d^2}{f_1 f_2} \end{pmatrix}$$

Hence:

$$r_{s+1} = Ar_s + B\theta_s \qquad [65]$$

$$\theta_{s+1} = Cr_s + D\theta_s \qquad [66]$$

Where:

$$A = 1 - \frac{d}{f_2}$$

$$B = d\left(2 - \frac{d}{f_2}\right)$$

$$C = \frac{-1}{f_1} - \frac{1}{f_2}\left(1 - \frac{d}{f_1}\right)$$

$$D = 1 - \frac{2d}{f_1} - \frac{d}{f_2} + \frac{d^2}{f_1 f_2}$$

From [65], we can write:

$$\theta_s = \frac{1}{B}(r_{s+1} - Ar_s) \qquad [67]$$

and therefore:

$$\theta_{s+1} = \frac{1}{B}(r_{s+2} - Ar_{s+1}) \qquad [68]$$

Equating [66] and [68] , we get:

$$Cr_s + D\theta_s = \frac{1}{B}(r_{s+2} - Ar_{s+1})$$

Substituting for θ_s using [67]:

$$BCr_s + D(r_{s+1} - Ar_s) = r_{s+2} - Ar_{s+1}$$

Re-arrange in terms of r_s:

$$r_{s+2} - (A + D)r_{s+1} + (AD - BC)r_s = 0$$

We can show that (AD - BC) = 1, therefore:

$$r_{s+2} - (A + D)r_{s+1} + r_s = 0$$

i.e. $r_{s+2} - 2br_{s+1} + r_s = 0$ [69]

where

$$b = \frac{(A + D)}{2} = 1 - \frac{d}{f_1} - \frac{d}{f_2} + \frac{d^2}{2f_1 f_2}$$

Equs. of the form of [69] have the solution:

$$r_s = r_0 \exp\{i(sq)\} \qquad [70]$$

sub [70] into [69]:

$$\exp(2\, iq) - 2b \exp(iq) + 1 = 0$$

$$\{\exp(iq)\}^2 - 2b \{\exp(iq)\} + 1 = 0$$

therefore:

$$\exp(iq) = b \pm i\sqrt{1 - b^2} \qquad [71]$$

let $b = \cos\psi$, [71] becomes:

$$\exp(iq) = \cos\psi \pm i \sin\psi \quad \exp(\pm i\psi)$$

sub into [70]:

$$r_s = r_0 \exp(\pm is\psi) \qquad [72]$$

Remember that r_s is the distance from the optical axis to the ray. For a 'trapped' ray r_s needs to remain finite, therefore ψ must be real.

$$b = \cos\psi$$

Hence for real ψ we have the condition:

$$-1 \le b \le 1$$

i.e.

$$-1 \le 1 - \frac{d}{f_1} - \frac{d}{f_2} + \frac{d^2}{2 f_1 f_2} \le 1$$

or

$$0 \le \left(1 - \frac{d}{2f_1}\right)\left(1 - \frac{d}{2f_2}\right) \le 1$$

This is the stability criteria for a periodic series of two lenses. By replacing focal length with mirror curvature we can obtain the stability criteria for a two mirror cavity:

$$0 \le \left(1 - \frac{d}{R_1}\right)\left(1 - \frac{d}{R_2}\right) \le 1$$

where

d = mirror separation

R_1 & R_2 = radius of curvature of mirrors

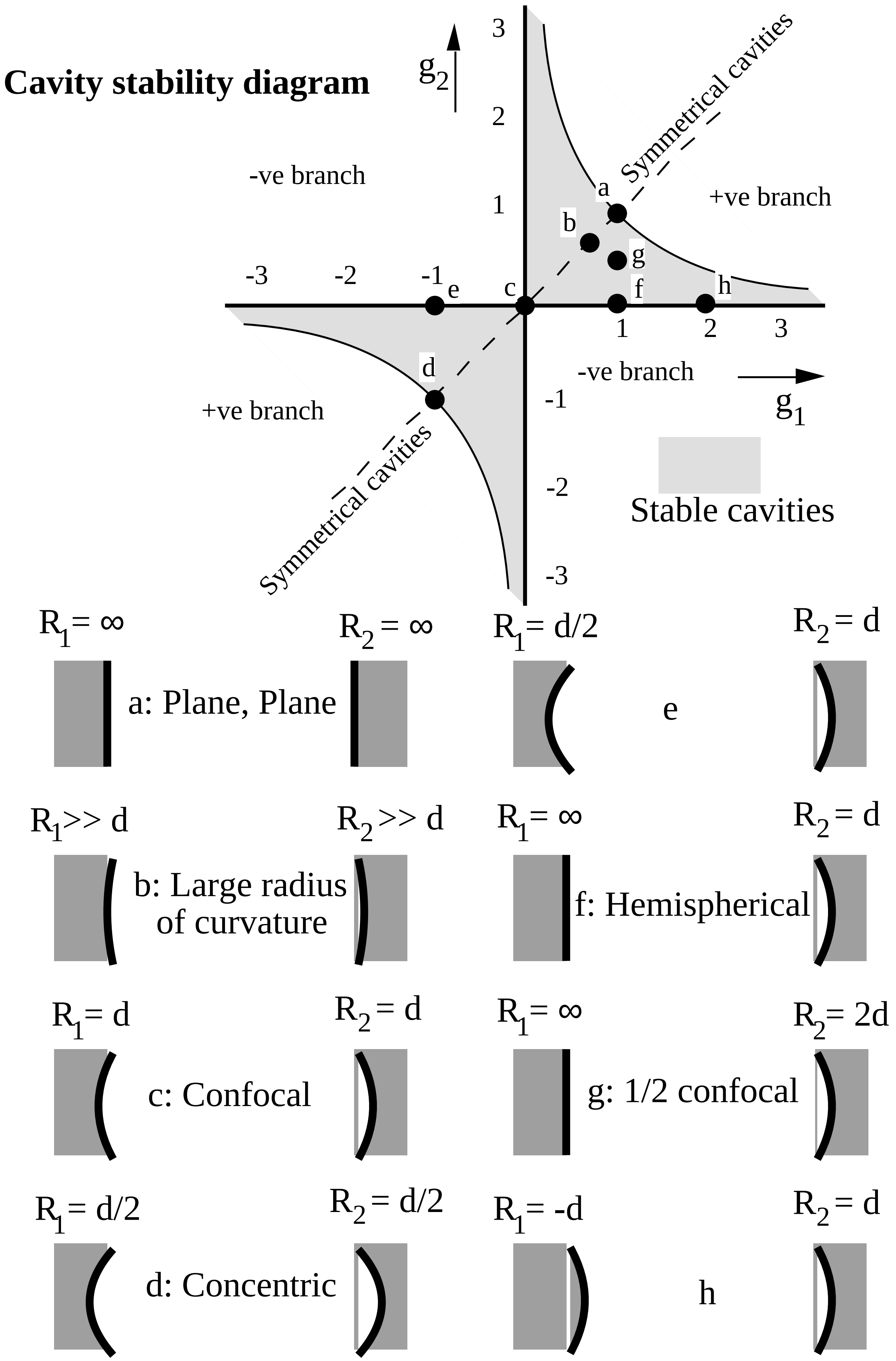
This is often written as:

$$0 \le g_1 g_2 \le 1 \qquad [73]$$

where

$$g_1 = \left(1 - \frac{d}{R_1}\right)$$

$$g_2 = \left(1 - \frac{d}{R_2}\right)$$

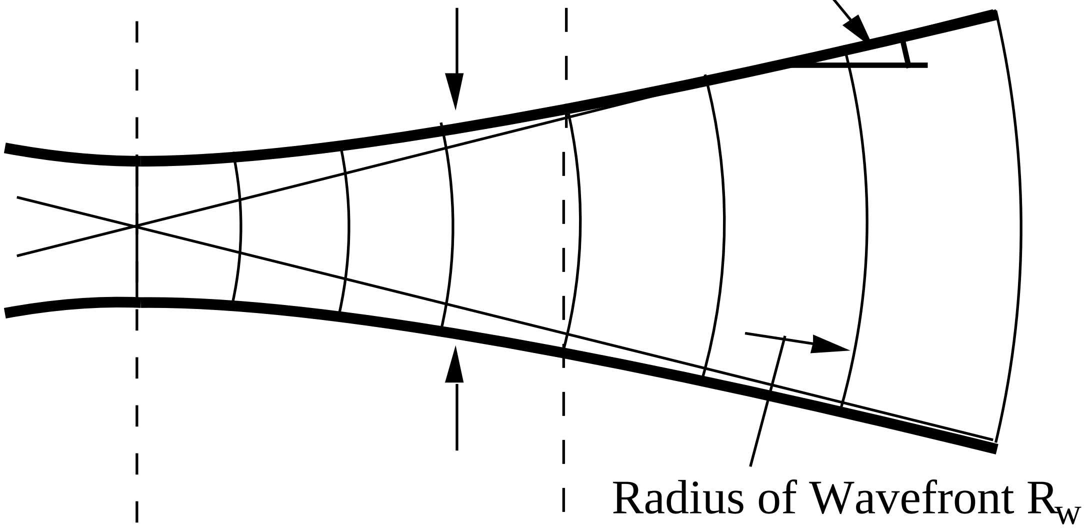

Cavity stability diagram
g_2
g_1
-ve branch
+ve branch
-ve branch
+ve branch
Symmetrical cavities
Symmetrical cavities
a
b
c
d
e
f
g
h
Stable cavities
$R_1 = \infty$
$R_2 = \infty$
a: Plane, Plane
$R_1 = d/2$
$R_2 = d$
e
$R_1 >> d$
$R_2 >> d$
b: Large radius of curvature
$R_1 = \infty$
$R_2 = d$
f: Hemispherical
$R_1 = d$
$R_2 = d$
c: Confocal
$R_1 = \infty$
$R_2 = 2d$
g: 1/2 confocal
$R_1 = d/2$
$R_2 = d/2$
d: Concentric
$R_1 = -d$
$R_2 = d$
h

Stable cavities in terms of Gaussian beams

A wave with a transverse Gaussian intensity propagates in a different fashion to a spherical beam.

Two key parameters are

Beam size & Wavefront curvature



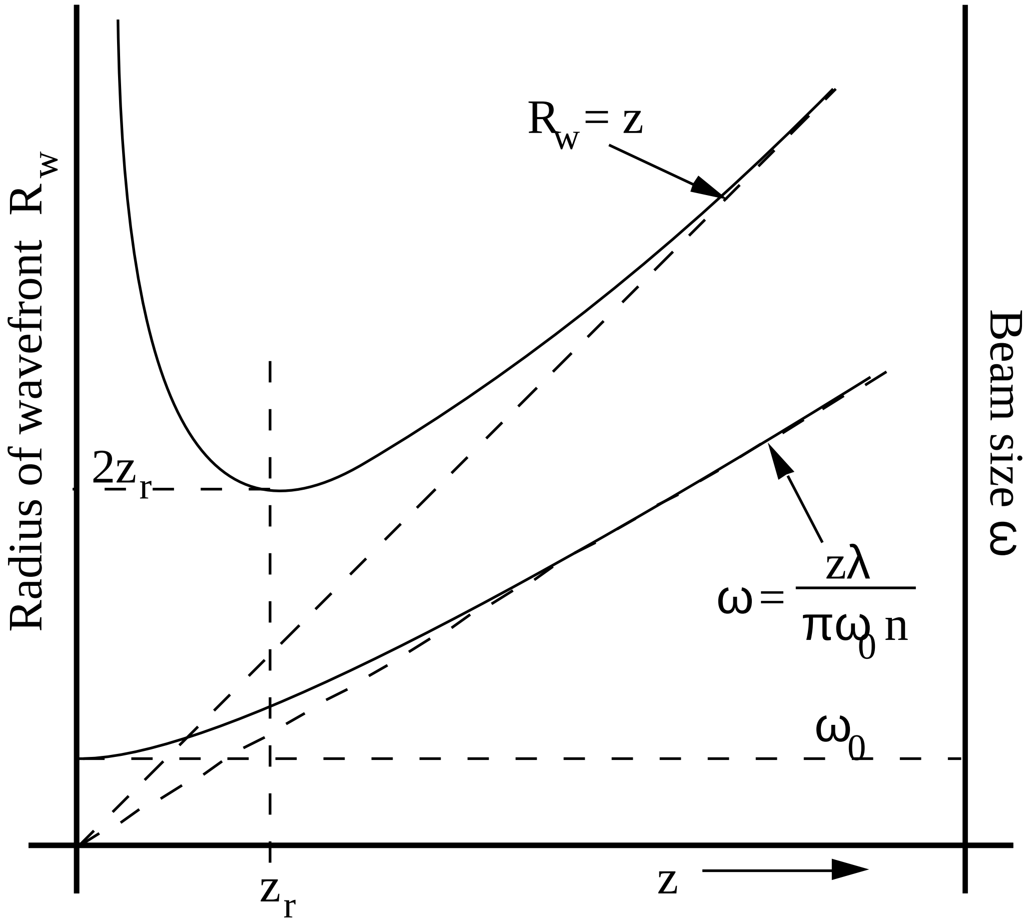
The position z = 0, where $\omega = \omega_{min} = \omega_0$, is called the **Beam Waist.**

$$\omega(z) = \omega_0 \sqrt{1 + \left(\frac{z}{z_r}\right)^2} \qquad [75]$$

$$R_w(z) = \frac{1}{z}\left(z^2 + z_r^2\right) \qquad [76]$$

where

$$z_r = \frac{\pi\,\omega_0^2\, n}{\lambda} \qquad [77]$$



z_r is called the **Rayleigh range**, or the **Beam parameter**

When $z = z_r$, $\omega = \sqrt{2}\ \omega_0$,

Also from differentiating [76] wrt to z we see,

When $z = z_r$, $R_w(z) = R_w(z)_{min} = 2z_r$, i.e.

Wavefront curvature is maximum at the Rayleigh range

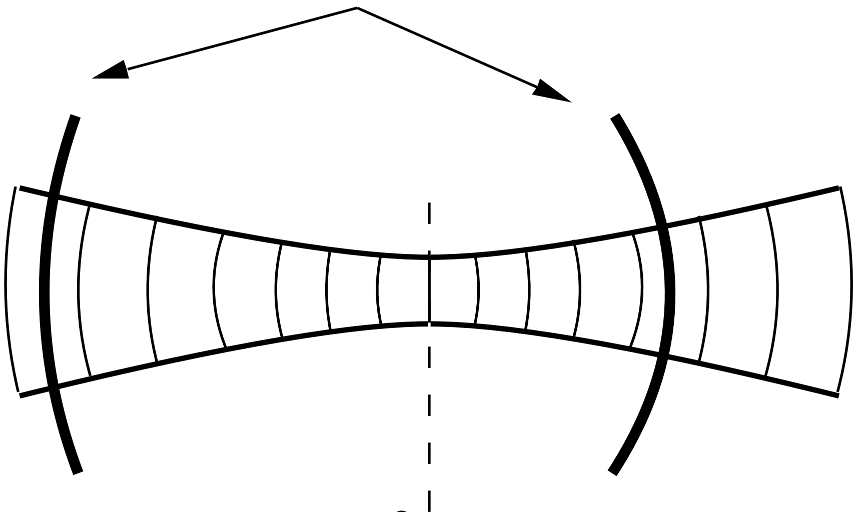
Note also that the far-field diffraction angle θ, is given by

$$\sin\theta \approx \frac{\lambda}{\pi\omega_0 n} \approx \frac{\lambda}{\text{beam dia.}}$$

Given that the beam within a stable resonator is a Gaussian, we need to position the mirrors so that they reflect the Gaussian wavefront without changing the beam parameter.

The simplest way to ensure this is to match the curvature of the mirror to the wavefront.

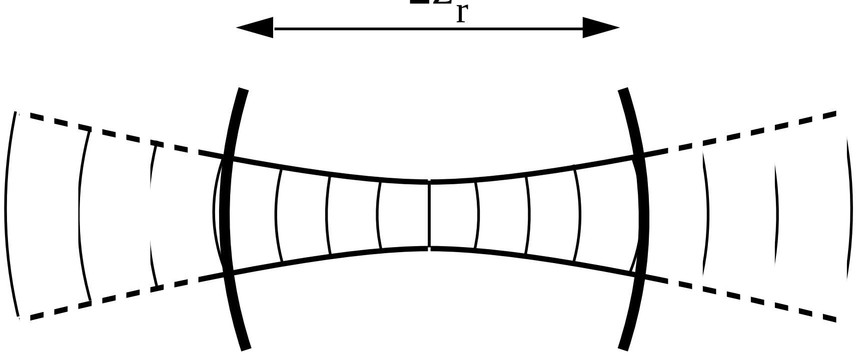
Possible mirror curvatures and positions



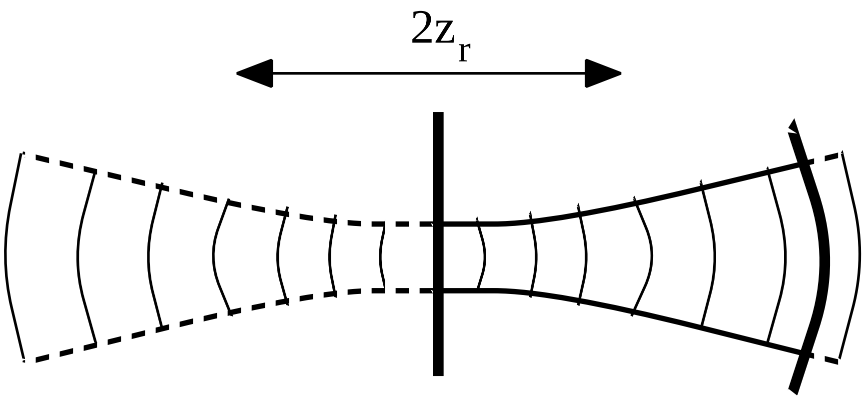
$z = 0, \omega = \omega_0$

Some examples of different cavity configurations

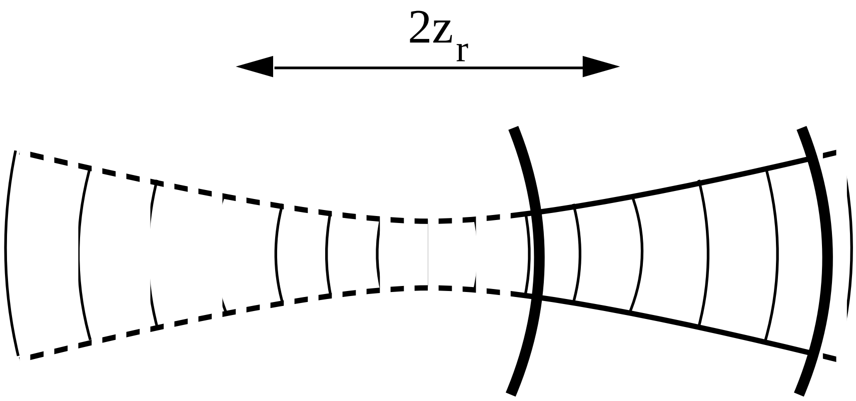
$2z_r$



$R_1 = R_2 = d = 2z_r = R_w(z=z_r)$, i.e. confocal cavity
Note in confocal cavity mirrors are placed at Rayleigh range.



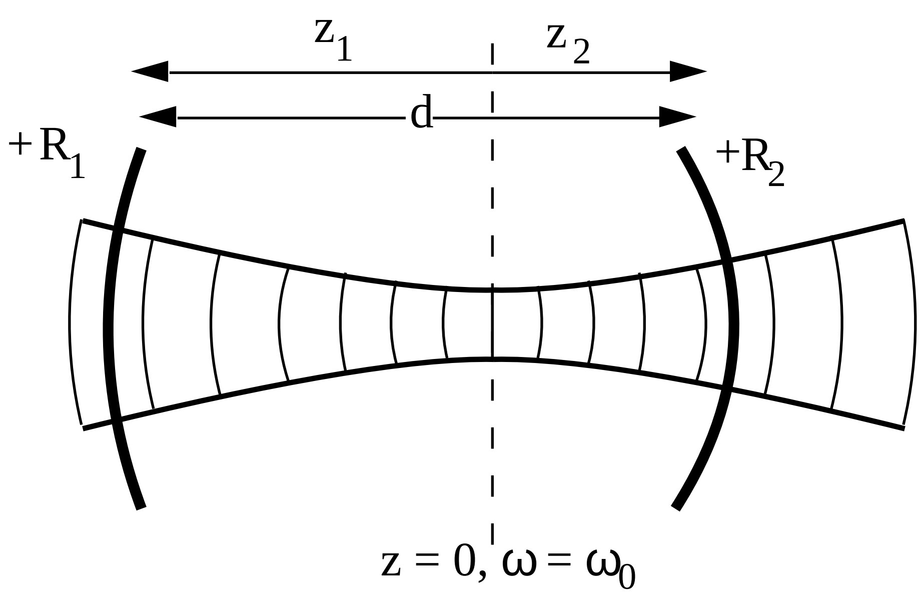
$R_1 = \infty$, $R_2 = d = R_w(z>>z_r)$, i.e. hemispherical cavity.



$R_1 = -d$, $R_2 = d$, i.e. $g_2 = 2$, $g_1 = 0$, see cavity h on stability diagram (note sign convention for R).

Calculation of beam waist size and position in a stable cavity

For a stable cavity, the Radius of curvature of the wavefront must match that of the mirrors.



From [76] we can write (care required to get the right sign!):

$$R_1 = -R_w(z = -z_1) = -\left(-z_1 - \frac{z_r^2}{z_1}\right) = z_1 + \frac{z_r^2}{z_1}$$

$$R_2 = R_w(z = z_2) = z_2 + \frac{z_r^2}{z_2}$$

Rewrite for z_1 and z_2:

$$z_1 = \frac{R_1}{2} \pm \sqrt{\frac{R_1^2 - 4z_r^2}{4}} \qquad [78]$$

$$z_2 = \frac{R_2}{2} \pm \sqrt{\frac{R_2^2 - 4z_r^2}{4}} \qquad [79]$$

From the diagram, we see that $d = z_2 + z_1$, from [78] and [79] we can derive an expression for z_r:

$$z_r^2 = \frac{d(R_1 - d)\,(R_2 - d)\,(R_1 + R_2 - d)}{(R_1 + R_2 - 2d)^2} \qquad [80]$$

where

concave mirrors are defined as +ve curvature
mirror separation = d (always +ve)

In the special case where $R_1 = R_2$ [80] becomes:

$$z_r^2 = \frac{d(2R - d)}{4} \qquad [81]$$

ω_0 is related to z_r by [77]:

$$z_r = \frac{\pi \omega_0^2 n}{\lambda} \qquad [77]$$

i.e.

$$\omega_0 = \sqrt{\frac{z_r \lambda}{\pi n}} \qquad [82]$$

Knowing z_r, the position of the beam waist can be calculated from [78] and [79]:

$$z_1 = \frac{R_1}{2} \pm \sqrt{\frac{R_1^2 - 4z_r^2}{4}} \qquad [78]$$

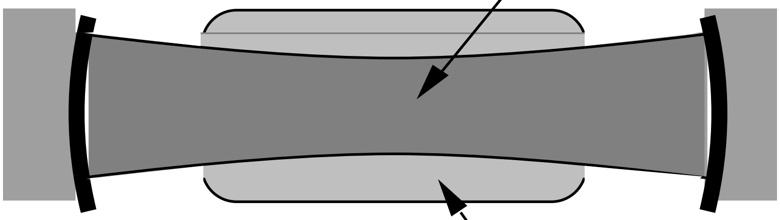
$$z_2 = \frac{R_2}{2} \pm \sqrt{\frac{R_2^2 - 4z_r^2}{4}} \qquad [79]$$

Which cavity to choose?

The best choice of cavity depends on the details of the laser system.

e.g.

High Power - gas discharge laser



To extract the most power from the population inversion we need to match the diameter of the Gaussian mode to that of the gain medium. i.e. the Gaussian mode should have a large **mode volume.**

For a symmetrical cavity [81] and [82] give:

$$\omega_0 = \sqrt{\frac{\lambda}{\pi n}} \sqrt[4]{\frac{d(2R - d)}{4}} \qquad [83]$$

A large ω_0 is obtained in a cavity where R>> d

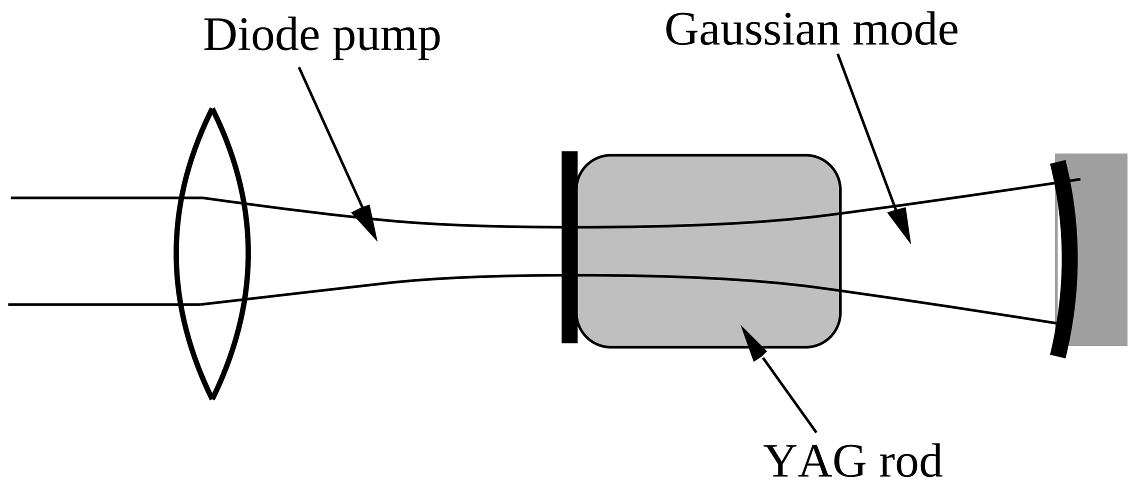
Therefore for maximum power extraction we should use plane:plane (not very stable) or better, a large radius of curvature cavity.

High power - flashlamp pumped solid state (Nd:YAG)

The poor directionality of the flashlamp results in the population inversion being distributed through the entire rod. Again we need to use a Gaussian mode with a large mode volume in order to 'extract' all the useful power.

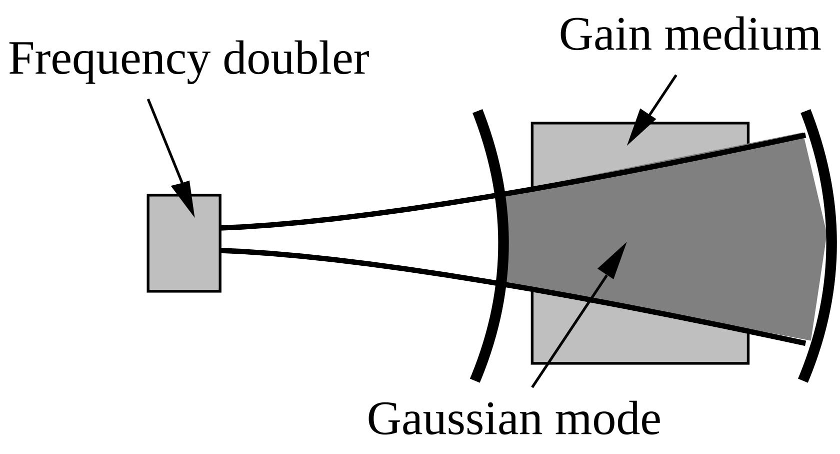
A further complication is that the thermal heating of the rod sets up a radial temperature gradient. This causes the rod to act a lens, which in turn changes the cavity parameter. We need to select the mirrors to compensate for the 'rod lens'

Diode Pumped solid state laser (Nd:YAG)



For maximum efficiency, we need to match the Gaussian mode of the diode pump beam to the Gaussian mode of the cavity. This is called **mode matching.**

High power pulsed solid state laser with external doubler



In pulsed systems, the intra-cavity power density can often reach a level where it will damage the material and coatings within the cavity, i.e. mirrors or laser rod.

This is most likely to happen near the beam waist where the power density is highest.

We can design a stable cavity with the waist external to the mirrors (see cavity h on stability diagram), which reduces the intra-cavity power density.

This design also focuses the beam external to the cavity without the need of an additional lens! Ideal for use with an external doubling crystal.

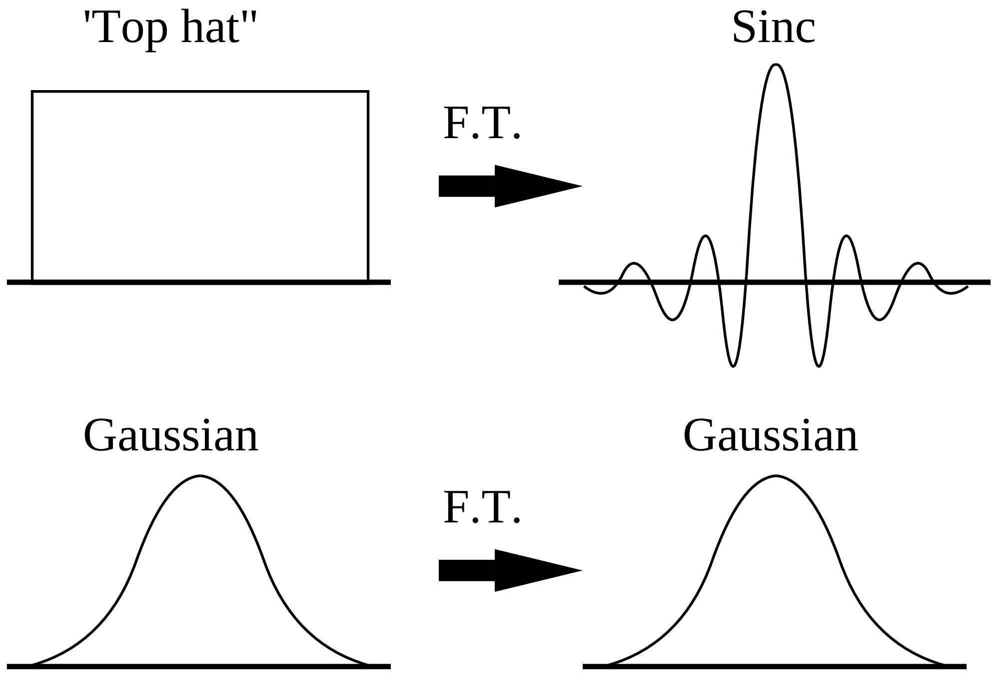
Why Gaussian?

Why do the transverse modes have a Gaussian profile?

In the steady state condition, light is reflected up and down the cavity many times. The electric field distribution has to be one that can "reproduce" itself upon reflection/diffraction from the mirrors.

The far-field diffraction pattern is the Fourier transform of the "object".

For a self sustaining solution we need a field distribution that is unchanged by a Fourier Transform:



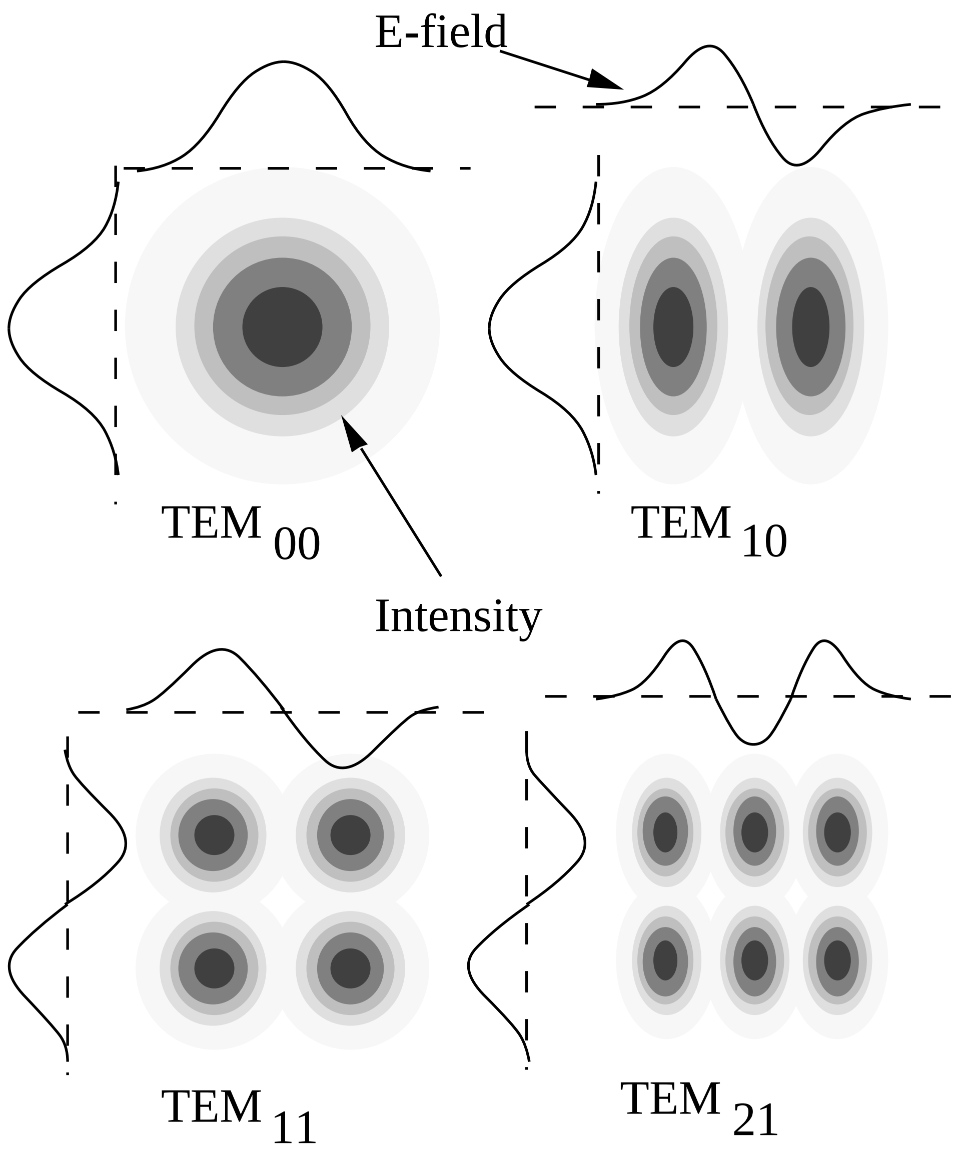
One function which satisfies this condition is a Gaussian.

