1
A Scenic Route through the Laser

1.1 The Meaning of »Laser«

»Laser« is one of the rare acronyms whose meaning has not been lost over five decades. It stands for *L*ight *A*mplification by *S*timulated *E*mission *R*adiation. These words are not sufficient to clarify the meaning, unless we have a picture associated to each of them in our mind. The next few sections will be devoted to unraveling the meaning of each term.

1.1.1 Light (Photon) is a Wave

The first letter »L« tells us that »laser« is »light«. Light is an old known entity originating from the sun and the moon. Once it was associated with fire and thought to be the first essential element. In modern language we say that light actually consists of photons, just as matter is made of atoms. Our intuitive picture of atoms is that they can be nicely classified by their mass in a table – the Mendeleev table. Atoms themselves are boxes filled with electrons, protons and neutrons, and there is a mass associated with each component.

Atoms can combine to make molecules, the ultimate component we expect to arrive at when grinding to its finest constituent any piece of material, from a live leaf to a piece of paper. In a way molecules and atoms are what we deal with on a daily basis, but at such a fine scale that it escapes our direct perception. Photons are as ubiquitous, but quite different from atoms and their constituents. Ubiquitous, because they are associated not only with visible light, but also with invisible radiation (infrared and ultraviolet), x-rays, gamma rays, radio waves, and even the radiation from our electrical network at 50 or 60 Hz. They are quite different because there is no mass associated to

Lasers, The Power and Precision of Light, First Edition. Ladan Arissian and Jean-Claude Diels

the photon. A wave is associated with the photon, which is an oscillation propagating at the speed of light.

What is a wave? There is always a pattern and motion associated to the wave; the ripples of a stone thrown in a pond or folds of a flag. One can imagine more and more examples that the word »wave« is applied to. As physicists we would like to pause and clarify some features of the wave with definitions that can be used to quantify similar observations. If you take a picture from the ripples on the pond you realize that there are regular patterns that repeat in water, and you can possibly count the number of peaks on the water surface, that are separated by a »wavelength«. You can also only consider a fixed point on the surface and monitor its motion as it goes up and down, or oscillates. It takes a »period« for each point on the pond to repeat its position. The pattern on the wave (for example, the peaks) have a certain »speed« or »wave velocity«, and the peaks that are created by the wave have an »amplitude«. It is reasonable to conclude that stronger waves have bigger amplitudes, but there is more to the strength or energy of a wave, as we will see in the following sections.

When a wave goes through a medium it does not mean that the medium is necessarily moving with it. In the case of a flag waving in the wind, there is a wave that goes through the flag, but the fabric itself is not carried away. The wave propagates for huge distances, while each particle responsible for the wave motion stays at the same average position, just inducing the motion of the next particle. In most cases, the wave starts from a local oscillation (Figure 1.1a), and propagates radially from there, like rings produced by a duck paddling on a pond (Figure 1.1b). In the case of light, it is the electric field produced by a charge oscillating up and down that starts off the wave. This is called *dipole* emission.

The velocity of a wave is a property of the medium in which the wave propagates. Sound waves propagate at 343 m/s (1125 ft/s) in dry air at room temperature and faster in denser media. The opposite holds for light waves that usually travel faster in air or vacuum. There are different types of waves. Mechanical waves like spring oscillations and sound waves are due to mechanical motion of particles. The oscillation of these waves is along the propagation direction. Light waves, however, are electromagnetic waves, which originate from the oscillation of charges (electrons, for example). This was the first dilemma in early attempts to interpret light waves: what is moving?

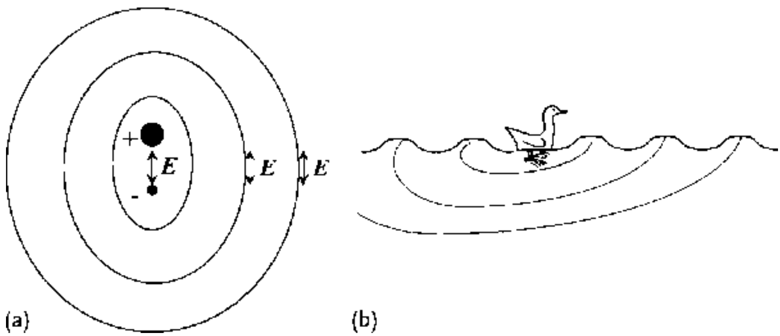


Figure 1.1 (a) An oscillating electric field is created by a pair of charges with a periodically varying distance (oscillating dipole). This periodically varying field creates an electromagnetic wave that propagates at the speed of light in vacuum, just as a wave is created in (b) by a duck paddling in a pond.

Our intuition is shaped by the observation of water waves in a pond, an oscillating spring, or the swing of a pendulum. These are all mechanical waves. Like sound waves, they require a medium; they need matter to exist. Hence was born the notion of the »ether«, a (fictitious?) medium to support the propagation of light waves. Today the »ether« has simply given way to vacuum, but it does not mean that the understanding of the nature of light has become simpler.

As will be explained in Section 1.1.2 below, quantum mechanics tells us that the amplitude of the positive–negative charge oscillation is restricted to discrete values. Consequently, the emitted oscillation also takes discrete values, to which is associated an energy: the photon energy $h\nu$, where ν the frequency of the oscillation, and h is called the »Planck constant« (see Eq. (1.1) in the next section). It is as if the duckling in Figure 1.1a had discrete gears to activate his webbed paws. What is more puzzling is that the »neutral« gear is missing. The minimum energy state of the quantum harmonic oscillator is not zero, but $(1/2)h\nu$. This is often referred to as *vacuum fluctuation* or zero point energy. The absence of vacuum (the ether concept) has been replaced by an absence of zero energy. Since, according to Einstein, there is an equivalence of matter and energy, the two concepts are not so far apart.

1.1.2 Photon Energy

Quantum mechanics tells us that a photon has dual characteristics, it acts both as a wave (Section 1.1.1) and as a particle. In a way, the photon is a wave that can be counted. This might be a bit hard to digest, since our common sense is restricted to our daily experience with objects that are not so delicate. What do we mean by acting like a particle? They can be counted. A photon is like a »currency«, and the light that we experience is like a sum of money, we never notice that tiny penny.

Let us take a closer look and see why we generally ignore single photons. A typical red laser pointer has an output power of 3 mW (3 mJ/s), which consists of individual photons having an energy of the order of 3×10^{-19} J. This means that every second there are 10 000 million million photons shooting out of a pointer. If we associate even a penny to each photon, in a second we get a sum of money that is more than the wealth of a country.

Just as not all currencies have the same value, photons have different energies. Here we need to use the wave aspect of the photon. The faster a wave oscillates, the more energy it possesses.

The longest (slower) electromagnetic wave that we encounter in our daily life is created by the 50 Hz electrical network covering the globe. As a result the earth radiates, making one oscillation over a distance of 6000 km. Radio waves are long too: it takes 3 m (3.3 yd) for a short wave (FM radio) to make an oscillation. For a long wave (AM) it takes about 300 m (330 yd) to complete one.

The visible light that we are used to also oscillates, but much faster. The green visible light consists of photons of 500 nm wavelength; meaning that over a thickness of a sheet of paper (which is 0.1 mm or 0.004 of an inch) it makes 200 oscillations. An x-ray with a wavelength of about 1 nm, oscillates 100 000 times over the same length. It thus appears that the following connection exists: photons that oscillate faster have a shorter wavelength, and more energy. Or in the simplified language of mathematics

$$E = h\nu = \frac{hc}{\lambda}\,, \tag{1.1}$$

where »E« stands for energy, »h« is the physical Planck's constant, »c« is the speed of light, »ν« is the number of oscillations of a pho-

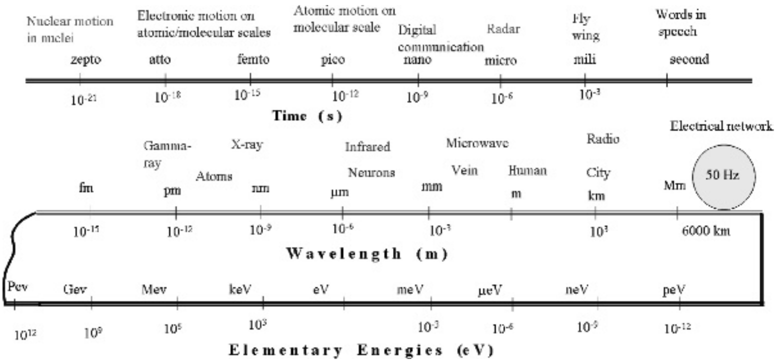


Figure 1.2 Different objects that radiate electromagnetic waves and the wavelength and elementary energy associated with them. Please find a color version of this figure on the color plates.

ton in a second, and »λ« is the wavelength, or the length in which a single oscillation takes place. For the photon associated with visible radiation, the elementary photon energy is too small to use the traditional energy unit of Joule. Instead, the energy unit used by physicists is the electronvolt (eV). 1 eV (1.602×10^{-19} J) is the energy acquired by an electron that is accelerated under the potential difference of 1 V. Infrared radiation at a wavelength of 1.24 µm has exactly the energy of 1 eV. As shown in Figure 1.2, our earth, due to the electric power network, radiates photons of 2.067×10^{-13} eV energy.

1.1.3 Energy and Size

Could »Spiderman« really have the strength of a spider, scaled up to his size? Is a cat that is 100 times more massive than a mouse 100 times stronger? In biology things will not scale linearly. Body mass increases linearly with volume in three dimensions, while muscle strength in arms and legs is proportional to cross-sections, and therefore increases only in two dimensions. If a human is a million times more massive than an ant, he is only 10 000 times stronger. In a way smaller animals are stronger relative to their masses. Physics scales in a simpler way than biology. In a musical instrument higher frequencies are generated by shorter strings, thus have more energy. Some physicists like to draw a box around the object that they study, and they know that as the box gets smaller they are dealing with higher

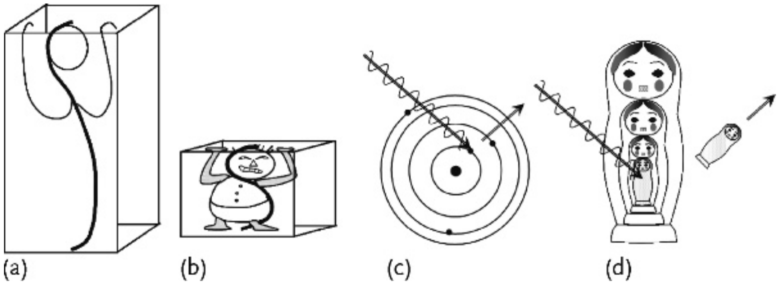


Figure 1.3 Particle and wave in a box: the longer wavelength fits in the larger box (a). A shorter wavelength fits in the smaller box (b), corresponding also to a larger particle energy. Electrons (c) orbiting around the nucleus are analogous to nested Russian dolls (d). A photon of sufficient energy can knock off the electron of the outer shell, as one can easily remove the outer layer of the nested dolls. More problematic is the removal of an inner shell electron. While it would be an unresolvable »Chinese puzzle« to remove the inner doll, the possibility to eject an inner shell electron exists with high energy photons.

and higher energies. The speed and energy of the electrons oscillating in an atom are much bigger than the ones traveling in a long wire loop.

