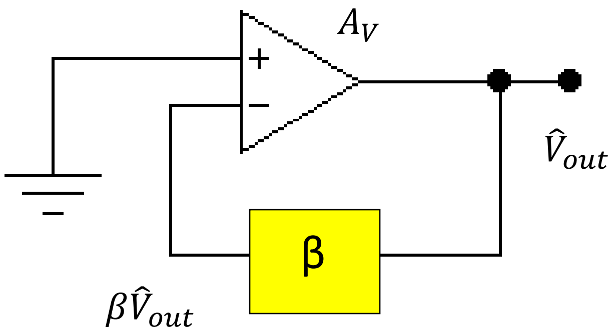
Amplifiers, oscillators and lasers

The laser pointer you hold in your hand is an oscillator, not just an amplifier. While lasers are used as amplifiers alone, the more common situation is oscillation, which requires both amplification and *feedback*. It's simpler to see with lumped circuits. The amplifier below looks at the voltage across its terminals and multiplies it by an open loop gain A_V. In phasor language,



$$\hat{V}_{out} = A_V(\hat{V}_+ - \hat{V}_-)$$

But if we feed a fraction β of the output back to the input then,

$$\hat{V}_{out} = A_V(0 - \beta\hat{V}_{out}) = -A_V\beta\,\hat{V}_{out}$$

For this to have a solution in which $\hat{V}_{out} \neq 0$ the circuit must satisfy the condition,

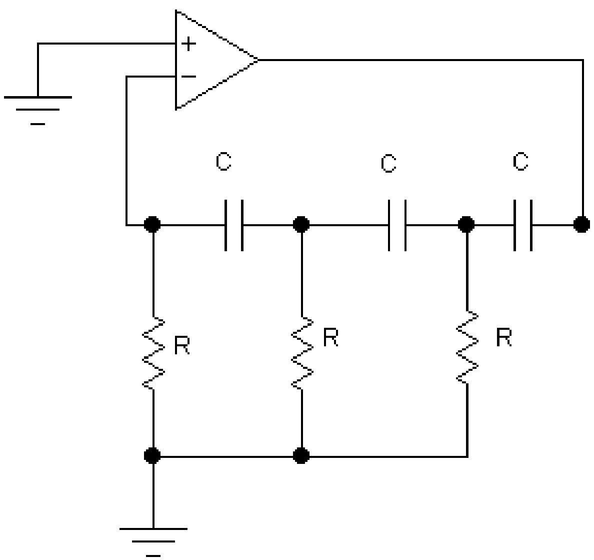
$$1 = -A_V\beta$$

In electronics this is called the *Barkhausen criterion*. Since both A_V and β are generally complex numbers, there are really two conditions,

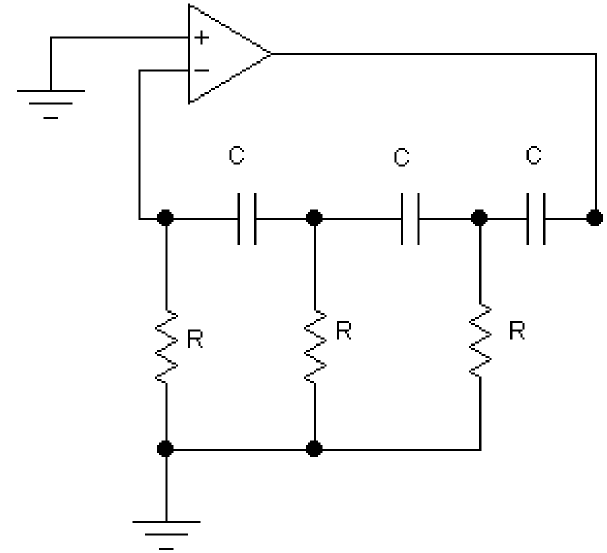
$$Re\,(A_V\beta) = -1 \qquad Im(A_V\beta) = 0$$

In general, both A_V and β both depend on frequency. There are two equations and these are usually satisfied at a particular frequency ω_0 and gain. If those conditions are met the circuit begins to generate a sinusoidal output even though there is no input voltage ($\hat{V}_+ = 0$). We're not violating energy conservation because the transistors inside the amplifier are driven into their active state by an external DC power supply, which is not shown in the diagram. In a sense, oscillators take an uninteresting source of power (the DC supply) and turn it into an interesting source of power (a sinusoidal output).

The circuit on the right, called a phase shift oscillator, is a concrete example. The β network consists of 3 capacitors and 3 resistors. Assume, for simplicity, that the open loop gain A_V is independent of frequency.



With a little circuit analysis we find that the feedback fraction is given by,



$$\beta = \frac{\hat{V}_-}{\hat{V}_{out}} = \frac{1}{1 - \frac{5}{(\omega RC)^2} + i\left(\frac{1}{(\omega RC)^3} - \frac{6}{\omega RC}\right)}$$

Since A_V is a real number, independent of frequency then the condition,

$$Im(A_V\beta) = 0 \;\rightarrow \left(\frac{1}{(\omega RC)^3} - \frac{6}{\omega RC}\right) = 0$$

$$\omega = \omega_0 = \frac{1}{RC\sqrt{6}}$$

That sets the frequency of oscillation. The second condition,

$$Re(A_V\beta) = -1$$

Can now be evaluated at ω_0 so,

$$\beta(\omega_0) = -\frac{1}{29} \qquad A_V\beta(\omega_0) = -1 \qquad A_V = 29$$

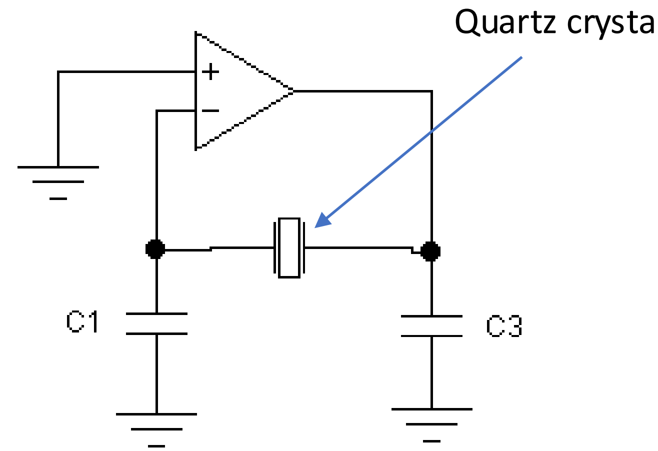
At the oscillation frequency a signal beginning at the (-) terminal picks up a 2π phase shift as it makes a round trip through the amplifier and feedback network and back. The gain has a minimum in order to make up for losses in the signal during this round trip.

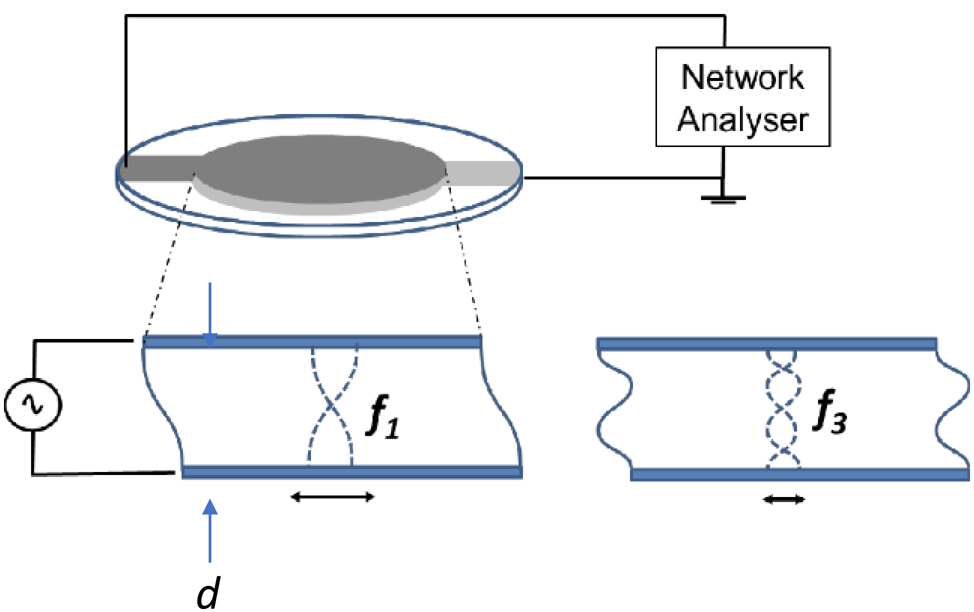
Nonlinearity

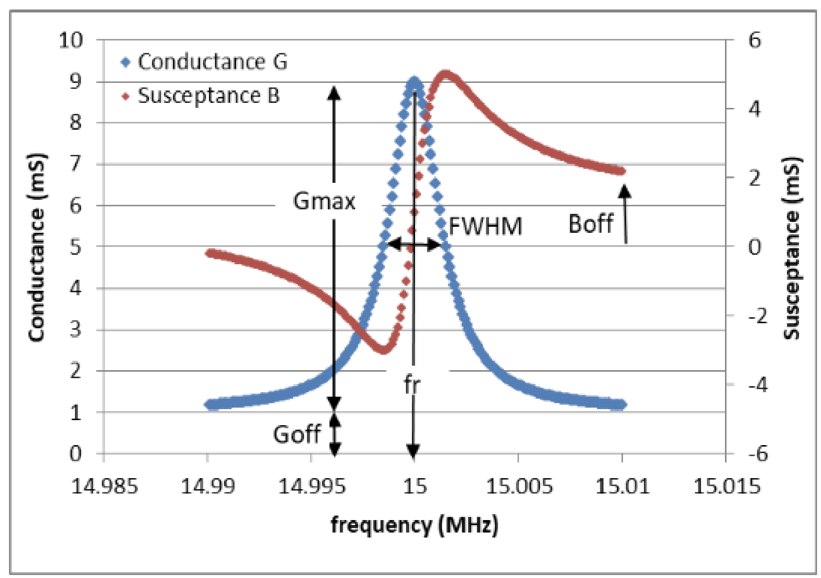
The equations we're written are linear, which means that we could multiply all the voltages and currents and any number and still come up with the same oscillation conditions. There is no way for the circuit to know what magnitudes of voltage and current to pick. In real life, once the gain and phase conditions are met the circuit begins to oscillate. As the voltages and currents grow, eventually they run into some nonlinearity. In the circuit above, the output stage transistors may eventually be driven beyond their linear response range. At that point the circuit equations have nonlinear terms and these *do* lead to limits on voltage and current. However, they generally also introduce harmonic distortion into the signal. Returning to the oscillation conditions, the gain condition should generally be regarded as a lower limit. The above circuit will oscillate for $A_V \geq 29$. In fact any sensible circuit would be designed with $A_V > 29$. If we set it at $A_V = 29$ then if might drift below that and turn off the oscillation. There are analogs to all of these concepts in laser physics.

Quartz Crystal Oscillators

The phase shift circuit illustrates the essential features of oscillators but it would be impractical as a stable frequency standard since the R's and C's will drift with temperature. A much more stable oscillator can be made by replacing the *RC* network by a quartz crystal resonator. As we discussed earlier, quartz is piezoelectric. An electric field across the faces of the crystal generates standing sound waves at extremely sharp, well-defined frequencies, typically in the 5 – 100 MHz range. With the correct choice of auxiliary components the circuit will oscillate at one of these acoustic standing wave frequencies. With careful temperature control this type of oscillator can be extremely stable, with short term frequency deviations of order $\delta f/f \sim 10^{-11}$ per day. For years, quartz resonators were the primary frequency standard at the National Bureau of Standards (now NIST).





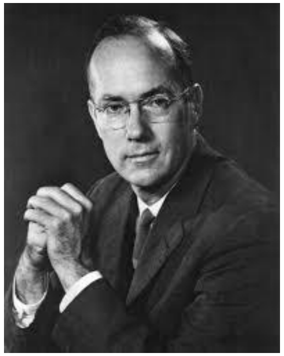


https://qcm-sensors.com/publications/the-principles-of-qcm-i/

Lasers

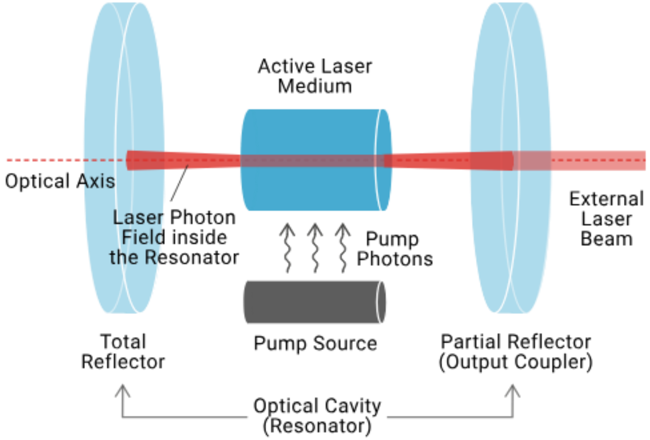
Laser stands for **L**ight **A**mplification by **S**timulated **E**mission of **R**adiation. Many people contributed important ideas but probably the three most notable were Charles Townes (USA) , Arthur Schawlow (USA) and Alexander Prokhorov (USSR). A laser has all the same components as the simple lumped circuit oscillator, albeit operating optical frequencies (1) an *active* or *gain medium* in which amplification takes place (2) an *optical cavity* which provides feedback as waves make multiple passes through the gain medium (3) a pump source that provides energy to the gain medium. Like the electronic oscillator, very specific conditions must hold for lasing to take place.