Therefore, a transverse Gaussian field distribution is one which can be sustained within a laser cavity!

Higher order Hermite-Gaussian and Laguerre-Gaussian modes are also 'unchanged' by Fourier Transform and therefore form other possible solutions.

Description of transverse modes

The transverse mode determines the beam shape.
All the allowed modes can be described by various Hermite-Gaussian polynomials

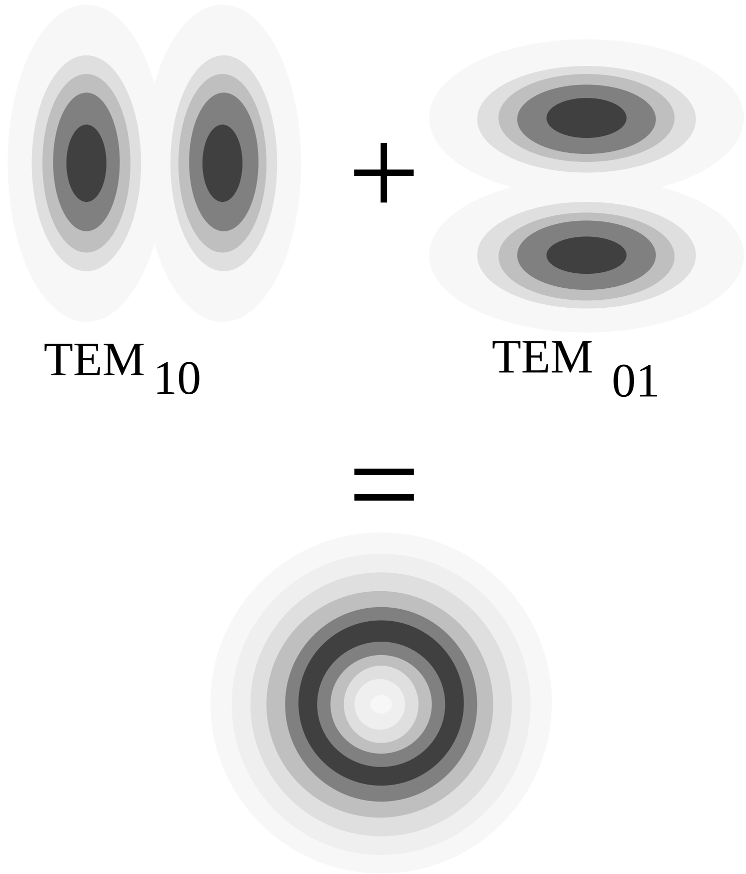


General form for transverse electric modes:

TEMpq

p - no. of nodes in x direction
q - no. of nodes in y direction

A frequently observed (and unwanted) transverse mode is the 'doughnut mode', a linear superposition of TEM_{01} and TEM_{10}



Selection of Transverse Modes

The preferred mode that oscillates within the cavity depends on the aperture of the gain medium and the radial dependence of the gain.

For most applications the TEM_{00} is the preferred mode.

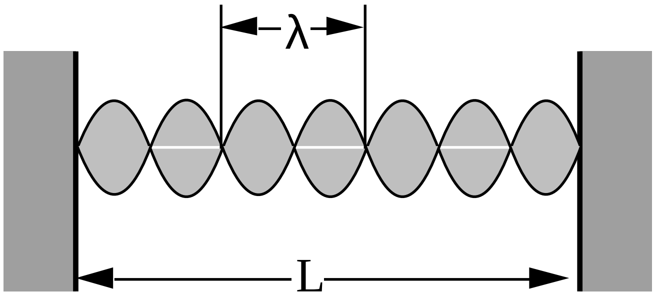
A **single transverse mode laser** is restricted to give TEM_{00} output.

For example, an intra-cavity aperture can be used to suppress the gain for the high order modes.

Longitudinal Modes

So far we have assumed that light of any frequency can 'bounce' up and down the length of the resonator, again we are wrong!

The reflections back and forth result in a standing wave field to be set up within the resonator (or cavity). As with other standing wave problems (e.g. violin strings, organ pipes) the wavelength of the standing wave has to 'fit' exactly within the cavity.



The condition is:

$$m\frac{\lambda}{2} = L \qquad [85]$$

where

m = an integer (for typical L and λ, m is very large!)

$\lambda = \lambda_0 / n$

In frequency terms:

$$\nu_{allowed} = \frac{mc}{2L} \qquad [86]$$

Therefore:

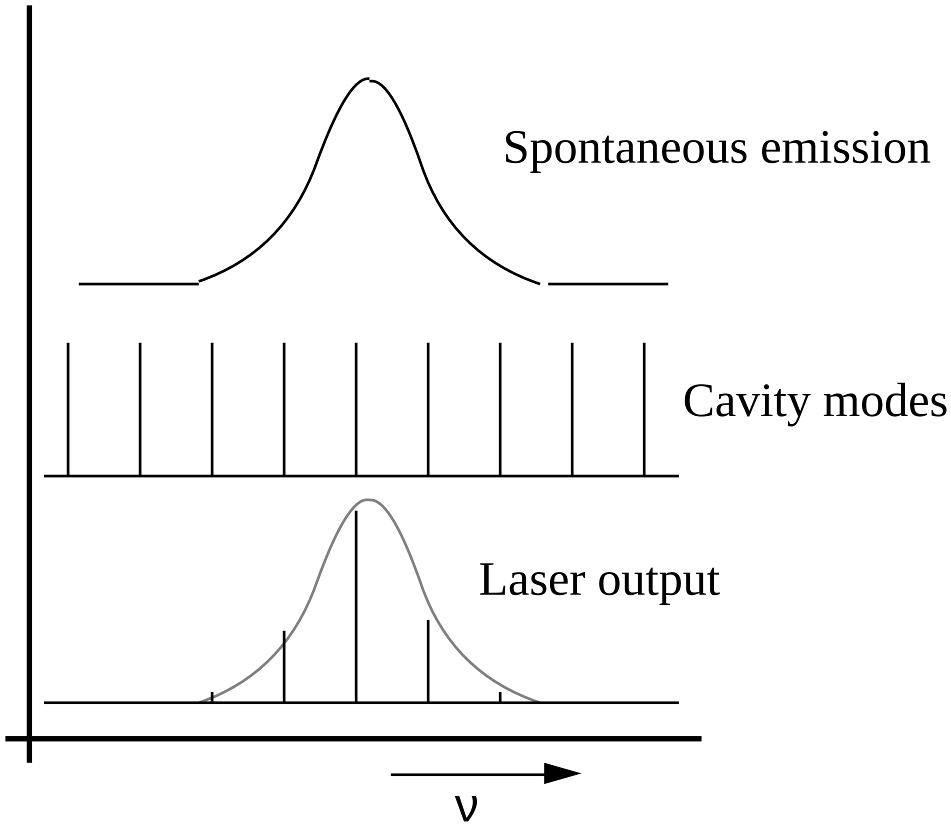
$$\Delta\nu_{m,\,m+1} = \frac{c}{2L}$$, the **intermode spacing** [87]

This is the frequency spacing between adjacent longitudinal modes of a cavity, called the

Free Spectral Range (FSR)

Laser Output Frequency

The homogeneous or inhomogeneous broadening means the gain medium will give optical gain over a continuous range of frequencies. However, the resonator will only provide feedback at the cavity mode frequencies. The output spectrum will be a combination of the two.



The output will be at one or more of the specific frequencies dictated by the cavity modes within the gain profile of the laser.

The number of modes depends on the broadening mechanism (homogeneous or inhomogeneous) and the exact cavity configuration.

A laser in which only one longitudinal mode oscillates is called a **single longitudinal mode (SLM)** laser.

Gain Saturation within the linewidth of the gain media

Under CW conditions, the circulating power within the cavity and associated stimulated emission, balances the pumping rate. The population inversion is maintained at the threshold inversion, corresponding to a gain of unity.

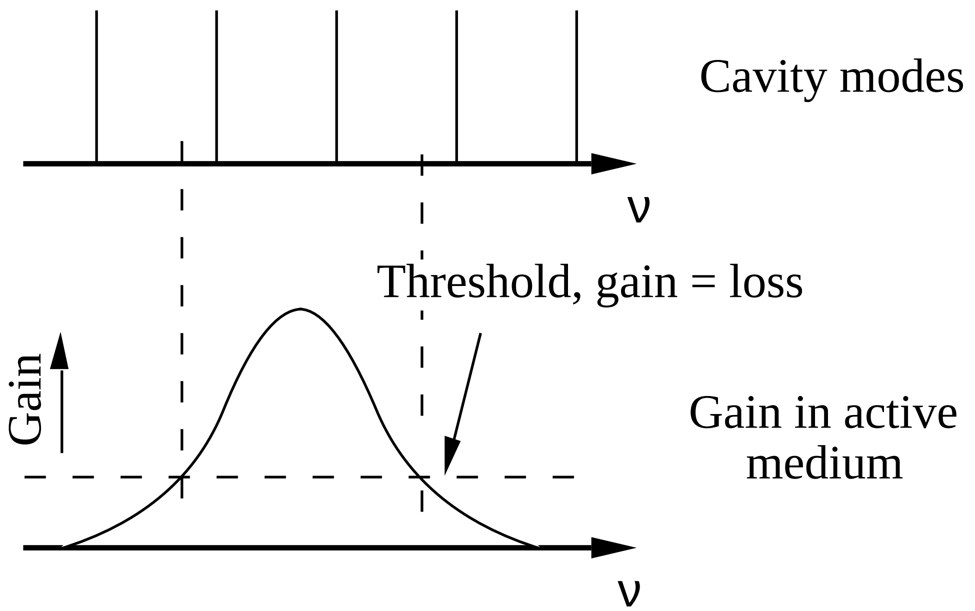
Under CW conditions the round trip gain is unity

The CW inversion equals the threshold inversion

We have also seen that the exact frequency of the laser output is determined by the longitudinal such that:

$$\mathbf{m}\frac{\lambda}{\mathbf{2}} = \mathbf{L} \qquad [85]$$

Let us consider the gain as a function of frequency, and how it interacts with the cavity modes.



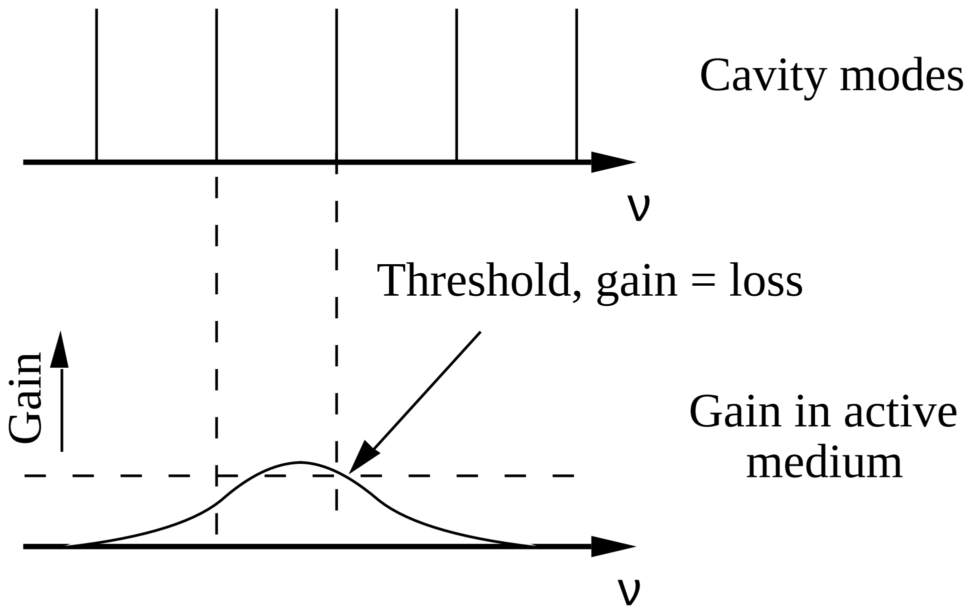
In principle, net gain is available at frequencies corresponding to two different longitudinal modes.

As the radiation density ρ_ν builds up inside the cavity the stimulated emission reduces the population inversion until the gain equals the loss.

Homogeneously broadened lasers

In a **homogeneously** broadened laser transition, **all** the **atoms** contribute to the gain at **all** the **frequencies.** As the population inversion is reduced, the gain is reduced at all frequencies.

The laser will oscillate at the frequency corresponding to the longitudinal mode closest to the gain maximum (i.e. closest to line centre)

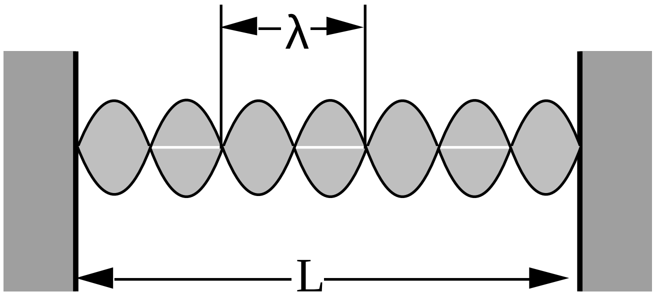


Once the steady state is reached, only one longitudinal mode will oscillate. No other modes have sufficient gain to reach threshold.

Homogeneously broadened Lasers oscillate with a single longitudinal mode

Spatial Hole Burning

In a standing wave cavity, the standing wave gives rise to a local position where the electric field is zero, i.e. a node



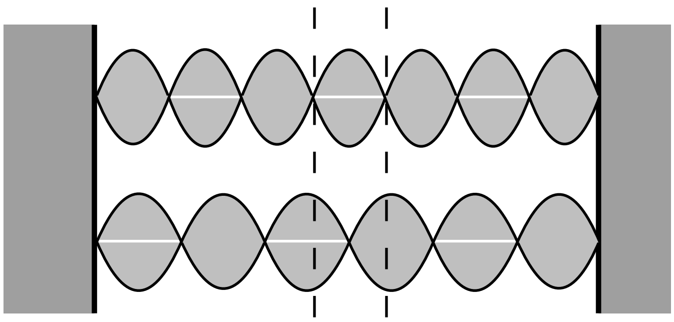
In the region of the node the ρ_ν is zero and therefore there is no stimulated emission and the local population inversion does not contribute to the gain of the laser.

This effect is called **spatial hole burning**

Spatial hole burning creates two problems
1) Some of the potential laser gain is wasted
2) It 'encourages' multiple longitudinal modes to oscillate

Previously we said that homogeneously broadened lasers only oscillated on a single longitudinal mode (SLM).

However, spatial hole burning means the population inversion left 'untouched' by the principal mode can be accessed by the second mode, since its electric field nodes fall in a different place.

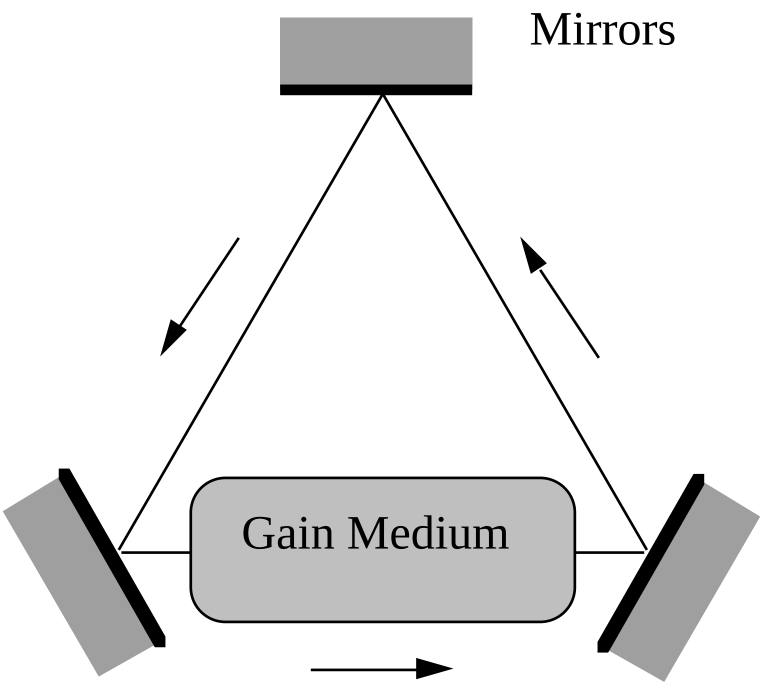


The anti-nodes of the second longitudinal mode fall at the nodes of the first mode. Therefore, even in a homogeneously broadened laser, both modes 'see' gain and can oscillate.

Ring cavities

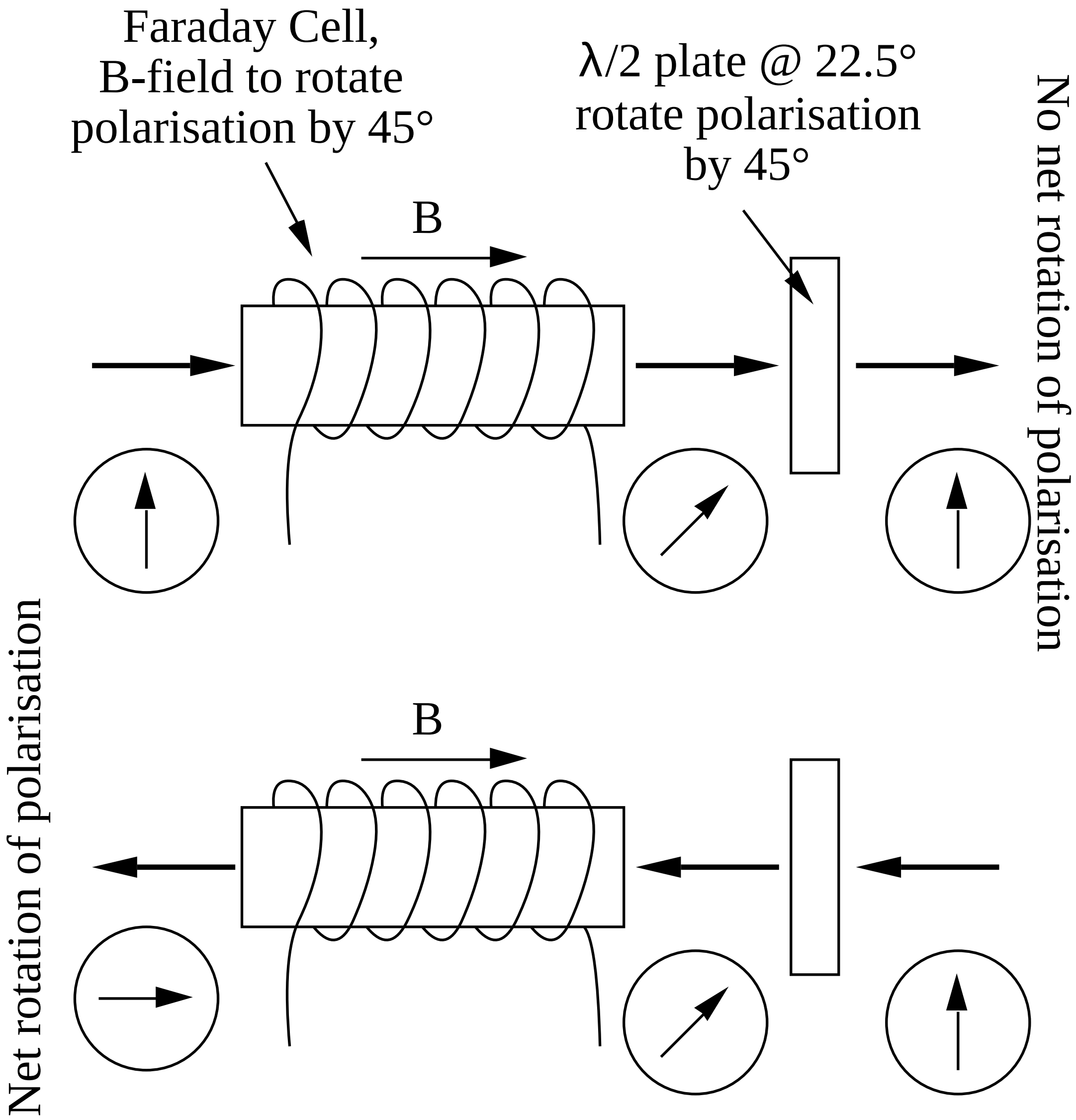
The problem of spatial hole burning stems from having a standing wave in the cavity.

An alternative to a standing wave cavity is a travelling wave ring cavity.

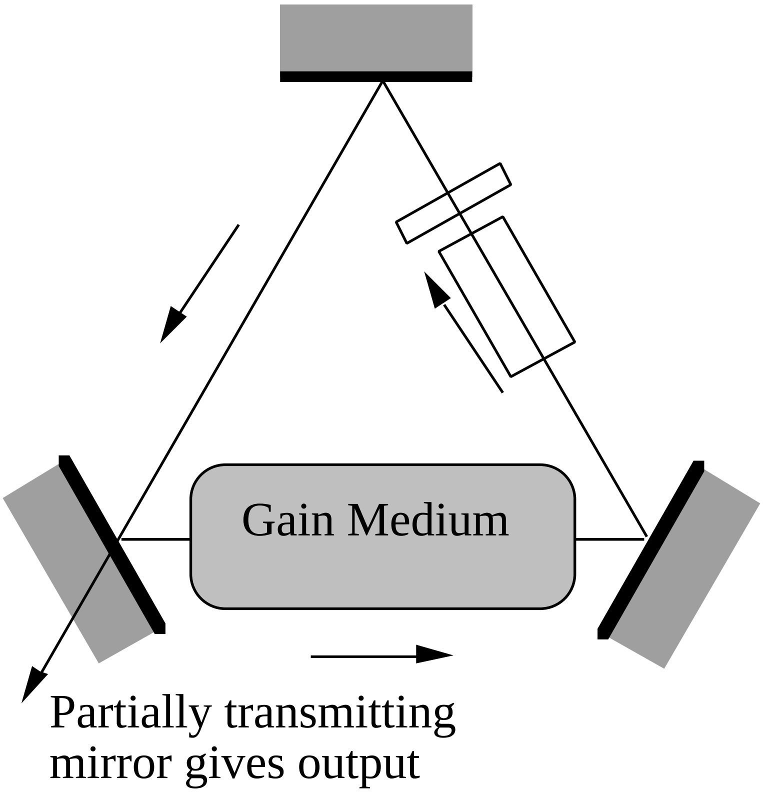


In order to get a travelling wave we need to restrict light to travelling one way round the ring

It is usual to use a **Faraday rotator** to stop light travelling the wrong way round the ring



The net rotation in polarisation gives high loss at a Brewster surface or polariser.



Light now passes only one way through the gain medium and there is no standing wave.

Ring cavities still have longitudinal modes. The electric field must 'repeat' itself after one round trip, i.e.

$$m\lambda = RT$$

where

m = an integer

RT = round trip length of cavity

$\lambda = \lambda_0 / n$

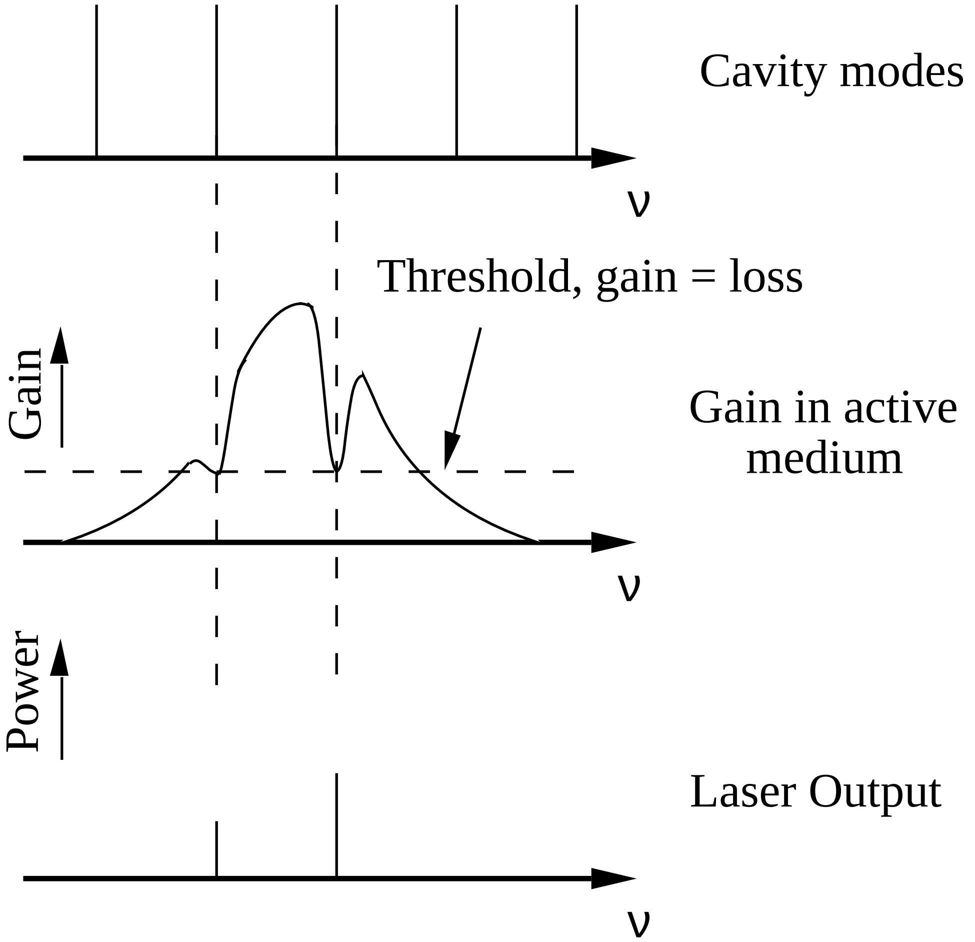
The free spectral range, intermode spacing is given by:

$$FSR = \frac{c}{RT} \qquad [88]$$

Inhomogeneously broadened lasers

In an **inhomogeneously** broadened laser transition, gain at **different frequencies** is provided by **different atoms.** As the gain at one frequency is reduced, the gain at other frequencies is unchanged.

The laser will oscillate all frequencies corresponding to the longitudinal modes where the gain exceeds the loss.



Once the steady state is reached, a number of longitudinal modes may have sufficient gain to oscillate.

Inhomogeneously broadened Lasers can oscillate on a number of longitudinal modes simultaneously.

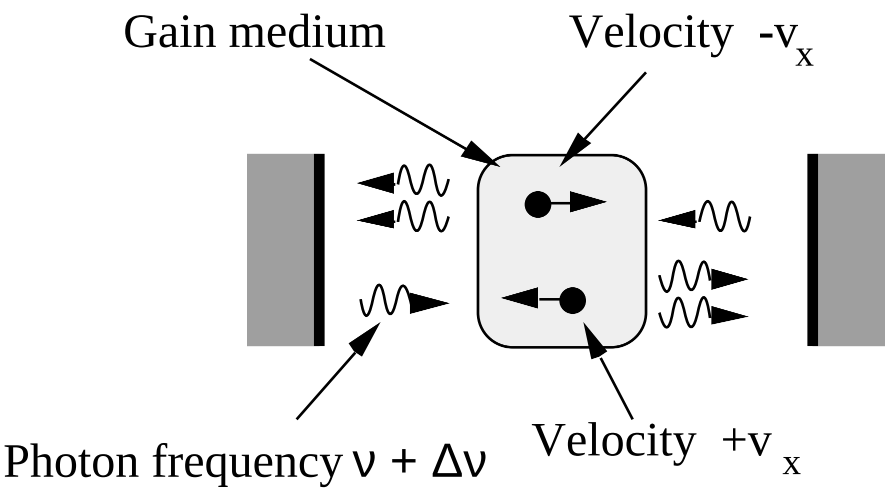
Frequency Hole Burning and Bennet Holes

The act of creating frequency specific reduction in the gain is called **hole burning**. Each of the holes is called a **Bennet hole**. Bennet holes are **only** formed within an **inhomogeneously** broadened system.

Homogeneous broadening within the inhomogeneous linewidth reduces the gain for frequencies near that of the oscillating longitudinal mode. The **width** and **shape** of the **Bennet hole** is the **homogeneous** linewidth/shape.

For Doppler broadened systems, the hole burning is complicated by the fact that the laser light passes both ways through the gain medium.

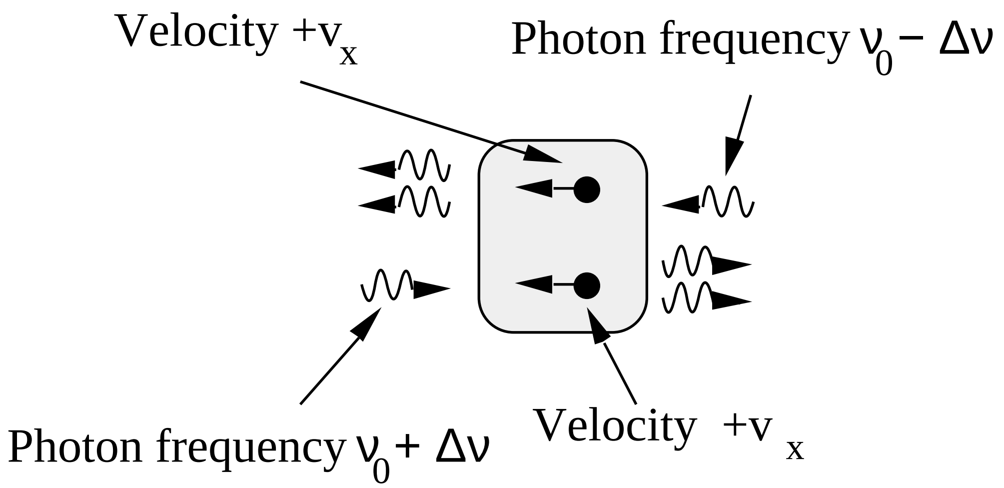
Light detuned by $+\Delta\nu$ from line-centre will interact with atom/molecules with a velocity v_x on the first pass through the medium and $-v_x$ on the return



Atoms Doppler shifted into resonance

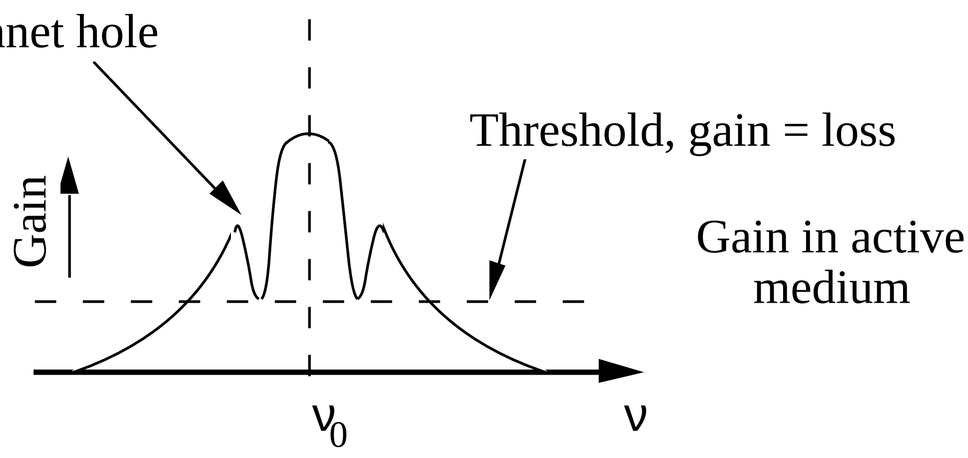
Alternatively, atoms with a velocity $+v_x$ can give gain at frequencies $\nu_0 \pm\Delta\nu$

For a Doppler broadened system, atoms/molecules with a particular velocity give gain at two different frequencies (one Doppler shifted up and the other Doppler shifted down).



Atoms Doppler shifted into resonance $\Delta\nu \approx \nu_0 \frac{v_x}{c}$

Therefore, laser oscillation at a frequency corresponding to a single longitudinal mode will burn **two Bennet holes,** symmetrically about line centre.