Using our wave picture and the equation of photon energy (1.1) we can look more closely at the size–energy relation. Consider fitting one full wave into two different size boxes. The wave that fits in the smaller box (Figure 1.3b) has a shorter wavelength than the one in the bigger box (Figure 1.3a). Using the photon energy equation (1.1), the wave with a smaller wavelength has higher energy. It seems that the more confined the wave, the stronger its elementary energy. This seems like an oppression force! Quantum mechanics tells us that the electrons around an atom are confined to well defined shells or electron levels like Russian nesting dolls (Figure 1.3c, d). The electrons in bigger shells have less energy and are loosely bound, which is why in most ionization processes the chance of knocking off an electron from an outer shell is the highest. This order is not as rigid as the order of taking out the Russian dolls: when dealing with higher energies photons, it is possible to scoop up the electron from an internal shell, leaving the external ones in place (not something you could to with the Russian dolls).

Looking at the waves in boxes, we only concentrated on the concepts of size, energy, and wavelength. There are other parallel manifestations of waves, such as in time and frequency. This means that in smaller boxes we are dealing with shorter time scales and faster frequencies. This is a very important fact for observing phenomena in nature: we need the proper time scale to watch an event. The speed that we take our data, or snapshots of an event, tells us what we can record and what we would miss. Imagine that we could only understand one word per minute. It would then be very difficult to get the meaning of sentences in a normal conversation. In cinematography 24 frames per second is sufficient to capture most daily events, but if we want to see faster events like a bullet passing through a target, we need to take more frames per second and then play it at a normal 24 frame per second rate. In Figure 1.2 the time scale of some events are shown. From this diagram one can see that in order to observe electron motion in atoms and molecules one should have a camera that takes picture every femtosecond. It is hard to imagine how fast a femtosecond is. If one could read the first Harry Potter book with 2.7 million words at the rate of one word per femtosecond, the whole book would be finished in a few nanosecond. Or, in other words, you could read Harry Potter a thousand million times in just 1 s. An attosecond is to the second, as a second is to the age of the Universe (0.4×10^{18} s). Figure 1.2 illustrates the various sources of radiation that affect our lives and how they are associated with energy, size, and time.

1.1.3.1 Light Energy

In the previous sections we talked about the energy of an individual photon; but only in a very sophisticated laboratory can one isolate a single photon. What we are exposed to usually consists of a large number of these photons. One can look at a light source as the wealth of a country and a photon as its currency unit. There is no country with a total wealth of only one unit of currency. For example, when we talk about wealth and debt of the United States of America, the numbers are in trillions, and it is quite easy not to mention a cent or two. That tiny cent which is the unit of currency is like a single photon. It exists, but is always lost in the total sum, just like a photon in a light source.

However, the analogy between currency and light stumbles at laser light. In our daily lives and economics there is only one way to add numbers. Luckily in science there are more. Let us assume not one duck as in Figure 1.1b, but two, paddling in synchronism as if competing in synchronized swimming. Along certain directions the wave will add »in phase«, along other directions they will cancel each other or »interfere destructively«. In the analogy of the coins, a large number N of photons of the same frequency will add up to N coins, for a total optical energy of $Nh\nu$. This is how daily white light of *incoherent light* adds up. In the case of laser light, or *coherent light,* N properly positioned and synchronized photons will produce a beam of $N^2 h\nu$. The reason is simple: it is the electric field of the waves that adds up coherently, which means overlapping the oscillations in phase (crest to crest and trough to trough). The total electric field is the sum of the electric field of each source element. However, the total *intensity* of the beam is proportional to the *square* of the total field and is thus proportional to the square of the number of emitters. When the waves are piled up randomly (incoherently) the phases do not match and only intensities are added linearly. Incoherent light is a crude summation over waves, it just adds the average intensity values and misses the phases. This property is being exploited by the US Air Force to create very high power lasers, by adding the beams of an array of fiber lasers.

One example of coherence can be found in the stock market. If we knew how to add coherently the phase fluctuations of the market, there would be no limit to our profit. It would take a financial wizard to buy stocks in a proper phase (exactly when they are at the lowest) and sell at the highest (peak). Such great wisdom is what fell upon physicists when they invented the laser; they found a way to add the photons coherently, stacking the waves on top of each other with equal phases.

How do we explain explain that incoherent »white light« adds up proportionally to the number of emitters N, rather than N^2? A similar addition takes place in daily life, at least for some of us. Suppose that, after having celebrated his graduation with numerous drinks, your inebriated friend starts heading for home, after hugging his favorite lamp post for the last time. The analogy of the photon field is his step. The total field is represented by the total distance from the lamp post to the end of the last step. How far will he have moved away from his starting point after $N = 100$ steps? The answer is: a distance of 10 steps away from the lamp post or $\sqrt{N}$ ft. If he had been sober, he would have moved 100 steps (N) straight towards his destination. Incoherent light fields add as incoherently as the steps of the drunkard: the sum of N elementary fields E of random direction adds up to a total field of magnitude $\sqrt{N}E$. An intensity I proportional to the square of the field is associated to each field E. The total intensity, in the incoherent sum, is $(\sqrt{N})^2 I = NI$. In the case of laser light or coherent light, the elementary fields line up, resulting in a total field NE and a total intensity $I_{\text{total}} = N^2 I$.

1.2 Radiation from an Atom

So far we have talked about photons as units of light energy that can be counted. They are not rigid; they act and move like waves. Now we want to know where they come from and how we can make them. In the following sections we learn how to add them properly (coherently) in a laser source. Photons are electromagnetic waves which are generated when charges oscillate. It is natural to turn our attention to atoms and molecules where there are as many electrons as the atomic number. Quantum mechanics tells us that not all the areas around the nucleus have the same chance of having electrons; electrons are stacked in shells just like the layers of Russian dolls. In the quantum mechanics world matter acts like waves, which means that it is hard to point at them as they are moving around. Electrons move around the atom nucleus in well defined shells. There are steps of energy and an electron can absorb a photon to change its energy status from one shell to another. Nothing gets lost in nature: if an electron goes from a lower energy shell to a higher energy one, it must have absorbed the energy difference. This is reversible: if an electron moves from a higher energy to a lower energy state, it must release the energy difference. That energy difference can be given entirely to a photon with an energy that satisfies the energy difference.

Just like the Brownian motion of particles, there is always fluctuations and excitements in the universe even in very remote places far from all galaxies. The electrons are just like children playing on slides in a playground as illustrated in Figure 1.4. Here, there are also clear energy levels; a lower energy on the ground and a higher energy up the slide. The children here use their muscle energy to go up the slide and they enjoy the release of energy when they slide down. This is similar to what happens around us, in all atoms. The electrons can pick up energy from photons around us, or even from thermal energy to climb the ladder and jump back down. That is how photons are absorbed and created.

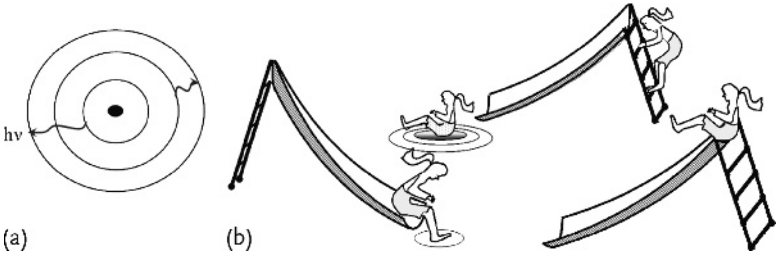


Figure 1.4 (a) A schematic of energy levels of an electron in an atom. (b) Children riding on slides; an analogy of electrons moving from one energy state to the other. We just have to imagine that the sound, momentum, and the joy of sliding down is the released photon energy. In the absence of coordination, this radiation is random or »spontaneous«.

The change in energy levels resulting in photon release is called »radiation«. This random radiation is everywhere and is »incoherent«. More than 40 years ahead of the realization of lasers, Albert Einstein formulated two types of radiation: »spontaneous« and »stimulated«. Spontaneous emission, like it sounds, is just the release of energy from higher to lower states without any coordination, just like the children in the playground of Figure 1.4. If the game is converted to a match by adding a referee on the ground giving the »start« signal, the motion of the children will be different. They will all wait at the top of the slide and start sliding at the gunshot. Here the radiation is stimulated. Another example of the comparison between stimulated and spontaneous emission can be found in shopping. Let us consider a number of people in the mall who are there with the intention of buying clothes and have the money in hand. They are all excited, just like our electrons in the upper state ready to emit. If there is a nice advertisement or a well dressed person walking by, our shoppers follow the signs and buy the same thing as they are being stimulated to. The stimulating advertisement or person puts their intentions in phase, inciting shoppers to walk in the same direction and simultaneously buy the same item. They are as hypnotized!

In the case of the emission from electrons in the excited states, a photon with similar energy difference in the vicinity of the atom causes the stimulated emission. One photon passing by one excited atom will stimulate the emission of another photon. Before the event, there was one photon and one excited atom. After the photon flew by, the atom energy has been reduced by the photon energy, and two identical photons are pursuing their course. The total energy of atom + radiation is the same before and after the interaction.

Having many atoms with electrons in the excited state, we need just a few passing by photons to make a cascade of photons with the same energy radiating at the same time, so well in phase. These photons also follow the path of the inciting photon and have the same direction. They have the same polarization, a concept that will be introduced in the following section.

1.2.1 Absorption and Dispersion

An electron can absorb a photon (or multiple photons in a nonlinear interaction) and move from one energy level to a higher level; it can also emit a photon and go to a lower level. The former is called *absorption* and the latter is associated to *gain*. The first one is like upgrading living standards by moving to a better house and putting down a sum of money; the second one is downgrading housing and releasing a sum of money for other causes. The house price is a fixed value and transition is only made when the money on the offer letter is provided. In the world of quantum mechanics, there is no fixed number and everything is described with *probabilities*. If the light frequency matches the separation the transition is more probable than when the frequency is off, but still even if the frequency is off, the transition in possible. The lines are not so sharp in quantum mechanics and there is always a shadow around them.

As seen in Figure 1.5, the absorption (which is the opposite of the gain) maximizes for the characteristic frequency »ν_0«. The absorption is not a delta function and has a characteristic *absorption width* in which the absorption is appreciated. Usually the absorption width is the range of frequencies over which the absorption is above half of its maximum value. In electromagnetism the *permittivity* »ϵ« is the resistance of the material to the light propagation, or how much it *permits* the light to go through. It is directly related to the electric *susceptibility* »χ,« which defines how easily the material is polarized by the electric field.