A. Schawlow



C. Townes

A. Prokhorov



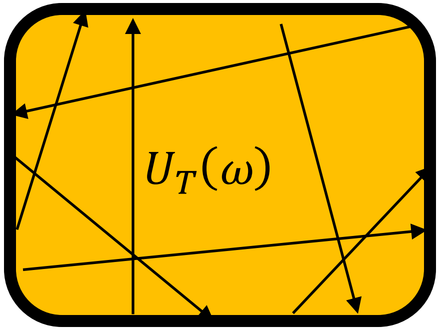
https://www.meetoptics.com/academy/laser-fundamentals#how-do-lasers-work

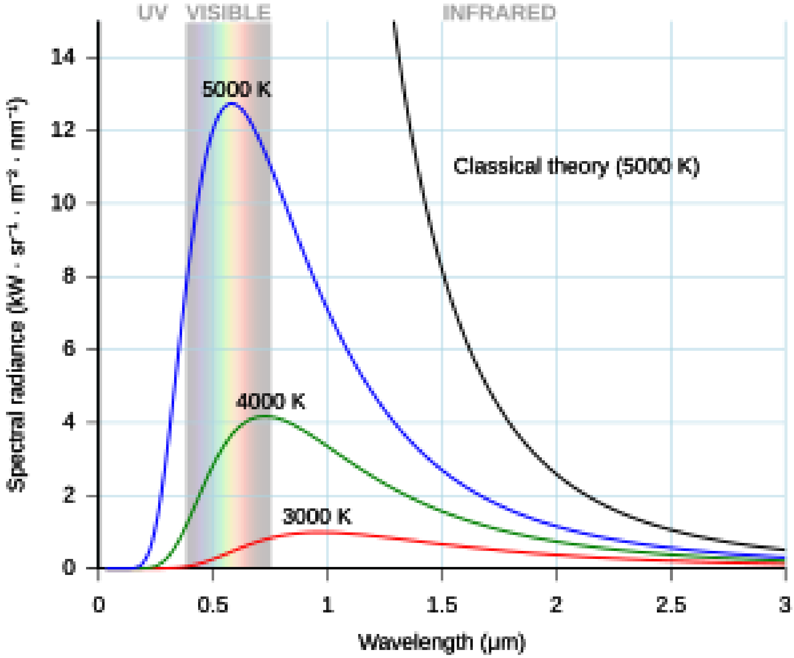
Planck spectrum

Consider a hollow cavity in thermal equilibrium at absolute temperature *T*. Inside the cavity, electromagnetic waves in a frequency range between ω and $\omega + d\omega$ will have an energy per cubic meter,

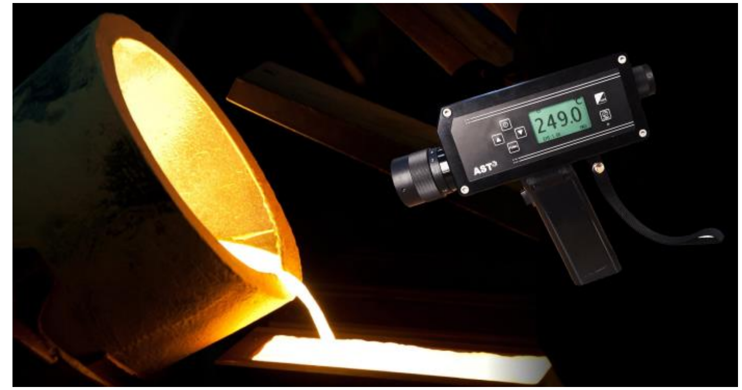
$$U_T(\omega)\, d\omega = \frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{e^{\hbar\omega/k_B T} - 1}\, d\omega$$

This is the famous Planck “black-body” radiation spectrum. Planck did not specify how the radiation was produced, beyond some mysterious “oscillators” in the walls of the cavity. Apparently he didn’t believe in atoms, didn’t like Boltzmann statistics and knew nothing about photons. Nonetheless, by accepting the quantization of energy levels in these oscillators he got the answer spectacularly correct. We now understand that cavity radiation is generated by quantum transitions between energy levels of whatever composes the cavity. If you drill a tiny hole in the cavity and look at the light coming out then it will have this same spectrum. In fact, even a hot piece of iron, with no surrounding cavity, generates this spectrum, which allows us to measure its temperature without touching it.



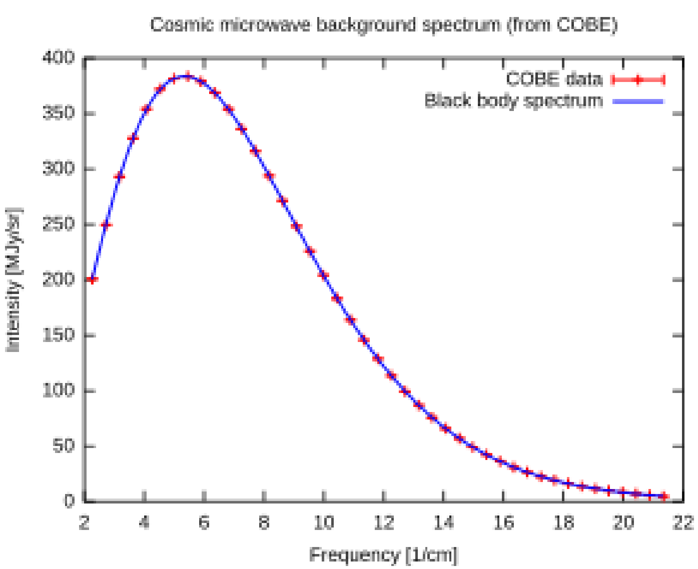


https://en.wikipedia.org/wiki/Black-body_radiation



https://pyrosens.com/what-is-a-pyrometer/

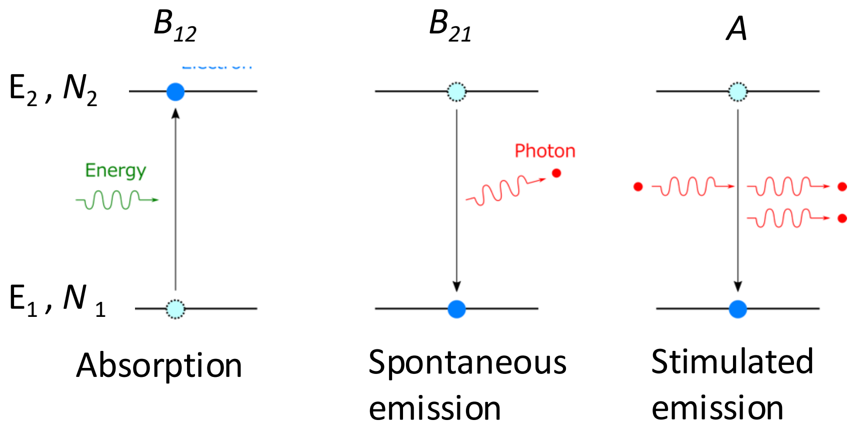
By looking at the spectrum of light coming from the sun we can use the Planck spectrum to deduce its surface temperature. Or, you can buy an optical pyrometer to measured temperatures of say, molten metal.



The Planck spectrum is so general that it holds even if we consider the entire universe as our cavity and fit the spectrum to a temperature of T = 2.73 Kelvin. The measurements are shown on the left. Rather amazing. In fact, the actual error bars are much smaller than the width of the theoretical line.

(https://en.wikipedia.org/wiki/Cosmic_microwave_background)

Einstein A and B coefficients



Einstein looked at the cavity radiation problem and introduced the idea of quantized units of light energy (photons, though that word came along many years later). Consider an atom and focus on two energy levels are E_2 , E_1 . He identified three processes that must be occurring to maintain thermal equilibrium (1) **Stimulated absorption**, in which the atom goes from the lower to the upper state and absorbs a packet of electromagnetic energy whose frequency obeys $\hbar\omega = E_2 - E_1$ (2) **Spontaneous emission**, in which an atom in the upper state emits a packet $\hbar\omega$ even if there is no existing field and (3) **Stimulated emission**, in which more packets are emitted in proportion to the energy density of the field that is already there. The packets are called photons. Imagine N_2 atoms are in the upper state and N_1 in the lower state. Then in equilibrium the principle of detailed balance says,

$$N_1 B_{12} U_T(\omega) = N_2 B_{21} U_T(\omega) + N_2 A \qquad \rightarrow \qquad U_T(\omega) = \frac{A}{\frac{N_1}{N_2} B_{12} - B_{21}}$$

But in thermal equilibrium N_1/N_2 is given by the Boltzmann factor,

$$\frac{N_1}{N_2} = e^{\hbar\omega/k_B T} \rightarrow U_T(\omega) = \frac{A}{e^{\hbar\omega/k_B T} B_{12} - B_{21}}$$

But this expression must reduce to the Planck distribution,

$$U_T(\omega) = \frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{e^{\hbar\omega/k_B T} - 1}$$

Therefore,

$$B_{12} = B_{21} \equiv B \quad A = B\frac{\hbar\omega^3}{\pi^2 c^3}$$

Inside the cavity we have,

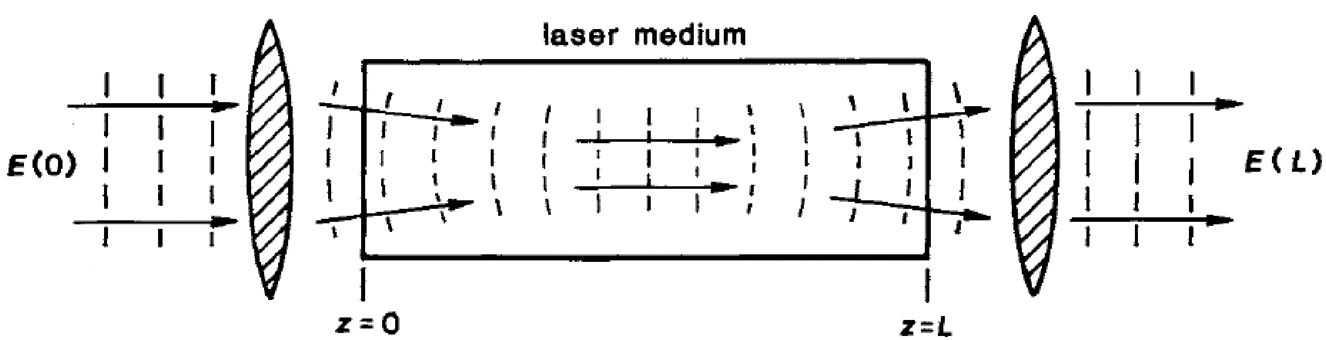
$$\frac{Rate\ of\ stimulated\ emission}{Rate\ of\ spontaneous\ emission} = \frac{N_2\, BU_T(\omega)}{N_2 A} = \frac{BU_T(\omega)}{A} = \frac{1}{e^{\hbar\omega/k_B T} - 1} = n(\omega) = average\ photon\ number$$

where $n(\omega)$ is the average number of thermally generated photons in the cavity mode with frequency ω. Although these relations were determined for a cavity in equilibrium, they are very general. In particular, N_1, N_2 and $U(\omega)$ may refer to values in a nonequilibrium situation like the inside of a laser cavity. The ratio of stimulated to spontaneous emission is a general feature of particles, like photons, obeying Bose-Einstein statistics. *A* represents the probability of some process involving bosons, in this case spontaneous emission. But if there are already *n* such identical bosons present, the *total* probability of the process, in this case spontaneous + stimulated emission, is (*n + 1*) *A*. It holds true whether *n* is a thermal equilibrium population, as above, or a highly nonequilibrium situation, as inside a laser cavity.

Light Amplification

The first step is amplification. For that to occur we need to emit more photons that we absorb. Referring to the figure, we need the electric field *E(L) > E(0)*. As light travels through a medium, the net power emitted per unit volume is the difference between the stimulated emission and stimulated absorption,

$$\frac{Power}{V} = (N_2 - N_1)\, BU(\omega)\, \hbar\omega = (N_2 - N_1) A \frac{\pi^2 c^3}{\hbar\omega^3}\, U(\omega)\hbar\omega = A(N_2 - N_1)\frac{\pi^2 c^3}{\omega^2}\, U(\omega)$$



https://www.fiberoptics4sale.com/blogs/wave-optics/laser-amplification-explained-in-detail

This equation for the power brings up a few important points.

1. To get any gain we need a *population inversion*, $N_2 > N_1$. In lasers that's generally provided by some external source of energy – flashlamps, chemical reactions, some other laser, etc.

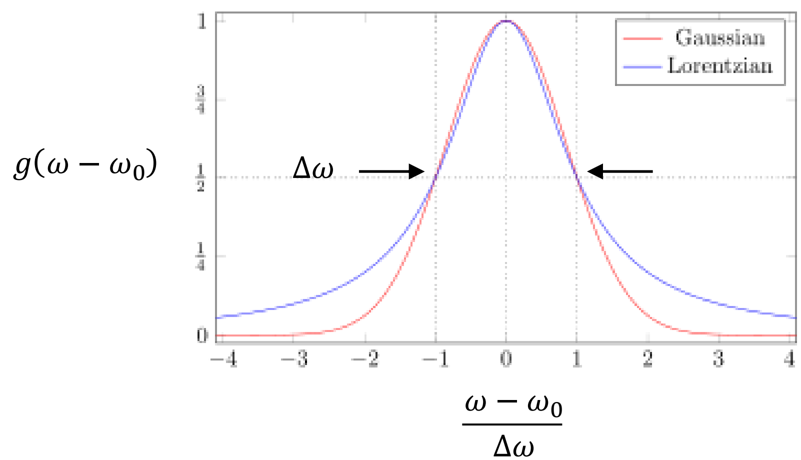
2. The power depends on the spontaneous emission rate, *A*. The greater *A* the greater the induced emission rate. However, spontaneous emission is noise.