Tuning the output frequency of the laser

The longitudinal modes set the allowed output frequencies.

$$\nu_{allowed} = \frac{mc}{2L} \qquad [86]$$

One way we can fine tune the output frequency of the laser is to change the cavity length. Differentiate [86] wrt L:

$$\frac{d\nu}{dL} = -\frac{mc}{2L^2} = \frac{-c}{2L}\frac{m}{L}$$

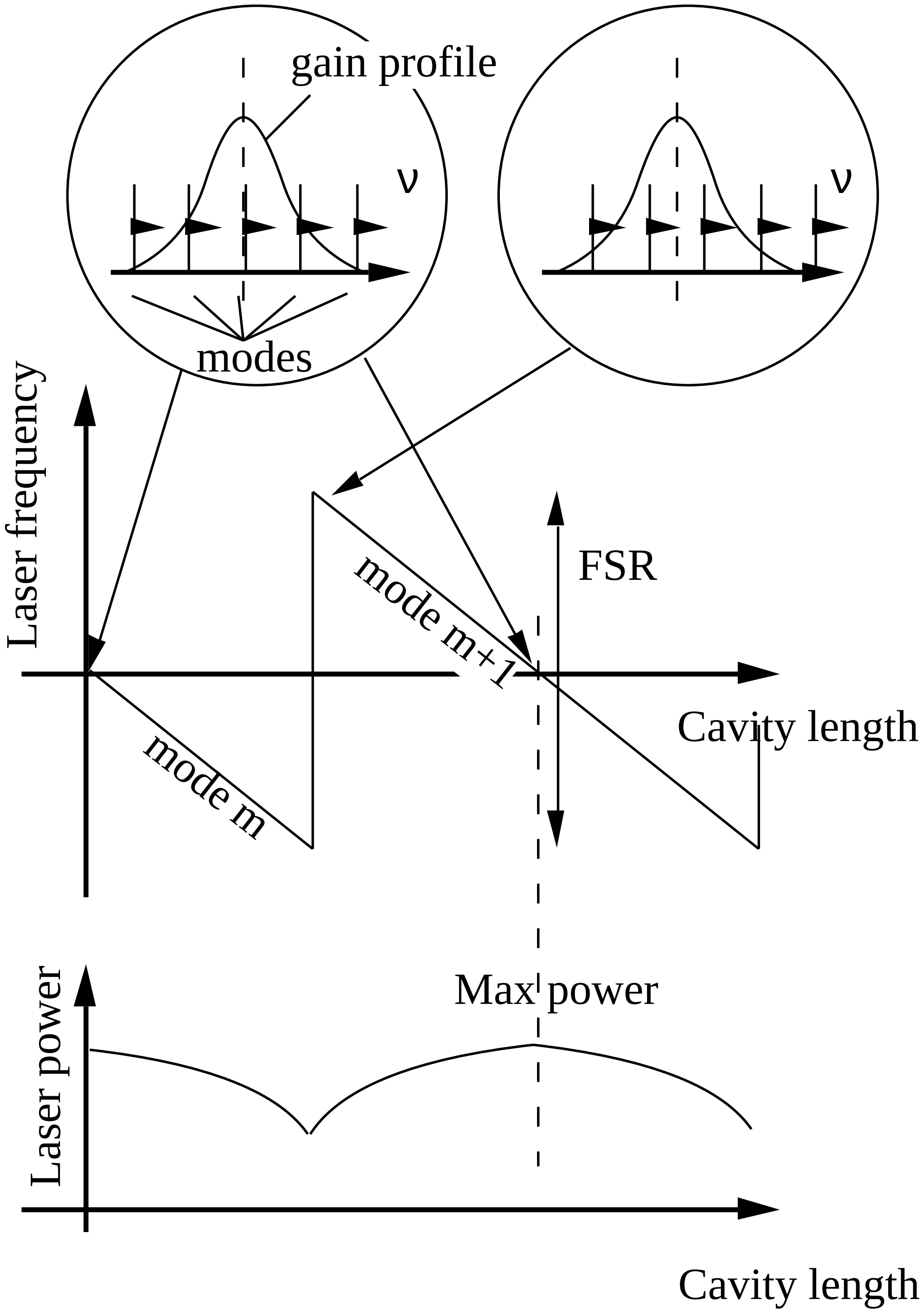
But $\frac{m}{L} = \frac{2}{\lambda}$, therefore:

$$\Delta\nu = \Delta L \frac{c}{2L}\frac{2}{\lambda} = -\Delta L.\ FSR.\ \frac{2}{\lambda} \qquad [89]$$

Small changes in L can be achieved by mounting one of the cavity mirrors on a piezo-electric transducer.

Obviously, the tuning range must lie within the gain linewidth of the laser transition.

Consider a homogeneously broadened laser, where only one longitudinal mode will oscillate at once. The mode that oscillates is the one closest to line centre (ν_0).



Initially the laser oscillates on the longitudinal mode nearest line centre.

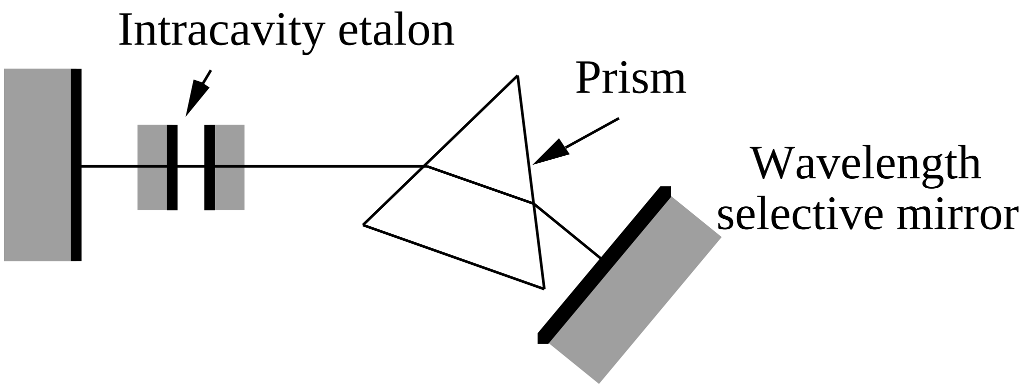
As the cavity length is increased, the frequency decreases.

Eventually, the neighbouring longitudinal mode is closer to line centre and the laser output hops to the new mode. A **mode hop** has occurred.

The tuning range of a simple cavity is limited to:

$$\Delta\nu_{max} = \pm \frac{FSR}{2}$$

To increase the tuning range and suppress unwanted laser transitions we need to introduce additional tuning elements.



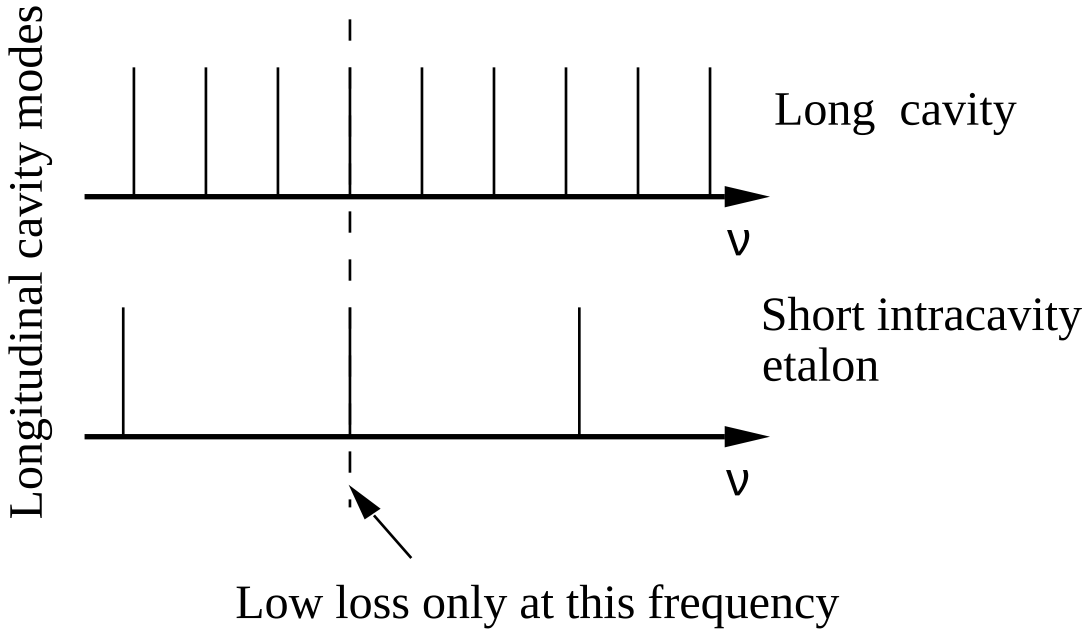
Prisms (e.g. selects 488nm or 514nm lines in an Argon Ion laser)

Mirrors (e.g. selects 633nm, 1.15μm or 3.39μm transitions in a helium neon laser)

Gratings (e.g. selects lines in a carbon dioxide laser)

Etalon (e.g. selects longitudinal modes in a dye laser)

By using an intra-cavity etalon, adjacent longitudinal modes can be suppressed, since low loss will only occur for frequencies that are longitudinal modes of both the main cavity **and** the etalon. In this way the tuning range of a laser can be extended beyond the FSR of the cavity.

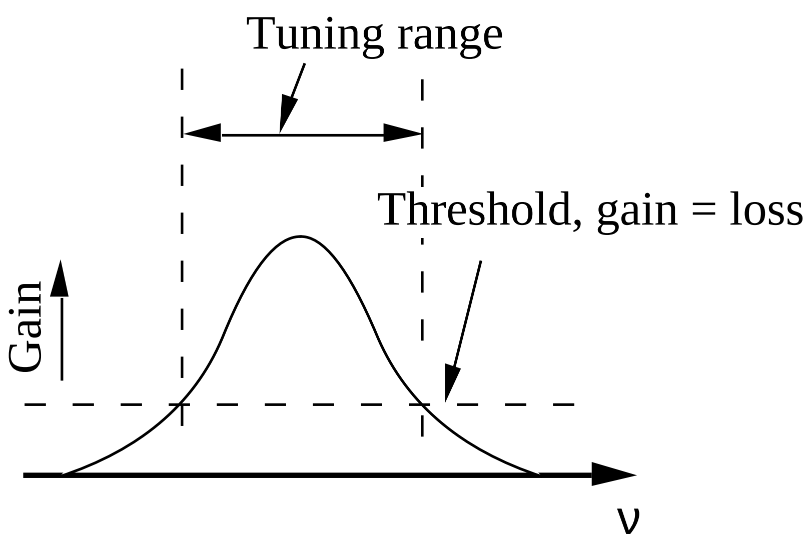


For continuous tuning, the length of the intra-cavity etalon has to be servo controlled to remain 'instep' with the main cavity.

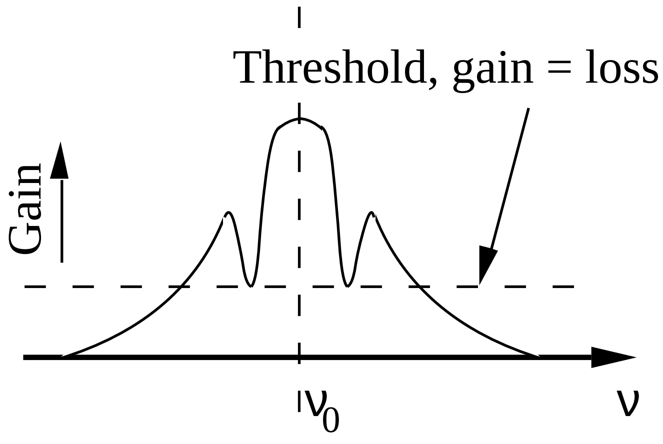
Ultimately, the tuning bandwidth is limited by the linewidth of the laser transition.

Bennet Holes and the Lamb Dip

What happens when we fine tune a Doppler broadened laser over its possible output frequencies?



As we saw early, two Bennet holes will be burnt into the population inversion, reducing the overall gain to unity.

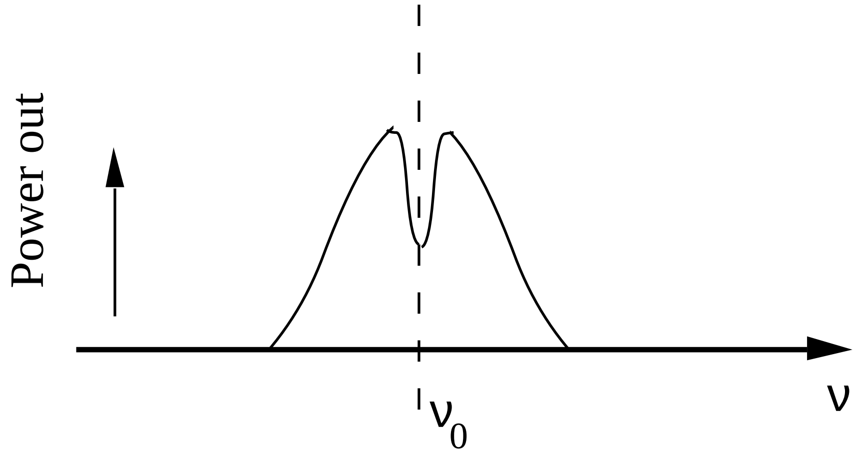


What is the power output as a function of frequency?

In general, when detuned from line centre, the two Bennet holes do not overlap. The forward and return light within the cavity interacts with different atoms ($+v_x$ and $-v_x$).

As the frequency approaches line-centre, a larger number of individual atoms/molecules can contribute to the gain and the output power increases.

However, a special case exists when the laser frequency is tuned exactly to line centre. The two Bennet holes overlap, both the forward and return light interacts with the same atoms/molecules in the cavity. In the extreme, this halves the number of atoms/molecules that can contribute to the gain at line-centre.



Laser output frequency (cavity length)

The reduction in power output at line-centre is called a **Lamb dip**.

A **Lamb dip** is **only** observed in **inhomogeneously** broadened systems where **symmetrical Bennet holes** are 'burnt' in the gain. As with the Bennet hole, the width of the Lamb dip is the **homogeneous linewidth**

The location of the Lamb dip can be used to stabilise the frequency output of the laser.

Frequency Stabilisation

A number of applications require a stable frequency output from the laser (e.g. distance measurement).

We will consider how to stabilise a He-Ne laser.

Output Frequency: 211 THz (632nm)

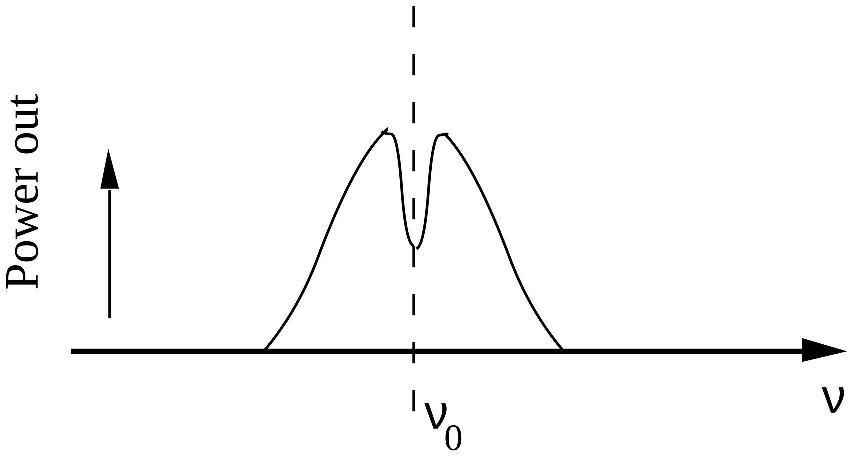
Doppler Broadened Linewidth: few GHz

Typically a Laboratory He-Ne (length 150mm) will oscillate on 2-3 longitudinal modes simultaneously.

Under normal operation, thermal expansion of the laser cavity will lead to a gradual tuning of the longitudinal modes through the gain profile, with periodic mode hops when the cavity has expanded by $\lambda/2$.

Lamb dip stabilised He-Ne

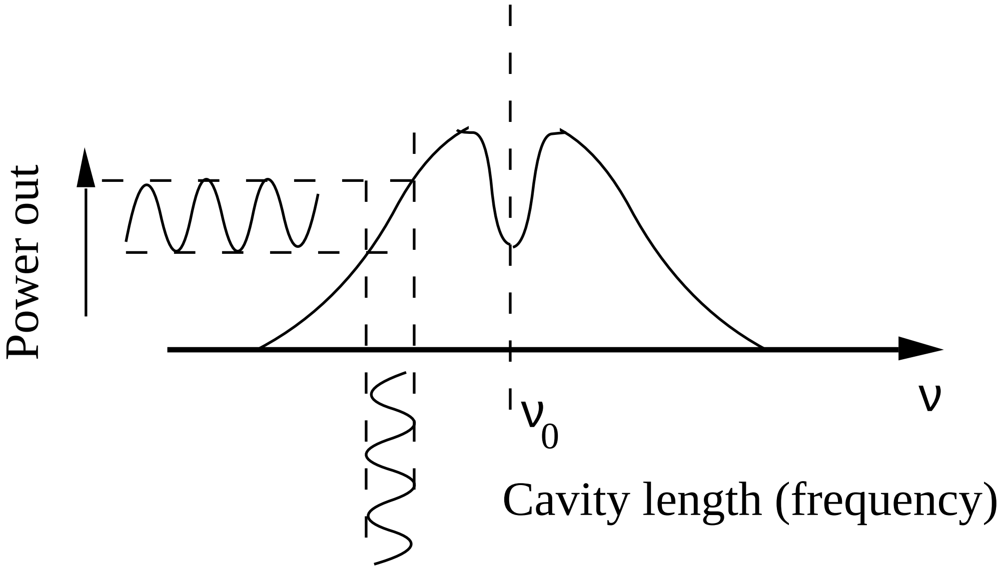
The position of the Lamb dip can be used to stabilise the frequency output of the laser.



Laser output frequency (cavity length)

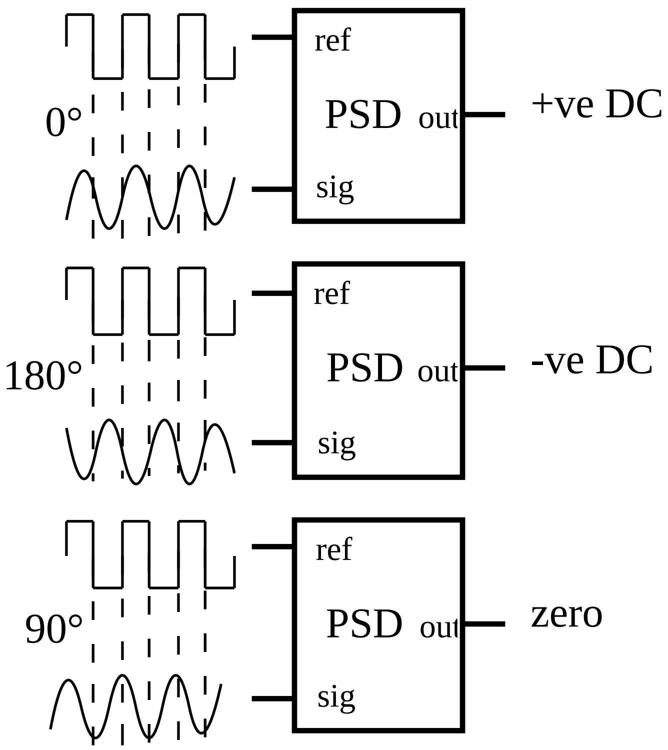
In many cases the Lamb dip is quite shallow and a sensitive method of detection is required.

The length of the cavity is modulated using a PZT mounted mirror. A corresponding modulation is observed on the power output of the laser.



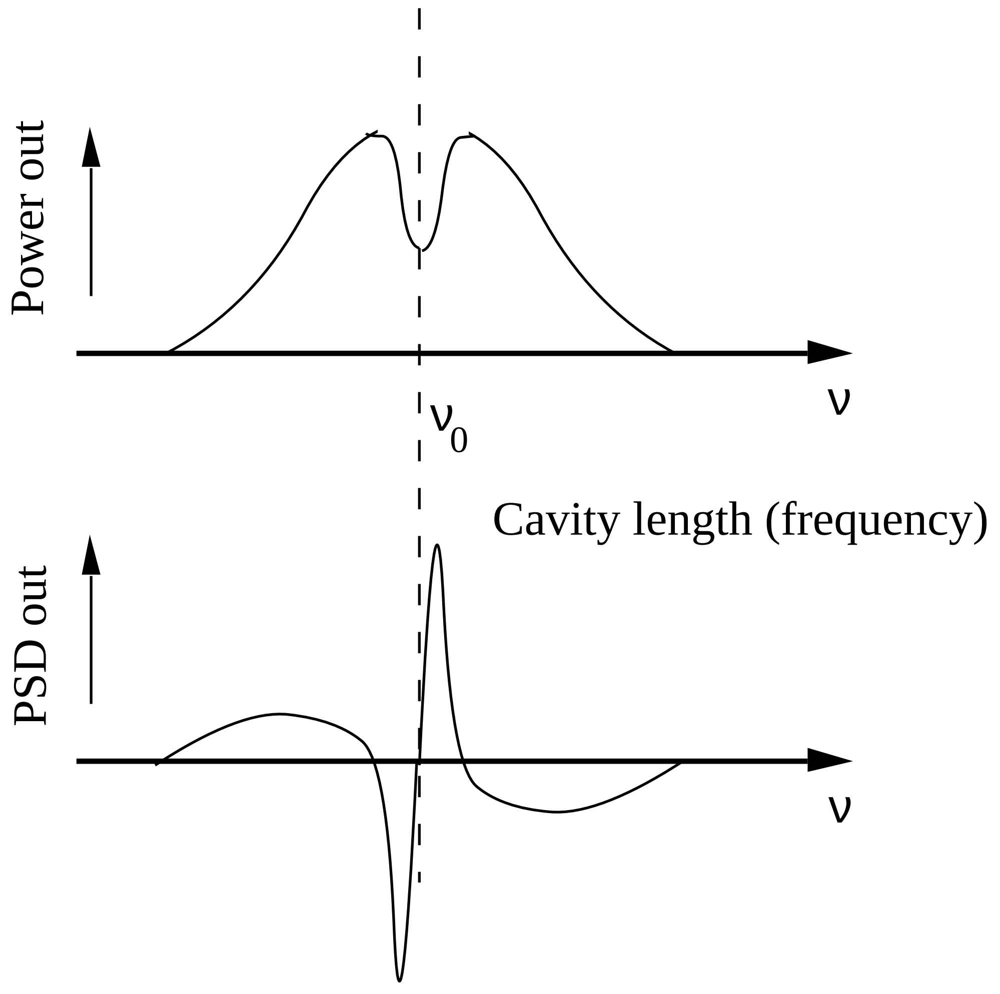
A phase sensitive detector (PSD, or sometimes called a lock-in-amplifier) is used to monitor the oscillation in power output.

The output from a PSD gives the size of the component of the input signal which is in-phase with a well defined reference frequency.



Using the mirror modulation as the reference frequency, the output of the PSD is proportional to the gradient of the power curve.

If the modulation is small, the output of the PSD is the first order differential of the power output.



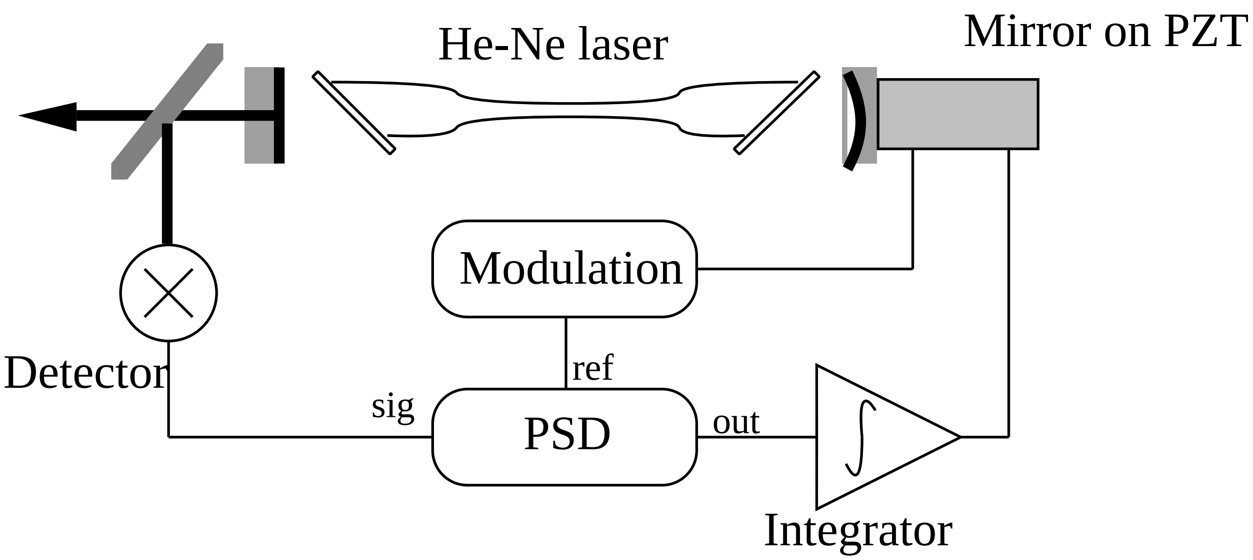
The PSD output is an 'error' signal 'telling' the laser its output frequency is above or below line-centre.

e.g. If laser frequency is too low, the PSD output is -ve
If laser frequency is too high, the PSD output is +ve
If laser frequency is correct, the PSD output is zero

The error voltage is an instruction to the cavity,

"cavity you need to get shorter"

By feeding the **error voltage** into an integrator we can derive a **control voltage** which will adjust the cavity length to keep the output frequency of the laser at line-centre.

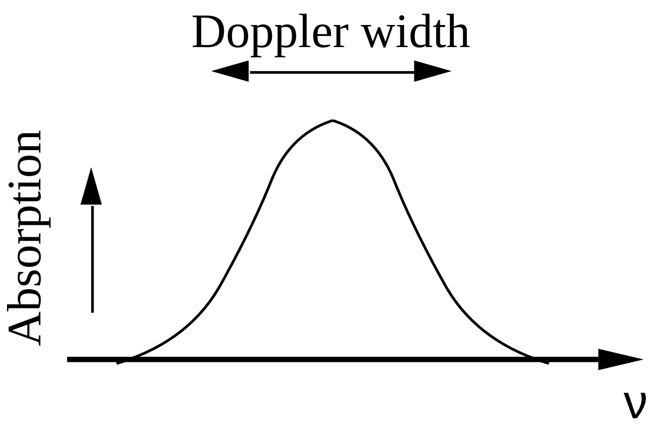


Iodine stabilised He-Ne

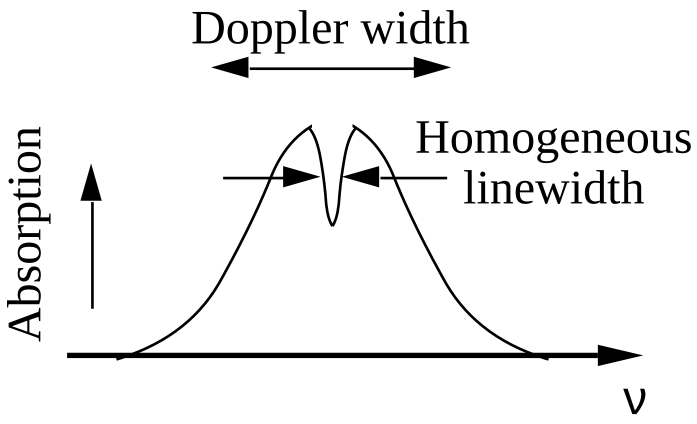
As we have discussed the width of the Bennet hole or Lamb dip is the homogeneous linewidth

In a He-Ne laser, the homogeneous linewidth is quite large due to 'pressure broadening' within the plasma.

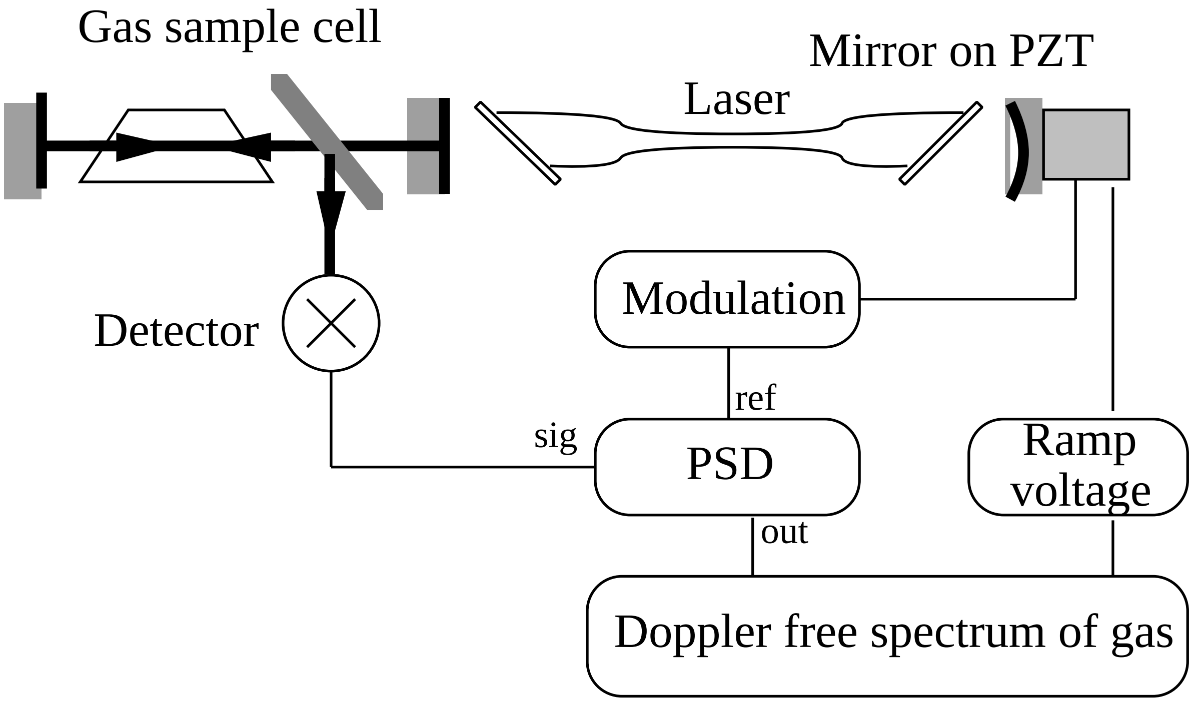
The formation of Bennet holes is not restricted to laser media. Any gas with a Doppler broadened absorption will exhibit a similar effect. A single frequency laser can be used to measure the absorption of the gas as a function of frequency.



If the laser is back-reflected through the gas, the overlapping Bennet holes will give a reduction in the measured absorption at line centre.



This is the basis of one form of **Doppler Free Spectroscopy.** It is called **Saturation Spectroscopy**

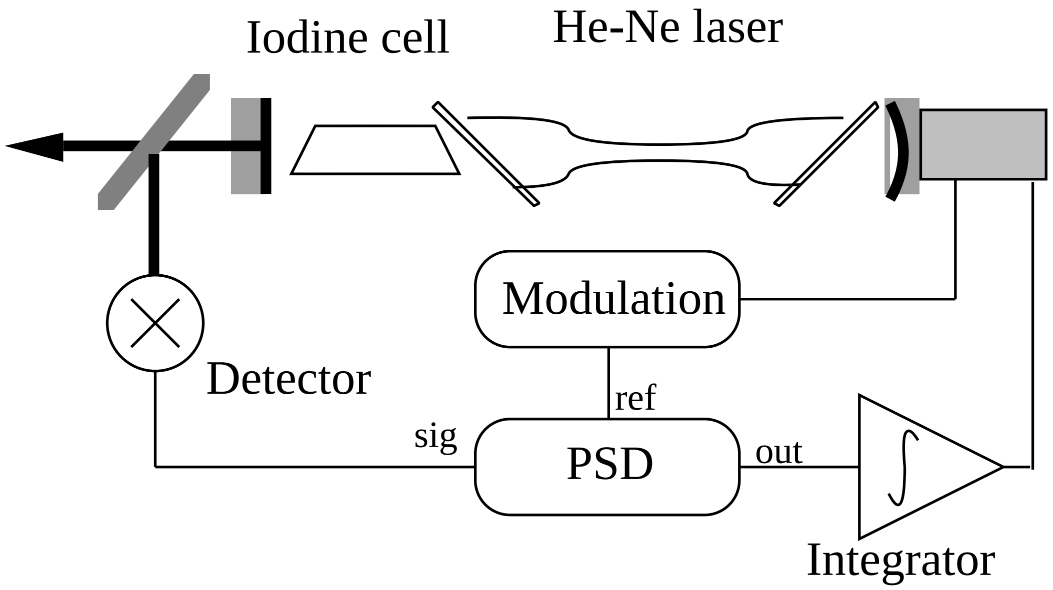


In this way, we can study spectra of gases which are not themselves laser gases. In addition, the lower temperature and pressure of the gas in the additional cell means the homogeneous linewidth is narrower than the width of the laser Lamp dip.

This enables accurate measurement of the transition frequencies and the study of closely neighbouring spectral lines.

We can also use it to 'lock' the frequency of the laser output to that of a well defined atomic transition.

In the case of the Iodine Stabilised He-Ne, the iodine cell is placed within the laser cavity, where the circulating power is higher.



The iodine stabilised He-Ne laser is one of the most stable lasers of all and is a frequency standard in the visible part of the spectrum. (Accuracy ≈ 10kHz).