Absorption and index of refraction are two aspects of the complex function of light permittivity, the former being the imaginary part and the latter the real part. They are related mathematically through Kramers–Krönig relations. Using these relations, the index of refrac-

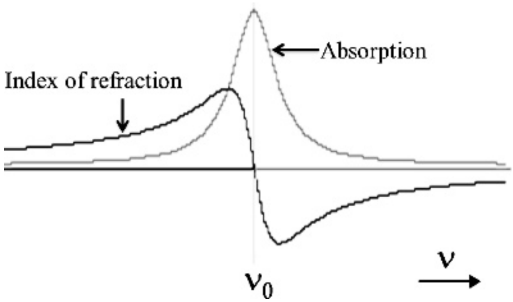


Figure 1.5 Absorption and index of refraction as a function of frequency ν. The absorption is maximized for the characteristic frequency ν_0. Any gain or absorption structure in the medium will effect the speed of light; as seen in the figure the light slows down for the frequencies less than the resonance and speeds up for the frequencies that are greater.

tion at a particular frequency can be calculated if the absorption at all frequencies is known, and vice versa.[1] The physical reason for this mathematical relation is »*causality*«; the fact that the material has a response time to the electric field. The response comes *after* applying the field. The speed of light »c/n« is directly related to the index of refraction »n«. No wonder the index of refraction curve in Figure 1.5 is known as the »*dispersion*« curve: the speed of light is lower for low frequencies and higher for frequencies higher than ν_0. A pulse covering all frequencies will be negatively chirped, with the frequency of the optical oscillation decreasing with time.

1.2.2 Polarization of Light

Polarization of light refers to the direction and path of the electromagnetic wave as it propagates. This property needs to be visualized in three dimensions. As for single photons, there is only one path for the wave: a circular path or »circular polarization«. A circular polarization propagates just like following the grooves of a screw, with one difference: most screws are tightened in a clockwise direction. They have only one helicity. A single photon has an equal chance of having either choices of helicity. Using the image of a screw, there are as many »left threaded« as »right threaded« screws. A ray of light con-

1) The opposite case, which is deriving the absorption from the measurement of the index of refraction at all frequencies, is mathematically possible but rarely practiced in reality.

sists of many such photons. By combining an equal number of photons with different helicity *in phase* one can get a linearly polarized light. A linear polarized light is a wave that oscillates perpendicular to the propagation direction and it can be visualized in two dimensions.

1.2.3 Beam Divergence and Resolution Criterion

Diffraction is a well known phenomenon in optics. When a wave passes through an aperture comparable to its wavelength, it gets distorted. When a bullet gets through a hole comparable to its size, it scratches the walls. For waves, the distortion is imprinted on the pattern as can be seen with water waves on a pond hitting a small obstacle. The reason for this distortion is wave addition or interference. The waves are generated at each point of the aperture. Looking at the two selected points at the edge of the aperture and at the center, we can add the wave amplitude coming from these two points on the observation screen. Since a wave completely changes its sign at half of its wavelength, if the difference in the path from the two beams is $\lambda/2$, there will be a shadow on a distant point on which the light transmitted by the aperture projected. The same difference is observed as we slide the two chosen points across the opening. They all result in the same path difference on a selected point on the screen.

Diffraction of a beam through an aperture is illustrated in Figure 1.6. The wave going through a circular opening is distorted. The black and white stripes of the beam are successive wavefronts, spaced by the wavelength. Being illuminated by a plane wave, all the points in the plane of the aperture are in the same phase of the oscillation (represented by the bright segment). As the wave from each point on the aperture propagates through different paths to the end screen, each one is in a different phase of its oscillation. The combination of all these rays, taking into account their interference, has an intensity distribution indicated by the red curve. This distribution depends on the shape and size of the hole that the wave has traversed. One can even follow the gray line that corresponds to the darkness on the right screen. The blue ray indicates a line corresponding to the first dark ring on the screen.

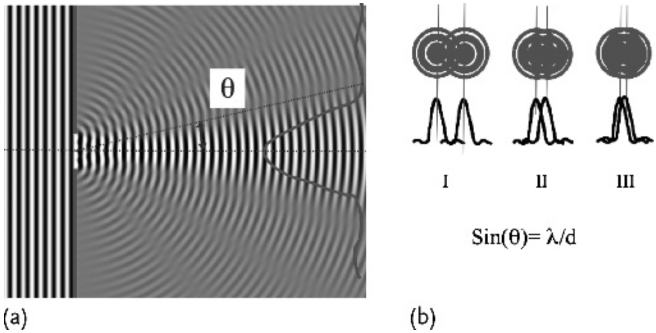


Figure 1.6 (a) Diffraction of a beam through an aperture. The path difference varies across the screen, showing more intensities where peaks from waves of different parts of the opening coincide and darker regions where peaks and valley coincide on the screen This results in an image (red pattern) of the aperture that is not so clear. (b) It is shown how far two points can be from each other in order to be distinguished. In order to resolve two images, the »Rayleigh criterion« defines the distance between the two points to be such that the first minimum ring overlaps the central maximum of the other image (case II). If the distance between the two points is larger (case I) the images are better resolved, while if the distance is smaller (case III) the resolution is lost. Please find a color version of this figure on the color plates.

For a plane wave[2] beam incident on a circular aperture, as shown in Figure 1.6, the angular observation of the dark rings on the screen follows the equation $D\sin(\theta_n) = n\lambda$, where $n = 1, 2, 3, \ldots$ is an integer, »D« is the size of the hole, and θ the angle of minimum brightness observed on the screen. For the first darkness »n« is 1.

A hole on the wave path is like a source emitting light through its extension, just like anything we want to observe by direct radiation or reflection. How close objects can be resolved is a major concern in all imaging techniques, from observing the stars to looking at tiny particles with a microscope. The resolution limit in imaging is called the Rayleigh criterion, which corresponds to having a minimum of one diffraction pattern overlap the maximum of the other one (II in Figure 1.6b). If the objects are closer (III in Figure 1.6b), the two peaks will merge, making it difficult to determine whether there are two objects. It is like having two people talking simultaneously. If there is a small delay between two words coming from the two persons, we might have a chance to catch the beginning of a one word and the

2) The plane wave is the simplest of all wave structures in space. It is an infinite sinusoidal wave with no transverse structure. It is represented as vertical, positive and negative stripes when the beam moves horizontally.

ending of the other one and record the two. However, when they speak within fractions of the second, it is not easy to distinguish both words clearly.

In imaging with lenses, from two objects, the Rayleigh criterion requires the distance between the two objects to be greater than $0.6\lambda/N_A$, N_A stands for numerical aperture. For imaging with a lens, $N_A = f/D$, where »*f*« is the focal length and, »*D*« is the diameter of the lens. Here the diameter of the lens corresponds to the hole size in the diffraction image. The Rayleigh criterion is a convenient concept used for any imaging technique from astronomy to microscopy.

1.3 The Anatomy of a Laser

Despite the simplicity of its acronym (*L*ight *A*mplification by *S*timulated *E*mission *R*adiation), the laser is more than just an optical amplifier. In general it has three main components: some »*pump*« mechanism to excite the radiation in a medium with optical »*gain*«, and a »*resonator*«, responsible for selecting the laser wavelength. The »pump« power can be electrical like a battery in a laser diode, or it can be another optical source. Its purpose is to excite the radiation. The »gain« medium can be a solid crystal or gas, prepared well to be excited and emit photons. We can use the photons radiated by the gain medium to induce »stimulated emission« by sending them back through the gain medium, provided it is still prepared in its excited state by the »pump«. A feedback mechanism is thus needed to return the waves, in phase, to the gain medium. This feedback is provided by a resonator, usually called the »laser cavity«. This resonator is not mentioned in the acronym, despite its essential role in the generation of laser light. There are some high gain pulsed lasers – such as the nitrogen laser or some excimer lasers, where the resonator is limited to one or no mirror. Because of the high gain, »lasing« occurs through the amplification of the spontaneous emission along the axis of the gain medium.

1.3.1 The Pump

The pump is the power plant for the laser and can take many different forms. It can be a source of incoherent light, like in many pulsed solid state lasers, it can be electrical, like in gas discharge lasers or in semiconductor lasers, it can be chemical energy, which is often the case in very high power infrared lasers used by the military, or it can even be another laser, which is the case for many crystal solid state lasers (the titanium sapphire laser, for instance) and dye lasers. It is often the pump that determines how efficient a laser system will be. There is a tremendous range of variation in the »wall plug efficiency« of laser systems. In a Ti:sapphire laser pumped by an argon laser, the argon laser will require over 30 kW of electrical power, to finally produce a Ti:sapphire laser beam of some 100 mW power! At the other extreme are the CO_2 laser and semiconductor lasers where the light output power is a sizable fraction of the electrical power input.

1.3.2 Gain

At the heart of a laser is a medium that has optical gain. It consists of an assembly of atoms or molecules that are in an excited energy state. In the analogy of the slides (see Figure 1.4) there are more children at the top than at the bottom of the slide. As a light pulse containing n photons[3] passes through such an excited medium, each photon will induce a certain number of stimulated emissions, resulting in an exponential increase of the number of photons with distance (in the medium). This is the reverse of the process of absorption, where the atoms are initially in the nonexcited state (»ground state«), and each photon will induce a certain number of upward transitions, resulting in an exponential decrease of the number of photons.

The higher the little girls stay on top of the slides before jumping, the more will be up on average. This is how a »population inversion«, necessary to achieve gain, is created. In laser jargon, an atom likely to be a good candidate as laser gain medium, is one where the upper level of the transition has a long lifetime. After each jump, the little child must quickly move away from the bottom of the slide, or there will be a traffic jam. Similarly, an essential condition for an efficient laser gain medium is that the lower energy state of the transition has a very short lifetime.

There is a process described in the next section by which atoms are put in the upper state: the »pump«. Given a constant rate of excita-

3) An optical pulse of a given energy in joules.
4) An optical beam of a given power in joules/s or watts.

tion into the upper state, in a continuous laser, n photons/s[4] passing through that medium will induce a certain number of stimulated emissions, resulting in an exponential increase of the photon flux (or beam power) with distance.

1.3.3 The Laser Body (Resonator)

Musical instruments are designed based on the fact that the size and geometry of the instrument defines what frequencies can be resonated in it (Section 1.1.3). However, it took a long time to apply this fact and combine it with all other elements to make a laser.

The analogy that we want to draw from in this section is that of an acoustic amplifier between microphone and speaker, when the microphone is put too close to the speaker. We have all heard the high pitched noise that drowns any other sound when the gain of the amplifier is turned up. Noise from the speaker enters the microphone, gets amplified, is blasted from the speaker into the microphone again, and so on. It is a case of a dog chasing its tail.