3. The energy density $U(\omega)$ will be strongly frequency dependent. Generally the light emitted is close to monochromatic. However, the spectrum of light from any real atomic transition will have some resonant line shape, just like the spectrum from a free induction decay in NMR. There will be homogeneous or inhomogeneous broadening or both.

The net result is that we need to replace the energy density $U(\omega)$ with something like,

$$U(\omega) \rightarrow U_0\, g(\omega - \omega_0)$$

U_0 is the energy density at the frequency ω_0 of the atomic transition. $g(\omega - \omega_0)$ is a line shape function that might be Lorentzian or Gaussian depending on the medium. We end up with,



$$\frac{Power}{V} = A(N_2 - N_1)\frac{\pi^2 c^3}{\omega^2}\, U_0\, g(\omega - \omega_0)$$

The light intensity is $I = c\, U_0$ so,

$$\frac{Power}{V} = A(N_2 - N_1)\frac{\pi^2 c^2}{\omega^2} g(\omega - \omega_0)\, I$$

Intensity is light energy crossing a unit area per unit time. For a wave moving in the +z direction its intensity will therefore change as,

$$\frac{Power}{V} = \frac{dI}{dz} = A(N_2 - N_1)\frac{\pi^2 c^2}{\omega^2} g(\omega - \omega_0)\, I \equiv \gamma(\omega)\, I \rightarrow I(z) = I(0)e^{\gamma(\omega) z}$$

$\gamma(\omega)$ denotes the exponential rate of change in intensity. Since $I \propto E^2$, $\gamma(\omega)/2$ will be the exponential rate of change of the electric field *E*.

Laser gain condition

In order to make the intensity increase as light passes through the medium we need the $\gamma(\omega) > 0$. That implies more atoms in the upper energy level than the lower, i.e., a *population inversion*:

$$N_2 > N_1$$

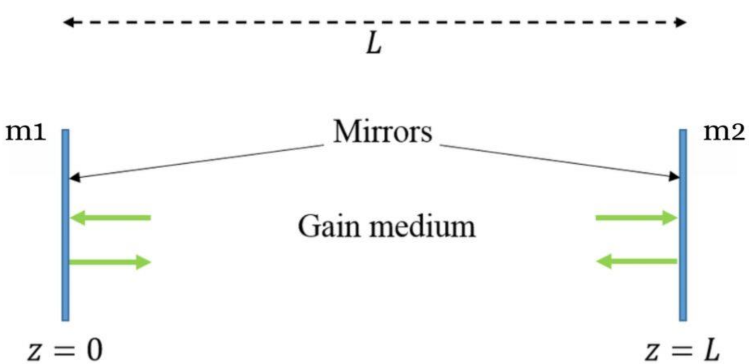
Light loses intensity as it passes through a medium, regardless of any stimulated emission. That's generally described by some coefficient α_{loss}. So, in the *absence* of any stimulated emission we'd have,

$$I(z) = I(0)e^{-\alpha_{loss}z}$$

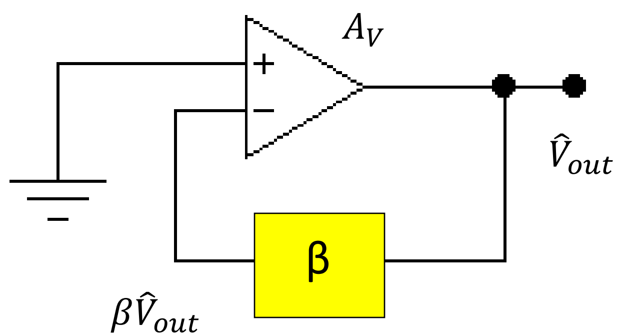
To have a net gain we need $\gamma(\omega) > \alpha_{loss}$ which implies a minimum gain. That, in turn implies a minimum value of the population inversion that satisfies,

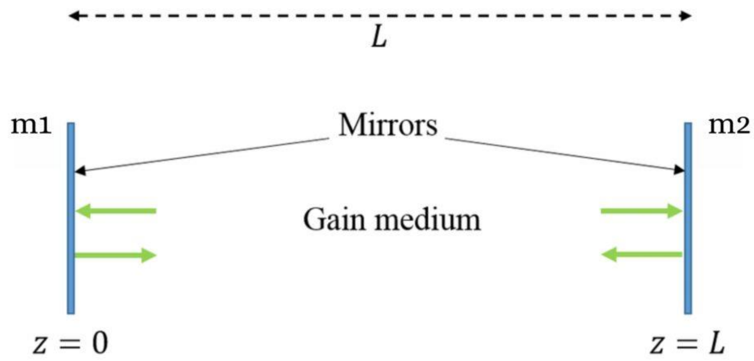
$$\gamma(\omega) = A(N_2 - N_1)\frac{\pi^2 c^2}{\omega^2}g(\omega - \omega_0) > \alpha_{loss}$$

If this condition is satisfied, we have an amplifier, but not an oscillator. For an oscillator, we also need feedback, just as shown for the lumped circuit case. At optical frequences the feedback consists of light waves making multiple passes back and forth between mirrors, except, of course, that a some of that light leaks out and you see it as a laser beam.



Fabry-Perot Cavity





Fabry-Perot Cavity

The design of the laser cavity is an interesting problem in optics but we can get the essential ideas with a simple Fabry-Perot device with two partially reflecting mirrors. Imagine an incident wave with electric field amplitude E enters the left mirror. A fraction $t_1 E$ goes through the mirror. It then propagates a distance L in which it acquires a phase factor so it has value $t_1 E e^{-ikL}$ before it strikes the right mirror. A fraction $r_2 t_1 E e^{-ikL}$ is reflected back and a fraction $t_2 t_1 E e^{-ikL}$ passes through. By the time the reflected wave reaches the left mirror is has value $r_2 t_1 E e^{-2ikL}$ and so on. The field that is transmitted through the right mirror is given by the geometric series,

$$E_{out} = t_2 t_2 E e^{-ikL}\left(1 + r_1 r_2 e^{-2ikL} + \left(r_1 r_2 e^{-2ikL}\right)^2 + \cdots\right) = E\frac{t_2 t_2 E e^{-ikL}}{1 - r_1 r_2 e^{-2ikL}}$$

If the denominator $1 - r_1 r_2 e^{-2ikL} = 0$ then the expression becomes infinite. This is the laser oscillation condition. It implies that E_{out} is nonzero even if $E_{in} = 0$. It's same idea with lumped circuit oscillators except that here, the energy for laser oscillation comes from the pump which maintains a population inversion, whereas in an electronic oscillator is comes from a power supply driving some kind of amplifier. The wavenumber k inside the gain medium has both real *and* an imaginary parts,

$$k = k_0 + k_R(\omega) + i\,k_I(\omega) \; - i\,\alpha_{loss}/2 \qquad k_0 = \omega n/c$$

k_0 is just the wavenumber if there were no stimulated emission. n is the index of refraction. α_{loss} accounts for attenuation of the electric field from processes other than stimulated emission. $k_R(\omega) + ik_I(\omega)$ come from the stimulated emission.

Oscillation conditions

The laser oscillation condition is given by,

$$1 - r_1 r_2 e^{-2i\left(k_0 + k_R(\omega)\right)L}\; e^{-\alpha_{loss}L + 2k_I(\omega)L} = 0$$

The expression on the left is complex so we have 2 equations to satisfy,

$$1 = r_1 r_2 e^{-\alpha_{loss}L + 2k_I(\omega)L} \qquad\qquad k_0 + k_R(\omega) = \frac{\pi\, n}{L} \qquad n = 1, 2, \ldots$$

As with the lumped circuit oscillator we have both a *gain* and a *phase* condition.

Gain

The gain condition can be rewritten as,

$$k_I(\omega) \geq \alpha_{loss} - \frac{1}{2L}\ln r_1 r_2$$

Since $r_1 r_2 < 1$, the lower the reflectivity of the mirrors the larger k_I must be to overcome generate an oscillation. To find $k_I(\omega)$ go back to our expression for the change of intensity a gain medium,

$$\frac{dI}{dz} = A(N_2 - N_1)\frac{\pi^2 c^2}{\omega^2} g(\omega - \omega_0)\, I \equiv \gamma(\omega)\, I \quad \rightarrow I(z) = I(0)e^{\gamma(\omega)z}$$

Since the intensity $I \propto E^2$ the electric field will grow as $E(z) = E(0)e^{-k_I z} = E(0)e^{\frac{\gamma(\omega)z}{2}}$ and therefore,

$$k_I(\omega) = \frac{\gamma(\omega)}{2} = A(N_2 - N_1)\frac{\pi^2 c^2}{2\,\omega^2} g(\omega - \omega_0)$$

So $k_I(\omega) \geq \alpha_{loss} - \frac{1}{2L}\ln r_1 r_2$ for the laser to oscillate. This linear theory tells us the *threshold* value of the gain. This leads to a minimum or critical population inversion to lase,

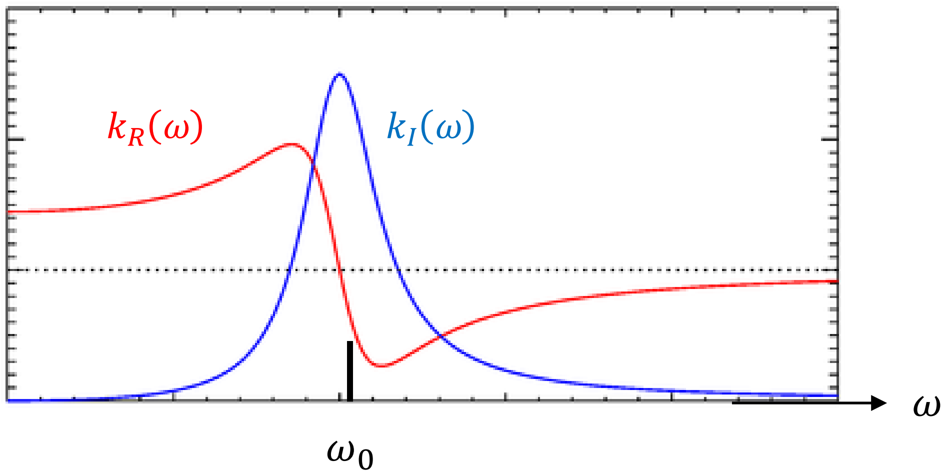
$$\Delta N_{crit} = (N_2 - N_1)_{min} = 2\,\omega^2 \frac{\alpha_{loss} - \frac{1}{2L}\ln r_1 r_2}{A\pi^2 c^2 g(\omega_0)}$$

Phase

The phase condition is,

$$k_0 + k_R(\omega) = \frac{\pi\, n}{L} \qquad n = 1, 2, \ldots$$

This is what you would expect: that the oscillations correspond to standing wave modes in the cavity. $k_R(\omega)$ has the dispersive form similar to all resonances, shown on the plot. It slightly modifies the standing wave frequencies from what they would be if there were no stimulated emission.



Homogeneous and Inhomogeneous broadening

The gain $\gamma(\omega)$ was worked out earlier. It determines the threshold population for laser oscillation. It's often called the *single-pass gain*, since it is the gain if a light wave travels once down the laser cavity,

$$\gamma(\omega) = A(N_2 - N_1)\frac{\pi^2 c^2}{\omega^2} g(\omega - \omega_0)$$

The function $g(\omega - \omega_0)$ is the lineshape function for the atomic transition. As in NMR, its width is determined by two different processes, homogeneous and inhomogeneous broadening. For inhomogeneous broadening, different atoms feel slightly different resonant frequencies due to Doppler shifts in a gas laser or inhomogeneity in the gain medium of a solid state laser. This is the case of a HeNe laser pointer. There is also homogeneous broadening. An example is an excited atom which decays by spontaneous emission. This emission has a lifetime t_{spont} which determines the width of the transition $\delta\omega \sim 1/t_{spont}$. Lasers can have both kinds of broadening.

The second point about the single pass gain is that it is proportional to the amount of population inversion $(N_2 - N_1)$. This must be maintained by some *external* source of power known as the pump. You can think of it like the power supply in an amplifier. In fact, for diode lasers, it *is* an external DC power supply. For low pumping power $(N_2 - N_1)$ is small to reach the critical value for oscillation. As the pumping power is increased, eventually some cavity mode has enough gain to oscillate. This is illustrated next.