Coupled Rate Equations

So far, we have only considered CW operation of the laser. In particular we have shown that, under steady state conditions, the stimulated emission increases to a level whereby the gain of the laser medium equals the total loss, i.e.

$$\text{Gain} = \text{loss}_{\text{intrinsic}} + \text{loss}_{\text{output coupling}}$$

Where

$\text{loss}_{\text{intrinsic}}$ = loss due to scattering, absorption etc.

$\text{loss}_{\text{output coupling}}$ = transmission of output mirror

For pulsed lasers the steady state condition is never reached and it is useful to understand the temporal evolution of both the population inversion and the light intensity.

For an ideal four level laser system ($\tau_2 > \tau_1$, $N_1 \approx 0$, $N = N_2$), we have:

$$\frac{dN}{dt} = R - \frac{N}{\tau_2} - N\,\sigma_\nu\,D_\nu\,\frac{c}{n} \qquad [90]$$

and $$\frac{dD_\nu}{dt} = N\,\sigma_\nu\,D_\nu\,\frac{c}{n} - \frac{D_\nu}{t_c} \qquad [91]$$

where

N = population, m^{-3}
R = pumping rate, $sec^{-1}\,m^{-3}$
τ_2 = lifetime of upper state, sec
σ_ν = stimulated emission x-section, m^2
D_ν = photon density, m^{-3}
t_c = cavity decay time, sec

In the steady state condition $\frac{dN}{dt} = \frac{dD_\nu}{dt} = 0$, [90] and [91] become:

$$\frac{dN}{dt} = 0 = R - \frac{N_0}{\tau_2} - N_0\,\sigma_\nu\,D_{\nu 0}\,\frac{c}{n}$$

$$\frac{dD_\nu}{dt} = 0 = N_0\,\sigma_\nu\,D_{\nu 0}\,\frac{c}{n} - \frac{D_{\nu 0}}{t_c}$$

This gives:

$$N_0 = \frac{n}{c\,\sigma_\nu\,t_c} \qquad [92]$$

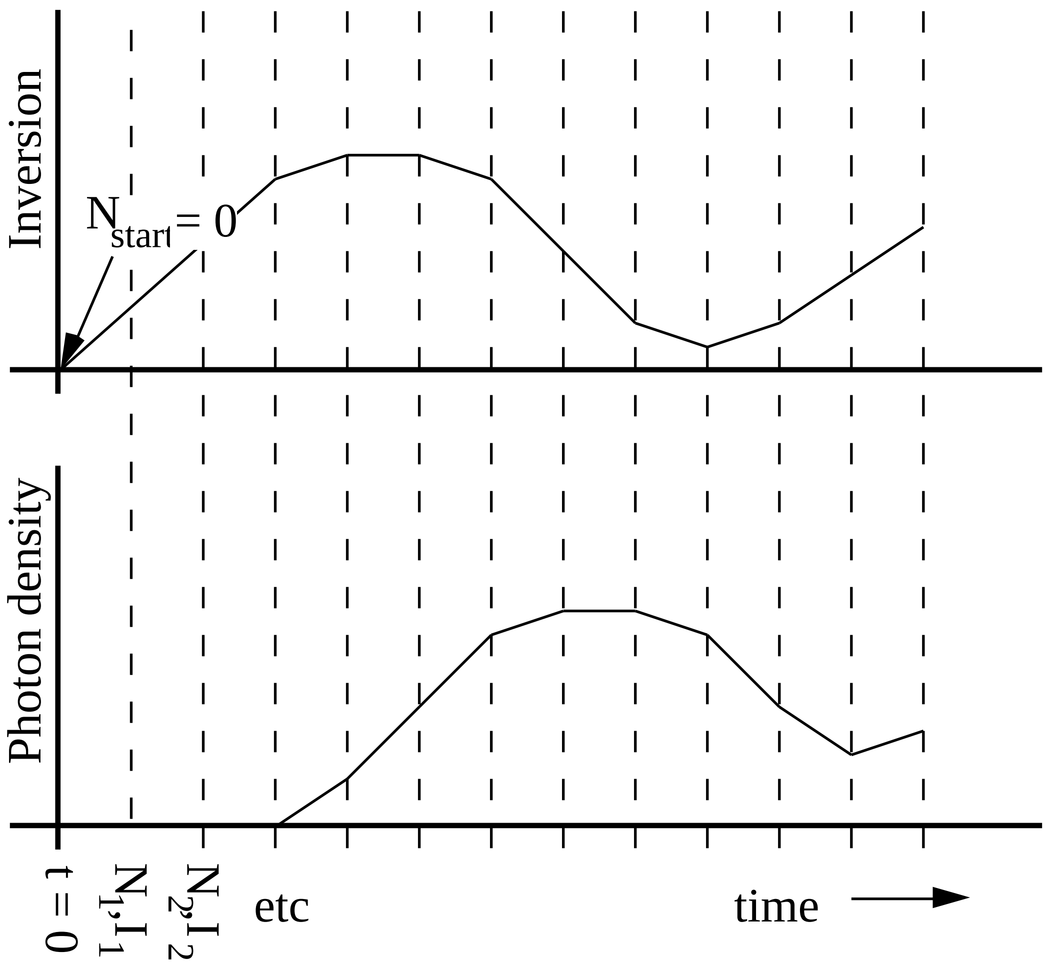
and,

$$D_{\nu 0} = \frac{R\,\sigma_\nu\,t_c - \frac{n}{c\tau_2}}{\sigma_\nu} \qquad [93]$$

More generally we should consider the case when we are not in the steady state such as when the laser is just switched on.

Equations of the type [90] and [91] are referred to as **coupled rate equations**. With real time varying parameters of pump rate and cavity decay time they are very difficult to solve analytically. Instead we can use a simple computer model.

By considering the change in N and D_ν over a short time we can build up the overall form for N(t) and D_ν (t) for any given R(t) and t_c(t).



From the starting values of N and D_ν we can calculate the values at t = 1

$$N_{t=1} = N_{t=0} + \Delta t \left(R - \frac{N_{t=0}}{\tau_2} - N_{t=0}\, \sigma_\nu\, D_{t=0} \frac{c}{n} \right)$$

$$D_{t=1} = D_{t=0} + \Delta t \left(N_{t=0}\, \sigma_\nu\, D_{t=0} \frac{c}{n} - \frac{D_{t=0}}{t_c} \right)$$

Likewise for t =2 etc., the general expressions to calculate the t = j+1 values from the t = j values are:

$$N_{t=j+1} = N_{t=j} + \Delta t \left(R - \frac{N_{t=j}}{\tau_2} - N_{t=j}\, \sigma_\nu\, D_{t=j} \frac{c}{n} \right)$$

$$D_{t=j+1} = D_{t=j} + \Delta t \left(N_{t=j}\, \sigma_\nu\, D_{t=j} \frac{c}{n} - \frac{D_{t=j}}{t_c} \right)$$

There is no restriction on R or t_c, and we can allow these to become time varying functions.

Using these equations and a computer we can model
Flashlamp pumping (i.e. R(t))
Q-Switching (i.e. $t_c(t)$)
Cavity Dumping (i.e. $t_c(t)$)

Q-Switching

For many applications, instead of a CW output from the laser we would prefer a short intense pulse, e.g.

Non linear optics (sometimes need very high powers)
Laser fusion
Laser drilling.

We could use our coupled rate equations [90] and [91] to model what happens when we pump the laser with a short pulse of pump energy.

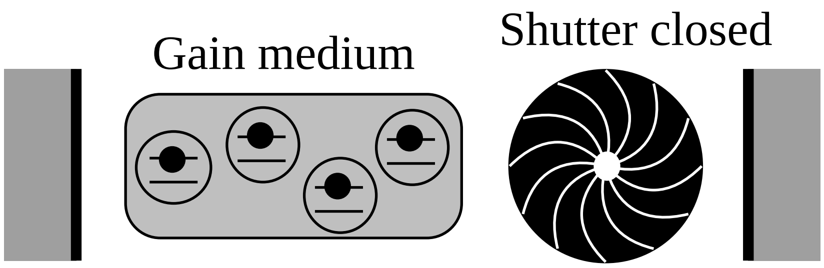
When N exceeds N_{th} the laser output rises, with or without relaxation oscillations, and the laser will remain "switched on" until the pumping rate falls below threshold.

We would like to "store" all the population inversion and get all the energy out in one short, high power pulse. One possible method for achieving this is **Q-Switching.**

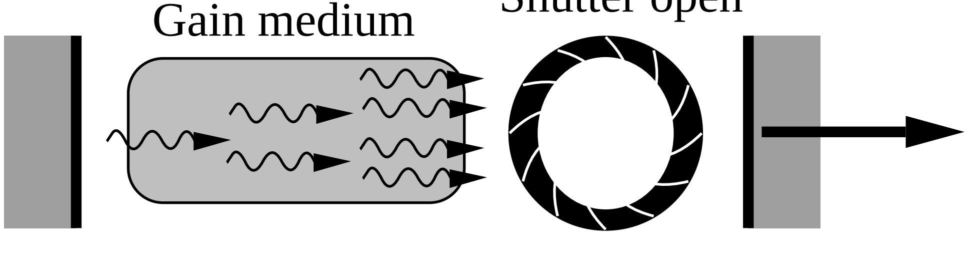
Q-Switching involves changing the cavity decay time t_c during the pumping pulse. This can be achieved by placing a "shutter' in the cavity.

Initially the shutter is closed, the cavity has a low Q, i.e. t_c is very short.

Even with heavy pumping, the laser is below threshold and all the pump energy is converted into a large population inversion.

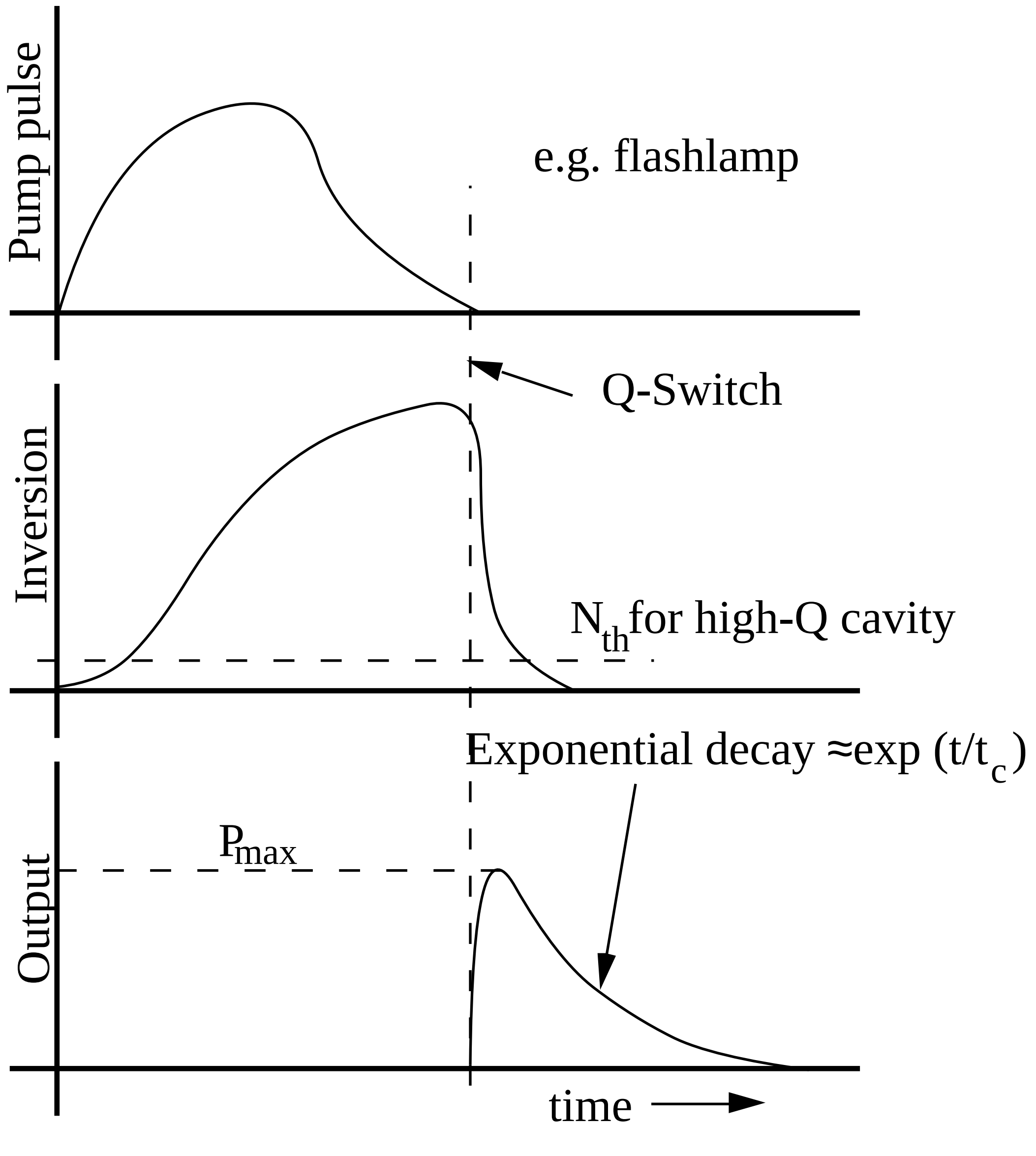


Once the pump pulse is over and the population inversion has reached its peak, the shutter is opened (i.e. high Q). The laser is now many times above threshold and the gain is very large.



The radiation density ρ_ν builds up very quickly within the cavity. Through stimulated emission, the large value of ρ_ν quickly destroys the population inversion and all the energy can be extracted from the gain medium.

The radiation now leaks out of the cavity through the partially transmitting mirror, its intensity decays with the cavity decay time constant of t_c.



Two things are required for a good Q-Switched system

1) A long upper state lifetime so that the inversion dose not leak away by spontaneous emission prior to Q-

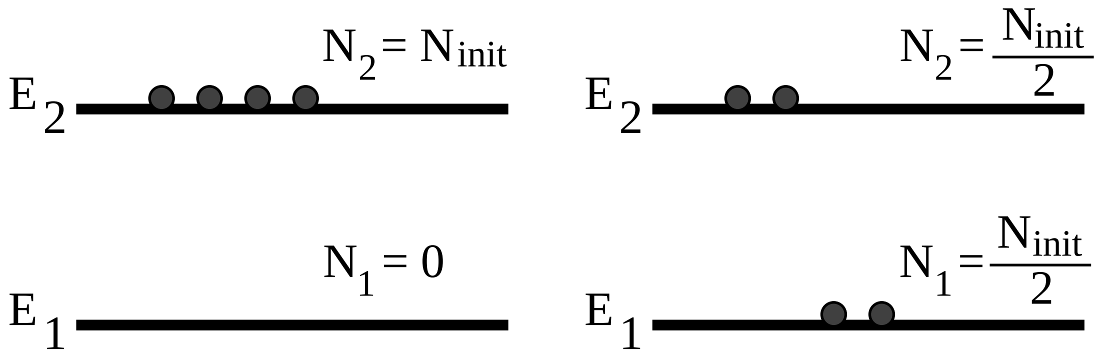
Switching

2) A very fast shutter (<1ns) !

Peak power from a Q-Switched system

Let's assume a perfect 4-level laser system , but one in which the pulse 'happens ' too quickly for the lower state population to empty.

The initial population inversion is N_{init}



Before Q-switch After Q-switch

The final population inversion is zero,

but $N_2 = N_1 = \frac{N_{init}}{2}$

The pulse duration is short compared to the upper state lifetime τ_2. Therefore <u>all</u> the reduction in upper state population is due to stimulated emission. Starting from no photons in the cavity prior to the Q-Switch, we can write:

$$D_{\nu \text{ after Q-Switch}} .V = \frac{N_{init}}{2} .V$$

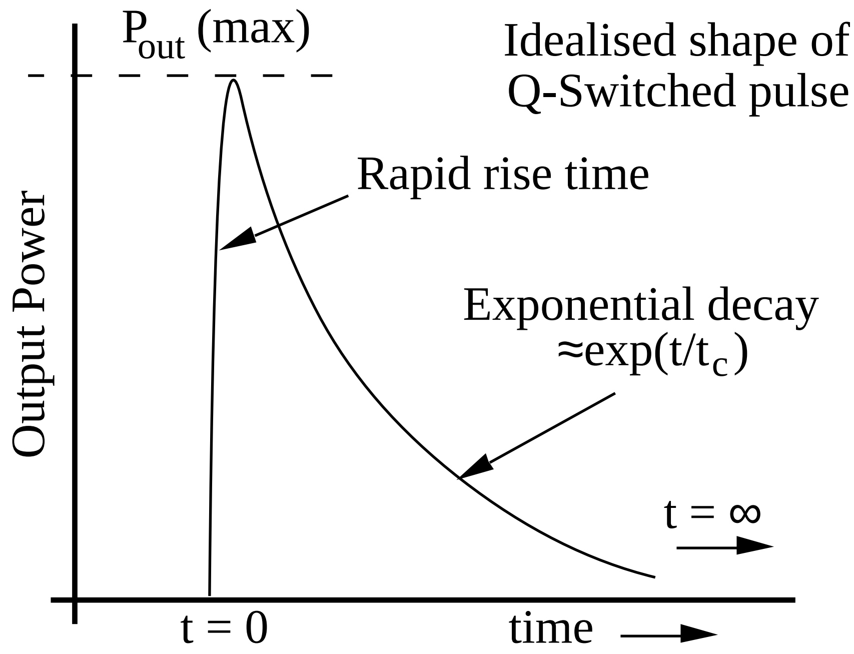
where

D_ν = photon number density

V = volume of laser cavity (=vol. of active medium)

All these photons will escape from the cavity as output, therefore the total energy in the Q-Switched pulse is given by:

$$E_{total} = \frac{N_{init}}{2} .V h\nu \qquad [102]$$



From [102],

$$\int_0^\infty P_{out}\, dt = E_{total} = \frac{N_{init}}{2} .V h\nu$$

where
P_{out} = the output power

The pulse shape is approximately a single sided exponential, therefore

$$\int_0^\infty P_{out}(max)\, \exp(t/t_c)\, dt \approx \frac{N_{init}}{2} .V\, h\nu$$

$$t_c\, P_{out}(max)\, \exp(t/t_c) \approx \frac{N_{init}}{2} .V\, h\nu$$

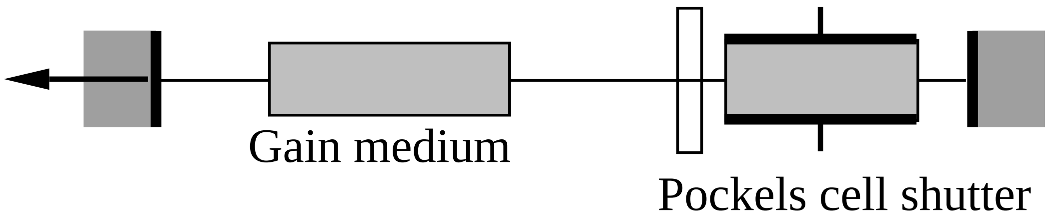
When t = 0, $P_{out} = P_{out}(max)$, i.e. for a Q-Switched pulse

$$P_{out}(max) \approx \frac{N_{init}}{2t_c} .V\, h\nu \qquad [103]$$

Active Q-Switching

A mechanical shutter, like in a camera, is not fast enough for good Q-Switching. Instead we use an electro-optic shutter. An electro-optic shutter uses a polariser and a Pockels cell.

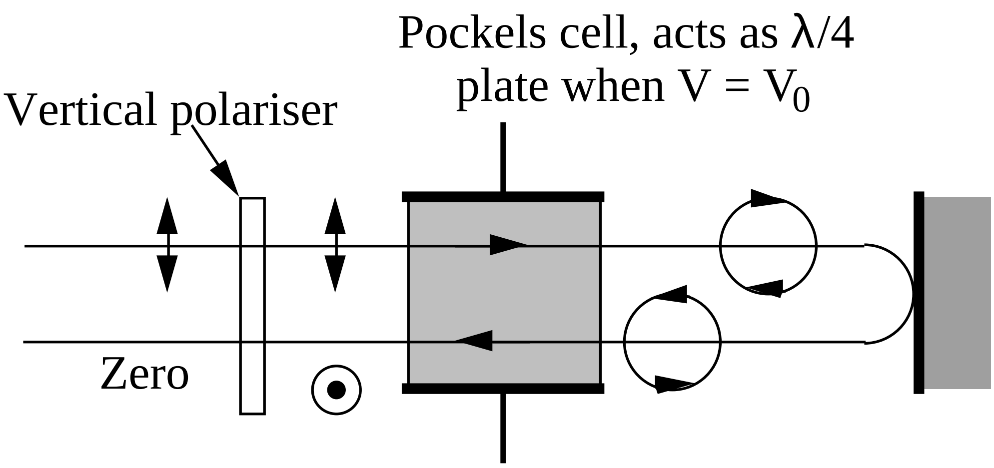
An active Q-Switch (switch Q from low to high)



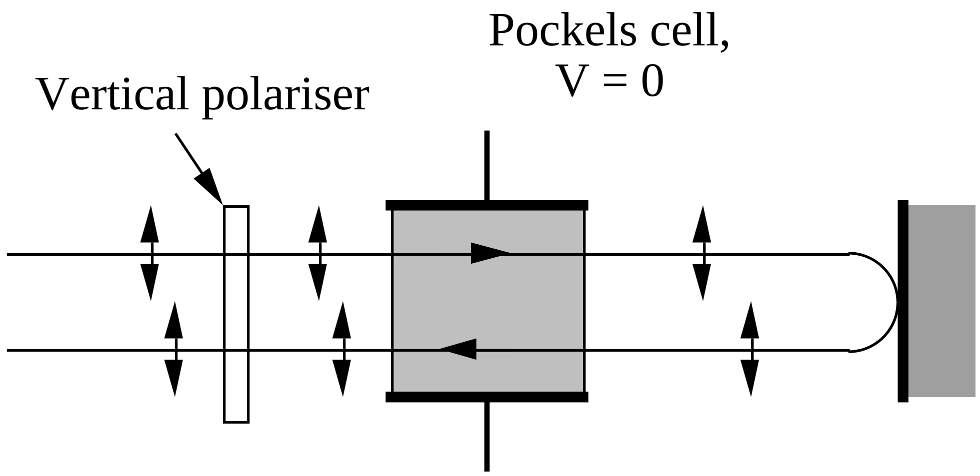
The Pockels cell is an electro-optic crystal, in which the e_{ray} - o_{ray} retardation depends on applied voltage.

When $V=V_0$, the cells retardation is λ/4, and the cell converts linearly polarised light into circular.

The reflected light is polarised in the wrong plane to be transmitted through the polariser, the cavity has low Q



With the voltage removed, the reflected light is transmitted back through the polariser, and the cavity has high Q.



Passive Q-Switching

Alternatively we can use a saturable absorber as a passive Q-switch.

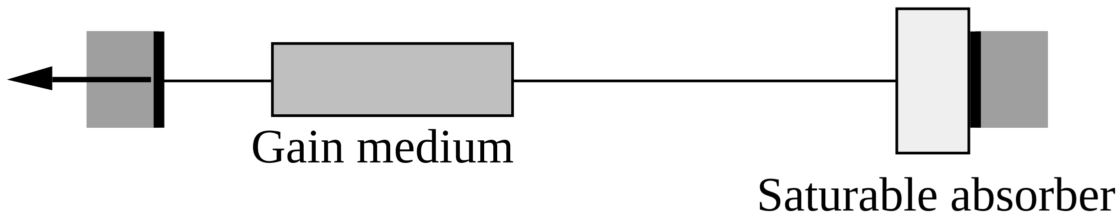
A saturable absorber (usually an organic dye) has an absorption coefficient that is a function of incident light intensity.

They can be understood in terms of a two level system that becomes transparent when all the molecules have been excited into the upper state. Once the light source is removed the upper state population relaxes and the dye, once again, becomes absorbing.

No external signal is needed to trigger a passive Q-Switch. As the population inversion grows, it is the increased level of spontaneous emission that saturates the dye and thus switches the Q.

Clearly, it is important the thickness/concentration of dye is adjusted so that the Q triggers at the optimum time (i.e. when the population inversion is at a maximum.).

A passive Q-Switch (switch Q from low to high)



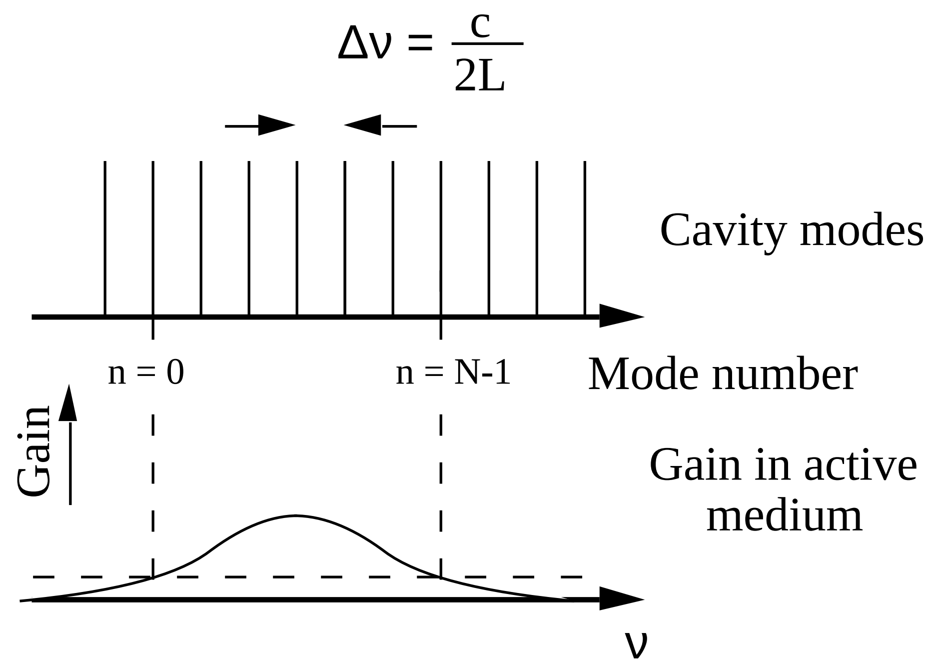
Q-Switching is capable of producing pulses of 10's nsec duration with megaWatt peak powers

Nd:YAG, Nd:Glass and Nd:YLF laser are often used in this way.

Mode-Locking

Previously, we have considered how to ensure that the laser oscillates on a single longitudinal mode.

What happens if many longitudinal modes are above threshold?



The resultant electric field is the sum of the field from all the oscillating modes. Assuming that N modes are above threshold and that they all oscillate with the same amplitude, we can write:

$$E_{tot}(t) = \sum_{n=0}^{n=N-1} E_0 \exp\{i\,(2\pi(\nu + n\Delta\nu)t + \varphi_n)\} \quad [105]$$

where

ν = frequency of the n = 0 mode

$\Delta\nu$ = intermode spacing

φ_n = relative phase of the n^{th} mode

Let's define the phase so that:

$$\varphi_n = 0 \text{ for all n} \quad [106]$$

[105] becomes:

$$E_{tot}(t) = \sum_{n=0}^{n=N-1} E_0 \exp\{i\,(2\pi(\nu + n\Delta\nu)t)\}$$

$$E_{tot}(t) = E_0 \exp\{i\,(2\pi\nu t)\} \sum_{n=0}^{n=N-1} \exp\{i\,(2\pi(n\Delta\nu)t)\}$$

[107]

The summation term in [107] is a geometrical series

$$\sum_{n=0}^{n=N-1} \exp\{i(2\pi n\Delta\nu\, t)\} = (1 + e^{-i\Delta\omega\, t} + e^{-i2\Delta\omega\, t} + \qquad)$$

where

$$\Delta\omega = 2\pi\,\Delta\nu \qquad [108]$$

It can be shown that:

$$1 + e^{-i\Delta\omega t} + e^{-i2\Delta\omega t} + \ldots + e^{-(N-1)i\Delta\omega t} = \frac{\sin(N\Delta\omega t/2)}{\sin(\Delta\omega t/2)}$$

Therefore, [107] can be written as:

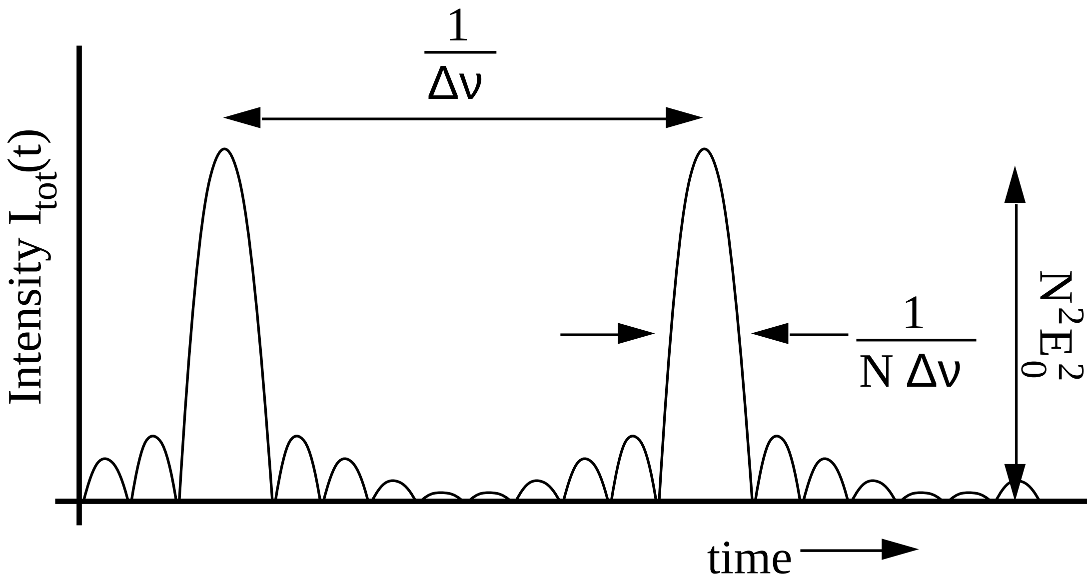
$$E_{tot}(t) = E_0 \exp\{i\,(2\pi\nu t)\} \frac{\sin(N\Delta\omega t/2)}{\sin(\Delta\omega t/2)}$$

The output intensity is E_{tot}. E_{tot}^*, hence

$$I_{tot}(t) = E_0^2 \frac{\sin^2(N\Delta\omega t/2)}{\sin^2(\Delta\omega t/2)} \qquad [109]$$

If N =1, i.e. single longitudinal mode, $I_{tot} = E_0^2$ and the laser has a 'DC' output.

When N>1 [109] has the form: (remember $\Delta\omega = 2\pi\,\Delta\nu$)



This is the output for a multi-mode laser where relative phase between the modes is fixed at zero, i.e. $\varphi_n = 0$.

We call this **Mode-Locking.**

The above is often referred to as the **frequency-domain explanation** of mode-locking.

The output of a mode-locked laser comprises of a series of short pulses with

$$\text{Pulse duration}_{\text{ time}} = \frac{1}{N\,\Delta\nu} \qquad [110]$$

$$\text{Pulse separation}_{\text{ time}} = \frac{1}{\Delta\nu} \qquad [111]$$

where

N = number of oscillating modes

$\Delta\nu$ = intermode spacing

The peak intensity for a mode-locked laser is related to the average, non mode-locked intensity by:

$$\text{Intensity}_{peak} = \text{Intensity}_{average} \times \frac{\text{pulse separation}}{\text{pulse duration}}$$

i.e. $I_{peak} = I_{average}\ N$

where

N = number of locked modes

[110] and [111] give an insight into an alternative view of mode-locking

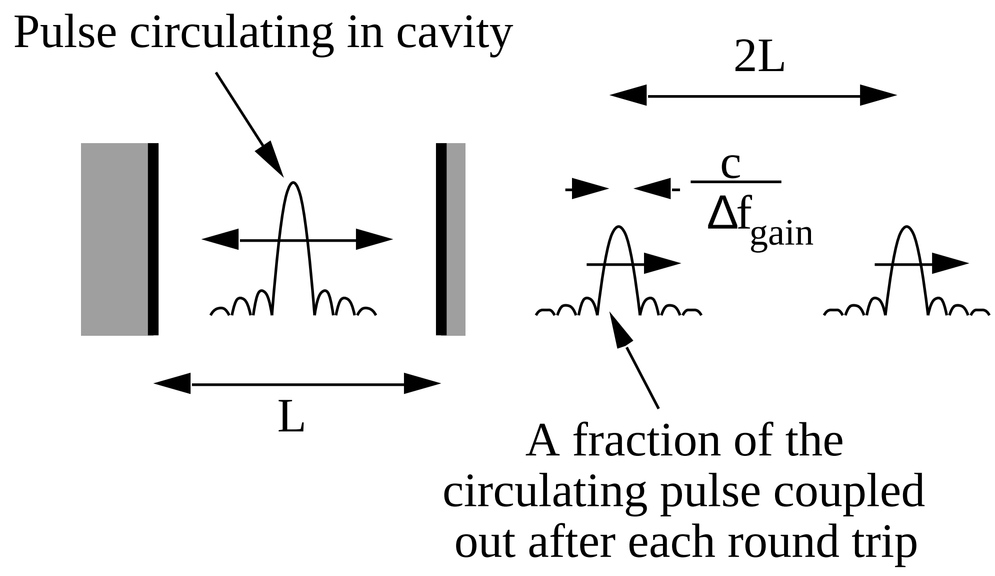
[111] the pulse separation, is equal to the round trip time of the light in the laser cavity, i.e.

$$\text{Pulse separation} = \frac{1}{\Delta\nu} = \frac{2L}{c} \qquad [111]$$

[110] the pulse duration, is equal to the reciprocal of the gain-bandwidth of the laser Δf_{gain}, i.e.

$$\text{Pulse duration} = \frac{1}{N\,\Delta\nu} = \frac{1}{\Delta f_{gain}} \qquad [110]$$

We can consider the output of a mode-locked laser to be due to a single pulse circulating within the cavity, which gives an output pulse each time it is reflected from the output coupler.