The same synergy exists between the optical gain medium and the optical resonator as between the oscillating amplifier and a musical resonance between microphone and speaker. The laser is a similar arrangement: an optical – rather than acoustic – amplifier where the output (i.e., the speaker) is fed back to the »input«. In the »acoustic amplifier«, the frequency, or pitch, of the oscillation can be set by any object with an acoustic resonance: a crystal vase, a glass, the string of a guitar, a tuning fork, a cymbal, a drum, ...

The closest analogy to the laser is that sketched in Figure 1.7a, where the microphone is inserted in a drum. It takes the least amount of energy to excite the note for which the drum is tuned. The noise from the speaker puts the membrane of the drum in vibration, result-

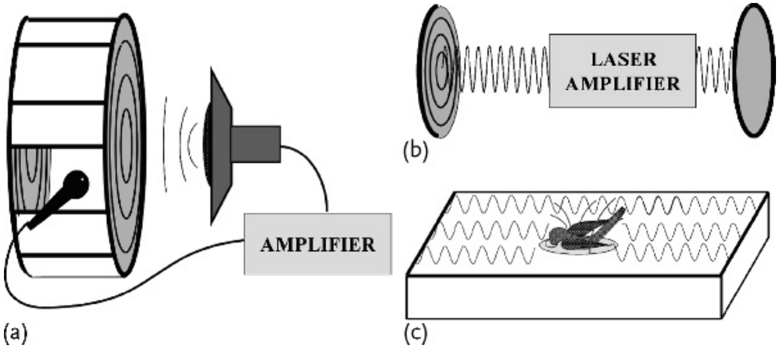


Figure 1.7 (a) An acoustic analogy of the laser: an amplifier with feedback. The pitch or wavelength of the sound produced is determined by the acoustic (mechanical) resonance of the drum. The laser is also an amplifier with feedback, where gain (Section 1.3.2) plays the role of an amplifier. A light beam that happens to be perpendicular to the two mirrors on either side of the optical gain medium is amplified. As the light wave makes a round trip from one mirror to the other and back to the initial point, it has to add upon itself, which imposes that the round trip length should contain an integer number of waves. The same situation occurs for the bird taking a bath: out of all the »noise waves« that it produces, the container is a resonator that will select the ones that add up upon themselves, hence have an integer number of crests between the walls.

ing in a plane compression wave of air propagating between the two faces of the drum, which is picked up by the microphone, amplifier, and fed back to the drum. A similar scenario takes place in the laser as sketched in Figure 1.7b. In the latter case, the resonance is provided by the mirrors guiding the light back and forth through the optical amplifier. Some of the sound is lost in going from the microphone to the amplifier. The gain has to be turned up to compensate for these losses. That minimum gain is called the threshold of oscillation, or the *lasing threshold* in the case of a laser.

An integer number of wavelengths has to fit in the »box« defined by the two mirrors of the laser resonator, otherwise the crest and the trough of successive waves will cancel each other. A similar situation occurs when a bird takes a bath in a perfectly rectangular container. No matter how wild it splashes in the water, it will produce standing waves that fit exactly the length of the box. This is a one-dimensional analogy – we should assume here that the transverse dimension of the box is infinite.

The role of the mirrors is twofold: they provide the wavelength selection, as well as the directionality of the laser beam. If the mirrors are perfectly plane, only the waves that are exactly, rigorously orthogonal to these two planes can provide the »feedback« to the optical amplifier. Only the first few lasers were built with flat mirrors. It was soon realized that curved mirrors made a laser considerably easier to align: a nonperfectly axial beam will still reflect back within a narrow cone, but it is no longer the »ideal resonator«.

The light amplifier and the resonator are not in perfect symbiosis, but thrive for different properties. As in acoustics, the resonator is there to provide purity of tone, eliminating any wavelength that does not fit as an exact integer fraction of the cavity length. Thus in a 0.5 m long laser cavity, there are $n_\lambda = 1\,000\,000$ wavelengths of 1 μm in a round trip (1 m). At a wavelength of 0.999 999 5 μm, the round trip corresponds to 1 000 000.5 wavelengths, meaning that, if the wave started with a crest at one mirror, it will come back with a trough and which will cancel the wave.

The most selective resonator will have mirrors of the highest reflectivity possible. In Figure 1.7b let us consider one end as a mirror with 100% reflectivity and the other mirror with a minor leakage (reflectivity 99.9%). A mirror designed for extraction of laser cavity power is called *output coupler*. If we were to turn off the amplifier and look at the evolution of stored energy in the resonator, we would see that after the first trip the light is reduced to 99.9% of its original value. In a second trip $(99.9 \times 99.9) = 99.8\%$ of the energy stays inside. How many trips does it take for the light to be reduced to half of its energy? After N_{RT} round trips, where $99.9^{N_{\mathrm{RT}}} = 0.5$ only half of the energy is stored in the resonator. With a help of a calculator in our example one finds this number to be 693 round trips. In other words, it takes 693 times the cavity length for the light to be reduced to half, and all the waves must perfectly overlap over this length. The effective length of the cavity is multiplied by 693. In the example of the 0.5 m cavity, even a wavelength off by $1/(693 \times 1\,000\,000)$ of a μm will be selected out, because it will interfere destructively with itself after N_{RT} round trips.

While a laser with a good resonator – called a »high *Q*« cavity,[5] provides a stringent wavelength selection, it does not produce large output power. Only a trickle (1.0%) of the power generated inside the resonator leaks out as useful output. In order to have large power out

5) The letter Q is used to designate a »quality factor« of a resonator – it is proportional to the number N_{RT} of round trips that survive in the cavity.

of a laser, one would need an end cavity mirror of large transmission, for instance 50%, requiring a gain of at least a factor of 2 per pass. Not only the wavelength selectivity is compromised by the large output coupling, but even the collimation of the beam. Indeed, since the lasing condition is achieved only after one round trip, any oblique ray that experiences two traversals of the gain volume will be part of the output beam.

It seems that lasers are not like a blast of energy as in a bomb, but rather a very conservative source of energy. Only a small percentage of the stored energy leaks out, similarly to the interest trickling out of a savings bank account.

It appears thus that, in order to create a laser of high output power, one has to compromise against the main qualities that characterize a laser beam: beam collimation and frequency selectivity. The solution to this dilemma is to inject the output of a low power, narrow band, and well collimated »seed« laser, into a high power »slave laser«. This is based on two types of radiation that we have already learnt.

We have seen that the stimulated photons are identical and undistinguishable from the stimulating photons. In general, a laser radiation is initiated by random spontaneous emission. The role of the seed photons is to overpower the spontaneous emission. As long as the seeded coherent radiation exceeds the spontaneous emission in the slave laser, the latter will clone the directionality, phase and frequency of the seeding radiation. This technique is also called »injection locking«.

1.4 Some Examples of Lasers

1.4.1 He-Ne Lasers

One of the first and most manufactured lasers has helium-neon as a gain medium. The first continuous He-Ne lasers were a technological challenge – in particular the ones built with a pair of flat mirrors for the cavity. Since the red He-Ne laser has a very low gain (of the order of 0.1% per pass), the radiation losses should be less than 0.1% in order for the laser to operate (or to reach the »lasing threshold«). The tolerances in flatness are incredibly difficult to meet ... and maintain. In addition, the gain of the red He-Ne laser emission line is very small;

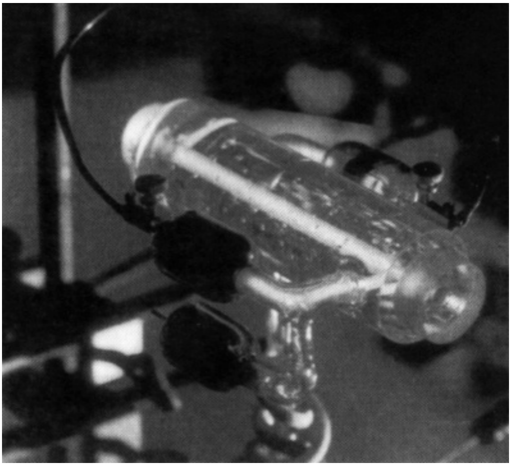


Figure 1.8 Short (12 cm) stable He-Ne laser with flat optical elements, directly bounded to the fused silica body (courtesy of Dr. Haisma, with permission from Koninklijke Philips Electronics N.V., Eindhoven from [1]). Please find a color version of this figure on the color plates.

it is not sufficient to compensate for the absorption of the best metallic mirrors. Good dielectric mirrors had to be developed for this laser. In the early 1960s, just after the first helium-neon laser was demonstrated, Philips Research Laboratories in Eindhoven (*Nat*uurkundige *Lab*oratorium) took up the challenge to make the perfect laser with a perfect flat resonator. The result was a thick wall cylinder of quartz, polished at both ends to better than 0.02 μm flatness, and a parallelism of better than 1 arcsec [2]. It thus required the development of polishing fused silica with extreme flatness and parallelism. In addition, coating techniques were developed to reach a reflectivity of 99.6%, a scattering of 0.1%, and a transmission of 0.3%. Mirrors and the polished quartz were of such perfect flatness and so polished that they formed a vacuum tight seal when put in contact – a technique still not very often used to this day because of the stringent finish requirements. There was no alignment possible and no room for error. The challenge was successfully met, leading to the laser sketched in Figure 1.8, probably the first and last He-Ne laser of its kind.

The perfectly cylindrically symmetric laser allowed putting some fundamental questions to test, such as the nature of the electric field produced by a laser. Is it always a linearly polarized beam and an electric field oscillating along a particular line perpendicular to the beam?

The answer was always yes, despite the fact that there was no preferential direction for the laser to oscillate.

1.4.2 Lasers for the Cavemen

Clearly, the He-Ne laser was not only a scientific innovation, but its construction required high vacuum technology, high quality dielectric coatings, and mirror polishing to a high degree of flatness. Other lasers, due to their high gain and/or long wavelength, did not pose such a technological challenge. For instance, the CO_2 infrared laser could have been built centuries ago, if only somebody had come across the recipe. Here the idea came far after the technology was available. Many »kitchen« versions of these lasers have been built and have probably cost their inventor a hefty electrical shock and a burned belly button. Lasers are typically more efficient at longer wavelength. Infrared lasers are easier to build and are more efficient than, for instance, ultraviolet lasers. One reason is that the mirror technology is simpler: the flatness only has to be comparable to the wavelength, which is considerably easier at 10 μm (10^{-5} m) than at 0.6 μm. Another reason is that inversion is easier to achieve, because there is less competition from spontaneous emission. Indeed, spontaneous emission rates are higher for more energetic transitions. As if the challenge of a higher slide in Figure 1.4 attracted children – a kind of reverse acrophobia?

1.4.2.1 The CO_2 Laser

An example of a continuous CO_2 laser is sketched in Figure 1.9. It is a tube of pyrex or quartz, typically 6–10 mm inner diameter, and 1–2 m long. Ideally, a second concentric tube is used for water cooling. An electric discharge requires a power supply of 15 to 25 kV, providing a current of 10–20 mA (current limiting resistances are essential, since the discharge has a negative resistance characteristic). A mixture of carbon dioxide (CO_2), nitrogen (N_2) and helium (He) is used, typically in the ratio 1 : 2 : 10, at a pressure of 10–30 torr.[6] Since CO_2 is the laser medium, one may wonder why the other gases are added. In fact, some gain exists in the discharge even with pure CO_2. However,

6) 1 torr is 1/760 of atmospheric pressure.