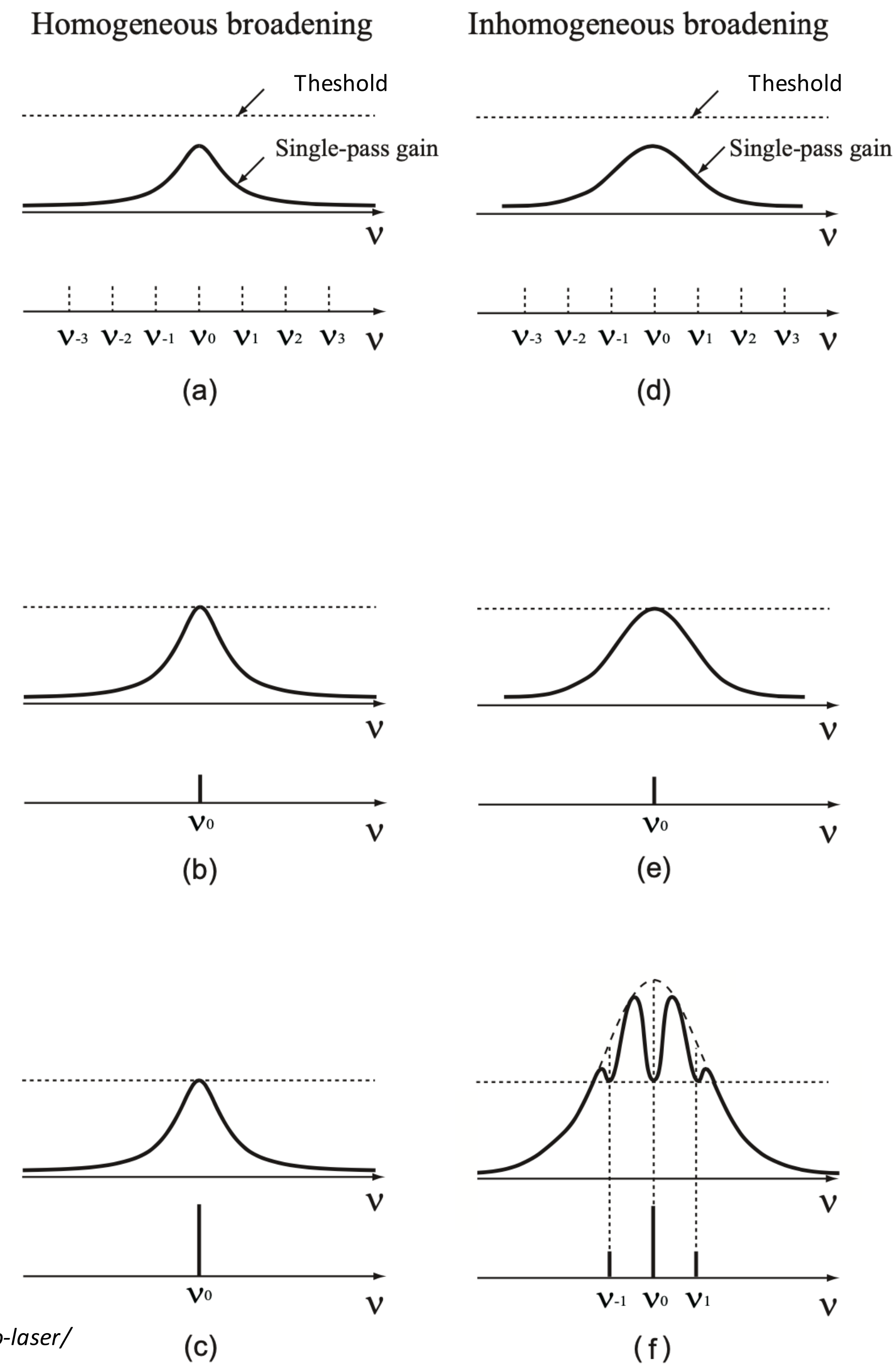
Laser gain and oscillation

Focus first on the left column labelled homogeneous broadening. Here, all the atoms have the same line shape. The threshold is the gain that must be reached for oscillation. The frequencies on the axis are the cavity modes that satisfy the frequency condition. As the pump power is turned up from (a) to (b) the population inversion (N_2 - N_1) and therefore the single pass gain increases until it reaches threshold. At this point, the mode at frequency ν_0 begins to oscillate. As we increase the pumping power to (c), it turns out that the gain stays *clamped* at the threshold level. This is analogous to electronic oscillators in which some nonlinear process limits the gain and therefore the amplitude of oscillation.

The second column shows the case for inhomogeneous broadening. Again, the pump power increases from (d) to (e) at which point the ν_0 mode begins to oscillate. But now, as the pump power increases to (f), modes ν_{-1} and υ_1 reach threshold and they too begin to oscillate. Once this happens *their* gain is also clamped at the oscillation threshold.

This clamping is called *gain saturation*. Initially, as we increase the pumping power $(N_2 - N_1)$ will increase until some mode reaches threshold. But as the light intensity increases further it induces more downward transitions ($E_2 \rightarrow E_1$) which tends to *reduce* the population inversion $(N_2 - N_1)$ and therefore the gain. The result is that the gain for each oscillating mode stays *clamped* at its threshold value where stimulated emission just makes up for the losses in the laser cavity.

The spikes in (c) and (d) indicate light intensity for each mode. Although the gain stays clamped, the light intensity <u>does</u> increase as the pumping power is increased.



https://en.symphotony.com/introduction-to-laser/

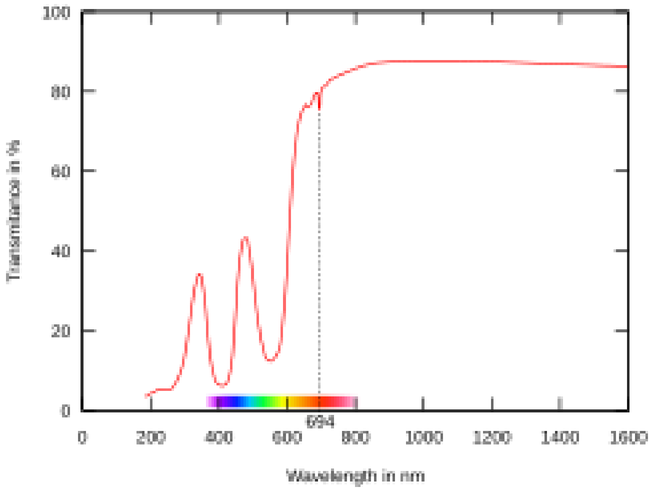
Real Lasers

The first person to build a working laser was Theodore Maiman, then working at Hughes Research Lab in Malibu, CA. The original device, shown in the figure, used ruby as a gain medium and unleashed its first burst of coherent red light on May 16, 1960. Charles Townes and Arthur Schalow , at Bell Labs, were close behind. Maiman's laser was pulsed but Gordon Gould later managed to make a continuous ruby laser.

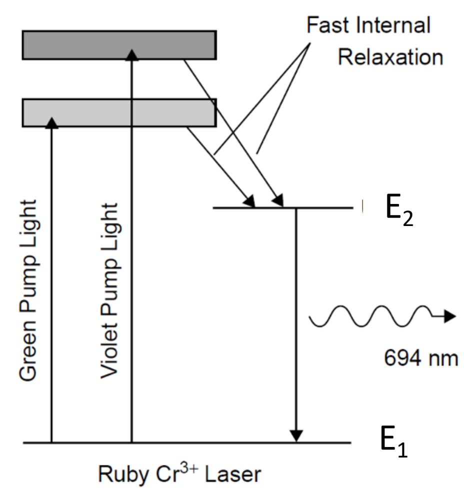
T. Maiman

Ruby is sapphire (Al_2O_3) with Cr^{3+} ions added. The left figure shows the transmission of light through ruby. There are broad dips in the green and violet and a narrow dip at 694 nm where ruby lases. The narrow dip implies a long lifetime for the upper laser level.

https://en.wikipedia.org/wiki/Ruby_laser



https://almerja.com/more.php?idm=40028



The relevant energy levels are shown on the right. Intense light from a flash lamp, which has a broad spectrum, pumps atoms up into the two broad energy bands. Atoms in these bands relax very rapidly down to the upper laser level. Atoms in this state have a long spontaneous lifetime so this state level builds up a population inversion and lasing takes place at a wavelength of 694 nm.

https://en.wikipedia.org/wiki/Ruby_laser

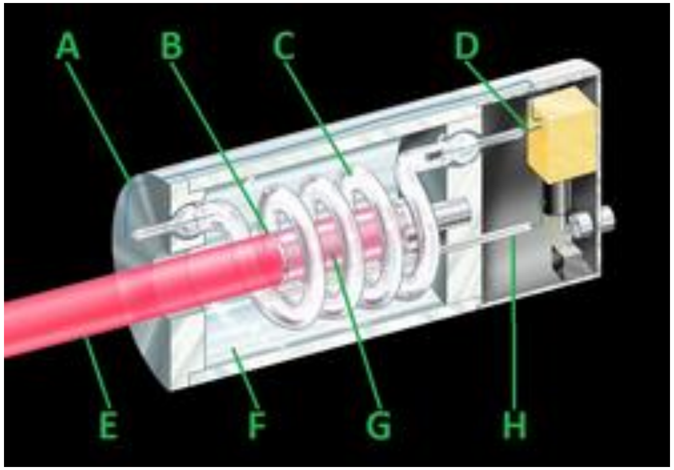
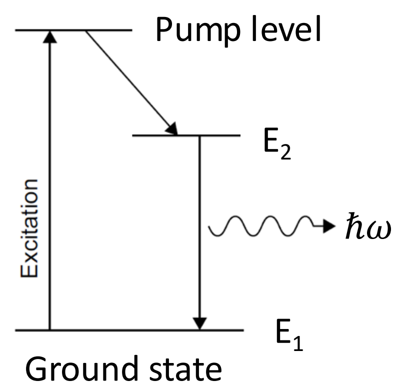


Diagram of the first ruby laser. **A** - Positive lead. **B** - Mirror coating. **C** - Xenon flashtube. **D** - Negative lead. **E** - Laser beam. **F** - Pumping cavity. **G** - Ruby rod. **H** - Trigger wire.

A diagram of the original ruby laser shows that the ruby rod itself formed a Fabry-Perot etalon. Its ends had to be polished with extreme precision. One end had a metallic coating to serve a mirror and the output end had a partially silvered mirror. Pumping was provided by a powerful flashlamp that surrounded the ruby rod in a spiral.



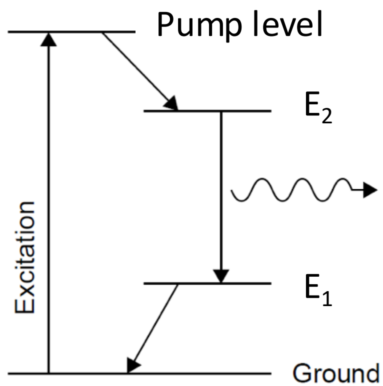
The ruby laser is one example of what is termed a 3-level laser. The lowest laser level (E_1) is the ground state. To lase, we need to maintain a population inversion of at least $N_2 > N_1$ so $N_2 - N_1 > N/2$ where $N = N_1 + N_2$ is the total number of atoms that lase. Assuming we pour a pump power P into the gain medium and the pump levels rapidly decay to N_2, the pump power required to maintain $N_2 > N_1$ is given by,

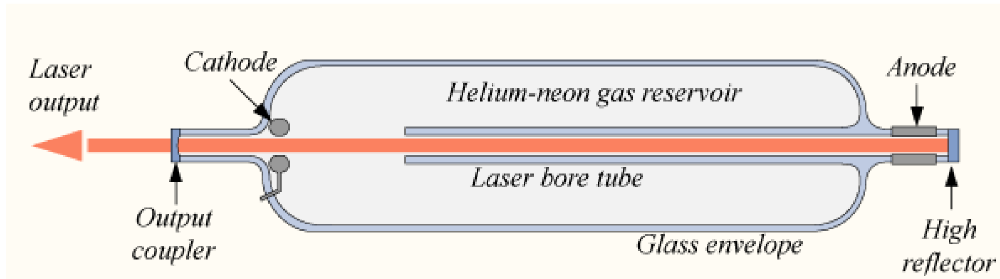
$$Pump\ power \approx \frac{N}{2}\ \hbar\omega\ A_{spont} = \frac{N}{2}\ \frac{\hbar\omega}{\tau_{spont}}$$

We need a long spontaneous lifetime τ_{spont} or otherwise the pump power gets too large to maintain the critical population inversion. For ruby it's about $\tau_{spont} = 3\ x\ 10^{-3}$ sec. Evidently Maiman was convinced by an IBM salesman that their flashlamp, which could melt steel wool, could deliver the required pumping power.

4-level lasers

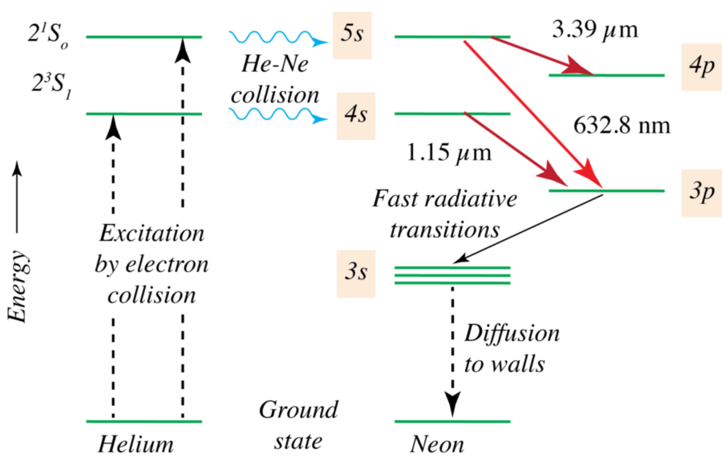
The scheme shown below, a 4-level laser, requires less pump power than the 3-level scheme. Here, the lower lasing level E_1 decays rapidly to the ground state by some other mechanism, perhaps by emitting phonons. To lase, the population inversion $\Delta N = N_2 - N_1$ must exceed the threshold ΔN_{crit} derived earlier. But in this case , since the lower lasing level empties out very rapidly we can have $N_1 \ll N$, $and\ \Delta N \approx N_2$. So fewer atoms must be pumped into the upper lasing level to maintain the critical population inversion and less pump power is required. This is basically the scheme used in the helium-neon laser you use for a laser pointer. It was invented at Bell Labs in 1960 by Ali Javan, William Bennet Jr. and Donald Herriott.