This is often referred to as the **time-domain** explanation of mode-locking

The duration, or length of the pulse is determined by the gain-bandwidth of the gain medium via the uncertainty principle i.e.

$$\Delta f \, . \, \Delta t \approx 1$$

The exact relationship depends on the pulse shape.

If the output from a mode-locked laser satisfies the above condition it is said to be **transform limited**.

Active Mode-Locking

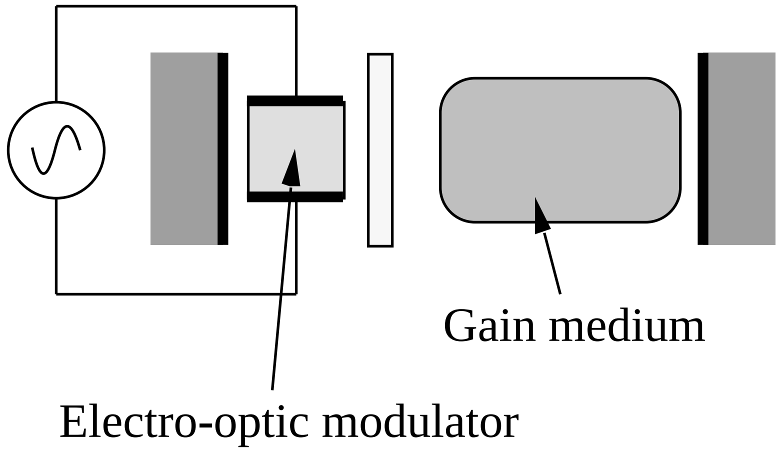
To obtain the pulsed output, the relative phase of each mode was set to zero, i.e.

$$\varphi_n = 0 \text{ for all n} \qquad [106]$$

This is why mode-locking is sometimes called **phase-locking**.

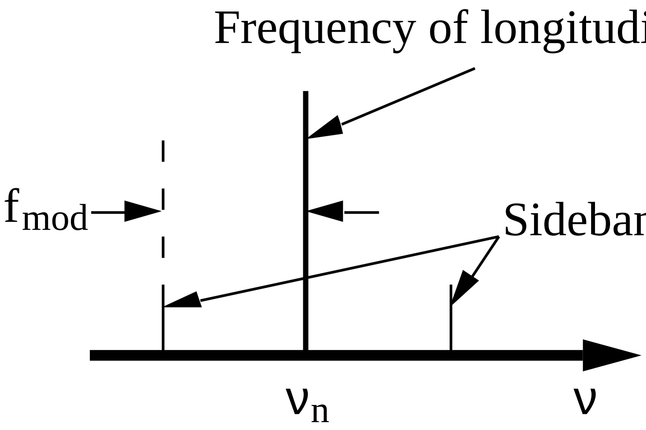
How do we set the phase of the modes?

One way of locking the phases is to modulate the gain or loss of the laser at the round trip frequency of the cavity.



We can use an electro-optic modulator (see Q-Switching) to modulate the Q of the cavity.

By modulating the amplitude of a single longitudinal mode, frequency side-bands are generated on the optical frequency.



The frequency of the side-bands is given by:

$$\nu_{sidebands} = \nu_n \pm f_{mod}$$

where

ν_n = frequency of longitudinal mode

f_{mod} = modulation frequency

If f_{mod} is matched to the longitudinal mode spacing then the side-bands have the same frequency as the adjacent longitudinal mode.

By exchanging energy between each adjacent pair of longitudinal modes, the individual phases are maintained to be 'in-step'.

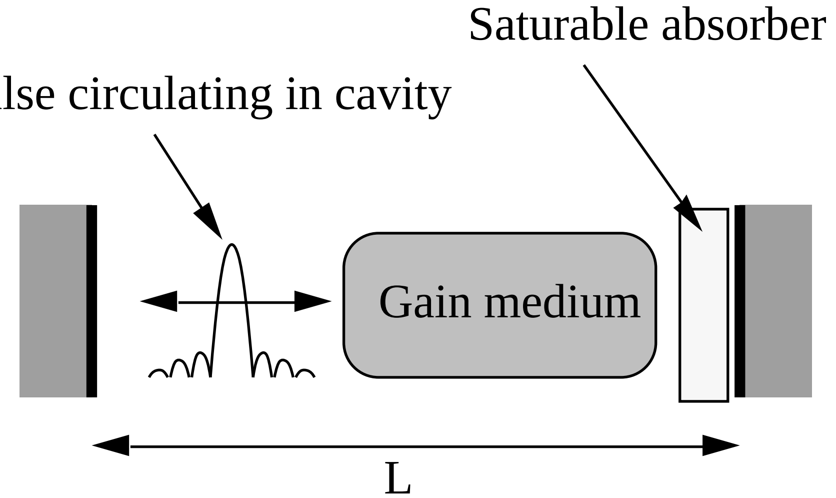
The condition for mode-locking is:

$$f_{mod} = \frac{c}{2L}$$

The action of applying an external modulation leads this technique to be called **Active Mode-Locking.**

Passive Mode-Locking

For Q-Switched operation of a laser we showed that the Q could be both actively and passively switched. In addition to Q-Switching, it is also possible to use a saturable absorber to mode-lock the output of a laser.



As with the Q-Switched system, the saturable absorber is usually an organic dye

1) Each time the intense circulating pulse 'arrives' at the dye, a few photons are absorbed and the dye becomes transparent.

2) The rest of the pulse passes through the dye unattenuated and is subsequently amplified in the gain medium.

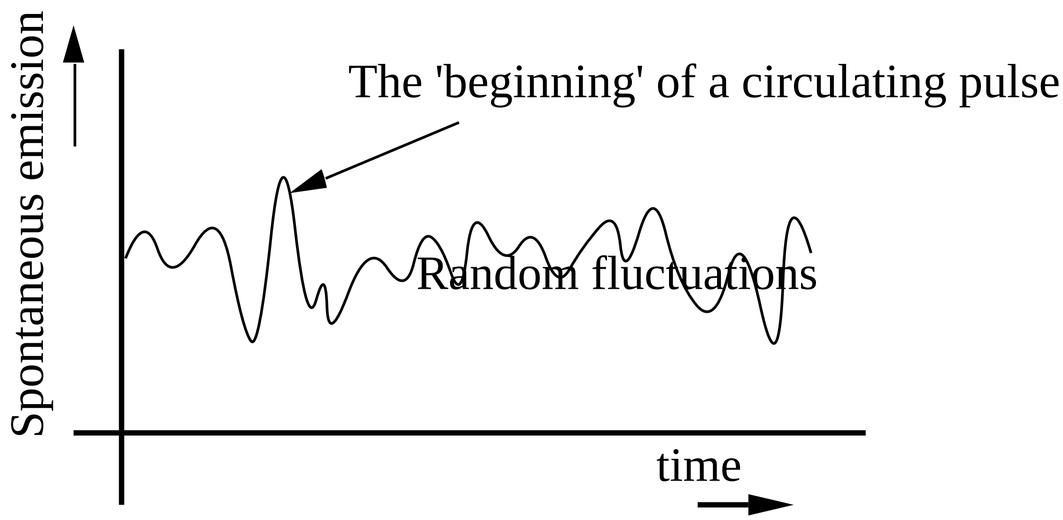
3) After the passage of the pulse, the upper state population of the dye relaxes so that the absorption of the dye 'recovers' prior to the next arrival of the pulse

The self modulation of the cavity loss and hence side-band coupling of adjacent longitudinal modes leads this technique to be called **Passive Mode-locking**

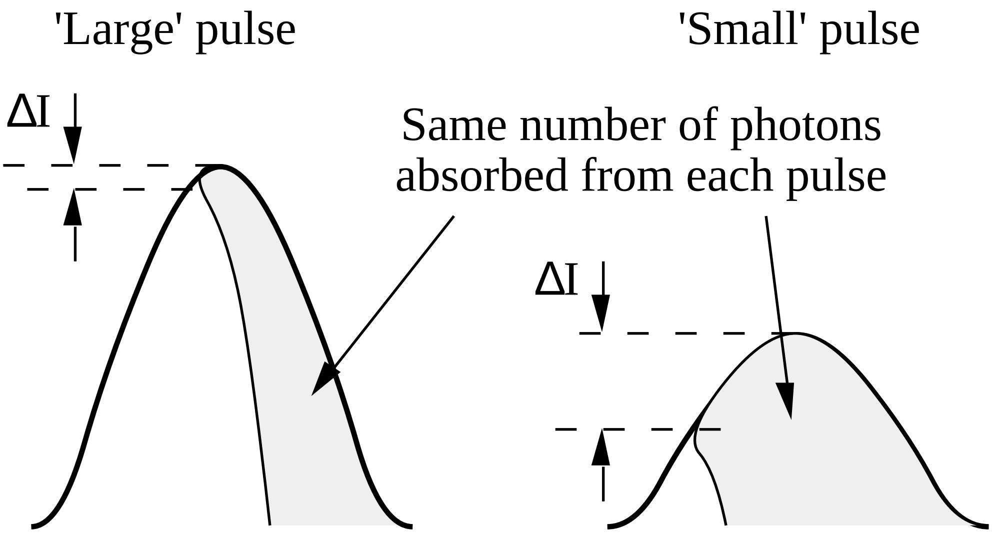
Pulse evolution in a passively mode-locked laser

We describe above how a saturable absorber can sustain a circulating pulse, but how does the pulse start?

In the case of a passively mode-locked laser, the pulse 'starts' from a 'chance' spontaneous emission

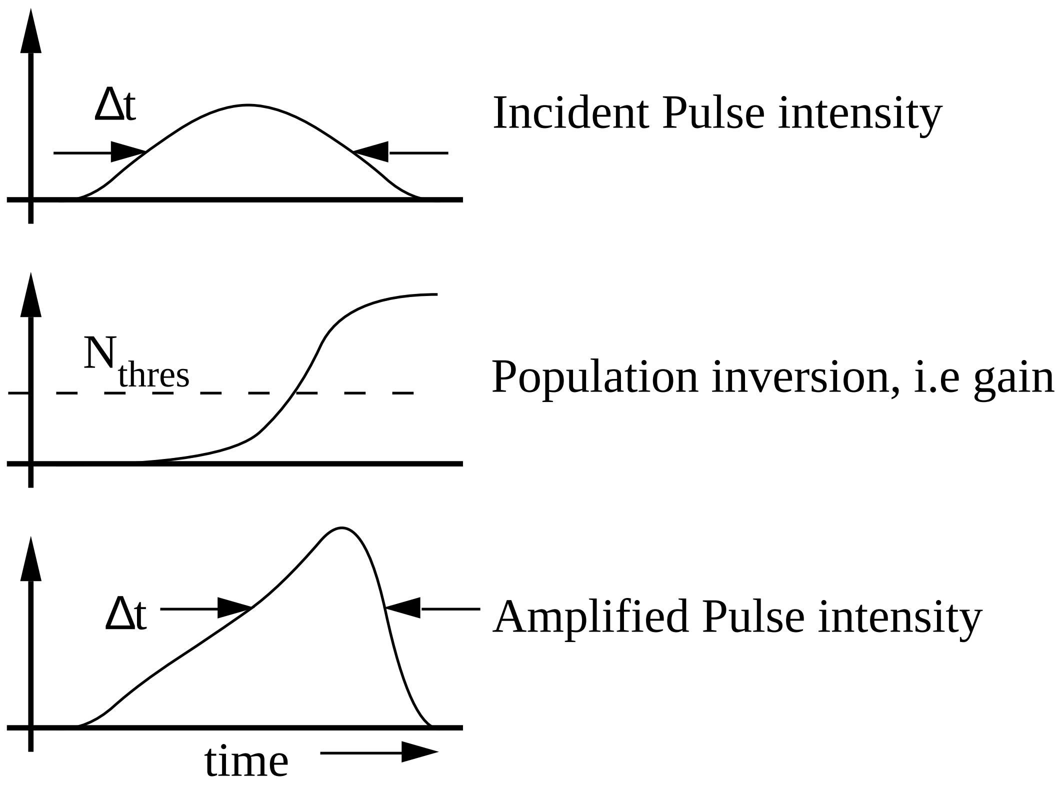


The saturable absorber 'discriminates' against small pulses which are completely absorbed. For each round trip, the larger pulses grow with respect to the smaller.



The saturable absorber attenuates the 'large' pulse less than the 'small' pulse.

The gain medium also plays a significant part in pulse shortening and selection:



When a pulse enters the gain medium the stimulated emission reduces the population inversion.

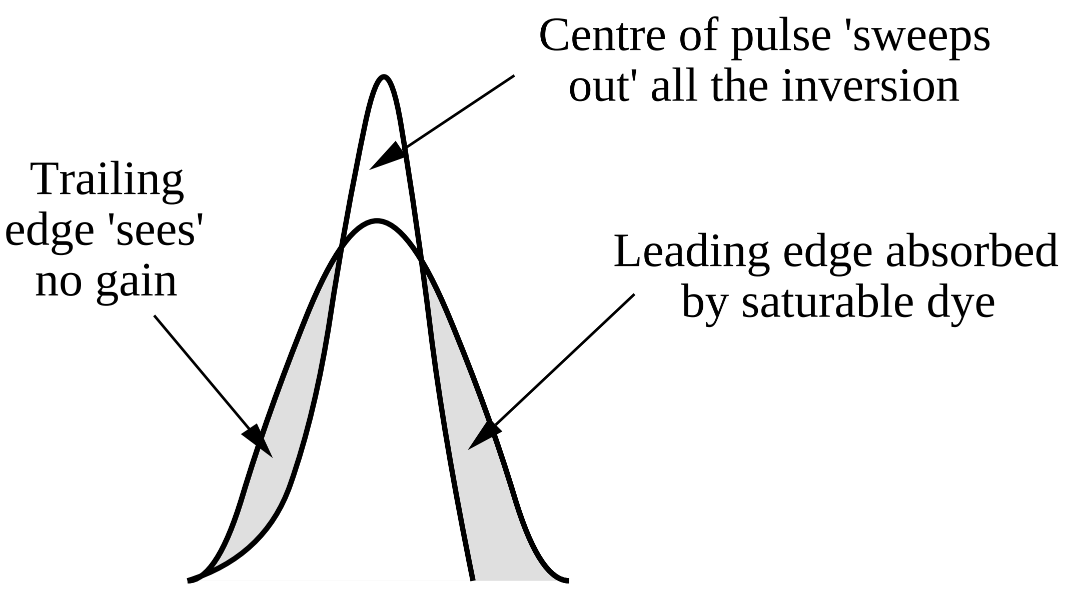
1) In the case of an intense pulse, the leading edge 'sweeps out' the gain so that the tail of the pulse is not amplified.

2) Pumping restores the population inversion prior to the next arrival of the pulse.

3) On successive passes through the gain medium, the leading part of the pulse becomes more intense leaving less and less gain for the tail.

4) Hence, the tail of the pulse is progressively removed and the pulse becomes shorter.

The 'sweeping out' of the gain coupled with the action of the saturable absorber acts to shorten and amplify the pulse as it circulates round the laser cavity.



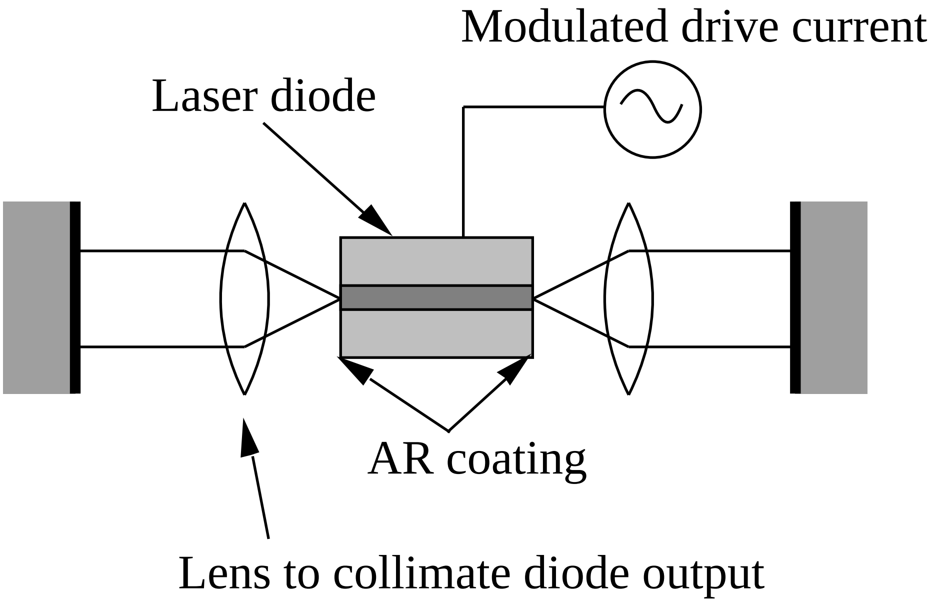
If the gain medium is over-pumped and the gain 'recovers' too quickly, it is possible to get a second pulse to simultaneously circulate in the cavity.

Mode-Locked diode laser

It is possible to actively mode-lock a diode laser by modulating the drive current (and hence the gain) at the round trip frequency.

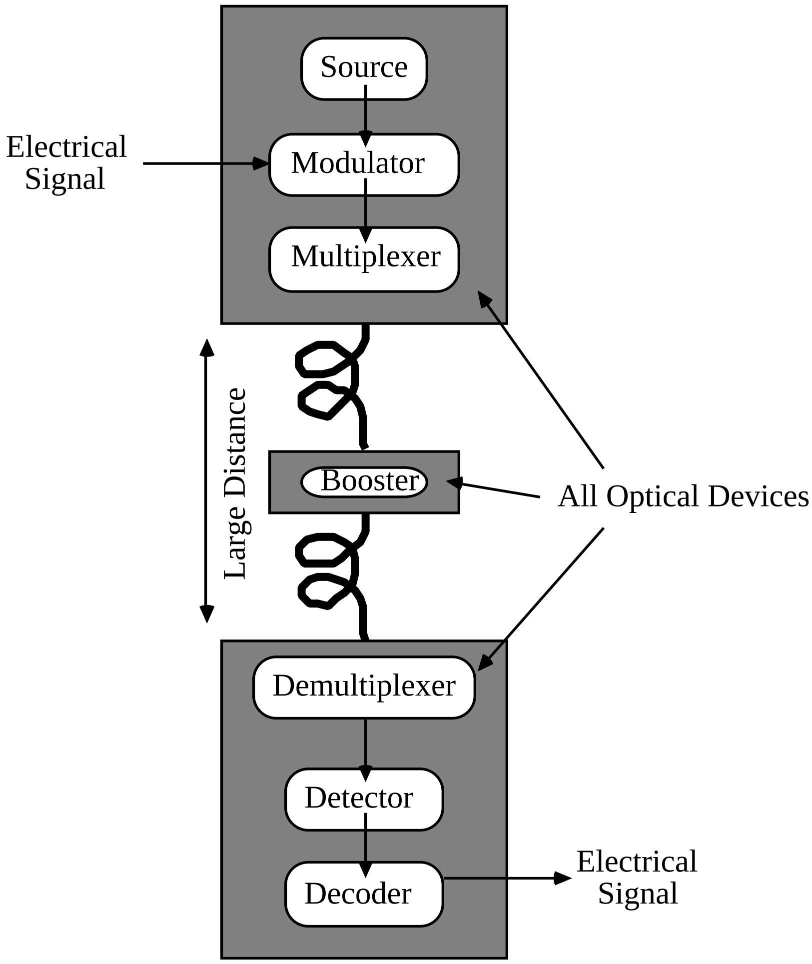
However, the short length of the diode cavity (i.e. the chip itself ≈ 100'sμm) would require modulation frequencies of 100's GHz (quite hard!)

Instead we place the diode within an external cavity



For a 1m cavity, the required modulation frequency is now only ≈150MHz.

Integrated optics

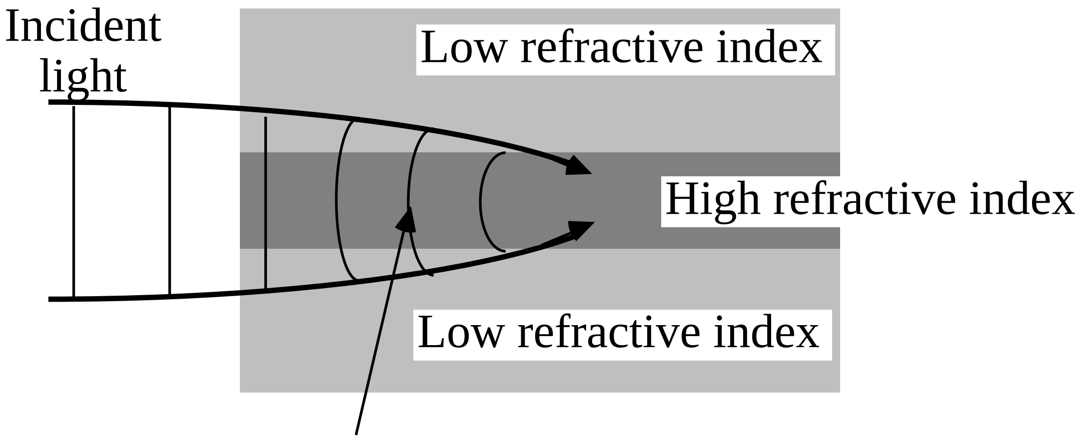


Most people are familiar with fibre optic. Fibre optic cable is ideal for external links between devices (more recent technology has led to fibre based devices too).

For links within the devices themselves we need ‘light wires’ that can be fabricated as part of the overall device. We would like to use semiconductor processing techniques for making these ‘light wires’

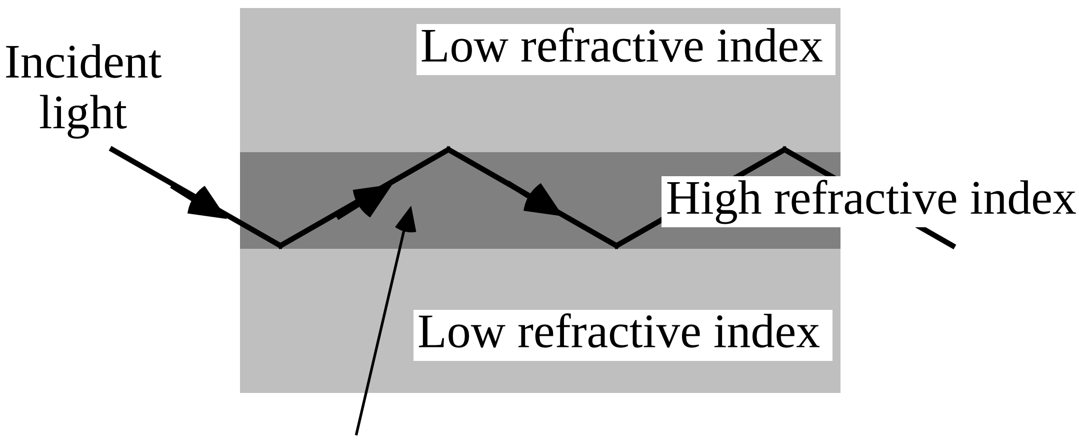
Light guiding
Consider what happens to light when it passes into a ‘layered’ material and the layers have different refractive indices.

Wavefront Picture



Centre of wave is slowed by high index media

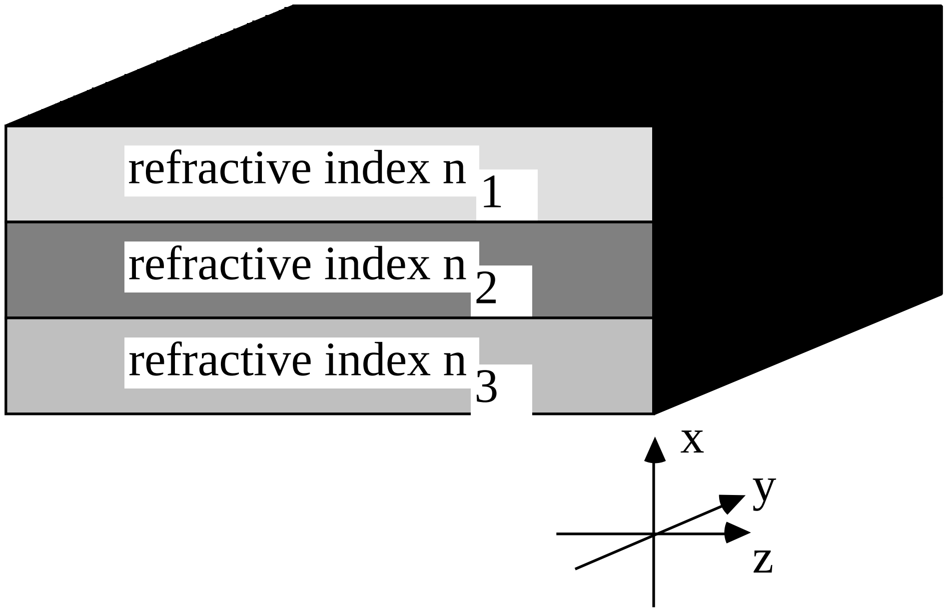
Ray Picture



Wave is reflected by total internal reflection

The light is contained within the layer of high refractive index. i.e. the light is guided.

Slab waveguides (e.g. “on-chip)



Consider a specific case and understand what happens in terms of ray optics:

Bottom layer, glass $n_3 = 1.5$

Middle layer, Zinc Sulphide (ZnS) $n_2 = 2.3$

Top layer, air $n_1 = 1.0$

Remember

if $\theta_i < \theta_c$ then some light is transmitted

if $\theta_i > \theta_c$ then all the light is reflected

& $\theta_c = \sin^{-1}\left(\frac{n_{low}}{n_{high}}\right)$

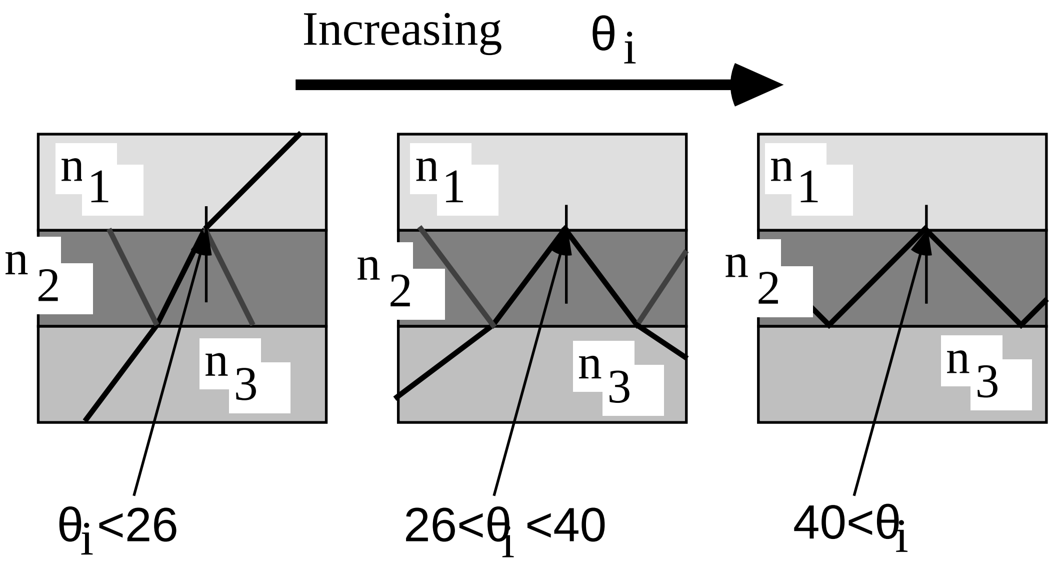
What is the critical angle is for Total Internal Reflection?

This gives:

for the ZnS:Air interface $\theta_c = 26°$

for the ZnS:Glass interface $\theta_c = 40°$

As the light ray is launched into the layers, three different outcomes can occur, see below:



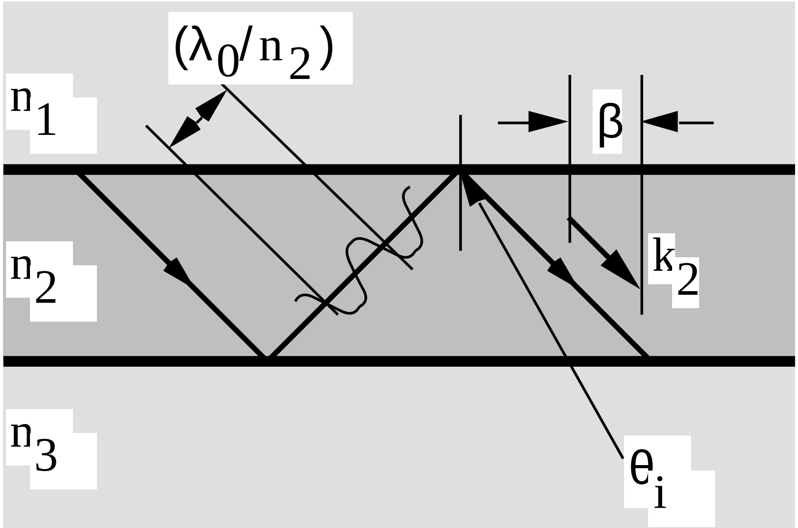
1) Ray launched such that $\theta_i < 26°$.
   No guiding takes place and the ray is free to travel through any layer

2) Ray launched such that $26° < \theta_i < 40°$.
   Total internal reflection occurs at the ZnS:Air interface, but the ray is still free to ‘escape’ into the bottom layer

3) Ray launched such that $\theta_i > 40°$.
   Total internal reflection occurs at both the ZnS:Air and the ZnS:Glass interface and the ray is confined to the ZnS layer.

This ray is now guided

Again, the ray optics model is the best way of understanding what's actually happening



The ray takes a zig-zag path, total internal reflection at the interface between the layers keeps it within the waveguide. If the zig-zags become too pronounced then the total internal reflection fails and the ray escapes!

β is the resolved component of k_2 in the z direction, i.e.