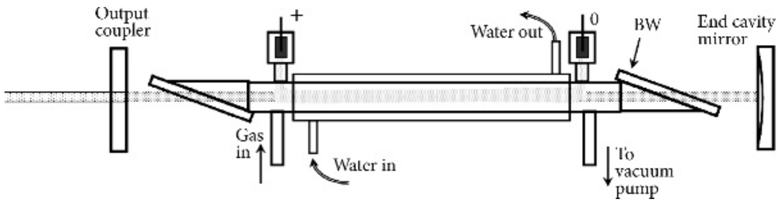


Figure 1.9 Typical low pressure CO_2 laser. The two windows (BW) can be made of germanium or salt (transparent to the wavelength of 10 µm emitted by the laser). They are at an angle such that there are no reflection losses for a beam polarized in the plane of the figure (called the Brewster angle). The electrodes can be small cylinders of nickel. The reflectivity of a flat plate of germanium (80%) is sufficient to achieve lasing action. The other reflector can be a flat metallic mirror (a concave mirror – typically 20 m radius of curvature – is generally used).Please find a color version of this figure on the color plates.

a marvel set of physical coincidences boost up the efficiency of that laser when the other gases are present.

The CO_2 laser is a *vibrational laser*. This means that the photon is emitted when the molecule switches from a higher energy mode of vibration to a different mode of lower energy. The CO_2 molecule is linear, with the carbon in the middle. Specifically, the upper state of that laser corresponds to the asymmetric motions of oxygen atoms with respect to the central carbon atom. To visualize this, one can picture an oxygen moving towards the carbon, while the oxygen at the other end is moving away from the carbon. The lower state is still an excited state, but corresponds to the »bending mode«, where the motion is that of a bird flapping its wings, with the carbon being the body, and the oxygen at the tips of the wings. When the electrons in the discharge collide with a CO_2 molecule, the latter gets excited somewhat preferentially in the asymmetric mode, the upper level of that laser. However, the electrons excite the vibration of a nitrogen molecule much more efficiently. Because the quantum vibrational energy of nitrogen is nearly equal to that of the CO_2 asymmetric stretching, that mode gets preferentially excited. Thus excited the CO_2 can emit a photon and in the process change to a bending mode or a bird flapping wing mode. How does the vibrating molecule come down to rest? Collisions with a helium atom are very effective in transferring the bending vibration energy of CO_2 into translation energy of the He, which is just heat. Helium is a very good heat conductor and will become cooled by collisions with the wall of the tube. Another approach for eliminating the heat is to replenish the medium quickly, in other words to increase the flow of gas.

The basic design of Figure 1.9 has evolved over the years to lead to multikilowatt lasers used for machining. The evolution towards higher power and efficiency has been quite straightforward.

One essential part of the efficient operation of a laser is the removal of the molecules in the lower state. For each photon created by stimulated emission in the beam, a CO_2 molecule is left in the bending excited state. When all the molecules have accumulated into that state, there can be no longer amplification. It is thus imperative to remove them. There are two methods: to remove them physically (high flow »gas dynamics« lasers) or to cool them more effectively. The latter approach involves increasing the (cooling) to surface ratio, which is the approach used in planar waveguide lasers, where the cooling surfaces are metallic plates also serving as electrodes. The other approach is to increase the energy of the laser, for a given size, is to increase the number of molecules participating in the gain, hence to increase the gas density. This has led to the atmospheric pressure pulsed CO_2 laser, where a strong electrical discharge is applied transversally across a tube containing a mixture of CO_2, He and N_2. These lasers are generally called TEA lasers (»transverse electrical discharge«). This type of laser produces nanosecond pulses of several joules of energy. High power – several kilowatts – CO_2 lasers are manufactured for industrial applications, welding, cutting, surface treatment, and so on …

1.4.2.2 The Nitrogen Laser

The nitrogen laser is another example of a laser that was simple enough to appear in peoples's kitchen, following a recipe published in the Scientific American in 1974 [3]. Surprisingly, in view of the previous discussion in the introductory paragraph of Section 1.4.2 about the infrared lasers being easier to produce, it is a UV laser operating at 337 nm. As stated previously, there is more competition from spontaneous emission at short wavelength. Indeed, the energy level responsible for lasing in the nitrogen laser empties itself in a few nanoseconds. It is, however, relatively easy to produce a strong electrical discharge transverse to the lasing direction in a shorter time. The energy storage is electrical: a thin sheet of mylar or other dielectric sheet able to withstand very high voltage without breaking down is sandwiched between two copper plates to constitute the energy stor-

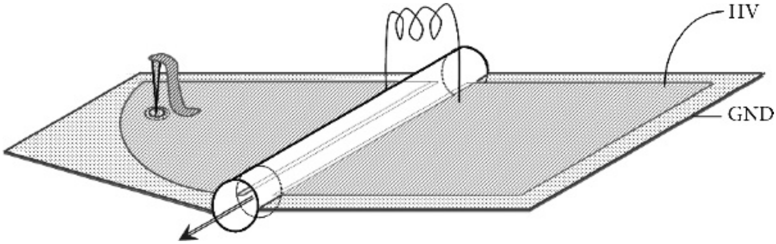


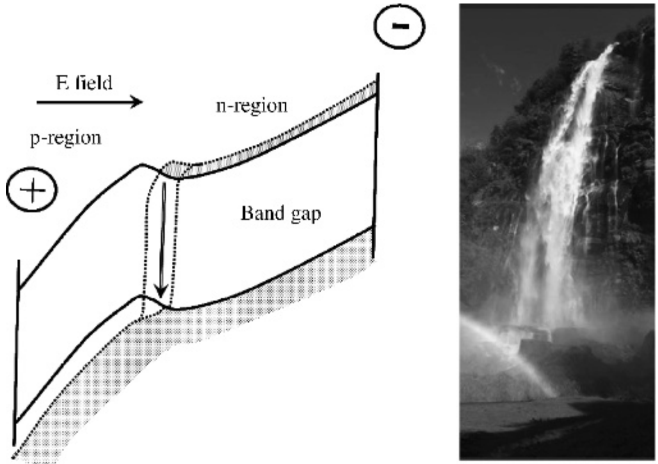
Figure 1.10 Sketch of a typical nitrogen laser. Copper foil is taped on both faces of a thin dielectric (circuit board, Kapton, or Mylar), creating a capacitive transmission line, which is charged via a resistance by a high voltage power supply (right side). The upper copper plates terminate as line electrodes in a discharge tube. The left capacitor/transmission line is connected to the right one via inductance of a resistance. On the far left: the spark gap that acts as a switch short-circuiting the capacitor (it can be a needle connected to the upper plate, pointing towards the ground plate, or a car spark plug). The short circuit propagates towards the right, brings the left electrode to ground potential, inducing a discharge across the tube, lasting until the complete capacitance is discharged (about 1 ns). The tube is closed by two windows, or one window and a mirror. Pure nitrogen at 25–50 torr pressure is pumped through the tube (air can also be used, but the performances are reduced). Please find a color version of this figure on the color plates.

age capacitor, as sketched in Figure 1.10. One of the plates is connected to a high voltage (in the order of 10–20 kV) and will also constitute the high voltage electrode of the laser. The other electrode is connected to a similar sandwich, terminated by a switch called the spark gap to the right of the figure. Some lasers have been built using just a car spark plug as a switch. Initially, the two capacitors are charged to the same voltage (they are connected by a resistance). As the spark gap (or spark plug) is fired, the short circuit established at that end propagates towards the left at a speed close to the speed of light, and a step function current of the order of a nanosecond in duration traverses the discharge tube. A little venturi pump can be used to flow the nitrogen at a low pressure (25–50 torr) through the tube. No mirror is needed: because of the laser spontaneous emission and very large gain, the geometry of the tube is sufficient to have light amplification by stimulated emission radiation without any mirror. Adding a mirror at one end increases the output energy, because the length of the amplifying medium is doubled. This simple technology grew in a small laboratory in Göttingen and led to one of the largest manufacturers of nitrogen lasers and excimer lasers. These lasers were extremely popular for three decades as a pump for visible dye lasers.

1.4.3 Semiconductor Lasers

Only two years after the birth of the first laser in 1960 (the ruby laser) IBM, GE, and the Lincoln labs of MIT simultaneously announced the realization of the semiconductor *p*-*n* junction laser. These first *diode lasers* (also called »injection lasers«) were made of gallium arsenide (GaAs). A semiconductor is characterized by having two energy bands instead of discrete energy levels as is the case for atoms and molecules. Photons are emitted by the transition of an electron from the higher energy band (the *conduction band*) to the lower energy band (the *valence band*). The two bands are separated by an energy difference called *band gap*. The laser consists of a small single crystal, with two halves of different physical properties, achieved by doping the crystal on either sides with different elements. The »*n*« region is doped with impurities labeled »donors« (for instance Se), because they tend to produce an excess of electrons giving this part of the crystal metallic like properties. The »*p*« region is doped with impurities labeled »acceptors« (for instance Zn), which tend to trap electrons, leaving no electron in the conduction band. The electrical properties of such a combination are asymmetric: the current will flow easily from the *p* side to the *n* side, while the junction will show a very high resistance if a positive voltage is applied to the *n* side and a negative one to the *p* side. Hence the label »diode« for these devices.

As voltage is applied to the diode, positive on the *p* side, negative on the *n* side, the electrons will cascade down from the high energy conduction band into the valence band at the location of the junction, where a population inversion (between the conduction band and the valence band) is created (Figure 1.11), provided that the current is sufficiently large. It is as if the electrons flow down a river, encountering a sudden drop at a waterfall. The radiation is thus emitted along a line that marks the junction between the *p* doped side and the *n* doped side of the semiconductor. While having a lot in common with standard lasers, a striking peculiarity of semiconductor lasers is their size. The *p*-*n* junction lasers are minuscule, with a size of about 0.1 mm. The junction thickness itself is of the order of 1 μm. The shorter the cavity, the less it will select a particular frequency, and the lesser its collimating properties. The gain of the junction laser can be very high. It is therefore not necessary to use a reflective coating on the facets of the



(a) (b)

Figure 1.11 Energy band structure of a semiconductor laser. The occupation of electrons shown by the gray area extends in the conduction band in the *n* region and does not reach the top of the valence band in the *p* region. As a field is applied to the semiconductor, the electric potential adds to the electron energies, resulting in slanted bands (a). The current of electrons flows down the »river« of the conduction band to cascade down to the valence band at the junction just as a waterfall (b). Please find a color version of this figure on the color plates.

cube. Because the index of the refraction of GaAs is very high (3.34), the uncoated facets have a reflectivity of 30%, which is sufficient for lasing. If better performances are needed (single frequency or wavelength, better directionality), one may put an antireflection coating on the facets and place the semiconductor gain medium in a traditional cavity with mirrors.