A 10:1 mixture of helium and neon gas are placed inside a glass discharge tube and a current is driven through the tube, exciting helium atoms to higher energy states. Through collisions, the excited helium atoms transfer energy to upper levels of neon. The first version lased at 1152 nm which is shown on the left. That's in the infrared and people wanted a visible light laser. Alan White and Dane Rigden at Bell Labs then worked to make it lase at 632.8 nm (red) which what you normally see today. It's also possible to make the HeNe system lase at 543.5 nm (green) , 594 nm (yellow) , 612 nm (orange) and 3391 nm.

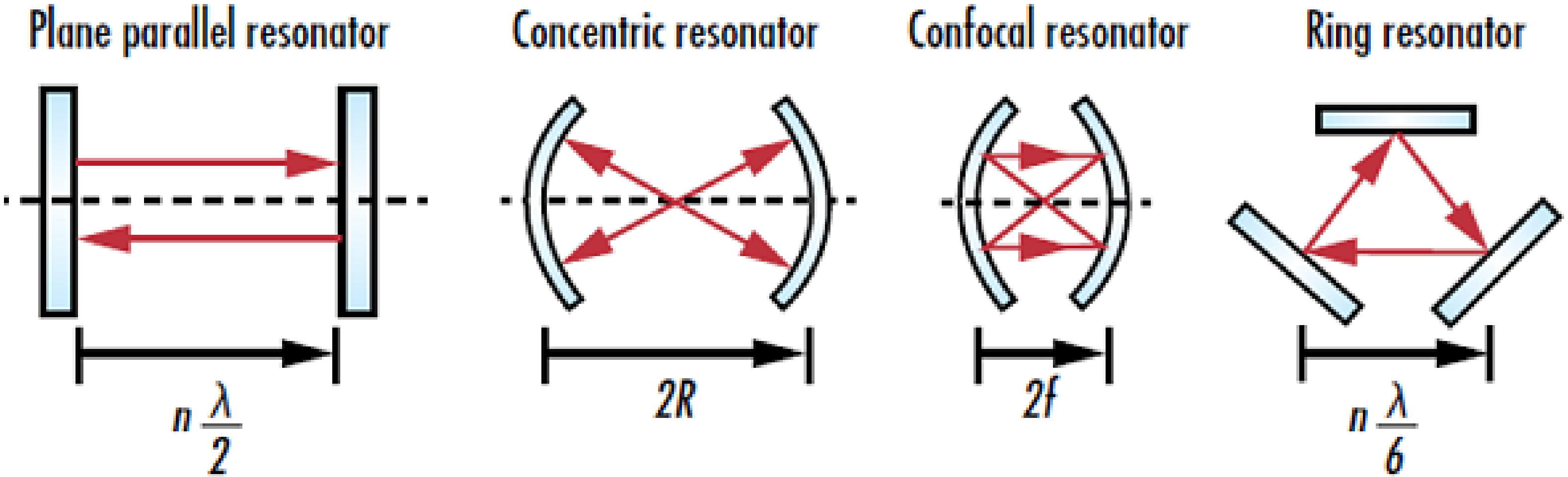
https://www.rpmclasers.com/blog/hene-lasers-bright-past-brighter-future/



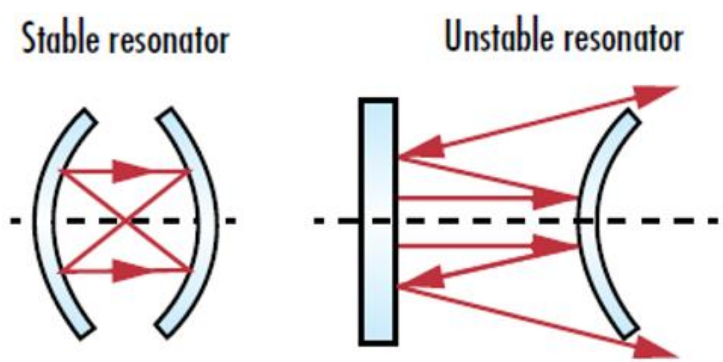
https://commons.wikimedia.org/wiki/File:HeNe_Laser_Levels.png

Laser Cavities

I discussed the simple Fabry-Perot parallel plate resonator but there are many other designs, some of which are shown below.



Resonators can be stable or unstable as illustrated below. In stable resonators a light beam bounces back and forth and eventually returns to itself. In an unstable resonator the beams eventually escape the cavity. Stable resonators are usually what's needed but un stable resonantors are used for high power lasers to prevent damage to the mirrors.



https://www.edmundoptics.com/knowledge-center/application-notes/lasers/laser-resonator-modes/

Laser beam shape and spreading

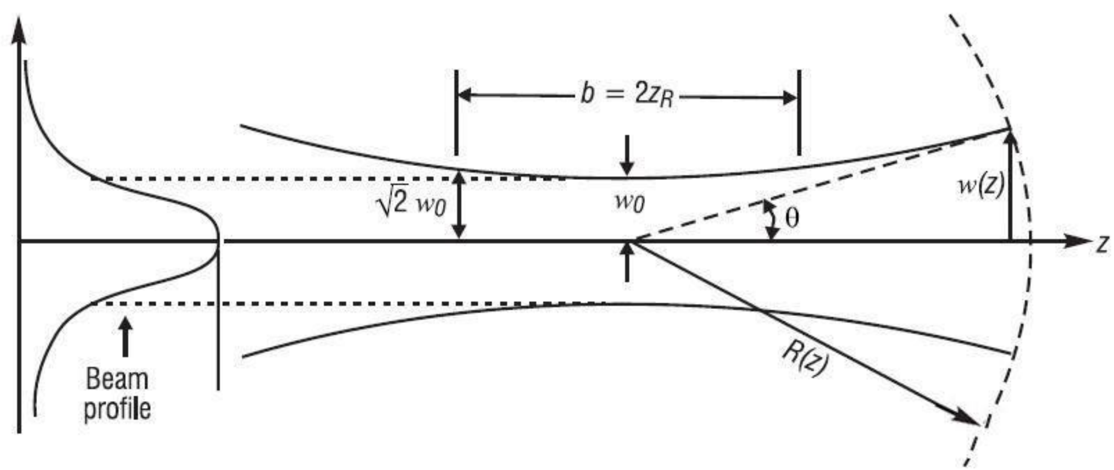
Since the laser cavity is an open resonator, confined solutions to the wave equation are considerably more involved than in a simple closed cavity. The most common case is the Gaussian beam whose parameters are shown. z is the direction of propagation. The intensity profile is Gaussian characterized by a beam spread function *w(z),*

$$I(z) \propto \left(\frac{w_0}{w(z)}\right)^2 e^{-(x^2+y^2)/w(z)^2} \qquad w(z) = w_0\sqrt{1 + (z/z_R)^2} \qquad R(z) = z(1 + (z_R/z)^2) \qquad z_R = \pi w_0^2/\lambda$$

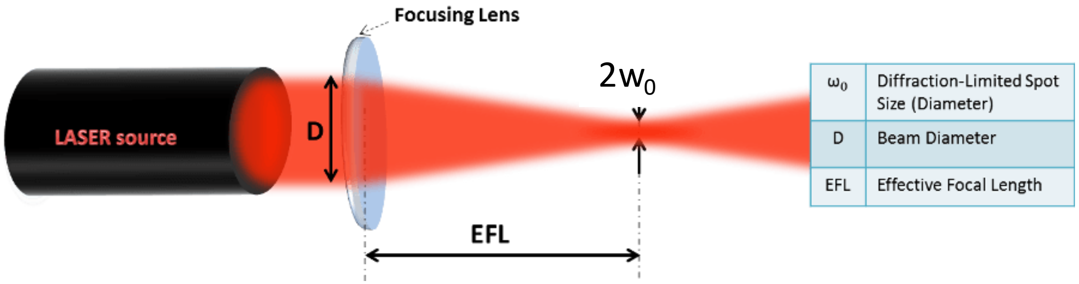
At the narrowest point $w(z=0) = w_0$ which is called the *beam waist.* z_R is called the *Rayleigh range* and *R(z)* denotes the radius of curvature of the constant phase front, as indicated by the dotted curve. Far to the right or left of this region the beam width diverges linearly with distance,

$$w(z) \approx \frac{\lambda z}{\pi w_0} \qquad \theta \approx \frac{\lambda}{\pi w_0} \qquad |z| \gg z_R \qquad (far\, field\, regon\,)$$

https://www.s-laser.com/info/what-do-you-mean-the-laser-beam-divergence-49491813.html



For the beam to reproduce itself the constant phase fronts of the beam need to coincide with the surfaces of the the spherical mirrors on either end of the laser cavity.



https://www.holoor.co.il/optical-calculator/diffraction-limited-spot-size/

There can be one or more beam waists inside the laser cavity. The emerging laser beam is still Gaussian and optics outside the laser will preserve the Gaussian character but can change the beam waist. The figure shows the laser focused to a diffraction-limited spot whose size is the beam waist. From there it diverges with the angle $\theta \approx \frac{\lambda}{\pi w_0}$.

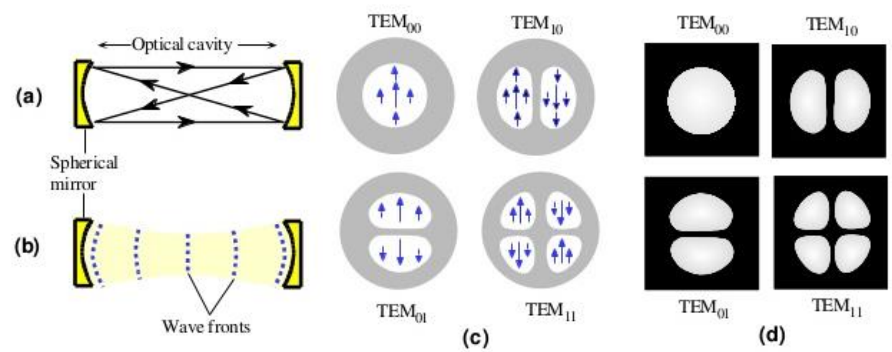
Transverse modes

In general the laser cavity will support more complicated transverse modes than the simple Gaussian. Some of these are shown for a confocal resonator. When I introduced the simple Fabry-Perot resonator the laser modes were very close to the standing wave modes of the cavity, separated by a frequency interval of,

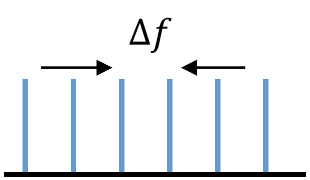
$$\Delta f = \frac{c}{2nL}$$

where L is the cavity length and n is the index of refraction. These are called the *axial modes* of the cavity. Transverse modes are clustered around each axial mode, as shown below.

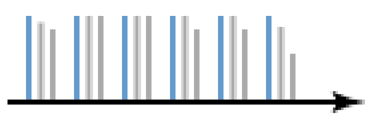
https://www.networxsecurity.org/members-area/glossary/t/transverse-mode.html



Laser Modes (a) An off-axis transverse mode is able to self-replicate after one round trip. (b) Wavefronts in a self-replicating wave (c) Four low order transverse cavity modes and their fields. (d) Intensity patterns in the modes of (c).



Axial modes alone



Axial and transverse modes

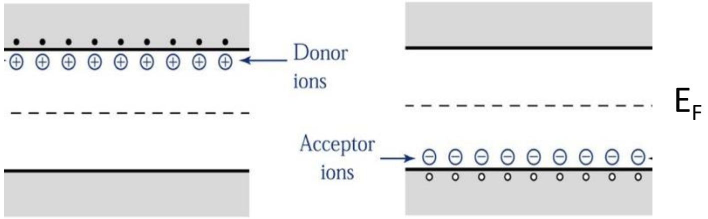
In principle, all the modes, axial and transverse, that satisfy the oscillation threshold can lase. Different modes, however, can have different oscillation thresholds so it's possible to ensure they don't lase. Sometimes the cavity is designed so most modes have too high a threshold to oscillate, i.e., too much energy leaks out as the wave bounces back and forth in the resonator.

Diode lasers

Diode lasers are everywhere these days. You can buy one for under $10. They come packaged like a small electronic component, complete with an internal sensor to regulate the output power. The basic device is indeed a diode, albeit an extremely sophisticated one. Current is driven though a PN junction causing a population inversion between energy bands. The laser cavity is part of the semiconductor itself.

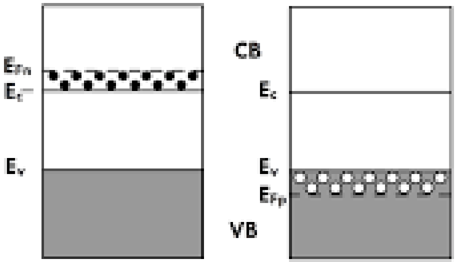
Nondegenerate semiconductors

We need to first go back and recall the band structure of ordinary N and P type semiconductor. The picture on the left is at room temperature. The N-type has donor atoms which release electron into the conduction band. In the P-type, acceptor atoms grab an electron from the valence band, leaving holes. Both the electrons and holes conduct electricity. At very low temperature, the electrons go back to the donor and acceptor states. The valence band is full, the conduction band is empty and no conduction occurs. The Fermi level is in the middle of the energy gap.