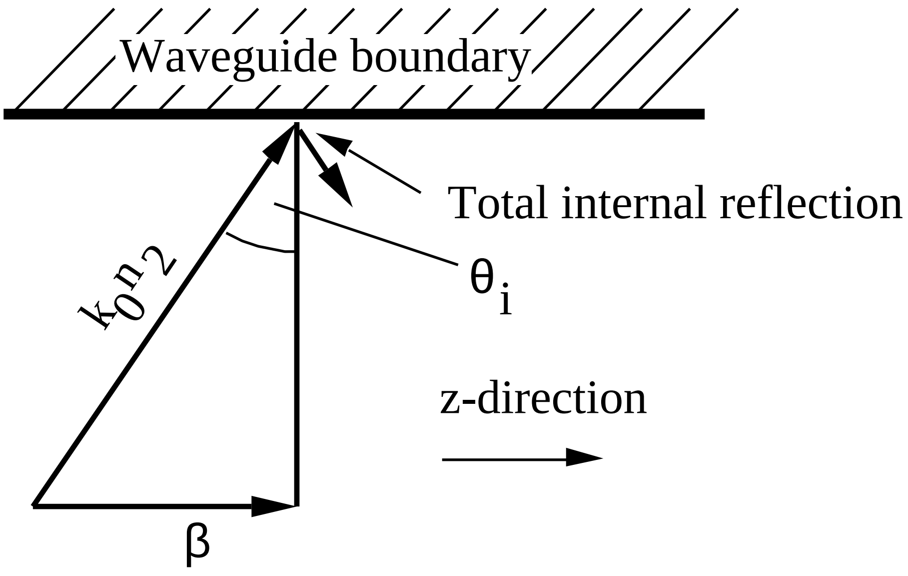
$$\beta = \sin\theta_i \, k_2 = \sin\theta_i \, k_0 n_2 \quad [108]$$

The guide wavelength is:

$$\lambda_g = \frac{2\pi}{\beta} = \frac{2\pi}{\sin\theta_i \, k_0 n_2} = \frac{\lambda_0}{\sin\theta_i \, n_2} \quad [109]$$

Re-arranging [108] for $\sin\theta_i$ we get:

$$\sin\theta_i = \frac{\beta}{k_0 \, n_2} \quad [110]$$



Total internal reflection fails when $\beta < k_0n_3$, or $\beta < k_0n_1$

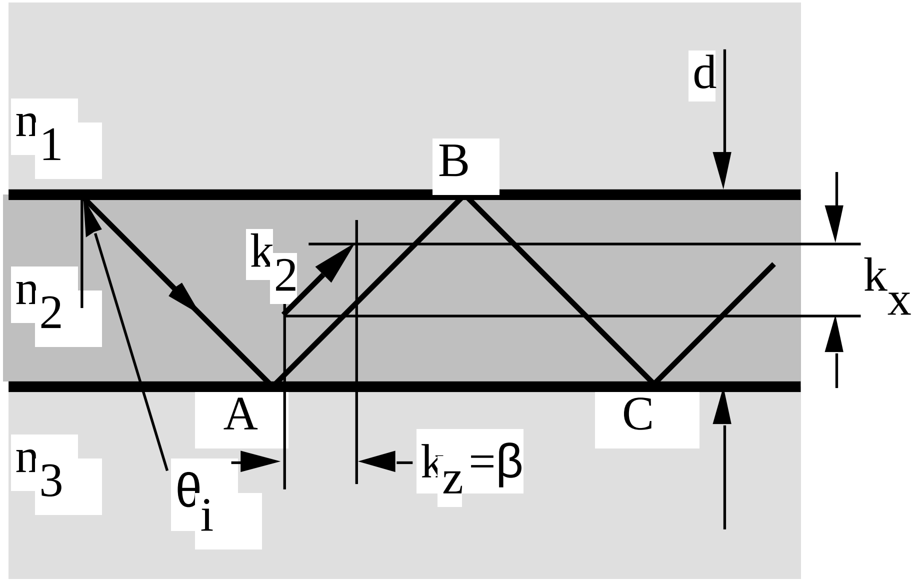
i.e. when $\theta_i < \sin^{-1}\frac{n_3}{n_2}$, or $\theta_i < \sin^{-1}\frac{n_1}{n_2}$

Modes in a slab waveguide

We have so far implied that providing $\theta_i > \theta_c$ then the ray will be guided

We are wrong

What is true is that if $\theta_i < \theta_c$ then no guiding will take place. However, even when $\theta_i > \theta_c$, there are only certain values of θ_i that give rise to allowed propagation modes. Let us again use the ray picture to understand what's happening:



The condition for propagation is that the phase of the wave at A must be equal to that at C (similar to etalon angle calculation for etalon rings). This sets a restriction on values of θ_i, such that:

$$2d\, k_x = m\, 2\pi$$

This problem is more complicated than it looks since in addition to the round trip path length we should also allow for phase shifts on reflection from the interfaces.

$$2d\, k_x + \Phi_{2\text{-}1} + \Phi_{2\text{-}3} = m\, 2\pi \quad [111]$$

where Φ is the phase change on reflection

but $\quad k_x = k_2 \cos\theta_i$

which, ignoring the phase shift on reflection, gives:

$$2d\, k_2 \cos\theta_i = 2d\, n_2 k_0 \cos\theta_i = m\, 2\pi \quad [112]$$

Thus modes of various “m”(m=0, 1, 2, etc) correspond to different θ_i

However, for a given guide, the maximum value of $\cos\theta_i$ is set by the limit of total internal reflection to be

$$\theta_i(\min) = \arcsin(n_1/ n_2) \qquad [113]$$

$$\text{but} \quad n_2\cos(\arcsin(n_1/ n_2))= \sqrt{n_2^2 - n_1^2} \qquad [114]$$

Subbing this into [112] gives

$$2d\, k_0\sqrt{n_2^2 - n_1^2} \quad > m\, 2\pi$$

or

$$m(\max) < d\, k_0\sqrt{n_2^2 - n_1^2} / \pi \qquad [115]$$

Thus if

$$d\, k_0\sqrt{n_2^2 - n_1^2} / \pi \quad < 1 \qquad [116]$$

Then only one mode can propagate and the guide is said to be “single mode”

Wave velocity in a slab waveguide

The phase velocity in the z-direction is given by:

$$V_\varphi = \frac{\omega}{k_z} = \frac{\omega}{\beta} = \frac{ck_0}{\beta} \qquad [120]$$

($\beta = \sin\theta_i\, k_0 n_2$)
we will define a mode index N where:

$$N = \frac{\beta}{k_0} \quad [121]$$

Therefore [120] can be written as:

$$V_\varphi = \frac{c}{N} \quad [122]$$

where N (the mode index) is the equivalent to the refractive index.

Note that for a guided mode:

$$k_0 n_1 < \beta < k_0 n_2 \quad [123]$$

Hence:

$$n_1 < N < n_2 \quad [124]$$

Therefore:

$$V_{\varphi,\,core} < V_{\varphi,\,guide} < V_{\varphi,\,cladding} \quad [125]$$

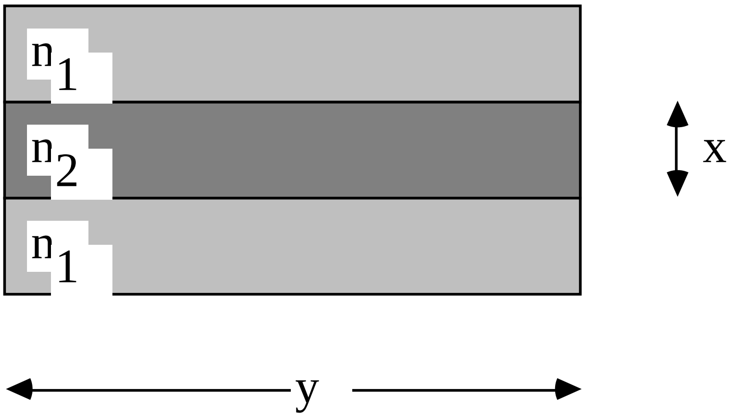
This isn't as surprising as it may seem. The guided mode partly penetrates into the cladding ($n_{core} > n_{cladding}$), therefore the wave energy spends some of it time travelling at a higher velocity than one would predict from the core alone.

The resultant phase velocity lies between the core velocity and the cladding velocity.

Fibre optic waveguides

So far we have only concerned ourselves with slab waveguides, where y >> x and y >> λ_0

Looking along a slab waveguide

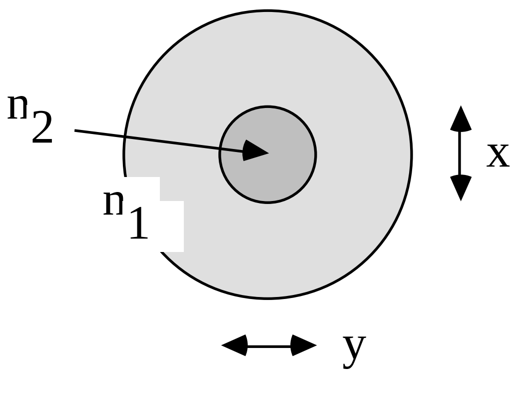


These are manufactured using semiconductor processes and devices are typically a few mm’s in size.

For linking devices together over cm’s or km’s some other technology is required.

Fibre optic waveguides are made from different types of glass or plastics. A high refractive index glass for the core with a low refractive index glass for the cladding.

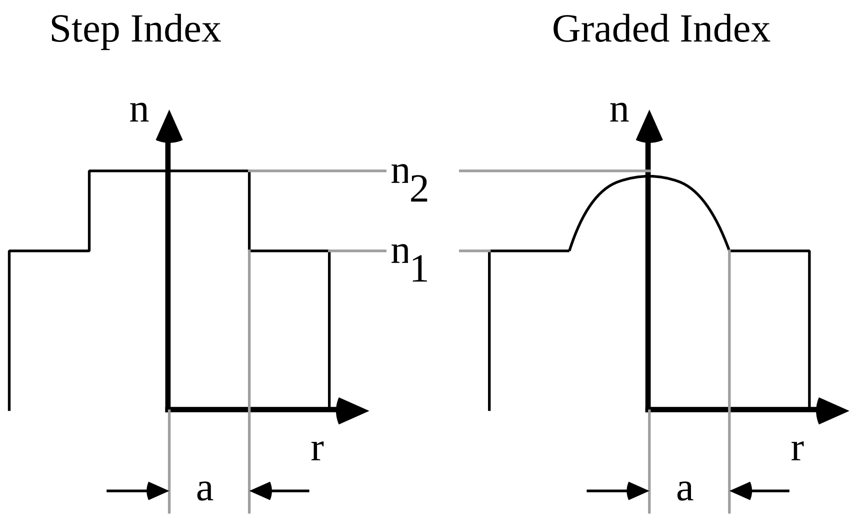
Looking along a fibre optic waveguide



There are two main types of fibre optic waveguides

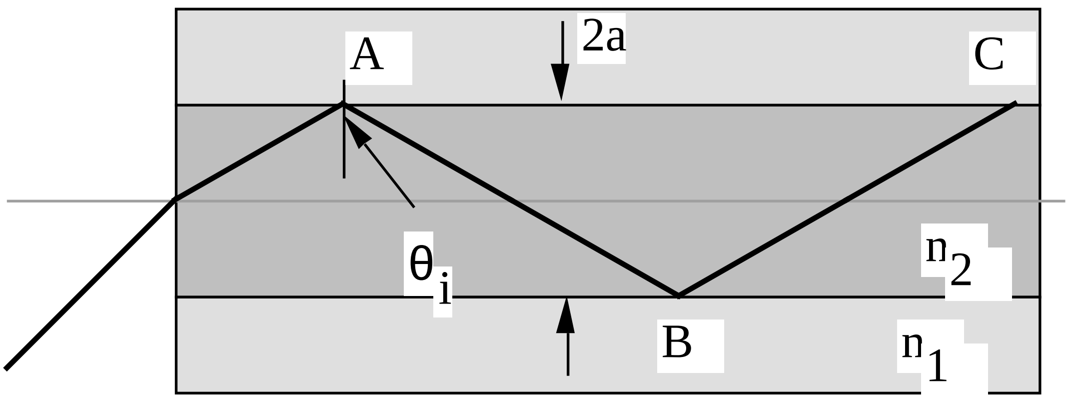
1) Step Index, where there is a sudden fall in the refractive index from n_2 to n_1 at a radius r=a

2) Graded Index, where the refractive index gradually falls from a maximum n_2 at the core to n_1 at a radius r=a



Fibre waveguide modes

Just as in the slab case, we rely on total internal reflection to keep the ray within the core of the fibre.



Shown above is a meridional ray (i.e. one that passes through the axis of the fibre).

For a contained waveguide mode we need the phase at A to be equal to the phase at C. from comparison to the slab waveguide we can write:

$$4a\, k_x + 2\Phi = m\, 2\pi \quad [130]$$

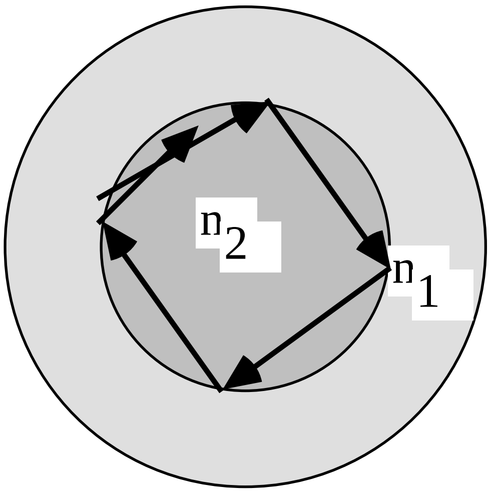
where Φ is the phase shift on reflection.

However, the rays are now bound in both the x and y directions and the order of the mode can be different in the two planes. Therefore, we need two mode numbers, i.e.:

$$TE_{lm} \text{ and } TM_{lm} \quad [131]$$

Sadly the problem is even more complicated.

In addition to meridional rays, skew rays also exist. These propagate down the fibre in a helical path without ever passing through the axis.



Looking along the fibre core

No longer can we split the modes into TE and TM, since the helical path ray contain a bit of both.

These modes are denoted by

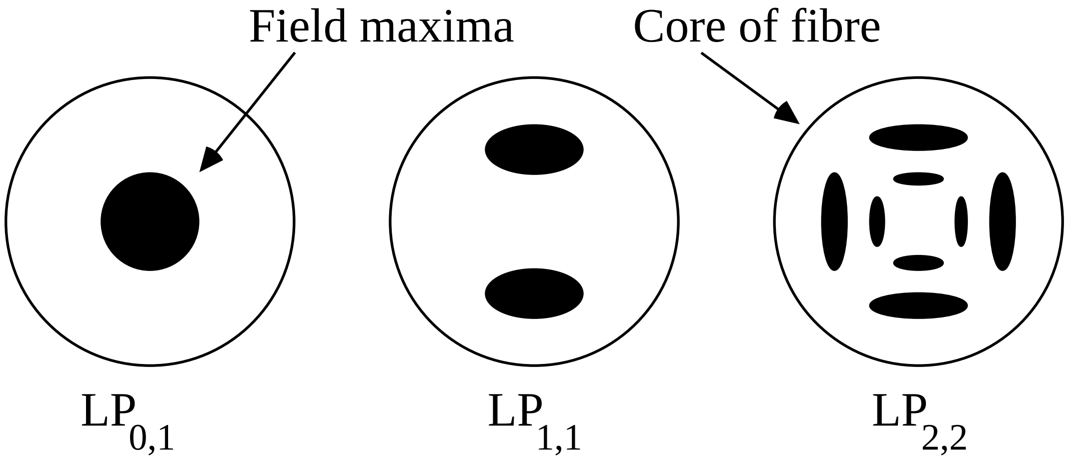
HE_{lm} and EH_{lm} [132]

In most fibre waveguides $n_1 \approx n_2$ and all the modes can be approximately represented by:

LP_{lm} [133]

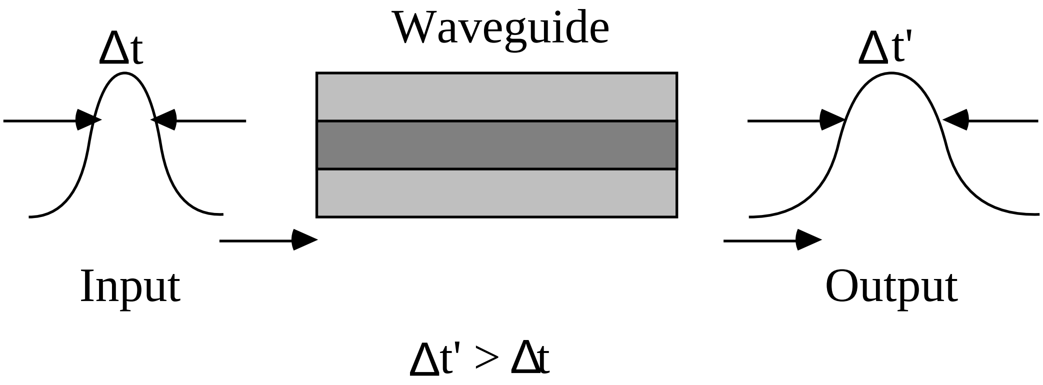
m is the no. of field max along the radius of the fibre (and relates to the angle θ_i)
2 l is the no. of field max along the circumference of the fibre (and relates to the degree of helicity)

Some example mode patterns



Dispersion

If we send a short pulse of light of duration Δt through a waveguide or fibre, we find that the output pulse has a duration Δt', where Δt' > Δt. This effect is known as dispersion.



Intermodal dispersion

Without knowing the detailed mode theory, it's hard to calculate a full expression for the dispersion.

However, we can make a calculated guess from knowing about the propagation of the modes.

The low order modes spend most of their 'time' travelling in the core, therefore:

$$V_{\varphi,\text{low order mode}} \approx \frac{c}{n_2} \qquad [140]$$

The high order modes spend a lot of their 'time' in the cladding, therefore:

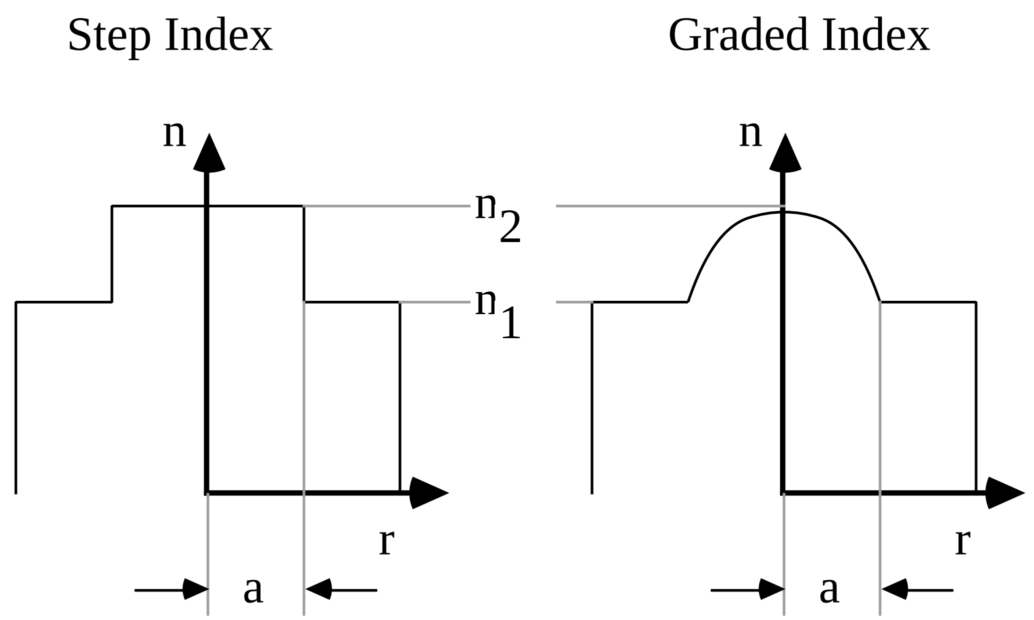
$$V_{\varphi,\text{high order mode}} \approx \frac{c}{n_1} \qquad [141]$$

For a short input pulse of duration Δt, launched into a guide of length L we can estimate the pulse duration on exit Δt' to be:

$$\Delta t' = \frac{Ln_2}{c} - \frac{Ln_1}{c} = \frac{L(n_2 - n_1)}{c} \quad [142]$$

for both slab waveguides and fibres, using a single mode guide removes this effect. Alternatively, a graded index fibre also shows reduced the dispersion.

Graded Index Fibre



To calculate the modes in a graded index fibre is even more complicated than for the step index fibre!

The ‘rays’ no longer follow zig-zag paths, instead, the gradual change in refractive index causes them to follow smooth curves through the fibre.

However, the resulting mode patterns are similar and are designated as before:

LP_{lm} [143]

m is the no. of field max along the radius of the fibre (and relates to the angle θ_i)

2 l is the no. of field max along the circumference of the fibre (and relates to the degree of helicity)

To gain an understanding into the intermodal dispersion, let’s consider two particular modes:

1) The on axis ray travels through the guide with a velocity governed by n_2

2) The helical ray has further to travel but does so in a region of lower refractive index and therefore travels more quickly.

With precise control of the refractive index profile the two modes can be made to propagate with the same velocity.

In practice, manufacturing tolerance and dispersion within the material make this condition difficult to maintain. However, intermodal dispersion can still be reduced by more than an order of magnitude by using graded index fibre.

Single Mode Fibre

For both slab waveguides and fibres, single mode operation is ensured by reducing the core dimension below a critical value, for slab waveguides we have

$$d\, k_0 \sqrt{n_2^{\,2} - n_1^{\,2}} \,/\, \pi \quad < 1 \qquad [116]$$

for fibres we have

$$a\, k_0 \sqrt{n_2^{\,2} - n_1^{\,2}} \,/\, 2.4 \quad < 1 \qquad [144]$$

where a = radius of core

However, for both form of guide, we will still get intramodal dispersion.

Intramodal Dispersion

Waveguide Dispersion
We showed for the slab waveguide, that different frequencies, propagate with slightly different mode

constant h & p. Hence, they have different phase velocities [95]. This is called waveguide dispersion.

Material Dispersion

In the case of well designed, single mode fibre, the real change of refractive index with wavelength is a far bigger effect.

$$\frac{\delta n_2}{\delta\lambda} \neq 0 \quad [145]$$

This has nothing to do with waveguide modes and would also be present in the bulk material. It is called material dispersion.

The velocity of a light pulse travelling in a material is given by the group velocity

$$V_g = \frac{\delta\omega}{\delta k} \quad [146]$$

Rewriting [146], for υ and λ, $(\omega = 2\pi\upsilon, k = \frac{2\pi}{\lambda})$ we get:

$$V_g = \frac{\delta\omega}{\delta\upsilon}\frac{\delta\upsilon}{\delta\lambda}\frac{\delta\lambda}{\delta k} = -\lambda^2\frac{\delta\upsilon}{\delta\lambda} \quad [147]$$

but $\upsilon = \frac{c}{\lambda n}$

therefore:

$$\frac{\delta\upsilon}{\delta\lambda} = \frac{\delta}{\delta\lambda}\left(\frac{c}{\lambda n}\right) = c\,\frac{-1}{\lambda^2 n^2}\left(n + \lambda\frac{\delta n}{\delta\lambda}\right) \quad [148]$$

sub into [147] and get:

$$V_g = \frac{c}{n}\left(1 + \frac{\lambda}{n}\frac{\delta n}{\delta\lambda}\right) \quad [149]$$

The finite bandwidth of the source $\Delta\lambda$, leads to a spread in group velocities ΔV_g, given by:

$$\Delta V_g = \frac{\delta V_g}{\delta\lambda}\,\Delta\lambda \qquad [150]$$

Using [149] we get:

$$\Delta V_g = \frac{c\lambda}{n^2}\left(\frac{\delta^2 n}{\delta\lambda^2} - \frac{2}{n}\left(\frac{\delta n}{\delta\lambda}\right)^2\right)\Delta\lambda \quad [151]$$

The duration $\Delta t'$ of a short pulse after passing through a length of material L is given by:

$$\Delta t' = \frac{L}{V_g}\,\frac{\Delta V_g}{V_g} \quad [152]$$

$$\Delta t' = \frac{L\lambda\left(\frac{\delta^2 n}{\delta\lambda^2} - \frac{2}{n}\left(\frac{\delta n}{\delta\lambda}\right)^2\right)\Delta\lambda}{c\left(1 + \frac{\lambda}{n}\frac{\delta n}{\delta\lambda}\right)^2} \quad [153]$$

The relative sizes of n, $\frac{\delta n}{\delta\lambda}$ & $\frac{\delta^2 n}{\delta\lambda^2}$ means that [153] simplifies to:

$$\Delta t' = \frac{L\lambda}{c}\left(\frac{\delta^2 n}{\delta\lambda^2}\right)\Delta\lambda \quad [154]$$

For pure silica (SiO_2), $\frac{\delta^2 n}{\delta\lambda^2}$ passes through zero at 1.3µm.

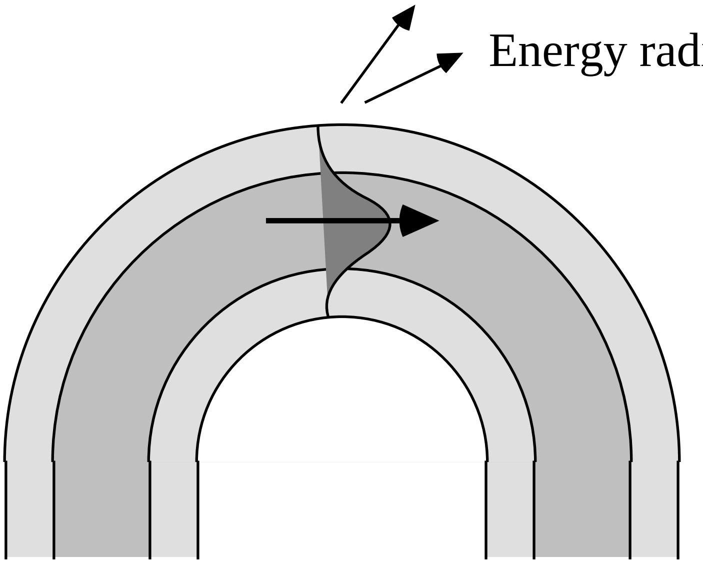
Therefore, at this wavelength there is no material dispersion. Alternatively, by control of the wavelength, deliberate material dispersion can be introduced to compensate for the waveguide dispersion!

Because fibres are used over large distance we need to consider not only the dispersion but the loss too.

Bending Loss

Loss can result from bends in the fibre.

As the mode travels round a bend we see that the energy on the ‘outside’ has to travel faster than defined by the mode velocity. The result is that this energy is radiated away from the guide.



If the radius of the bend is comparable to the fibre radius then large amounts of power can be coupled out. This can be used to make a fibre power meter.

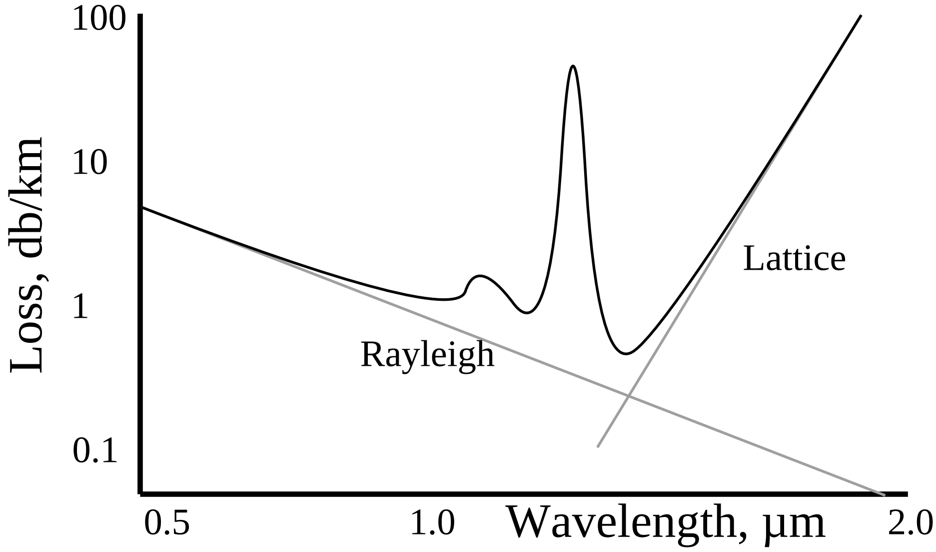
Intrinsic Loss

Even with a perfect fibre, losses are still present.

Rayleigh Scattering:
The fibres are made from glasses which are themselves composite materials (or at least have a disordered structure). This gives rise to small fluctuations in the refractive index which act as scattering centres. This is known as Rayleigh Scattering and scales as λ^{-4}.

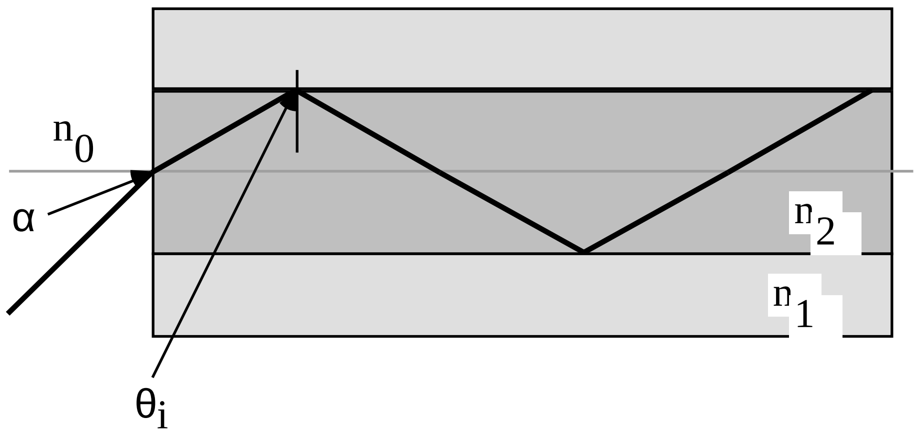
Absorption by Impurities:
Impurities in the fibre material lead to absorption, in particular - OH ions have strong absorptions.

Interaction with lattice vibrations:
As the wavelength of the light gets longer, photon-phonon interactions take place leading to a net absorption.



Coupling

Let’s look at a ray launched into a fibre



For a fibre mode to propagate, θ_i must be large enough to ensure total internal reflection, i.e.

$$\theta_i > \theta_c = \sin^{-1}\frac{n_1}{n_2} \qquad [160]$$

Remembering that $\sin^2 + \cos^2 = 1$, we have:

$$n_0 \sin\alpha_{max} = n_2\sqrt{1 - \sin^2\theta_c} \qquad [161]$$

$$n_0 \sin\alpha_{max} = n_2 \sqrt{1 - \frac{n_1^2}{n_2^2}} \quad [162]$$

$$\alpha_{max} = \sin^{-1}\left(\frac{\sqrt{n_2^2 - n_1^2}}{n_0}\right) \quad [163]$$

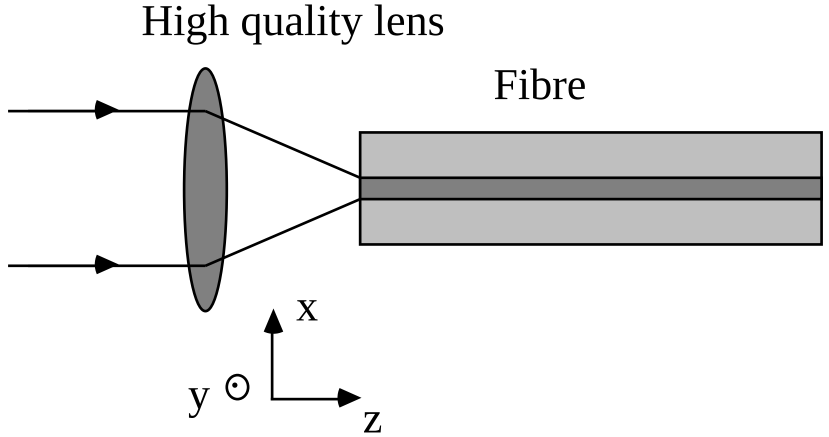
$\sqrt{n_2^2 - n_1^2}$ is defined as the numerical aperture of the fibre, therefore we can write [163] as:

$$\alpha_{max} = \sin^{-1}\left(\frac{NA}{n_0}\right) \quad [164]$$

α_{max} is the fibre acceptance angle.

Lens coupling

Light is usually coupled into a fibre using a lens. The focal length/aperture combination is selected so that the beam waist produced matches the size of the fibre core and α_{max} is not exceeded.

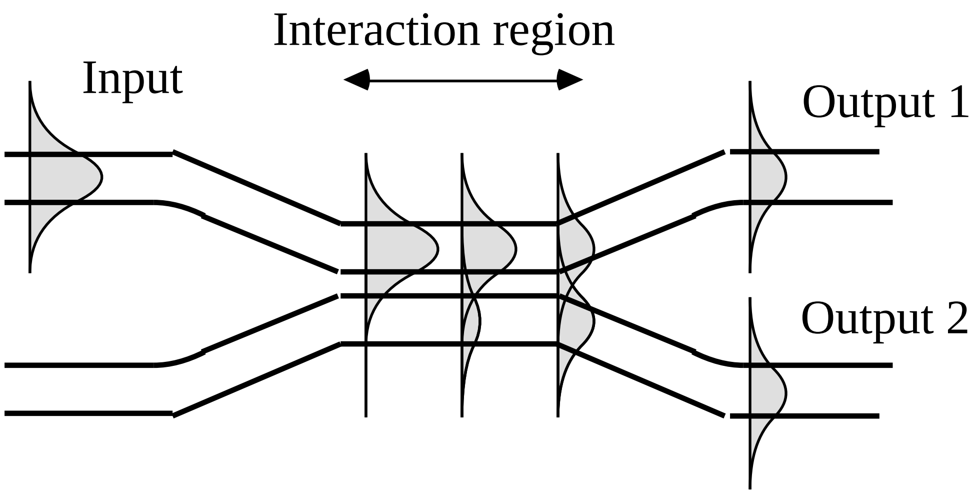


In the case of single mode fibre the x-y stability is of the order of 1µm (cf. fibre core size ≈ few µm's).

Auto fibre aligners are available which constantly adjust the x-y position of lens to maximise the coupled power.

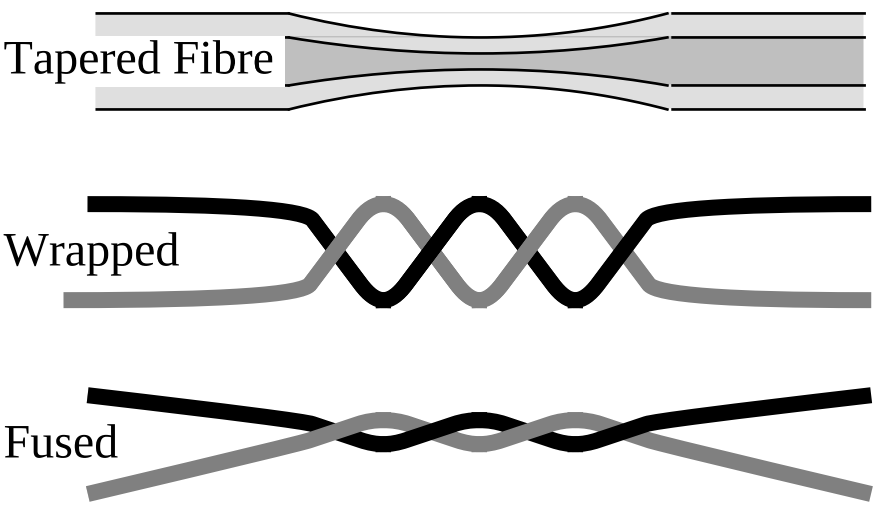
Evanescent wave couplers

With both slab waveguides and fibres it is possible to use the evanescent wave to couple energy between two neighbouring waveguides



For high efficiency, we need to match the values of β in the two guides (use identical guides).