Another limitation associated with the minuscule size of the *p-n* junction laser is the optical power that can be extracted. With a gain volume that is in the μm^3 range, the intensity can reach the damage threshold with only a few tens of milliwatt of optical power. The other limitation to the output power is saturation. It will be shown in Section 1.5.2 on »saturation« that the media that have a stronger gain and larger spontaneous emission, will see a drop in gain at lower intensities. Therefore, one would expect that semiconductor lasers do not have a future in high power applications. There is, however, a solution to the power limitations brought on by the microscopic emitting cross-section (a narrow line at the junction) of the semiconductor lasers: a total change in geometry leading to a large emitting area. These are the vertical-cavity surface-emitting lasers (VCSEL) and the vertical-external-cavity surface-emitting lasers (VECSEL). Instead of emitting in the plane of a junction, these lasers emit along the normal to the plane. Instead of being pumped by an electrical current, they are pumped optically (although electrically pumped devices are under development). With a large emitting area, one can expect to extract more power out of the semiconductor laser for the intensity corresponds to complete inversion. Most often the limiting factor is the optical pump power available.

1.4.4 Fiber Lasers

1.4.4.1 Introduction

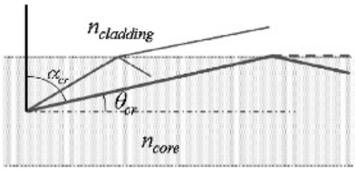
The first few years following the discovery of the ruby laser were full of surprises. After 50 years, the ruby laser has nearly fallen into oblivion. As early as 1961, another laser was discovered, which was left largely ignored for the next three decades: the fiber laser [4, 5]. The gain medium was a fiber of barium crown glass, doped with the neodymium ion Nd^{3+}. This laser did not promise a bright future. Citing a review of the laser technology near the end of the first decade of laser development: »A number of fiber optics lasers have been constructed using Nd^{3+} in glass (wavelength 1.06 μm) as active medium in glass, but these devices have thus far not played a conspicuous role in the development of laser technology« [6]. Fiber lasers have started their comeback since the development of ultralow loss fibers for communication in the 1980s. Fibers are slowly extending their tentacles over areas where gas lasers or solid state lasers have been well entrenched. As a commercial ultrashort pulse laser, the fiber laser has even taken a big chunk of the market from the Ti:sapphire laser. It is now considered to be one of the most promising weapons for the Air Force, having replaced the development of chemical lasers.

Two main breakthroughs in fiber technology (i) the manufacturing of near lossless fibers and (ii) speciality fibers dubbed »Holey fibers«, »photonic crystal fibers«, or »photonic band gap fibers«, are contributing to progress in fiber lasers. The first is a material technological advance, the manufacturing of ultrapure silica, which has enabled the production of silica fibers with such a low loss that, after propagation through 1 km of fiber, 1 W of optical power would only be attenuated to 0.96 W.[7] The second advancement results from progress in manufacturing, which has made it possible to produce fibers where the inner core consists of a complex pattern of hollow and filled structures.

1.4.4.2 Optical Fibers

An optical fiber has very much the appearance and size of a hair. In fact, after a long haired graduate student had completed a complex fiber laser system, it became very difficult to disentangle hair from fibers. Optical fiber in fact exists in nature: the hair of polar bears not only provides insulation, but conducts the sunlight to the skin. A polar bear might even get a suntan under his white coat! The principle of light conduction in fibers is total internal reflection, as sketched in Figure 1.12a. If a light ray is incident on an interface between a medium of a high index of refraction n_{core} and a lower index n_{cladding} there will be a critical angle of incidence α_{cr} beyond which no light is transmitted, and all light is reflected by the interface (hence the name »total internal reflection«). The critical angle is given by $\sin \alpha_{\text{cr}} = n_{\text{cladding}}/n_{\text{core}}$. An optical fiber consists essentially of a filled spaghetti where the core would be denser than the outer part. The material of the core has a higher index of refraction n_{core} than that of the surrounding called cladding (n_{cladding}). Rays that are incident on the interface at an angle larger than α_{cr}, – or that make an angle with the axis of the fiber smaller than θ – are trapped in the core. Of course, the light energy will propagate at a different velocity for all the rays that zigzag through the fiber core at a different angle. This effect is called *modal dispersion* (the larger the angle, the smaller the light energy traveling along the fiber). We have seen in Section 1.2.3 that a beam of diameter D spreads (diffracts) with a half angle $\theta = \lambda/D$. If the diameter of the core is large, a large number of rays with different angles can be trapped by total internal reflection. If the diameter of the fiber is reduced, the value of λ/D increases until that angle reaches the critical value. There is at that point only one ray that can be trapped: the fiber is said to be *single mode* for that wavelength. Fibers with a larger core diameter are called *multimode fibers*. A single mode fiber at a given wavelength can be a multimode fiber at a short-

7) The fiber manufacturers express the fiber loss in dB/km, where the decibel is defined as 10× the logarithm in base 10 of the ratio of the transmitted power to the input power. The Corning fiber SMF 28 ULL has a loss of 0.17 dB/km.



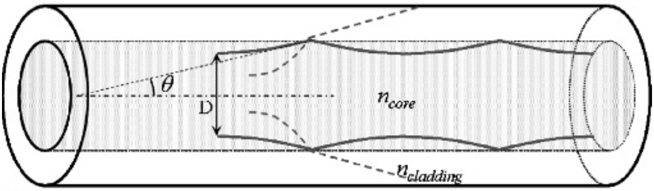


Figure 1.12 (a) Beam incident from the bottom on the interface between two glasses of different index of refraction $n_{\text{core}} > n_{\text{cladding}}$. Rays incident at an angle larger than or equal to α_{cr} are totally reflected, without any loss in the medium of index n_{cladding}. (b) Total internal reflection takes place between the core of an optical fiber and the cladding, provided that the rays make an angle with the axis of the fiber smaller than the critical value $\theta = 90° - \alpha_{\text{cr}}$. The smallest beam diameter D that can be trapped is the one for which the diffraction angle λ/D (see Section 1.2.3) corresponds to the critical θ. If the core diameter is equal to D, the fiber is said to be single mode. Please find a color version of this figure on the color plates.

er wavelength. The single mode fiber is to the fiber optics what the fundamental ideal Gaussian beam, least diffracting, is to free propagation.

The high quality of single mode fiber optics does not come cheap or easy. A fundamental problem is that the glass homogeneity and accuracy with which the index of refraction can be controlled implies a small core size – of the order of 10 μm. Such a small core size is incompatible with high optical power, since the large intensity (power/core area) will result at best in nonlinear effects (intensity dependent index of refraction, generation of unwanted wavelengths), at worst in the destruction of the fiber. Other related problems are:

- Rapid saturation of a fiber laser amplifier.
- Difficulty to effectively couple the fiber to the outside world. An extreme positioning accuracy is required to effectively couple light from a Gaussian beam into a fiber and vice versa.

- Difficulty in coupling fibers to fiber.
- Difficulty in controlling the polarization along a fiber.

The latter point is due to the fact that any motion, bending, temperature change, and acoustic disturbance of a single mode fiber affects the polarization of the beam that is propagated through the fiber. The development of fiber lasers and fiber laser amplifiers is intimately linked to progress in single mode fiber optics. The polarization instability and uncontrollability has been solved through the development of polarization preserving fibers, with a slightly different index of refraction along two orthogonal axes. Light linearly polarized along one of these two axes will conserve its polarization. Remarkably sophisticated instruments have been developed to accurately splice two fibers together (single mode or polarization preserving) with minimum loss.

A significant breakthrough came when researchers at the University of Bath created a new family of single mode fibers, by using as a preform[8] an array of capillary and rods [7]. This new fiber was fabricated by putting together glass tubes of hexagonal cross-section, forming a regular pattern, with a full rod in the center. By pulling and stretching this structure, the dimensions of the cross-section were reduced by a factor of 10 000, until the central rod was reduced to a diameter of the order of 1 μm, the pitch (spacing between holes) reduced to 2 μm, and the holes had a diameter of between 0.2 and 1.2 μm. A considerable number of variations have been made on these structure, leading to an equally large number of exotic effects. Because of the better guidance of the light by the pattern of holes surrounding the core, the »single mode« guidance is observed over a wavelength range covering the visible to the near infrared [7]. The fiber can be engineered to have the light pulse traveling at the same velocity over a broad range of wavelengths [8], which is an important property for the transmission of ultrashort signals (cf. Section 1.7.3). Because of the high intensity at the core rod, a rainbow of new wavelengths can be generated [8]. Other exotic effects are being investigated by inserting vapors or gases in the capillaries surrounding the core [9, 10].

8) The fiber are typically drawn from a preform consisting in a cylinder of glass. In this particular case, the solid cylinder is replaced by a bundle of tubes, of which the inner diameter is progressively reduced in successive stretching steps.

1.4.4.3 Fiber Amplifiers

The fiber laser amplifier is the simplest device of all. The core of the fiber is generally doped with some rare earth (erbium for 1.55 μm lasers, neodymium and/or ytterbium for 1.06 μm fibers, to cite only a few). In very early designs the fiber was wrapped around a flashtube. This practice was totally abandoned with the availability of efficient semiconductor lasers. The pump light from a semiconductor laser is coupled to a fiber which is coupled directly into the doped fiber with a wavelength division multiplexer (WDM), the fiber equivalent of the railroad switch, or beam splitter mirror. The limit to the amplified power is set by the cross-section of the core: the laser beam cannot be amplified beyond a certain »saturation intensity« (power divided by cross-section), which is discussed in Section 1.5.2. The minuscule cross-section of the fiber core sets an upper limit to the power that can be amplified. There has been considerable progress recently in the manufacture of large core fibers, with special cladding guiding the pump light, such that the amplified power has approached the kW level. It is anticipated that a very high power device can be realized by adding the *field* (i.e., adding the laser beams as waves in phase) of several of these devices.

The semiconductor laser pump pulse loses its energy as it propagates down the doped fiber to create a population inversion. For a given pump laser, how can one determine the useful (optimal) length of fiber to be used, given that the IR beams (pump and laser) are in-

Figure 1.13 Erbium doped fiber glowing in the dark, upon excitation by the pump radiation at 980 nm. Please find a color version of this figure on the color plates.

visible and hidden inside the core of the fiber? Figure 1.13 offers an answer. It shows and erbium doped fiber pumped by a 980 nm semiconductor laser. The pump photon excites the erbium to the upper state of the laser transition, as it should to create a population inversion for 1.55 μm. However, other pump photons excite the erbium from that level to a higher energy one, from which green fluorescence is emitted. Hence the green glow can be observed in the dark, emanating from an erbium doped fiber laser or amplifier.