N-type P-type

Nondegenerate



N-type P-type

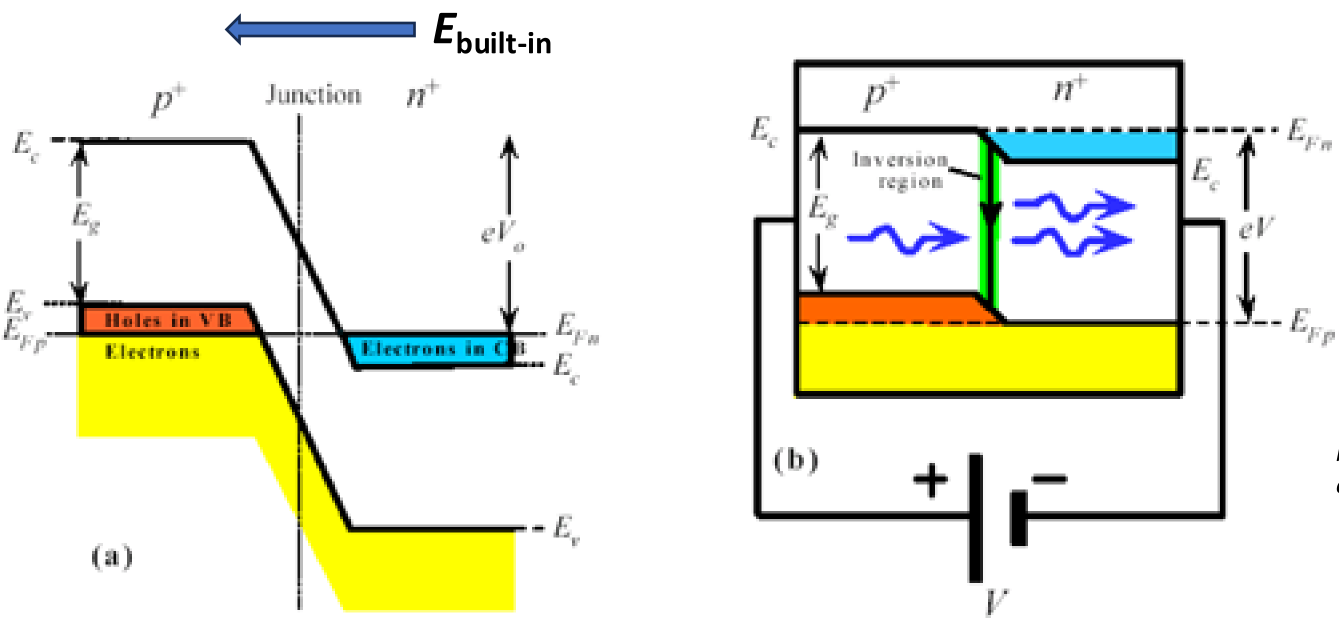
Degenerate

Degenerate semiconductors

Imagine we add many more donors and acceptors to each type. Eventually the donor levels broaden into a band of their own and overlap the original conduction band. Now, the Fermi level for the N-type is up in the conduction band so it's like a metal. And, the acceptor levels merge with the valence band and the Fermi energy for the P-type moves down into the valence band. Now, even at $T = 0$, there are mobile electrons in the conduction band and mobile holes in the valence band. These are call *degenerate* semiconductors and they are used in laser diodes. GaAs, rather than silicon, is the current semiconductor of choice for laser diodes.

PN junction in laser diodes

Recall that in the depletion region of a PN junction, a built-in potential develops whose electric field points from N to P, impeding the flow of holes from P to N and the flow of electrons from N to P. The electron must climb a potential hill to move from N to P, so the conduction and valence bands have a step, as shown. With no bias voltage the equilibrium situation is shown on the left. The Fermi energies on each side must line up, otherwise current would flow. But since these are now *degenerate* semiconductors, the Fermi level intersects the valence band in the P region and the conduction band in the N-region.



(a) p-n junction without bias

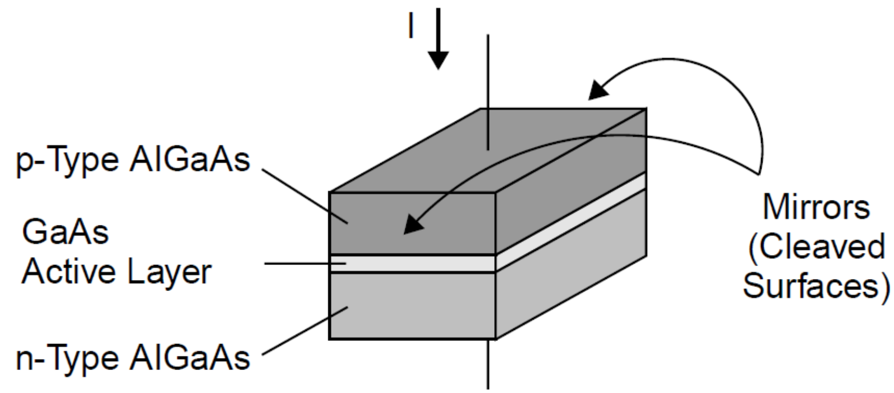
(b) p-n junction after forward bias

https://www.laserdiodesource.com/laser-diode-technical-overview-three

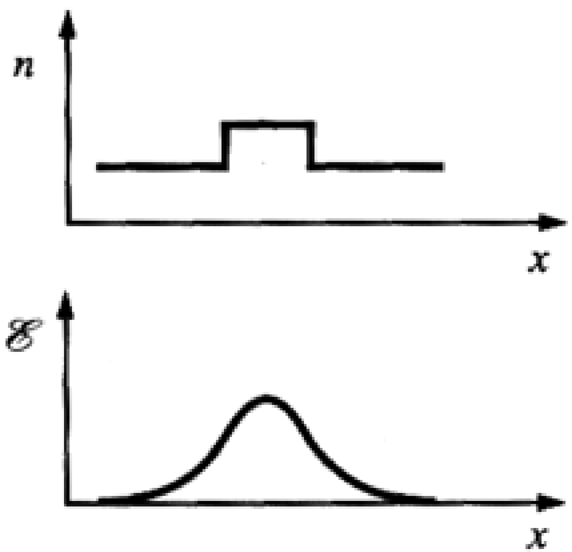
Now, apply a forward bias voltage, as shown on the right. This *opposes* the built-in potential and the energy bands are changed as shown. In the very narrow region were there is still a built-in potential, albeit smaller, a population inversion can develop between the upper and lower bands. This is the now the gain medium and, with the correct amount of feedback the device will lase. Unfortunately, in this simple arrangement light is not confined so very large currents are needed to achieve lasing and the early devices needed to be cooled to cryogenic temperatures. [An excellent reference is *https://www.laserdiodesource.com/laser-diode-technical-overview-three* .]

Mode confinement

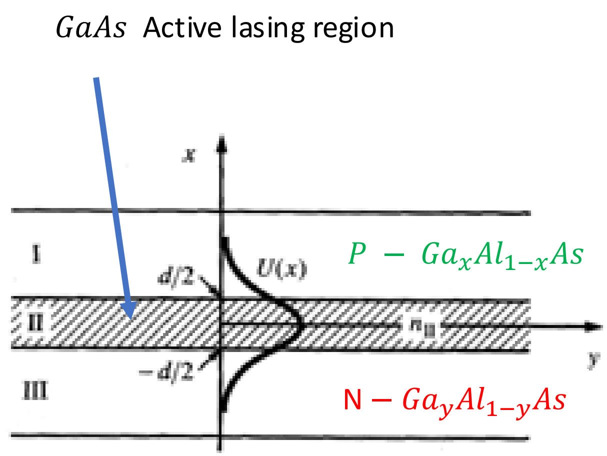
The diode serves as the cavity so light bounces back and forth and in the gain region of the junction, lasing can take place. However, in any laser, the light you need to amplify needs to be confined so intensity can build up. The simple PN junction, requires to much current to lase. To better confine the light, layers of lower refractive index are grown on either side of the junction. Then by total internal reflection, light will be confined in the central layer, as shown. One of the most common materials is GaAs. When Al replaces some of the Ga the index of refraction is lowered. The result is better mode confinement. The structure, shown on the right, is called a *heterojunction laser.*



https://mail.almerja.net/more.php?idm=44418



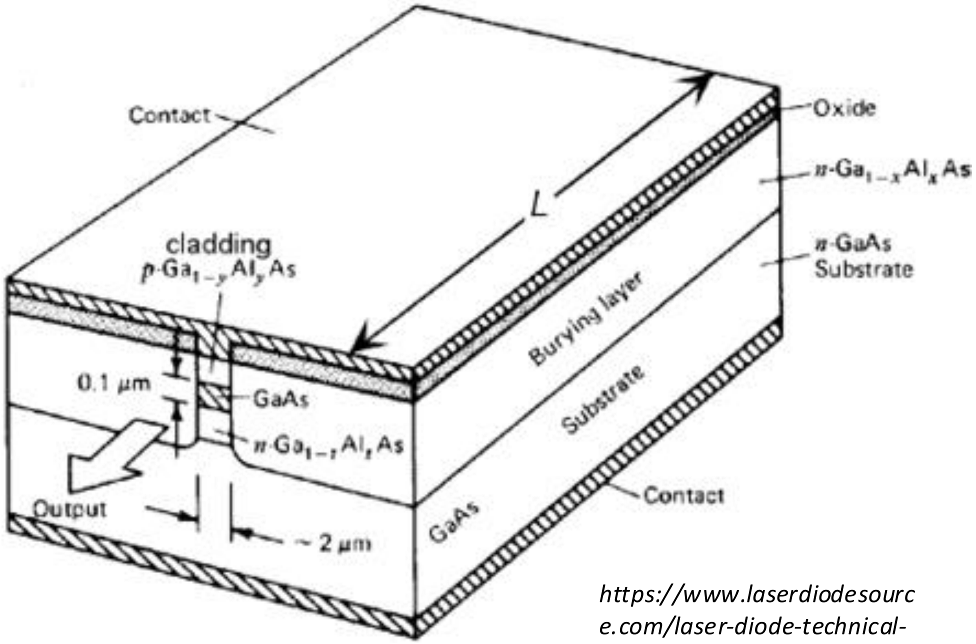
(a) Refractive index



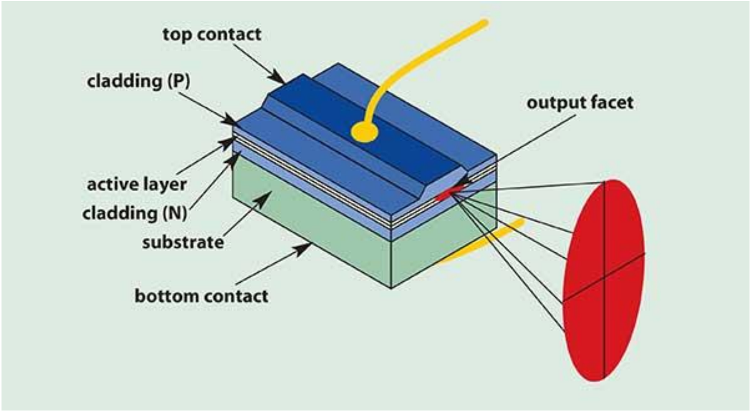
(b) Field distribution

https://www.laserdiodesource.com/laser-diode-technical-overview-three

To further confine the light to the active region, different material is also grown on either side of the junction. The device shown below is one example. All of this engineering has made it possible to operate laser diodes with much lower currents and therefore dispense with cryogenic level cooling.



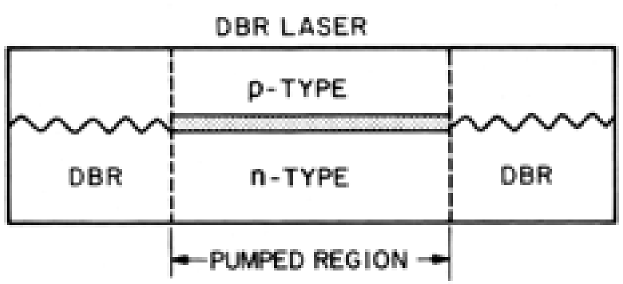
https://www.laserdiodesource.com/laser-diode-technical-overview-three



https://www.rpmclasers.com/blog/laser-diode-fundamentals-beam-properties/

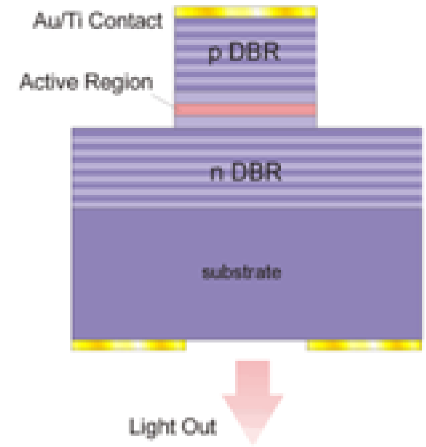
Resonant cavity

With just polished ends the diode laser is essentially a Fabry-Perot cavity and, as we saw, this will lase at all the standing wave frequencies that satisfy the gain condition. It's usually better to lase in one mode. You could make the cavity shorter to spread out the frequencies and make it more selective but that has its own problems. The more common way to do it is to incorporate a *distributed Bragg reflector* (DBR) - basically a periodic grating structure fabricated into the chip. The grating will preferentially reflect light whose wavelength is twice the grating period and let other wavelengths leak out so it's a frequency-selective mirror. The figure shows the idea.