Evanescent wave coupling is also possible with fibres. Two tapered fibres are ‘wrapped’ together and fused so that the two cores are coupled.



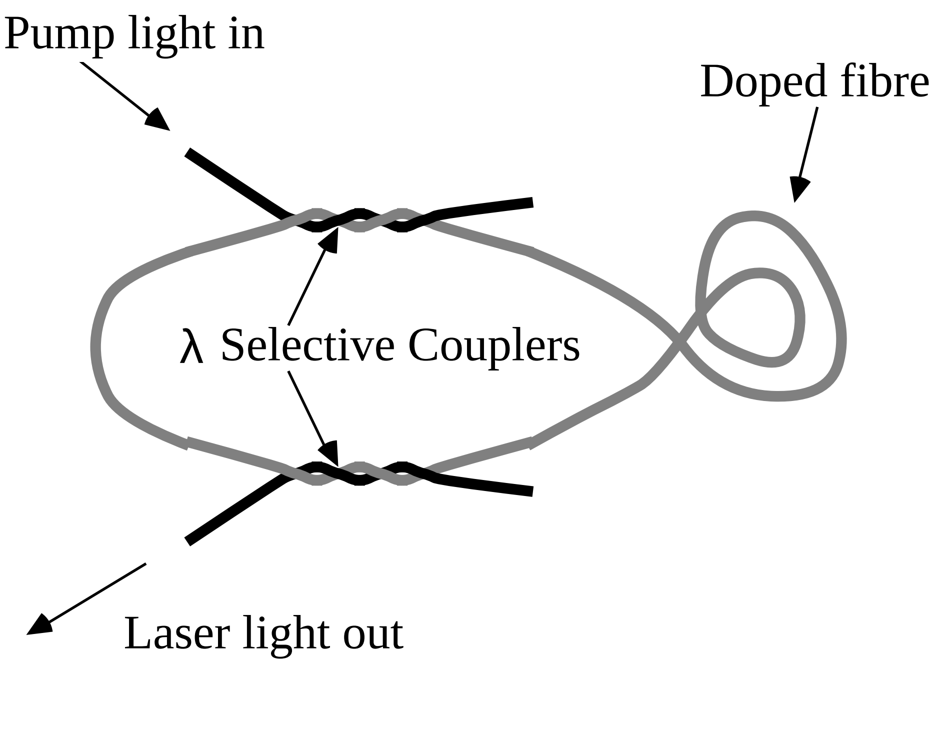
As with the slab waveguides the ratio of light at the two outputs is a function of the interaction length and the coupling coefficient K.

$$K = \frac{2h^2 p\, e^{-ps}}{\beta w(h^2+p^2)} \quad [170]$$

Note, the coupling coefficient is a function of β and hence λ. Therefore, we can make frequency dependent couplers.

Fibre Lasers

By introducing dopants into the fibre core and using frequency dependent couplers we can make fibre based laser systems.



Nonlinear Optics

When an electric field is applied to a dielectric material, the material becomes polarised (i.e. the electrons shift with respect to the nuclei). The resulting polarisation depends on the strength of the electric field and the dielectric susceptibility, χ

$$P = D - \varepsilon_0 E = \chi\, \varepsilon_0 E$$

where

$$D = \varepsilon_0\, \varepsilon_r\, E, \quad \text{i.e.} \quad \chi = \varepsilon_r - 1$$

In isotropic materials the above relationship may be scalar (i.e. P parallel to E). However, in general, crystals are anisotropic and the relationship between P and E is best described by a tensor.

$$\left(P_x, P_y, P_z\right) = \varepsilon_0 \begin{pmatrix} \chi_{1,1} & \chi_{1,2} & \chi_{1,3} \\ \chi_{2,1} & \chi_{2,2} & \chi_{2,3} \\ \chi_{3,1} & \chi_{3,2} & \chi_{3,3} \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix}$$

note that the polarisation in the x-direction depends on the applied fields in the x , y and z-directions.

Although the relationship between E and P is somewhat complicated it should be noted that it is a linear relationship. The internal field of the crystal is $\approx 10^{11}$V/m, one might expect it is only when $E_{applied} << E_{internal}$ that this linear relationship will hold.

At optical wavelengths, a power density of 10^{12}W/m^2 corresponds to an electric field strength of $\approx 10^9$V/m which compares to 100V/m for sunlight. We will assume that the relationship between E and P can be treated as a power series.

$$P_i = P_i^{DC} + \sum_j \chi_{ij} E_j^{\omega} + \sum_{jl} \chi_{ijl} \nabla_l E_j^{\omega} + \sum_{jl} \chi_{ijl} E_j^{\omega_1} E_l^{\omega_2}$$

$$+ \sum_{jlm} \chi_{ijlm} E_j^{\omega_1} E_l^{\omega_2} E_m^{\omega_3} + \sum_{jl} \chi_{ijl} E_j^{\omega_1} B_l^{\omega_2}$$

term 1, the DC component
- rectification of the light field, leaving the medium with a net DC polarisation

term 2, the normal linear response
- P is linear to and has the same frequency as E

term 3, the Curl E term
- responsible for optical activity

term 4, second order processes (rest of this course)
- proportional to the product of the two E-fields e.g.
- $\omega = \omega_1 + \omega_2$, sum frequency mixing
- $\omega_1 = \omega_2$, $\omega = \omega_1 + \omega_2$, second harmonic generation
- $\omega_2 = 0$, $\omega = \omega_1$, the Pockels effect
- $\omega = \omega_1 - \omega_2$, difference frequency mixing

term 5, third order response
- third harmonic generation
- Raman effect,
- optical Kerr effect
- four-wave difference frequency mixing

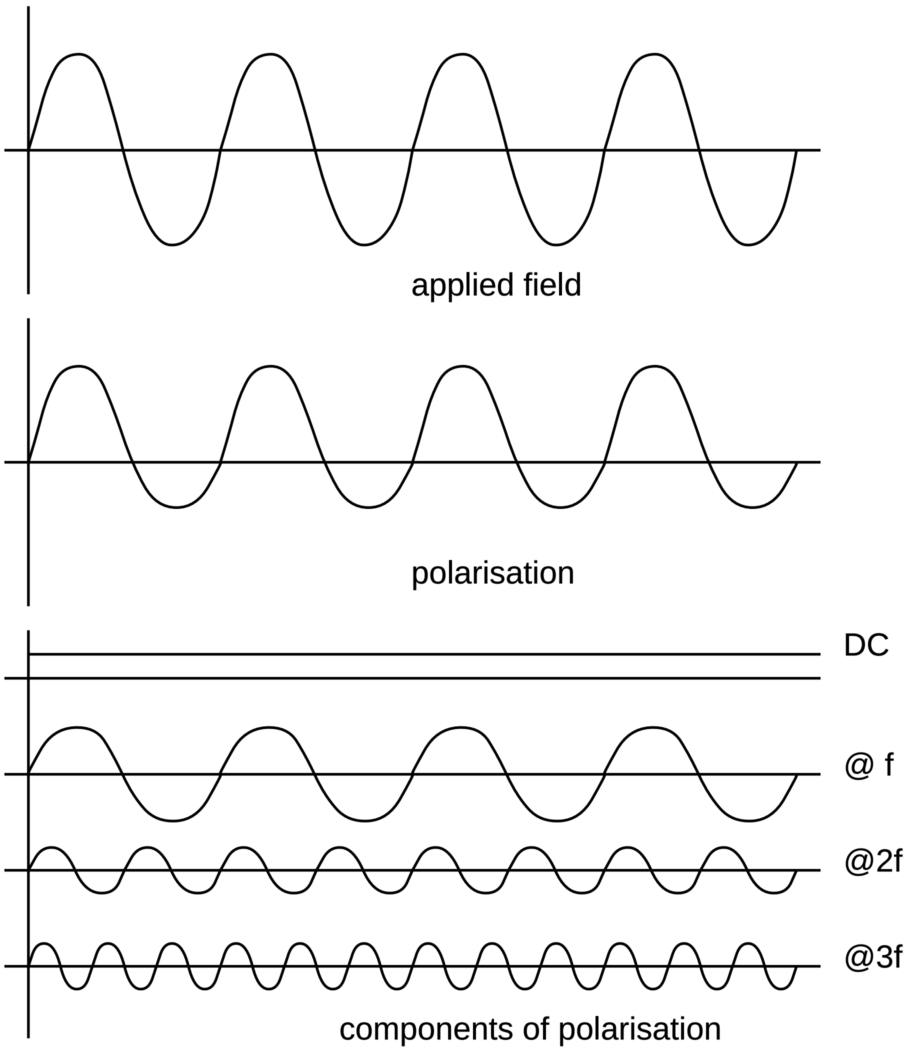
term 6, magneto optical effects
- $\omega_2 = 0$, the Faraday effect

For the remainder of this course we will consider the implications of term 4, the second order processes, i.e.

$$P_i = \sum_{jl} \chi_{ijl} E_j^{\omega_1} E_l^{\omega_2}$$

Generation of harmonic frequencies

As stated above, an electromagnetic wave polarises the medium through which it propagates. The polarisation oscillates at a fundamental frequency equal to that of the applied field. However, the nonlinear response implies that the oscillation of the polarisation cannot be described in terms of a pure sine wave.



The oscillating polarisation of the medium implies oscillating dipoles.

An oscillating electric dipole re-radiates light at the frequency of the oscillation and consequently the radiated light will have the same frequency components as the oscillations in the polarisation.

The form of the nonlinearity dictates the relative strength of the various components. For example, in crystals with a centre of symmetry the amplitude of the polarisation is symmetrical about E = 0 and consequently there are little/no second order effects. This makes the third order effects easier to observe. For high second order effects an anisotropic media is required, in which the polarisation about E = 0 is not symmetric.

The second order interactions are given by

$$P_i = \sum_{jl} \chi_{ijl} E_j^{\omega_1} E_l^{\omega_2}$$

E_j and E_l are sinusoidal, therefore in general we can write

$$P_i \alpha E_j \sin(\omega_1 t) \times E_l \sin(\omega_2 t)$$

This yields terms in

$$\sin\left((\omega_1 + \omega_2)t\right)$$

and $$\sin\left((\omega_1 - \omega_2)t\right)$$

The re-radiated light will have the same frequency components as the oscillations in the polarisation, i.e. the sum and difference frequency of the two incident fields.

The nonlinear coefficients

The second order contribution to the polarisation is given by

$$P_i = \sum_{jl} \chi_{ijl} E_j^{\omega_1} E_l^{\omega_2}$$

The constant of proportionality χ_{ijl} is a third rank tensor. It relates the polarisation in x, y and z-directions to the product of the incident fields in any combination of the x, y or z-direction. Hence the tensor has 27 components (3x3x3).

Rather than deal in terms of χ, most nonlinear crystal properties and experimental results are quoted in terms of the nonlinear optical coefficient d. The magnitude of the second order nonlinear optical coefficient can be expressed in terms of the product the magnitudes of χ(linear) at the three frequencies involved

$$d(\omega_3) = \frac{m a_2 \varepsilon_0^2}{N^2 e^3} \times \chi^{lin}(\omega_1)\chi^{lin}(\omega_2)\chi^{lin}(\omega_3).m / V$$

The constant of proportionality depends on the form of potential well and the electron number density. It is experimentally observed to exhibit a constant value over a wide range of materials (Miller's rule)

The polarisation in the x-direction can be related to the incident fields in the x, y and z-directions by

$$\begin{aligned} P_x(\omega_3) = {} & d_{xxx}(\omega_3)E_x(\omega_1)E_x(\omega_2) + d_{xyy}(\omega_3)E_y(\omega_1)E_y(\omega_2) \\ & + d_{xzz}(\omega_3)E_z(\omega_1)E_z(\omega_2) + d_{xzy}(\omega_3)E_z(\omega_1)E_y(\omega_2) \\ & + d_{xyz}(\omega_3)E_y(\omega_1)E_z(\omega_2) + d_{xzx}(\omega_3)E_z(\omega_1)E_x(\omega_2) \\ & + d_{xxz}(\omega_3)E_x(\omega_1)E_z(\omega_2) + d_{xxy}(\omega_3)E_x(\omega_1)E_y(\omega_2) \\ & + d_{xyx}(\omega_3)E_y(\omega_1)E_x(\omega_2) \end{aligned}$$

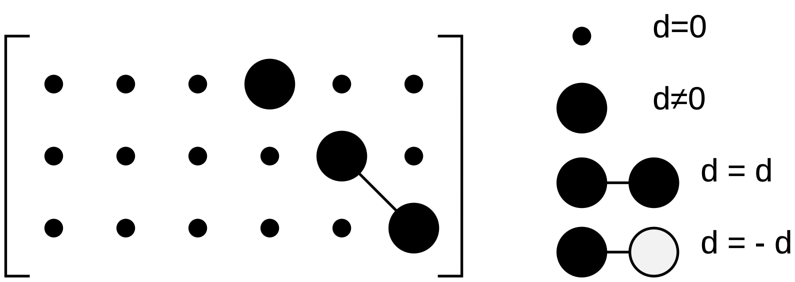
In a lossless medium, the order of the multiplication of the fields is not significant, therefore

$$d_{ijl} = d_{ilj}$$

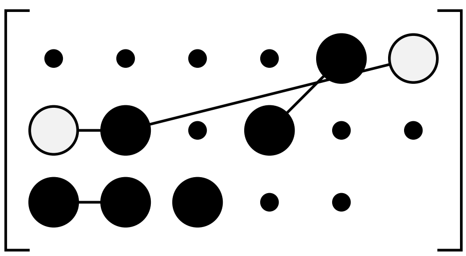
Consequently the matrix expression for P_i can be simplified (slightly) to

$$\begin{pmatrix} P_x \\ P_y \\ P_z \end{pmatrix} = \begin{pmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} \\ d_{21} & d_{22} & d_{23} & d_{24} & d_{25} & d_{26} \\ d_{31} & d_{32} & d_{33} & d_{34} & d_{35} & d_{36} \end{pmatrix} \begin{pmatrix} E_x(\omega_1)E_x(\omega_2) \\ E_y(\omega_1)E_y(\omega_2) \\ E_z(\omega_1)E_z(\omega_2) \\ E_y(\omega_1)E_z(\omega_2) + E_z(\omega_1)E_y(\omega_2) \\ E_x(\omega_1)E_z(\omega_2) + E_z(\omega_1)E_x(\omega_2) \\ E_x(\omega_1)E_y(\omega_2) + E_y(\omega_1)E_x(\omega_2) \end{pmatrix}$$

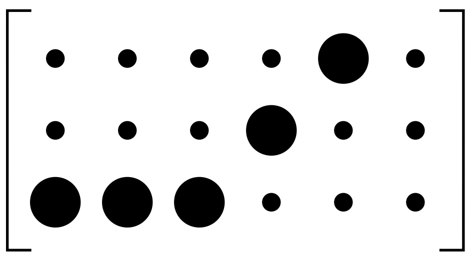
The symmetry of the crystal class determines which of the d elements are non-zero. For example, of the 32 point groups 12 have inversion symmetry and the second order nonlinear coefficients are zero. Of the remaining 20, further symmetry arguments can be applied to obtain relationships between the various d elements, for example.



$\bar{4}2m$ eg. KDP



3m eg. Quartz, $LiNbO_3$

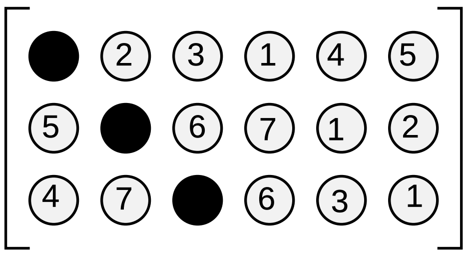


mm2 eg. KTP, LBO

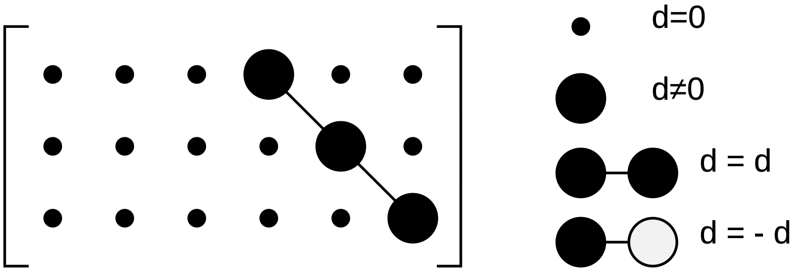
If the nonlinear polarisation is due solely to electronic motion and there are no absorptions with the range of ω_1, ω_2 and ω_3 all frequencies act in the same manner a further simplification can be made

$$d_{ijl} = d_{ilj} = d_{lij} = d_{lji} = d_{jil} = d_{jli}$$

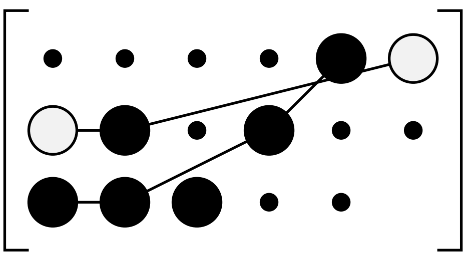
This is called Kleinman's symmetry condition and further simplifies the d tensor. Under Kleinman symmetry we have the following equalities between the d elements



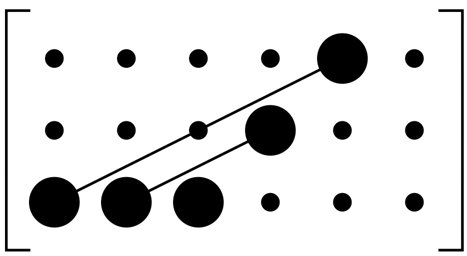
When Kleinman symmetry is applied to the previous crystal classes we get



$\bar{4}$2m eg. KDP



3m eg. Quartz, $LiNbO_3$



mm2 eg. KTP, LBO

The strength of any one interaction geometry is related to a particular d element e.g. the strength of $P_x(\omega_3)$ given an incident $E_x(\omega_1)$ and $E_z(\omega_2)$ depends on d_{15}. Note that $\bar{4}2m$ $d_{15} = 0$. Therefore KDP would be unsuitable for this geometry.

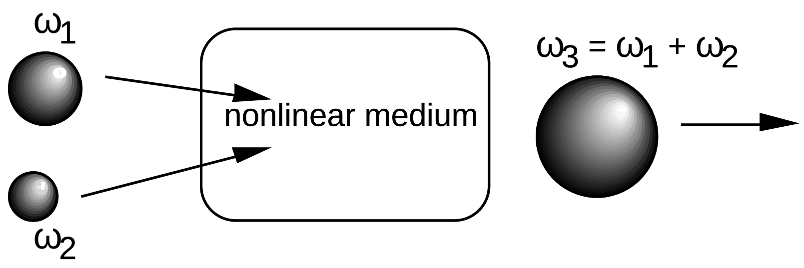
In general, the tensor representation of d can be dropped and replaced with a single scalar quantity called effective d, d_{eff}. d_{eff} is calculated from d with reference to the interaction geometry.

Conservation of energy

The conservation of energy is a general principle in physics and must be true within nonlinear optics.

The simplest interpretation is within a photon picture of nonlinear interactions. The energy of a single photon is given by $\hbar\omega$. Consider the case of sum frequency mixing,

$$\omega_3 = \omega_1 + \omega_2$$



By considering the nonlinear interaction in terms of photons we see that energy is conserved providing we assume that two photons of frequencies ω_1 and ω_2 “join” within the nonlinear medium to produce one photon of frequency ω_3.

Constraints within nonlinear interactions

We can see from the relationship between the polarisation and the applied fields

$$P_i = \sum_{jl} \chi_{ijl} E_j^{\omega_1} E_l^{\omega_2}$$

that a large number of second order interactions are possible within a particular crystal. Even allowing the restriction on the values of the d elements imposed by symmetry considerations a range of interactions are possible. For example given input fields of ω_1 and ω_2, will the output be at $(\omega_1 + \omega_2)$ or $(\omega_1 - \omega_2)$?

The additional constraint which we have not yet considered is the conservation of momentum. Within nonlinear optics this is termed phase-matching.

Phase-matching

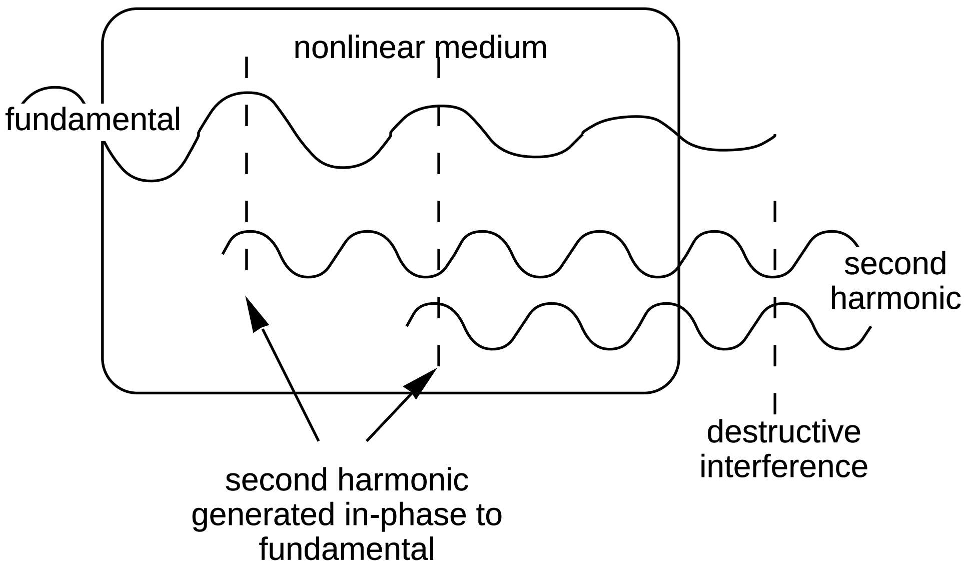
Let us consider the case of second harmonic generation,

$$\omega_1 = \omega_2 = \omega$$

$$\omega_3 = 2\omega$$

A wave of frequency ω propagates in the z-direction through the crystal. At all z positions energy is transferred into a wave at 2ω. For maximum efficiency, we require that all the newly generated light interferes constructively at the exit face of the crystal.

The nature of the nonlinear interaction ensures that at the point of generation, the generated light (the second harmonic) has a well defined phase relationship with respect to the incident light (the fundamental).



However, dispersion within the crystal means that the fundamental and second harmonic light travel with different phase velocities. It follows that the second harmonic light generated at a second position within the crystal may not be in phase with the second harmonic light generated earlier.

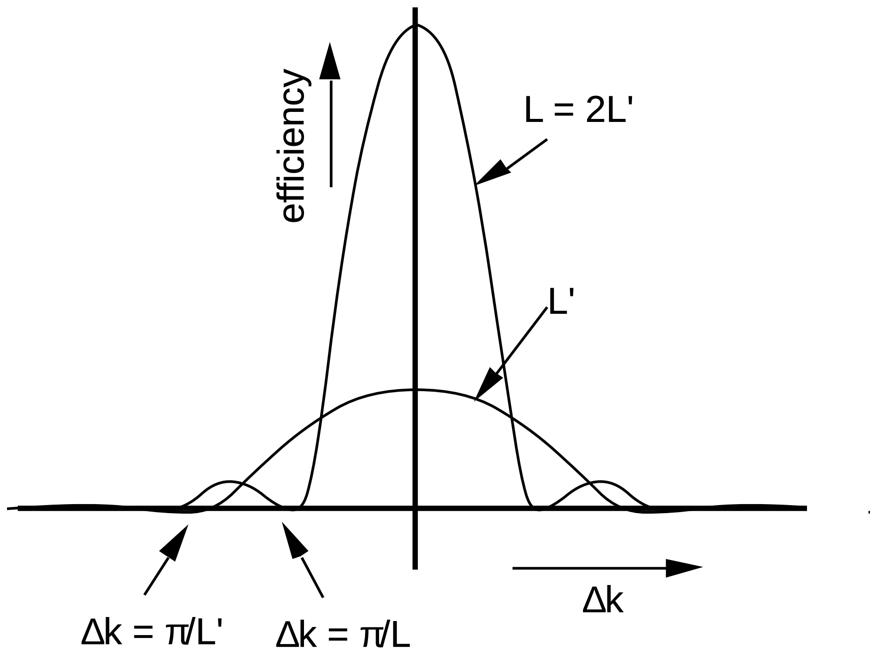
Clearly, the "phase-matching" of the second harmonic generation becomes more difficult to maintain as the crystal becomes longer.

The effect of imperfect phase-matching on the efficiency of the nonlinear process is given by

$$\text{efficiency} \; \alpha \; L^2 \left(\frac{\sin^2(\Delta kL/2)}{(\Delta kL/2)^2} \right)$$

where

$$\Delta k = k_3 - k_2 - k_1 \quad (c = \omega/k)$$



For high overall efficiency we need a large L and a small ΔkL, ie. $\Delta k \approx 0$.

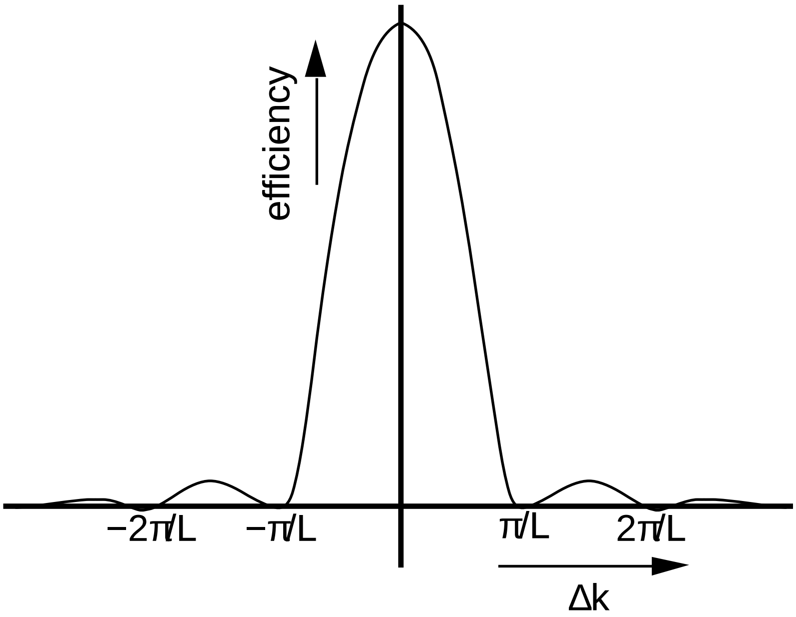
The momentum of a photon is $\hbar k$. Therefore the $\Delta k = 0$ condition implies that

momentum of the generated photon
= the sum of the momentum of the incident photons

Hence the phase-matching condition is equivalent to the conservation of momentum.

Coherence length

The graph of conversion efficiency against Δk shows well defined minima, where the overall efficiency falls to zero.



Consider the first minima at $\Delta k = \pi/L$. It follows that for a given Δk there will be no nonlinear output for a crystal length of

$$L_c = \pi/\Delta k$$

This is because the relative phases of the generated light throughout the length of the crystal has resulted in complete destructive interference. The length of crystal for which this occurs is called the coherence length.

Returning to our earlier equation for the efficiency of the nonlinear conversion

$$\text{efficiency} \propto L^2 \left(\frac{\sin^2(\Delta k L/2)}{(\Delta k L/2)^2} \right)$$

we see that for a given value of Δk we have

$$\text{efficiency} \propto \frac{1}{\Delta k^2} \sin^2(\Delta k L/2)$$

The nonlinear output varies sinusoidally with the length of the crystal with a period given by $2\pi/\Delta k$

The coupled differential equations for the interactions between fields

The second order nonlinear interaction between the three fields can be written in terms of three coupled differential equations.

Assuming that each of the three fields (propagating in the z-direction) can be written in the form

$$E_i = E0_i \exp i(\omega_i t - k_i z)$$

the three coupled equations are

$$\frac{dE0_1}{dz} = i\,\kappa_1\, E0_2^*\, E0_3 \exp(i\,\Delta kz)$$

$$\frac{dE0_2}{dz} = i\,\kappa_2\, E0_1^*\, E0_3 \exp(i\,\Delta kz)$$

$$\frac{dE0_3}{dz} = i\,\kappa_3\, E0_2\, E0_3 \exp(-i\,\Delta kz)$$

where

$$\kappa_i = \omega_i d_{eff} / n_i\, c$$

The direction of energy flow ie. 1 & 2 to 3 of 3 to 1 & 2 depends on the relative phase of the three fields. These equations can be used to calculate conversion efficiencies and threshold pump powers for nonlinear optical experiments.

Methods for phase-matching

For very short lengths of non linear material phase-matching is not important since the product ΔkL is still small. However, for most nonlinear systems this length of material is insufficient to observed efficient non linear interactions. Phase-matching is important because

- increases efficiency of nonlinear process
- acts to select the nonlinear process of interest

In all nonlinear materials, dispersion in the phase velocity (a refractive index which depends on wavelength) ensures that the requirements for energy conservation

$$\omega_3 = \omega_1 + \omega_2$$

and momentum conservation

$$\Delta k = k_3 - k_2 - k_1 = \frac{1}{c}\left(n_3\omega_3 - n_2\omega_2 - n_1\omega_1\right) = 0$$

cannot be satisfied simultaneously.

A number of methods for phase-matching can be employed

- dielectric waveguide phase-matching
- non colinear phase-matching
- birefringent phase-matching
- quasi phase-matching

One of the surprising points to arise from the tensor nature of the nonlinear coefficient is the fact that the generated light need not have the same polarisation state as the incident light. Phase-matching types are characterised in term of the relative polarisation states of the three fields. This is particularly significant with respect to birefringent phase-matching.

Type-I phase-matching ($\omega_3 = \omega_1 + \omega_2$)

- $E(\omega_1)$ and $E(\omega_2)$ have parallel polarisations
- $E(\omega_3)$ is orthogonally polarised with respect to $E(\omega_1)$ and $E(\omega_2)$

Type-II phase-matching ($\omega_3 = \omega_1 + \omega_2$)

- $E(\omega_1)$ and $E(\omega_2)$ have orthogonal polarisations
- $E(\omega_3)$ has parallel polarisation with respect to $E(\omega_1)$ or $E(\omega_2)$

Dielectric waveguide phase-matching

The propagation of light within a dielectric waveguide is more complicated than that through bulk material.

The transverse boundary conditions within the solution of Maxwell's equations lead to allowed transverse modes of propagation. These modes have characteristic intensity profiles (which decay to zero outside the guiding layer) and travel with a phase velocity which differs from that of the bulk material.

Assuming that the refractive indices of the guiding and cladding layers are n_g and n_c respectively then the effective mode index, N, of the transverse mode can take on a range of values

$$n_c < N < n_g$$

Different modes within the guide have different values of N. By appropriate design of the waveguide geometry, the mode index of two or more different modes with different frequencies can made such that

$$N_{\text{mode A}}\,(@\ \omega_3) - N_{\text{mode B}}\,(@\ \omega_2) - N_{\text{mode C}}\,(@\ \omega_1) = 0$$

ie. phase-matching.

Birefringent phase-matching

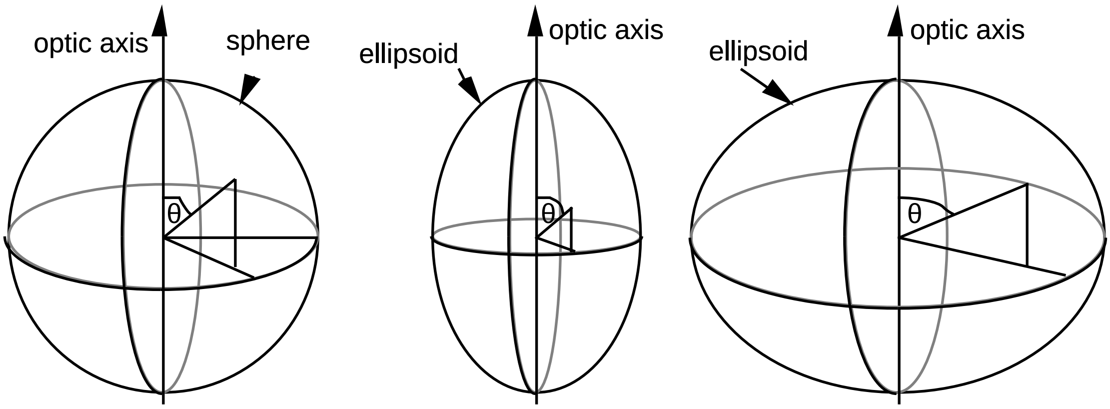
The 20 crystal classes for which $d_{eff} \neq 0$ can be broken down into three groups

- isotropic
  cubic (23 and $\bar{4}2m$)
- anisotropic, uniaxial (a single symmetry axis)
  trigonal (3, 32, 3m)
  tetragonal (4, $\bar{4}$, 422, 4mm, $\bar{4}2m$)
  hexagonal (6, $\bar{6}$, 622, 6mm, $\bar{6}m2$)
- anisotropic, biaxial (two symmetry axes)
  triclinic (1)
  monoclinic (2, m)
  orthorhombic (222, mm2)

The isotropic crystals exhibit a refractive index which is independent of polarisation state and direction of propagation of the light through the crystal.