Since there is no limit to the length of a laser amplifier (provided one can continuously add pumped sections), one might wonder how a laser pulse will evolve after a considerable distance. Will it evolve towards a steady state entity, like a light bullet speeding through the core, and without any shape distortion? We will see that this can indeed be the case, in the study of short pulse propagation in Section 1.7.

1.4.4.4 Fiber Lasers

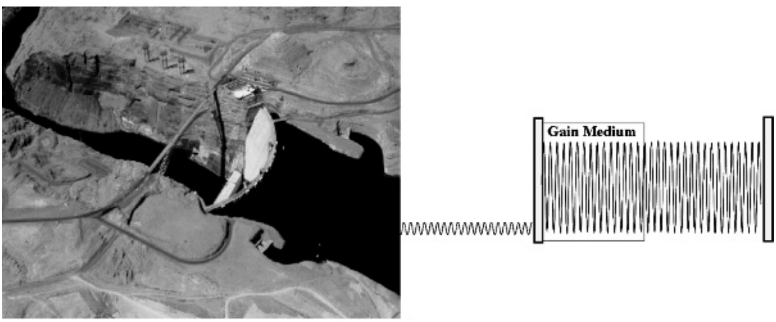
As we have seen in Section 1.3, the laser is simply an optical amplifier sandwiched between two mirrors. The same high reflectivity, low losses, and high power mirrors that have been designed for other lasers do not apply to the fiber: the cross-section is too small to make precise multiple layer deposition on the fiber end. A simple technique employed in the past was to dip the fiber end in mercury, but the use of mercury has lost its popularity. Another technique still in use is metallic coating on the fiber end, but the intensity at the fiber core may be sufficient to vaporize the thin film. The fiber equivalent of the dielectric mirror is the fiber Bragg grating. A dielectric mirror consists of a stack of transparent materials of different indices of refraction. The material and layer thickness are chosen in such a way that the reflections produced at each interface add in phase, resulting in a near prefect mirror if a sufficient number of layers is used. The fiber Bragg grating consists of periodic changes of the index of refraction over a considerably longer distance than the thickness of a dielectric mirror. The basic principle, however, is similar: the changes of index are engineered in such a way that the backscattered radiation adds in phase. A solution that circumvents the need for mirrors is to have the laser literally as a dog biting its tail: splice the input end to the output end to form a ring. A short pulse that is let to evolve in this amplifier in a ring configuration, looks in vain for the end; the ring is in this case equivalent to the infinitely long amplifier of the previous section. In certain conditions discussed in Section 1.7, the short pulse will evolve as a light bullet circulating endlessly in the ring. Here again the equivalent of railroad switch, or WDM, will come handy in extracting the pulse at every round trip.

One of the most successful commercial fiber lasers is the femtosecond source of IMRA America, which produces a train of pulses of the order of 100 fs in duration at a repetition rate of 50 MHz. This can be seen as a train of light bullets of only 30 μm in length ($c \times 100$ fs) spaced by 6 m.

1.5 Pulsed Lasers

1.5.1 The Laser as Energy Storage

A low gain laser requires a very high quality cavity, with no more losses than the gain. The laser oscillation starts when the gain per path exceeds the loss. As the light rattles back and forth between the two mirrors of the laser cavity, the optical power increases, just as the level of water, and the pressure increases when a river is dammed (Figure 1.14). The power input to the laser eventually comes in equilibrium with the power released through the end mirror, just as the river flow into the lake eventually balances the release at the bottom of



(a) (b)

Figure 1.14 (a) A dam: most of the energy is stored in the lake in front of the dam. Only a small fraction of the stored water flows through the release of the dam. (b) Optical power is stored between the two mirrors of a laser cavity. The laser beam issued from this device is actually the trickle transmitted by the output mirror.

the dam, which increases with the increase in pressure or energy contained by the dam. In the case of the laser, the gain per pass – which initially exceeded the loss – decreases as the power inside the cavity increases, until it exactly balances the losses at the output mirror. The decrease in gain with power is called gain saturation.

1.5.2 Saturation

Saturation is a nearly universal phenomenon. Unlike the banker, for whom there is no limit to the number of coins that he can absorb, matter has a limited greed. There is only so much light that our eye can take. Beyond a given intensity, we do not see any difference in brightness, whether the power is multiplied by 2 or 10. The same applies to the temperature of a radiator, which results from a balance between the power applied to it and the cooling by radiation and conduction to the surroundings. Doubling the power applied to it will not double the temperature, because the cooling rate to the surroundings will increase.

The process described in the Section 1.3.2 assumes that the number of photons added is small compared to the total population in the upper state. Only a few of the lazy girls on top of the slides (Figure 1.4) are »stimulated« to go down. However, the total gain, which corresponds to the release of energy in emitted photons, is proportional to the number of excited atoms, which is going to decrease as a result of the transitions down. The girl that has gone down the slide cannot stimulate another transition before she has climbed the ladder again.

In the case of a constant rate of excitation into the upper state, the stimulated transitions will compete with the pump rate that creates atoms in the upper state and reduce the upper state population, and hence the gain. The *saturation rate* is the number of photons per second that results in a decrease of population difference by a factor of 2. Why population difference? This might not be an easy concept to grasp. The girls down on the ground have a potential to be up the ladder and those who are up will be pulled down. A good inversion made by a pump laser brings as many electrons as possible to the upper state. If we start with 100 girls up the ladder and none on the ground, the population difference is 100. Now only if 25 of these girls go down on the slide is the population difference reduced to 50, which is half of initial difference. In this example 25% of girls going down corresponds to the saturation rate of the transition. At the same time a gain (maybe a trampoline in our example) pumps the electrons to the upper state.

The same applies to an absorber, where the atoms are initially in the lower state. There is only so much light that an absorber can take. Once an atom has been excited by a photon, it cannot absorb another photon before it has relaxed back to its initial absorbing state. The more power in a laser beam, that is, the more photons/second are sent to an absorbing medium, the lesser the number of atoms/molecules that are available to absorb a photon, and the medium becomes transparent to the light. It can also be seen as a gate that gives way under the pressure of a crowd (in this case of photons). This property is used quite often in laser technology as a fast gate that gives way to an excess of optical energy.

1.5.3 Releasing Stored Energy

1.5.3.1 Q-Switching

How do you release the maximum energy that can be stored behind a dam in the shortest amount of time? If the release van of the dam is totally open, the water level may never rise. In the laser, if the output coupling mirror has a too large transmission, the pumping radiation may never get the population inversion to the »threshold level« that makes the laser oscillate. In other words, there is no dam anymore, just a leveled river.

To obtain a maximum flow in a short time, the release valve should be closed, to let the water level rise to the maximum, then suddenly opened to let the reservoir empty in a giant surge of water. In the Q-switching mechanism the van is open but the dam is not eliminated. In the laser, the gain is first accumulated to the maximum value that the pump source (a flashlight in the case of a ruby laser, a discharge current in the case of a gas laser) can provide. In the radiation picture of sliding girls, they are all up the ladder to create the maximum population inversion but they cannot go down. Then the resonator is suddenly formed, allowing the laser oscillation to start at a very fast pace, because the gain has reached a high value.

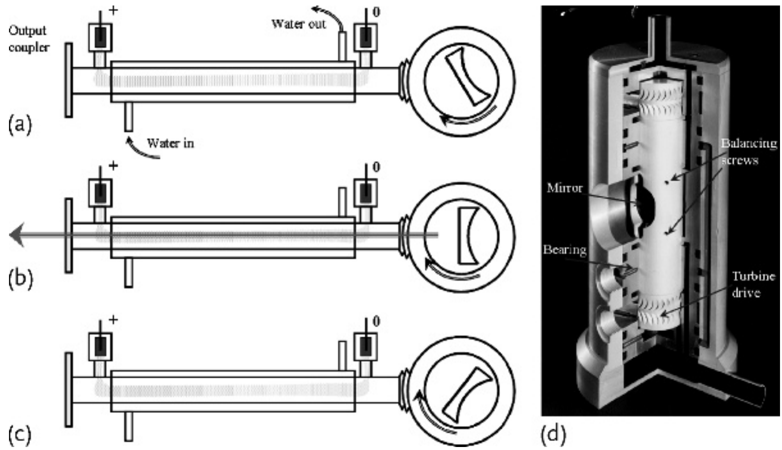


Figure 1.15 (a) A CO_2 laser tube is closed at one end by a partially reflecting mirror, an uncoated germanium plate, or a coated ZnSe plate. The other end is sealed with an enclosure containing the rotating mirror. In position (a), with the rotating mirror at a nonzero incidence with respect to the tube axis, the gain develops inside the discharge tube. As the rotating mirror arrives in position (b), the output mirror and the rotating mirror are parallel, forming a resonator in which a laser pulse forms, and depletes the gain. As the mirror pursues its rotation (c), there is no longer a resonator, and the gain can recover until the next position (b). (d) Exploded view of an air turbine used for Q-switching. The rotation speed could reach 60 000 revolutions/min, limited by the centrifugal expansion of the shaft (courtesy of Koninklijke Philips Electronics N.V., Eindhoven, 1970). Please find a color version of this figure on the color plates.

1.5.3.2 Mechanical Q-Switching

In the early years of the laser, the mirror was put physically in and out of place by placing it in the shaft of a rotating turbine. An example of such a turbine [11], that was used for Q-switching a CO_2 laser at Philips Research Laboratories in Eindhoven, the Netherlands, is shown in Figure 1.15. The alignment of the rotating mirror is rather easy: there is only one degree of freedom (around an horizontal axis), since the mirror scans all angles about the vertical axis. Such spinning mirrors are particularly convenient for Q-switching lasers with continuous gain, such as the continuous discharge CO_2 laser described in Section 1.4.2. During most of the revolution of the shaft, the laser cavity is not aligned, and the gain accumulates. For a few seconds of arc during the revolution, the two cavity mirrors come into parallelism, and the laser wave builds up from noise, extracting all the energy that was stored in the gain medium, in a pulse that lasts between 10 and

100 ns. As the mirror continues its spin into the next cycle, the gain medium recovers, in preparation for the next pulse after another revolution. The disadvantages of the rotating mirror are (i) the stringent requirement of near perfect shaft balancing, and (ii) the deafening noise (which, in the case of the CO_2 laser, was eliminated by operating in vacuum). Therefore, while this method of creating »giant pulses« had some popularity in the late 1960s, it was abandoned in favor of either electronic Q-switching or passive Q-switching.

1.5.3.3 Electronic Q-Switching

Instead of moving the mirror successively out- (to prevent laser emission, in order for the gain to evolve to its maximum value) and in-position (to suddenly dump all the stored energy in a giant laser pulse), it is easier to place an ultrafast obturator in front of the mirror, which opens at the time that the gain has reached its peak value. This technique was conceived and demonstrated in the early days of the ruby laser [12]. The principle involves the manipulation of the laser polarization discussed in Section 1.2.2.