https://www.laserdiodesource.com/laser-diode-technical-overview-three

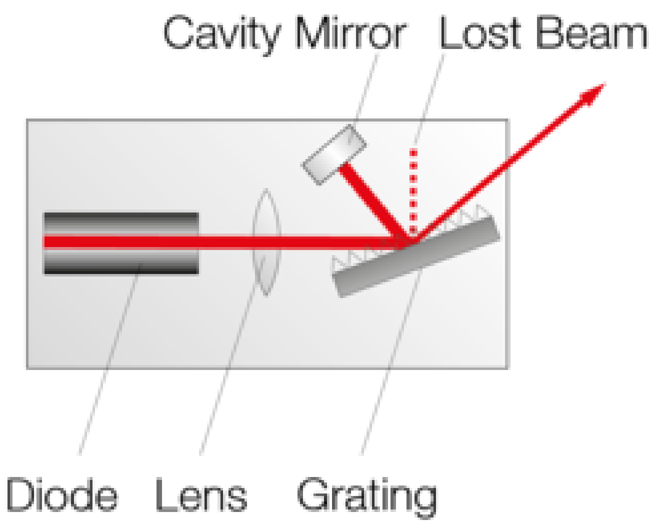
The devices shown so far are planar. The light is emitted from the active region out the side of the chip. An alternative way to do it is called a vertical cavity surface emitting laser (VCSEL) . Here, the light is bouncing back and forth *perpendicular* to the semiconductor layers. The Bragg reflectors are deposited as separate layers of semiconductor. Lateral confinement of the light comes about because the post is surrounded by air which has a lower index of refraction than the semiconductor. Light now comes out the bottom. Evidently these lasers are easier to make and more efficient than planar designs. Also, the cavity is very short, containing only a few wavelengths so they tend to oscillate in a single mode.



https://www.laserdiodesource.com/laser-diode-technical-overview-three

Wavelength tuning

Rather than put the Bragg mirrors inside the semiconductor it's also possible to make a diode laser with an *external* grating, as shown. Now, by changing the angle of the grating, it's select which wavelength will satisfy the frequency condition and lase. The grating can be tuned piezoelectrically to incredible precision so it's possible to fine-tune the laser output frequency over a broad range. This is called an external cavity diode laser.



https://www.toptica.com/application-notes/phase-and-frequency-locking-of-diode-lasers/actuators-for-laser-phase-and-frequency-control

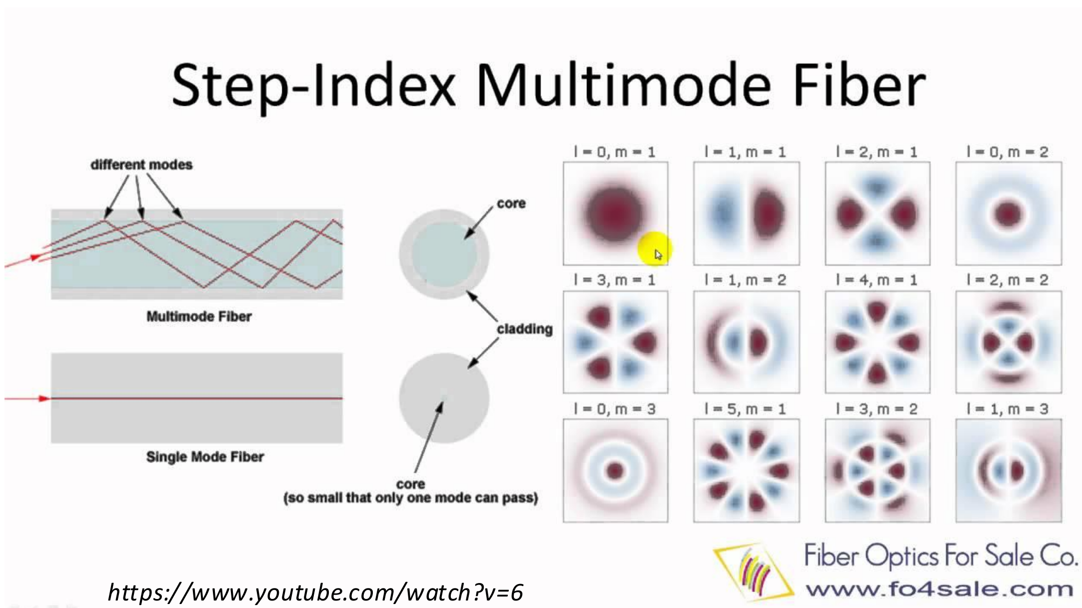
Optical fibers

Modern communications revolves around the transmission of light through optical fibers. This history is fascinating and goes back to the 1840's (https://en.wikipedia.org/wiki/Optical_fiber) when Daniel Colladon and Jacques Babinet showed that light could be guided around by refraction. An early "light fountain" is shown. Light fibers were used for medical examinations in the early 1900's and the Apollo moon mission used a TV camera containing optical fibers. In 1966 physicist Manfred Borner (Germany) patented the first fiber optic transmission system and 1960 Charles Kao (China) and George Hockman theorized that silica glass fibers could have attenuation as small as 20 dB/km. The manufacturing breakthrough came at Corning Glass Works in 1970 where they produced a fiber with 17 dB/km attenuation. These days, fibers for communication have 0.3 dB/km attenuation at 1550 nm (the so-called C-band). It all comes from total internal reflection.

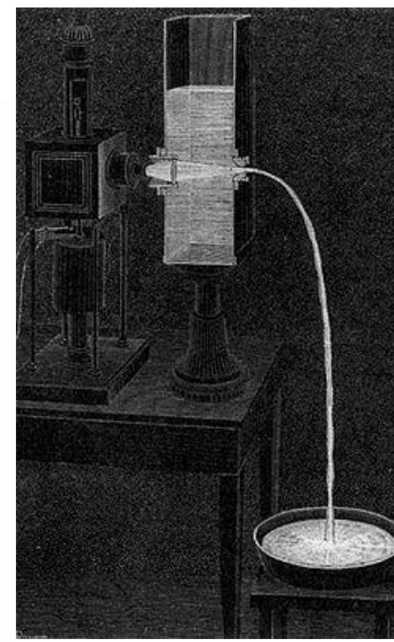
Manfred Borner

Charles Kao, Nobel Prize in Physics 2009

These days you can buy fiber that will support a single mode, multiple modes multiple cores, all of which can be doped with various elements to induce lasing. Modern fibers can carry communication signals 100 km before they need a boost.



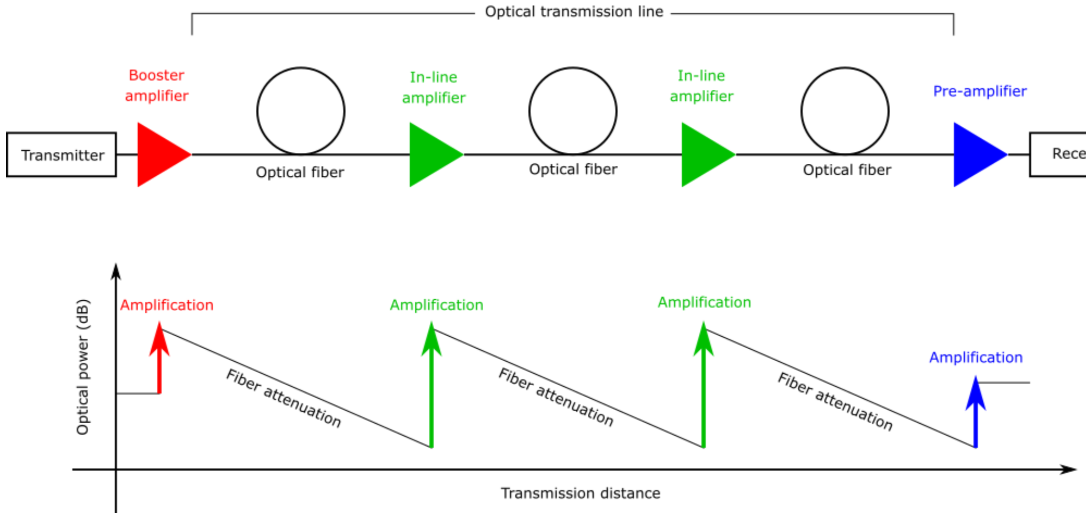
https://www.youtube.com/watch?v=6xYOzY4zj0o



A light fountain

Optical Amplifiers

Optical fibers have huge bandwidth and modern fibers can carry signals over long distances but they *do* attenuate the signal so it needs to be amplified along the way. In earlier incarnations the optical signal had to be converted to an electrical one, boosted by a conventional amplifier and then reconverted to an optical signal. The need for this complicated conversion and reconversion went away with the the invention of the erbium-doped fiber amp in 1987. It has the advantage of several closely spaced levels that will amplify over a range of wavelength from 1525 – 1575 nm. This is called *C-band*, used for optical communication because of its low attenuation in fibers.

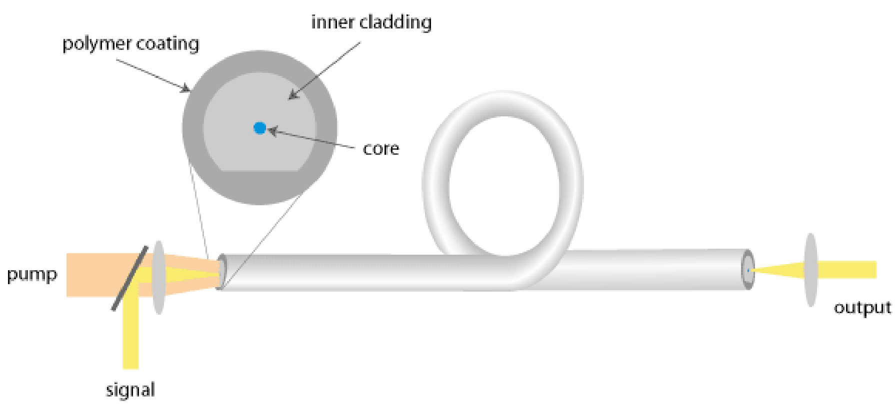


https://www.fiberlabs.com/glossary/erbium-doped-fiber-amplifier/

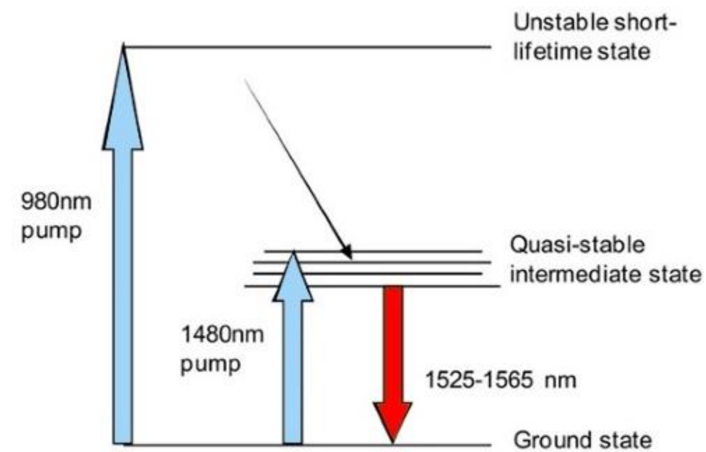
Er-doped fiber amps are commercial items. The one shown sells for around $1300 and has a gain of 35-45 dB.

https://www.civillaser.com/index.php?main_page=product_info&products_id=1148

The basic scheme is shown below. The gain medium is the core of an optical fiber doped with rare earth atoms – the most common is erbium (Er). The pump source is typically one or more diode lasers working in tandem to feed pump light into inner cladding of the fiber. The erbium ions are in the fiber core, which might be only a few microns in diameter. The erbium energy levels show strong absorption bands are at 1480 nm and 980 nm. Since this is basically a 3-level laser, a lot of pump power is needed, thus the need for a laser as a pumping source. By far the most common source is a single diode laser or several of them working in tandem.