The anisotropic crystals exhibit birefringence (double refraction) ie. a refractive index which depends both on polarisation state and direction of propagation. Strictly speaking, birefringence is a term applied to uniaxial crystals, but a similar argument to that presented below can also be applied to biaxial crystals.

The refractive index experienced by a light wave within a birefringent crystal depends both on the polarisation state of the light with respect to the symmetry axis (optic axis) and the direction of propagation.



n, as a function of propagation direction for the o-ray

n, as a function of propagation direction for the e-ray, $n_e < n_o$ (-ve birefringence)

n, as a function of propagation direction for the e-ray, $n_e > n_o$ (+ve birefringence)

Ordinary-wave

- an o-ray is polarised perpendicularly to the optic axis of the crystal
- refractive index, n_o, is independent of the propagation direction, θ

Extraordinary-wave

- an e-ray has a component of its polarisation parallel to the optic axis
- refractive index, n_{eff}, is dependent on the propagation direction, θ
- $$\frac{1}{n_{eff}^2} = \frac{\cos^2\theta}{n_o^2} + \frac{\sin^2\theta}{n_e^2}$$
- usually n_{eff} is termed n_e

In all optically transparent crystals, the dispersion is such that the refractive index increases with increasing ω. The basic principle of birefringent phase-matching is to use the change in refractive index with polarisation to compensate for the dispersion within the nonlinear materials. The phase-matching geometries can be divided into type-I and type-II.

Positive birefringence ($n_e > n_o$)

- type-I $\omega_3 n_o(\omega_3) = \omega_1 n_{eff}(\omega_1) + \omega_2 n_{eff}(\omega_2)$

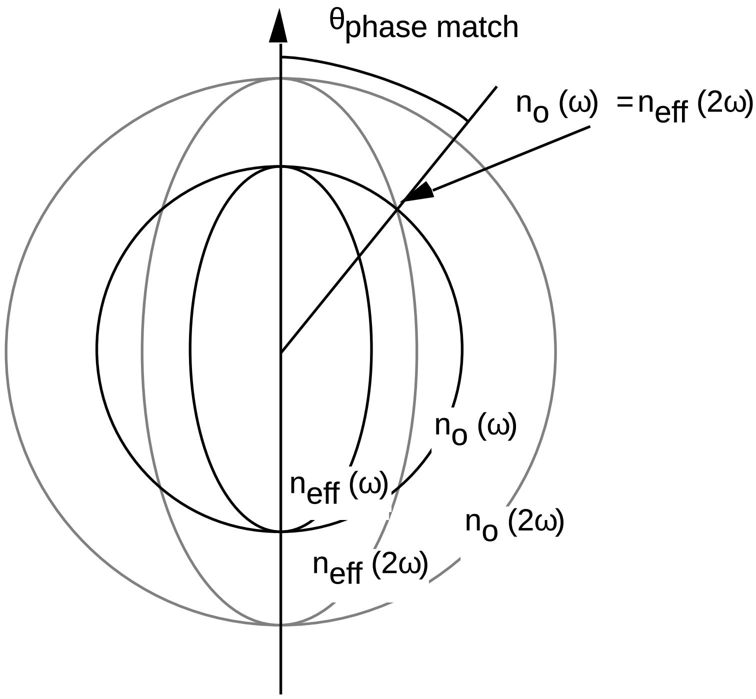
- type-II $\omega_3 n_o(\omega_3) = \omega_1 n_o(\omega_1) + \omega_2 n_{eff}(\omega_2)$

Negative birefringence ($n_e < n_o$)

- type-I $\omega_3 n_{eff}(\omega_3) = \omega_1 n_o(\omega_1) + \omega_2 n_o(\omega_2)$
- type-II $\omega_3 n_{eff}(\omega_3) = \omega_1 n_{eff}(\omega_1) + \omega_2 n_o(\omega_2)$

Angle-tuned phase-matching

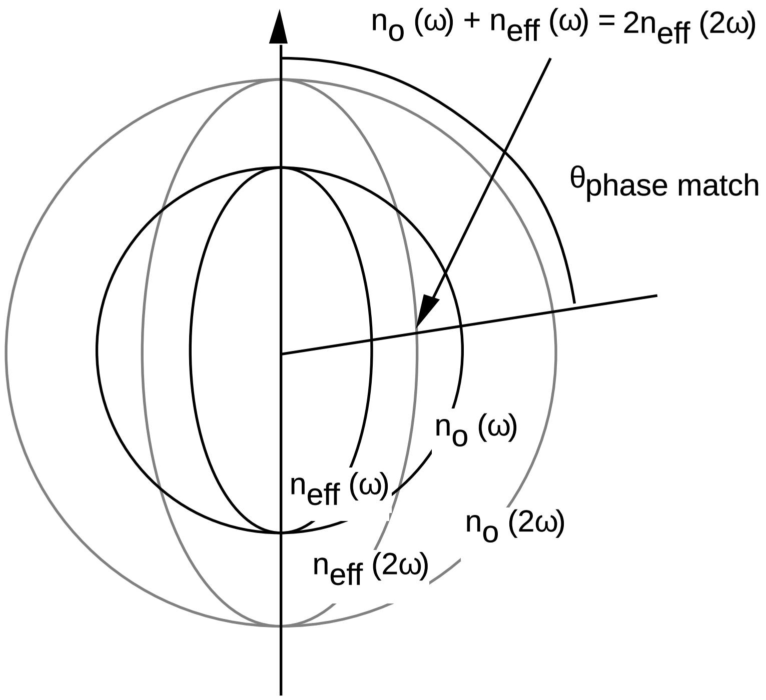
One approach to birefringent phase-matching is to set the orientation of the crystal such that one of the above conditions is satisfied. Taking frequency doubling in KDP as an example and reducing the 3-D representation of the index ellipsoid to 2-D we have



For a fundamental wavelength of 1064nm, type-I, phase-matched frequency doubling in KDP occurs for a propagation direction of θ = 42.4°, ie.

$$2\omega\, n_{eff}(2\omega, @\theta=42.4°) = 2(\omega\, n_o(\omega))$$

Type-II frequency doubling is more complex but can also be represented on a similar diagram.



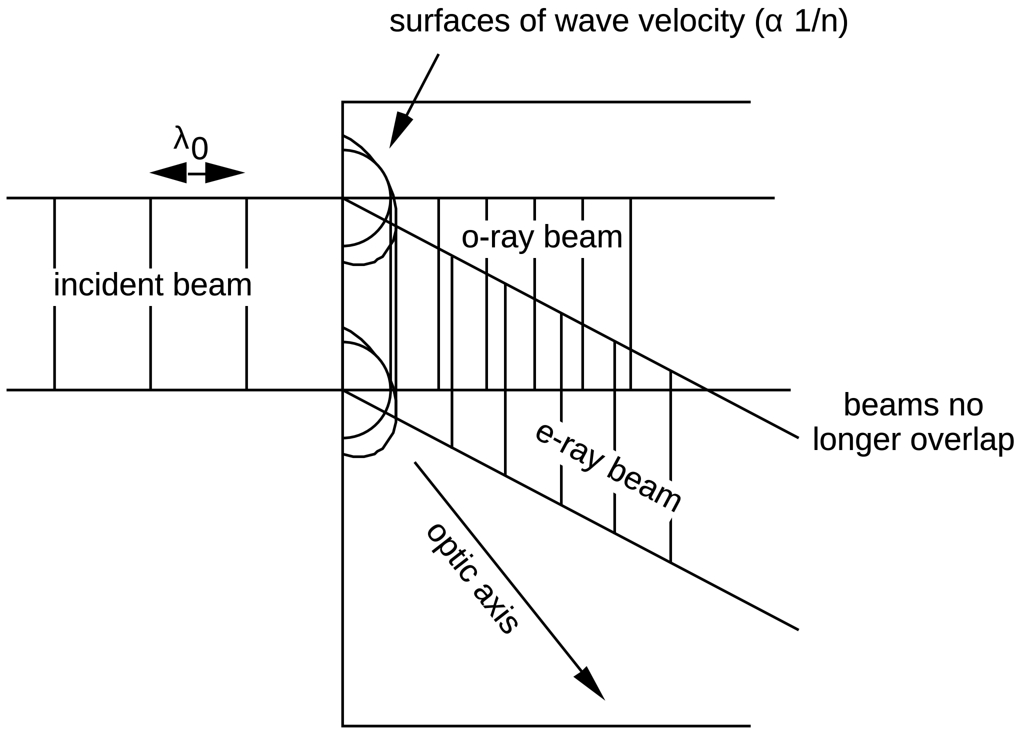
For a fundamental wavelength of 1064nm, type-II, phase-matched frequency doubling in KDP occurs for a propagation direction of θ close to 90°, ie.

$$2\omega\, n_{eff}(2\omega) = \omega\, n_o(\omega) + \omega\, n_{eff}(\omega)$$

However, in the case of KDP the d_{eff} for this combination of E-fields is zero and hence this particular configuration is unsuitable.

Walk-off

One limitation of angle-tuned phase-matching is walk-off. In birefringent materials the elliptical nature of the variation in n_{eff} with angle results in an angular separation between the o-ray and e-ray as they propagate through the crystal.



Consequently since all birefringent phase-matched geometries involve combinations of e-rays and o-rays, the beam overlap can only be maintained for a short distance into the nonlinear crystal.

This problem is made worse by the desire to focus the optical radiation to a small spot within the crystal to maximise the field strength and the resulting nonlinear interaction.

When $\theta \neq 0° \neq 90°$, angle-tuned, birefringent phase matching is termed critical phase-matching.

Noncritical phase-matching

A special case of angle-tuned, phase-matching occurs when $\theta = 90°$. This is termed noncritical phase-matching (NCPM). When the e-ray propagates with $\theta=0$ (perpendicularly to the optic axis) there is no walk off between the o-ray and e-ray. Consequently there is no limit to the length over which the interacting beams can be made to overlap and the efficiency of the nonlinear interactions can be increased.

Obviously, the chance of finding a crystal that exactly satisfies the condition (e.g. for type-I phase-matching in a +ve birefringent crystal)

$$\omega_3 n_o(\omega_3) = \omega_1 n_e(\omega_1 \ @ \ \theta = 90°) + \omega_2 n_e(\omega_2 \ @ \ \theta = 90°)$$

is extremely remote. Some fine tuning of n_{eff} is inevitable.

The most common way of fine tuning the refractive index is changing the temperature of the nonlinear material. If the temperature coefficients of n_o and n_e are non zero and different then phase matching with θ =90° can be achieved by setting the crystal temperature accordingly.

For example, noncritical, temperature-tuned, type-I phase-matching can be obtained for frequency doubling light at 1064nm in $MgO{:}LiNbO_3$ at ≈ 100°C, ie.

$$2\omega n_{eff}(2\omega, @\theta=90°, T\approx 100°C) = 2(\omega n_o(\omega, @T\approx 100°C))$$

Non colinear phase-matching

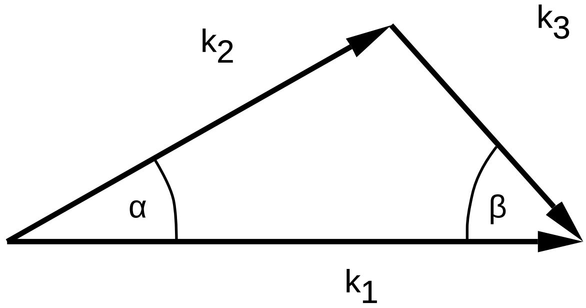
As stated earlier, the condition for perfect phase matching is

$$\Delta k = k_3 - k_1 - k_2 = 0$$

So far we have considered this to be a scalar relationship. However, k is the wavevector and consequently should be considered in vector form, i.e.

$$\overline{k_3} - \overline{k_1} - \overline{k_2} = 0$$

Clearly, non colinear solutions to this equation are possible, e.g.



applying the cosine law we get

$$k_3^2 = k_2^2 + k_1^2 - 2k_1k_2 \cos\alpha$$

re-arranging for α

$$\cos \alpha = \frac{n_1^2k_1^2 + n_2^2k_3^2 - n_3^2k_3^2}{2n_1\omega_1n_2\omega_2}$$

This method of phase matching is particularly useful for materials which are not birefringent, e.g. GaAs.

Quasi phase-matching

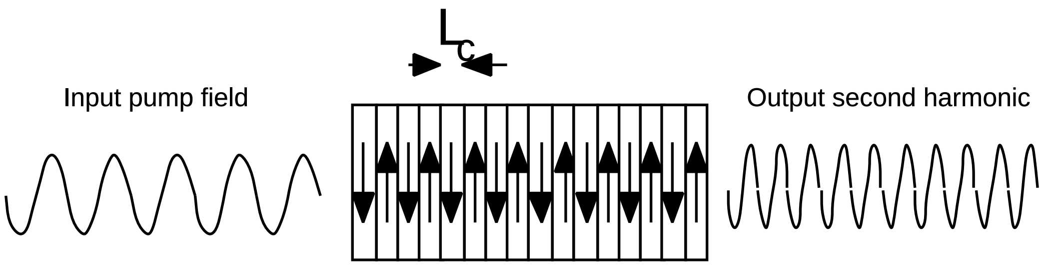
Until recent years birefringent phase matching has been by far the most common form of phase matching in both commercial and laboratory laser systems. However, birefringent phase matching imposes a number of limitations

- only works for birefringent materials
- only works for certain geometries
- may require angle tuning to a critical phase matching condition and associated walk off

Over the past couple of years another technique has been developed, quasi phase matching.

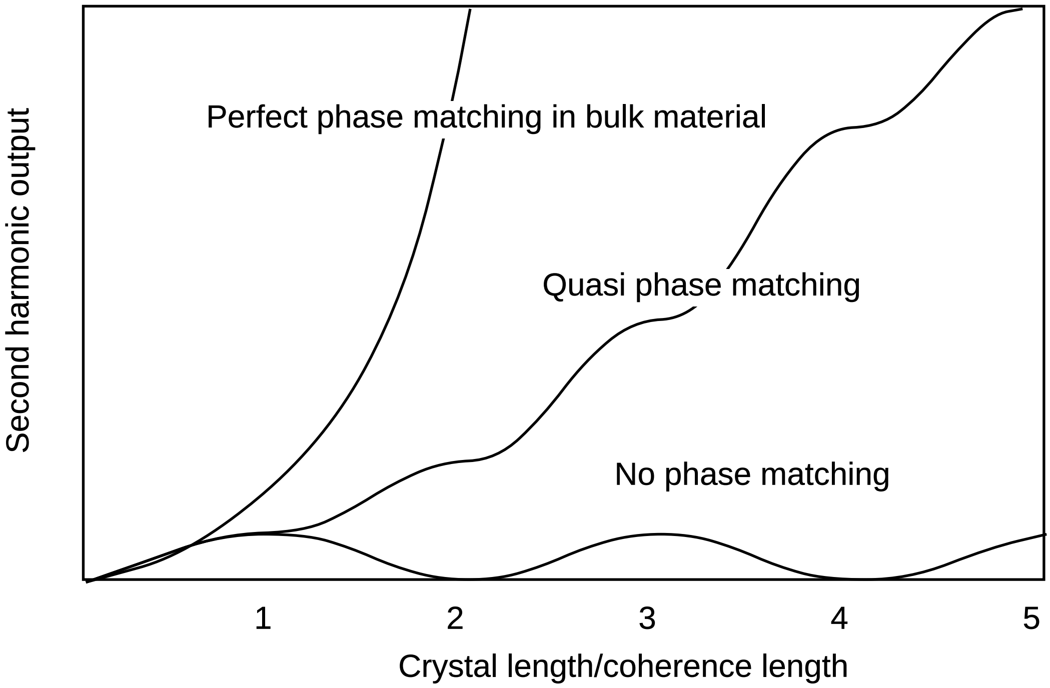
Consider a short length of nonlinear crystal used for frequency doubling. For all lengths upto the coherence length, L_c, no phase matching is required. Beyond L_c the pump and generated waves become out of phase, the direction of energy flow is reversed and the net amount of second harmonic light is reduced.

In quasi phase matching the crystal is effectively divided into sections L_c long. By turning each successive section upside down the phase relation between the pump and the second harmonic can be maintained.



Quasi phase matching can give high efficiency over a long length of crystal

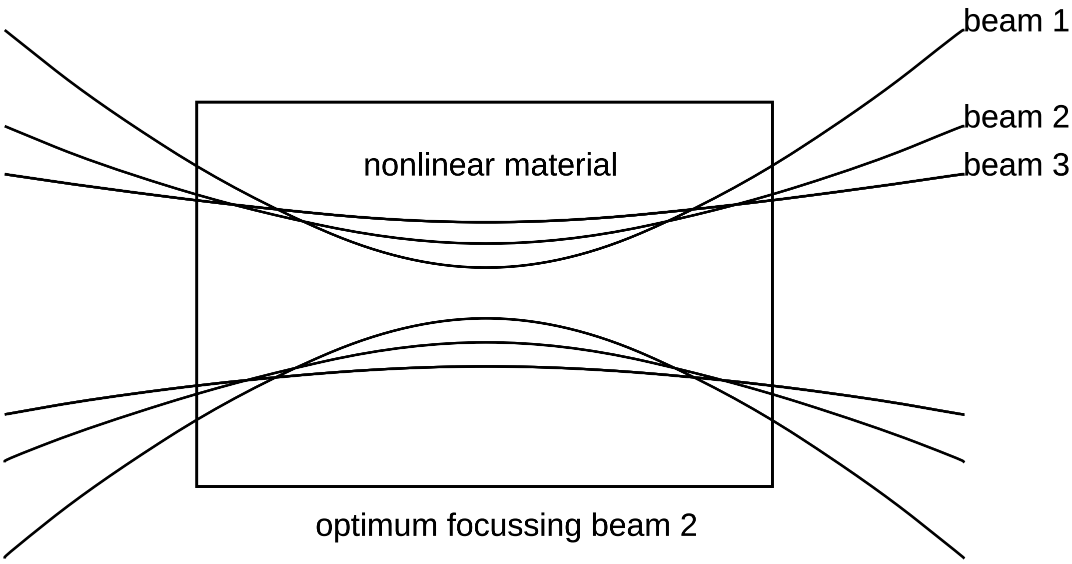
Quasi phase matching (QPM) can also be applied to OPO’s and other nonlinear interactions.
Rather than physically cut the crystal into short sections it is usual to process bulk crystals to orient the crystal structure. This orientation is frequently performed using an electrode pattern on the crystal and applying an electric field at elevated temperatures. After cooling the crystal is “poled” and no further fields are required. Such crystal are said to be “periodically poled". Amonst others Lithium niobate and KTP have both been widely used (PPLN, PPKTP).



Optimum focussing for nonlinear interactions

As a general rule, tight focussing of the interacting beams results in a high power density and consequently high nonlinear efficiency. However, two factors may limit the degree of focussing

- high power densities may damage the nonlinear material (this is only a practical limit and should not be confused with the following)
- the divergence of a tightly focussed beam limits the length of crystal over which the high power density can be maintained



- beam 3 is not tightly focussed, the resulting power density is reduced leading to low nonlinear conversion
- beam 2 is correctly focussed giving a reasonable power density over the entire length of the crystal
- beam 1 is very tightly focussed giving a high power density at the centre of the crystal. However, the divergence of the beam means that the power density at the edges of the crystal is reduced

Optimum focussing can be shown to be when the beams are focussed to the centre of the crystal such that

$$\sqrt{2} \text{ x beam dia.}_{\text{crystal centre}} = \text{beam dia.}_{\text{crystal edge}}$$

This is called confocal focussing.

Optical Parametric Oscillators (OPO)

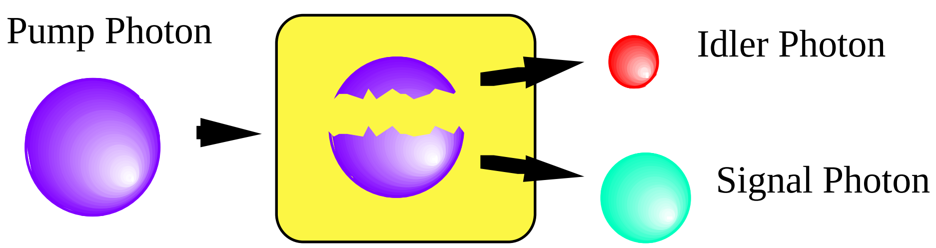
Another example of a second order nonlinear process is optical parametric down conversion. In optical parametric down-conversion, an input pump wave at frequency ω_3 (ω_{pump}) is converted into two outputs at frequencies ω_1 and ω_2, these are termed the signal and idler frequencies ($\omega_{sig} > \omega_{idl}$). As with other nonlinear interactions, these frequencies must obey the energy conservation relation, i.e.

$$\omega_{pump} = \omega_{sig} + \omega_{idl}$$

For a given ω_{pump} the ω_1 and ω_2 are determined by the phase-matching within the crystal, ie.

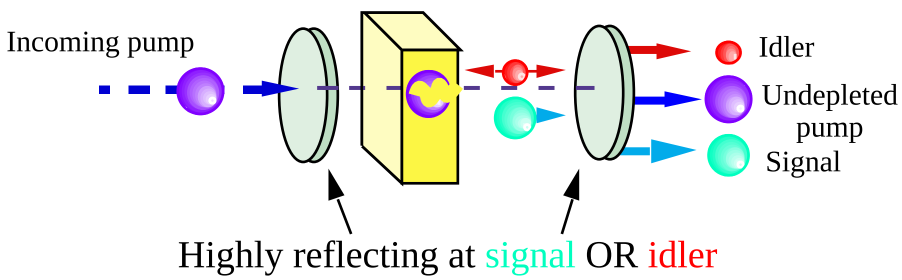
$$k_{pump} - k_{sig} - k_{idl} = \Delta k = 0$$

Within a photon picture, the OPO can be thought of as a photon splitter

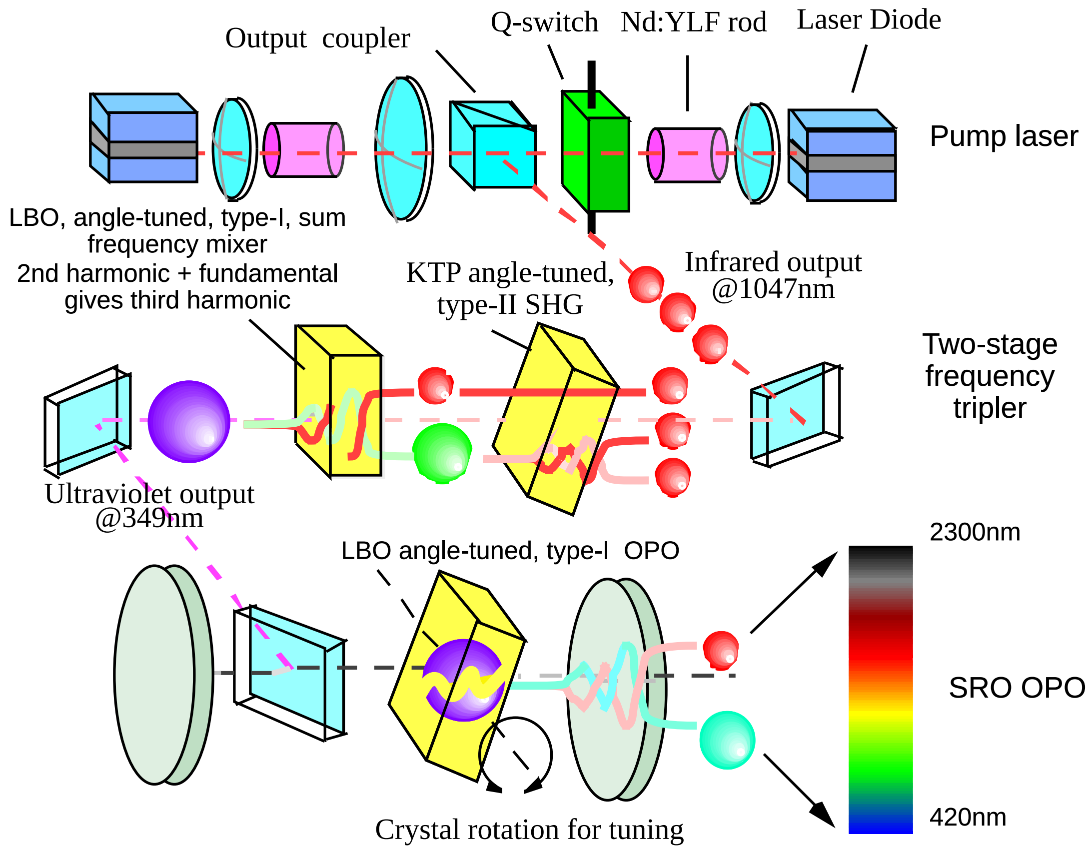


To form an optical parametric oscillator, resonance is provided by feedback at cavity mirrors for either, or both, the signal and idler frequencies. This significantly increases the efficiency of the conversion from pump frequency to signal and idler frequencies.

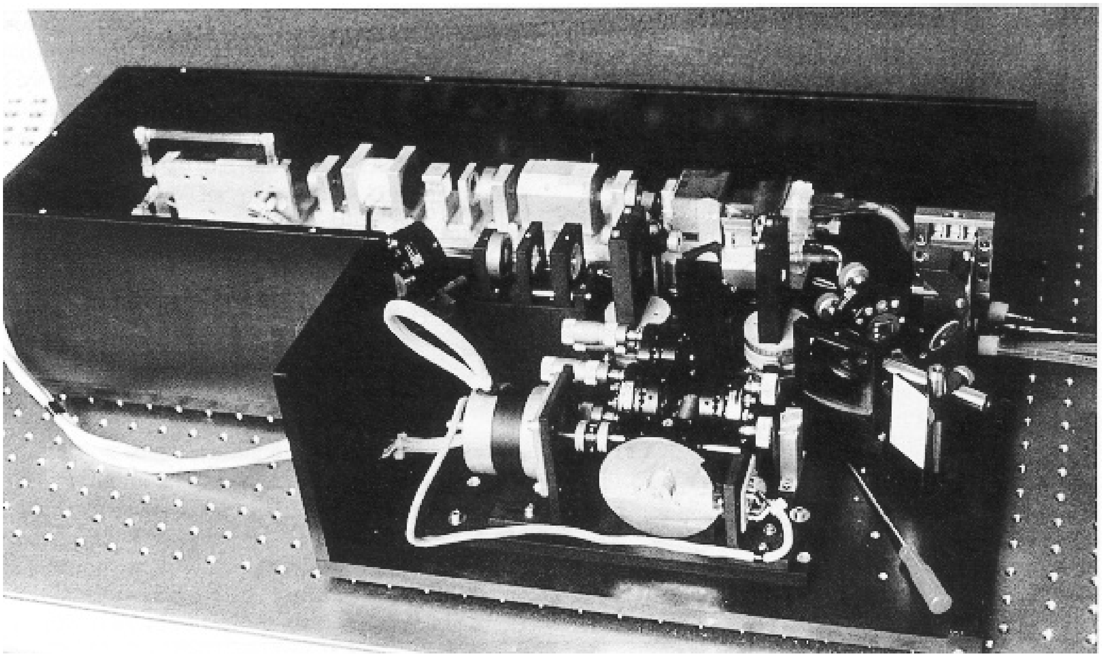
A device with feedback at only one of the signal or idler frequencies is referred to as a singly-resonant OPO (SRO).



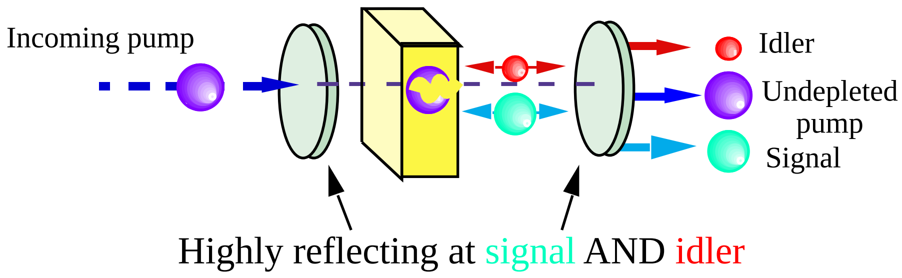
An example of an SRO OPO incorporating a frequency tripled diode-pumped laser as the pump source is shown below



A photograph of the device designed by C F Rae and M H Dunn (St Andrews).



To further increase the conversion efficiency the OPO cavity can be made resonant at both signal and idler frequencies, a doubly-resonant OPO (DRO)

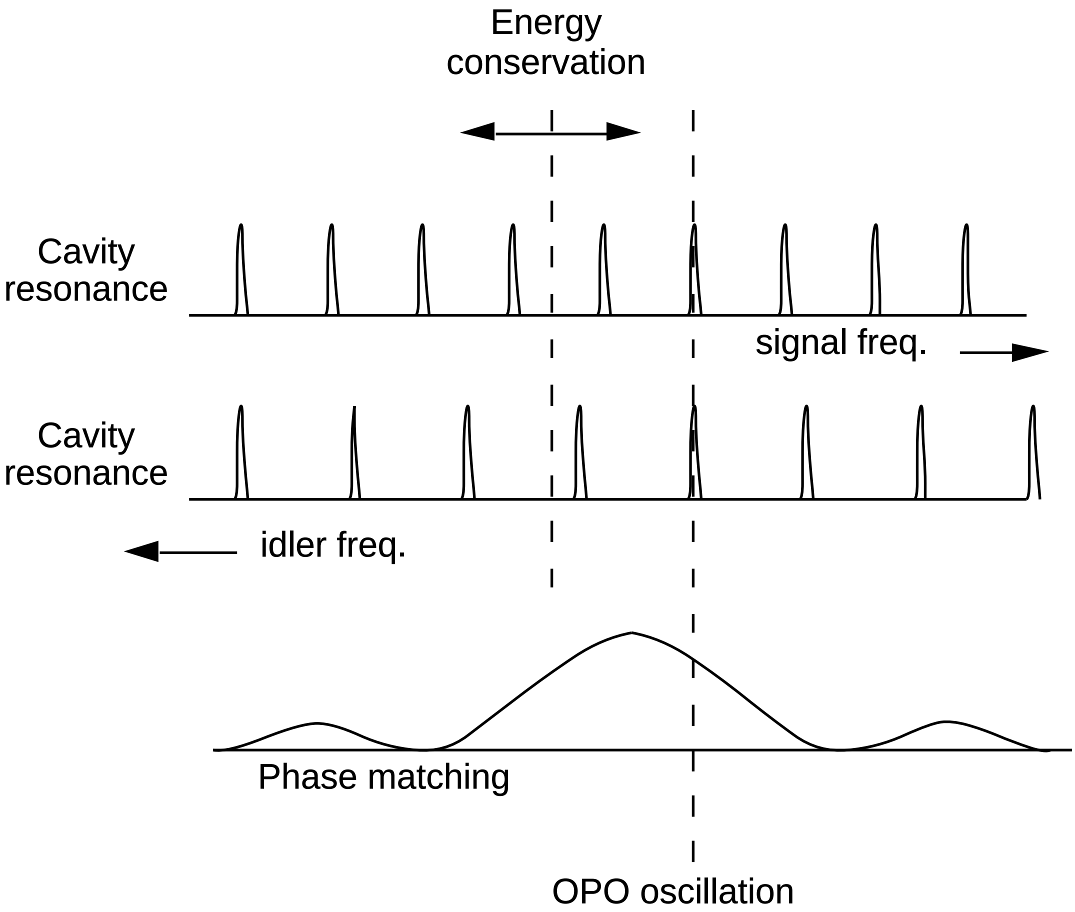


However, to operate a DRO, four conditions must be satisfied simultaneously

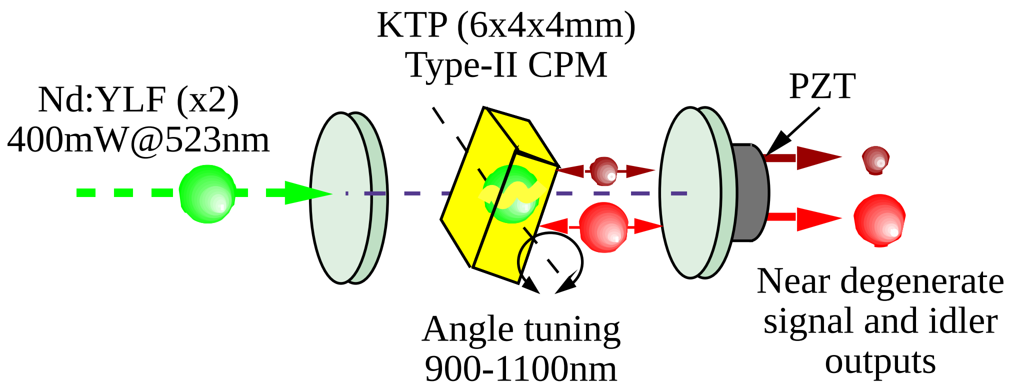
- energy-conservation
- phase-matching
- cavity resonance for the signal field
- cavity resonance for the idler field

In general, a DRO is over-constrained, and this introduces complications in the tuning of these devices.

The figure below shows the interrelation between the various constraints. (Note that the axes for the signal and idler frequencies are reversed.)



An example of a doubly resonant OPO based on KTP and pumped by a frequency-doubled, diode-pumped laser is shown below.



Index