https://www.rp-photonics.com/high_power_fiber_lasers_and_amplifiers.html

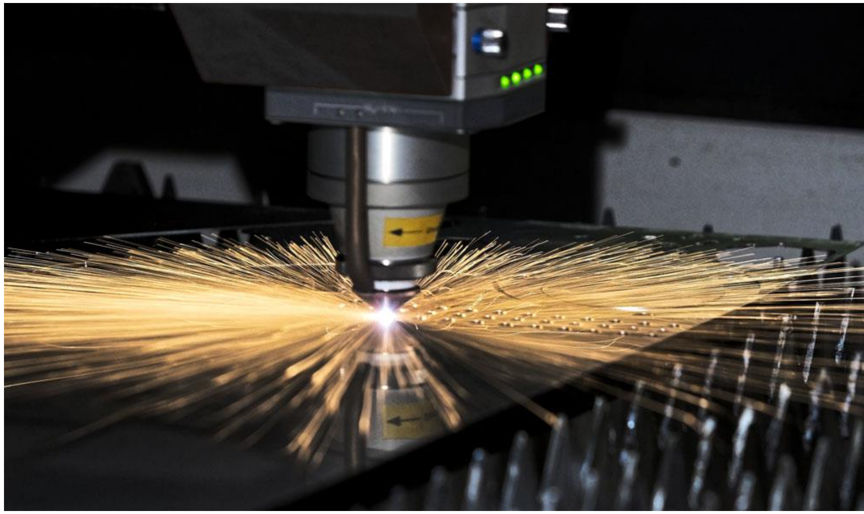


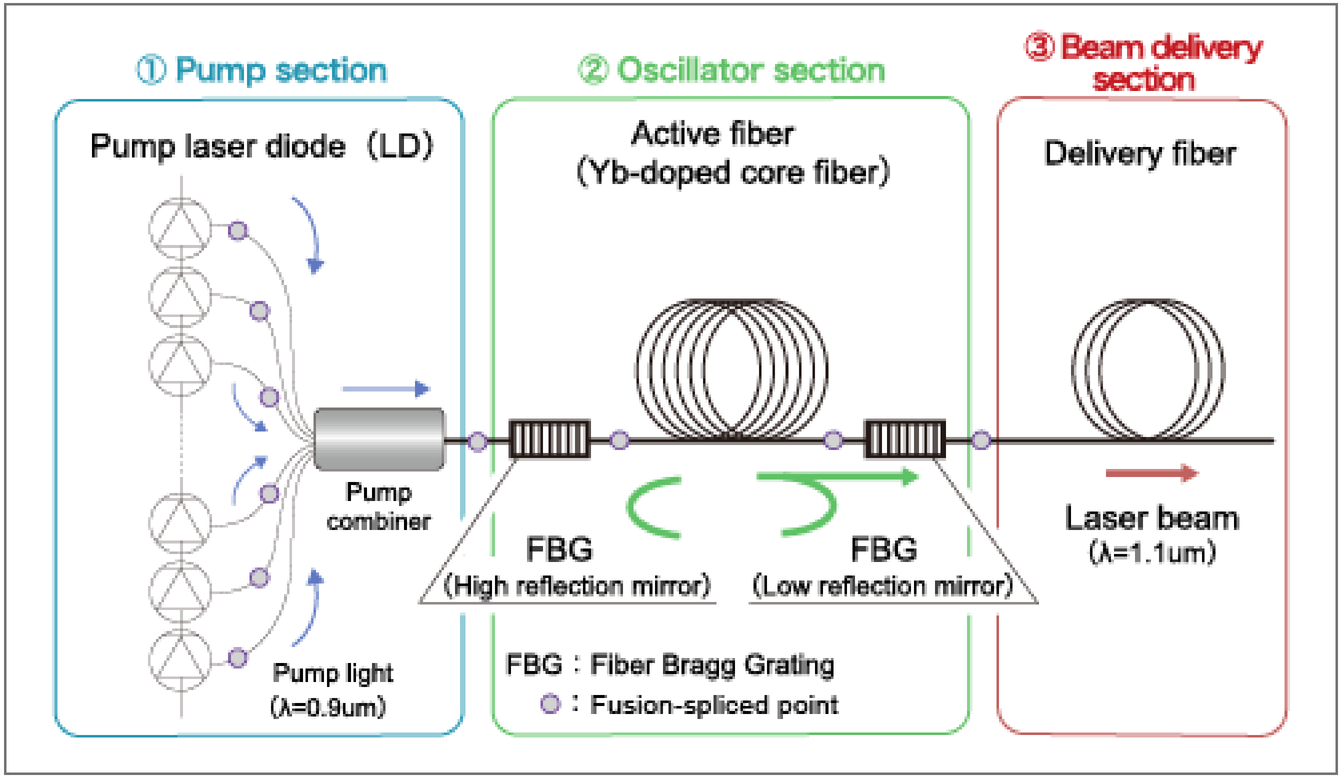
https://www.gophotonics.com/community/what-is-an-erbium-doped-fiber-amplifier-edfa

Why use fibers in the first place? They are more convenient than laser cavities since they are stable, compact, can be wound around, they easily dissipate heat and can be made to support single or multiple modes. The fibers have low attenuation so the gain region can be long and the core can be doped with many different atoms that will provide gain. Fiber amps are not particularly low noise. About the best noise figure is around 5 dB, which is nowhere near as good at typical electronic low noise amplifiers. Instead, fiber amps are used as moderate to high power amplifiers. Erbium-doping seems to be the current choice for optical communication booster amps. For much high power it turns out that Ytterbium is a better dopant.

Fiber Lasers

For all the reason fibers are good for amplifiers, they are also good for lasers. If you need the kind of power to weld or cut through a metal plate, it used to be a CO_2 gas laser. These days, your best choice is probably a fiber laser. Ytterbium is the rare earth dopant of choice for the gain medium. Since this is a laser it needs a cavity and for fibers that's done with a Bragg grating, similar to what we discussed in diode lasers. For serious power the fiber laser is pumped by several a laser diode array and the fiber itself may have more than one core to handle the high intensity.





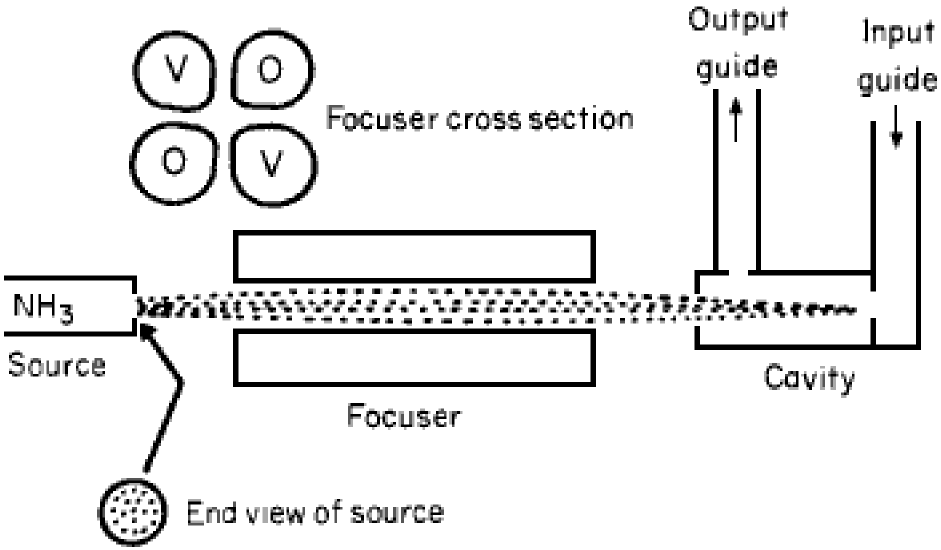
https://www.fiberlaser.fujikura.jp/eng/products/about-fiber-laser.html

The Maser

Before the laser came the maser, an idea put forth by N. Basov, N. Prokhorov and J. Weber in 1952. The first functioning maser used a microwave transition (f =24 GHz) in ammonia - NH_3. It was developed at Columbia U. by J.P. Gordon, H.J. Zeiger and C.H. Townes. The basic scheme is shown below. A beam of ammonia molecules streams from a source into the focuser, which consists of 4 long rods with voltages in a quadrupole arrangement, producing an electric field *gradient.* Molecules in the upper and lower energy states now feel a different force and are thereby spatially separated. It's like a Stern-Gerlach experiment but with electric fields. Only those in the upper state make it into the microwave cavity. Energy fed into the cavity at f =24 GHz now stimulates emission from those excited molecules, thus providing amplification. Unlike lasers, here the population inversion takes place spatially, *outside* the resonant the cavity.

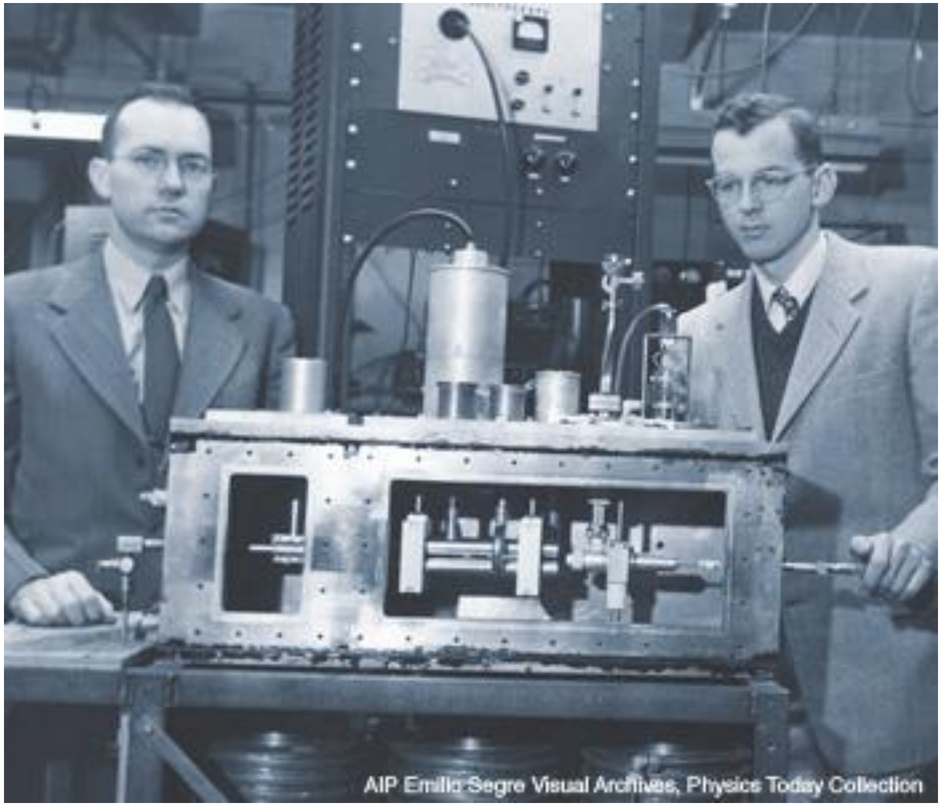
N. Basov

J. Weber



https://laserstars.org/history/ammonia.html

Microwave Amplifier



The ammonia maser is shown along with C.H. Townes (left) and his PhD student J.P. Gordon (right). Gordon is adjusting the microwave cavity. Masers are particularly interesting for their low noise properties. Remember back to the definition of noise temperature in an amplifier. That's the temperature of a resistor that would generate the same noise as an amplifier connected to that same resistor held at absolute zero. it turns out that masers can theoretically operate in the *quantum limit* of Nyquist noise,

$$T_{noise} = \frac{hf}{k_B \, ln2}$$

For f = 10 GHz that corresponds to $T_{noise} = 0.7\ K$, which is quite low. Unfortunately masers are, by their nature, extremely narrow band so they have somewhat limited application. These days, hydrogen masers acting as oscillators are used for extremely precise atomic clocks, i.e., frequency standards.

https://www.slideserve.com/donnadackow/donna-dackow?utm_source=slideserve&utm_medium=website&utm_campaign=auto+related+load#google_vignette

There are indeed masers that generate a lot of energy but they aren't man-made. They were first discovered in 1965 and attributed to some unknown substance termed "mysterium" . By now there appear to be many sources of maser action out in the universe. It's believed to come from stimulated emission because the spectrum is non-thermal and narrow band. Some of the sources include SiO, H_2O and even methanol. The figure shows radiation from the pole of Jupiter, observed by the Hubble telescope and believed to be cyclotron maser radiation.

https://en.wikipedia.org/wiki/Astrophysical_maser

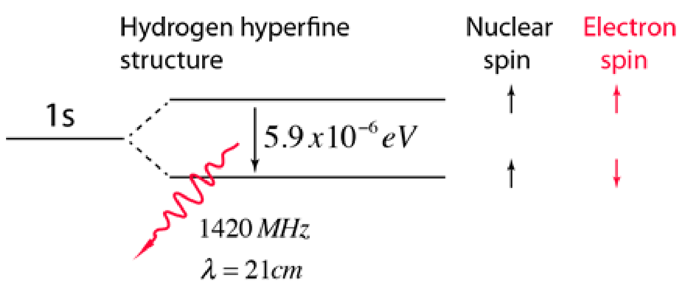
Hydrogen Maser

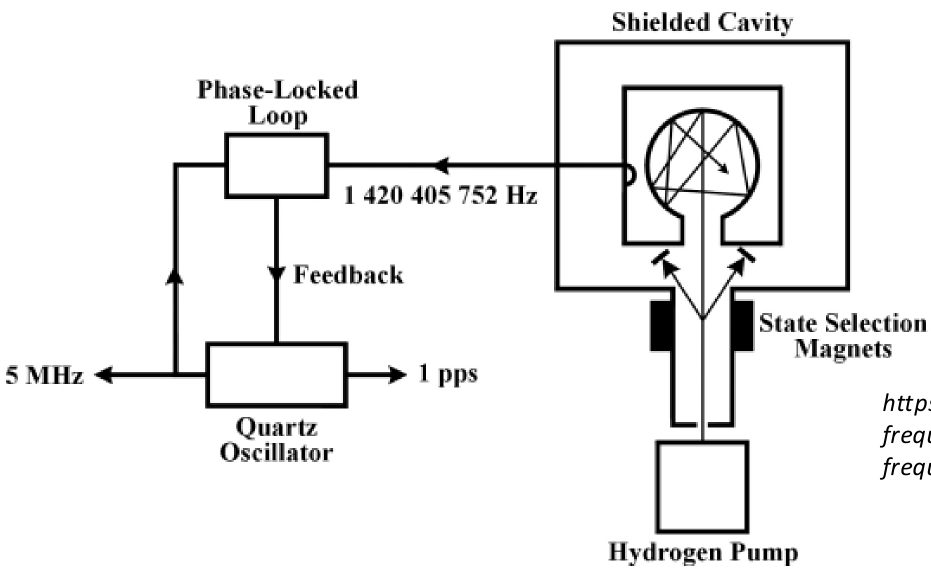
The hyperfine interaction splits the 1s ground state of hydrogen into two levels separated by 1420 MHz. This is the famous "21-cm" line first observed by E.M. Purcell and H.I. Ewen in 1951. (The same Purcell responsible for NMR.) The figure shows the horn antenna they used to observe it radiating from the Milky Way. It's of enormous significance to astrophysics and cosmology. Hydrogen masers based on this transition are used as atomic clocks, i.e., extremely stable frequency standards. Some of the microwave signal from the cavity goes into a phase-locked loop that locks to a quartz oscillator operating at 5 MHz, which is then available for stabilizing other frequency sources.

H. Ewen

E.M. Purcell

https://en.wikipedia.org/wiki/Hydrogen_line





https://www.nist.gov/pml/time-and-frequency-division/popular-links/time-frequency-z/time-and-frequency-z-h