A linear polarization can be rotated, or changed to elliptical or circular, by propagating the light in a medium which has a different propagation velocity along two directions at $\pm 45°$ from the polarization direction (»birefringent medium«, or »waveplate«). After some propagation distance, the components of the electric field oscillation along these two directions will no longer be in phase, resulting in elliptical polarization. Or they may have opposite phase, resulting in linear polarization again, but rotated by 90°. There also exist components that selectively deflect a particular polarization direction (polarizing beam splitters) or transmit only one direction of polarization (polarizers).

The birefringence property can be created by application of an electric field, since the material structure and electronic response could vary from one direction to the other. The key to optical switching is the *Kerr effect,* which is a change of the wave velocity (or index of refraction[9]) of a material under application of an electric field. The original implementation involved a liquid. Today's Q-switch elements are nonlinear crystals called *Pockels cells*. These elements are based on the

9) The index of refraction is the fraction of the speed of light in vacuum that a light wave can take inside the medium.

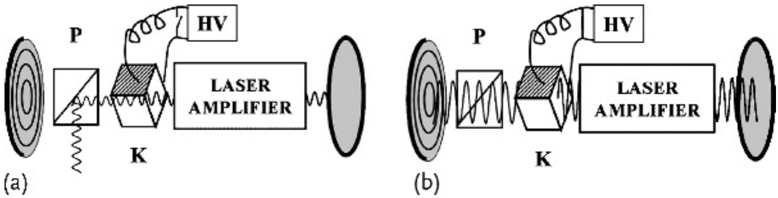


Figure 1.16 (a) The gain of the laser is activated (a flashlamp is fired, in the case of the ruby or Nd:YAG laser). A high voltage (HV) is applied to the Kerr cell *K*, such that the polarization of the beam issued from the amplifier is rotated 90° and rejected by the polarizing beam splitter *P*. The combination of Kerr cell and polarizer essentially eliminates the end mirror from the cavity. (b) After a time such that the gain has reached its maximum value, the voltage on the Kerr cell is shorted; the end mirror is part of the cavity and the laser oscillation can develop with full gain.

electro-optic, by which the index of refraction of a beam is changed upon application of an electric field. This change is different along different axes of the crystal. It results in a linear polarization being distorted in elliptical or circular polarization. Under application of a sufficient electric field, the polarization of the beam can be rotated by 90°. A »switch« consisting of a Kerr cell (or a Pockels cell) combined with a polarizing beam splitter is inserted in the cavity, as in Figure 1.16. A voltage is applied to the Pockels cell, rotating the polarization by 90°, such as to eject any light that would be generated in the cavity as the laser medium (amplifier) is being activated (Figure 1.16b). After a time such that the maximum population inversion – that is, the maximum gain of the laser element – has been reached, the high voltage (HV) on the Kerr cell is shorted suddenly, the laser cavity is closed, and a »giant pulse« develops (Figure 1.16b).

1.5.3.4 Passive Q-Switching

The advantage of electronic Q-switching is that the timing of the laser emission is well controlled, which is important when several laser systems have to be synchronized. The disadvantage is that the pulse to pulse reproducibility is not ideal. A preferred technique when synchronization is not an issue is to let the laser itself decide the opportune moment to open its cavity. This is done by replacing the electronic switch of Figure 1.16 by a »saturable absorber«. Saturation is a ubiquitous phenomenon. There is only so much light that an absorber can take. Once an atom has been excited by a photon, it cannot absorb another photon before it has relaxed back to its initial absorbing state.

As shown in Section 1.5.2, the more power in a laser beam, that is, the more photons/second are sent to an absorbing medium, the smaller the number of atoms/molecules that are available to absorb a photon, and the medium becomes transparent to the light.

One can think of a saturable absorber as a closed gate of a football stadium. After the game the crowd presses on to exit. The pressure builds up at the gate, as more and more people squeeze against it, as the light level increases inside a cavity trying to make the absorber transparent. Finally, the gate gives way, the saturable absorber opens, and scores of people are ejected through the opening.

The saturable absorber can be a dye, a crystal, a color glass (RG-8 filters of Schott glass for example in the case of ruby lasers) which is totally opaque to the laser light. Inserted in the cavity, it prevents lasing, since the amplifier does not »see« the end mirror. When the amplifier is fired, there is so much radiation from the spontaneous emission that, after a certain dose, the absorber »cannot take it anymore« and bleaches (becomes transparent). This simple technique provides the most reliable operation with the shortest pulses, but with a timing that cannot be controlled.

One can compare the »saturable absorber« and »electronic switching« to taking money from a bank account. In the saturable absorber case, the money is taken when it reaches a certain value, in the electronic method it is withdrawn at equal time intervals. As a result, one has certain values but may not be necessary timed equally and the other is timed but may not have the same value. In cavity dumping the whole bank account is emptied and then closed for the new build up.

1.5.3.5 Cavity Dumping

In our analogies of the dam, the previous Q-switch techniques amounted to close the dam, let the water level rise to the maximum, before opening the valves to maximum. Once the Kerr or Pockels cell is opened, the laser oscillation develops as fast as the »net gain« (gain minus transmission through the output mirror) allows. If the output mirror has a very high reflectivity, the optical pulse will develop fast, but the power will remain trapped in the cavity. If the output mirror has high transmission, more output will be released, but the oscillation will develop more slowly. The solution to maximize the extraction of energy from the laser and from the dam is to let the water level rise to its highest level and blow up the dam, as in Figure 1.17.

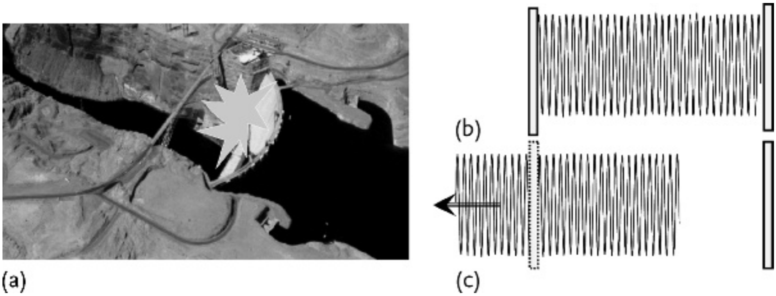


Figure 1.17 (a) A dam is filled slowly to capacity, ideally by closing all water release. The energy stored by the dam can be released in a minimum amount of time by blowing up the dam. (b) Similarly, in a laser cavity with totally reflecting mirrors, the optical power increases swiftly with time. (c) To release the energy stored in that cavity in the shortest amount of time, one of the cavity mirrors is »removed« in a time short compared with the time it takes light to make a complete round trip in the cavity. Please find a color version of this figure on the color plates.

This technique is called cavity dumping because all the energy stored between the mirrors is extracted in a pulse that will have exactly the length of the cavity. Instead of »blowing up« the output mirror, it is eliminated electronically. In Figure 1.16, the output mirror is chosen to have maximum reflectivity. As soon as the power in the cavity has reached its maximum value, a pulse is applied to the electronic switch to release all the energy via the polarizing beam splitter. The switching time has to be short compared to the cavity round trip time of light, hence it must typically have a rise time shorter than 1 ns.

1.6 Properties of a Laser Beam

1.6.1 Directionality

A simple and useful characteristic of a laser beam is its directionality: it propagates with minimal diffraction. Most light sources will expand through space, which is generally a desired feature. House light is made to illuminate a whole room without sharp exposures. As we move towards virtual words and replace material with data and images, our interest in using light as an appliance increases: we can replace a knife or a ruler with a laser beam. In Section 1.7.5.3 and in Chapter 7 we will learn about a new medium created by laser

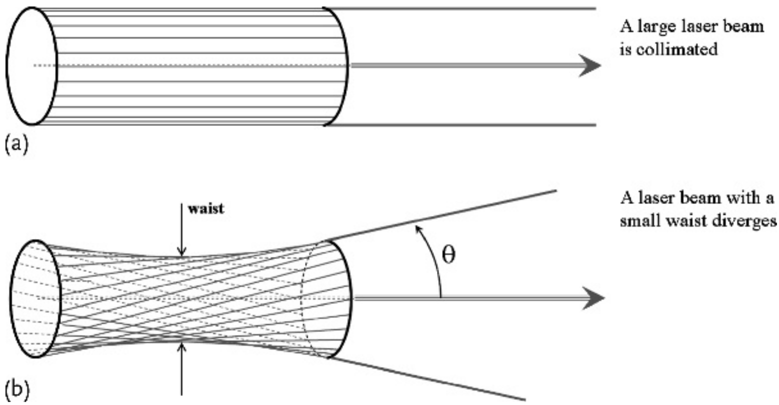


Figure 1.18 (a) A well collimated output of a laser. Rigorously speaking, such a perfect collimation can only take place for beams of infinitely large diameter. In general, the diameter of a laser beam will increase from the smallest value (the waist) as in (b). The diameter of the waist defines the angle θ of the divergence. The situation of (a) applies to a laser with flat mirrors, where the light rays are the strings spanned between the two mirrors. Curved mirrors lead to the situation of (b): the »rings« are twisted with respect to each other. The envelope of the rays (that are perpendicular to the surface of the mirror) draw an hyperboloid which can often be seen with a laser beam in a dusty environment. If the curvature of the mirrors is equal, the minimum diameter »D« – the waist – is at the midpoint of the cavity. Please find a color version of this figure on the color plates.

light that has a unique properties and even less divergence than laser beams.

When we compare a flash lamp with a laser pointer, we notice that the laser seems to have a well collimated output (Figure 1.18a), while the output of the flash lamp expands quickly in the space. Why can a laser beam can be collimated better than the light of a flashlamp?

The resonator is the main body of the laser (Section 1.3.3), and it defines the size of the beam. In that body, laser engineers use various components to adjust the wavelength, amplitude, or in the case of pulsed lasers, the length and shape of the pulse. One can look at the »resonator« as the cooking pot of the laser, where you can spice the beam as desired. Often the population inversion required for laser action is confined to a small volume, which imposes to design the resonator with the smallest beam size (waist) at the location of the gain medium. From the waist, the beam diverges as in Figure 1.18b, with a half angle θ:

$$\theta = 0.6\frac{\lambda}{D}\,, \tag{1.2}$$

where »λ« is the laser wavelength and »D« the minimum beam size or *beam waist*. Mirrors or lenses are generally used to recollimate the diverging beam. Another way to quantify the spreading of a laser beam with distance is the *Rayleigh length* ($\approx 2.3\,D^2/\lambda$), which is the length over which the beam diameter has increased from D to $1.4\times D$. Clearly, the smaller the wavelength, the smaller the divergence angle.

The simple Equation 1.2 tells us basically what can and cannot be achieved with a laser beam. Given the diameter D of a laser beam, it tells us how large the spot will be at a large distance. For instance, in the early years of the laser (1962), a pulsed Ruby laser was sent to the moon. The beam was expanded to fill the 48 inch diameter of the telescope at MIT Lincoln Observatory. Only a few photons were able to return back to earth, but that was sufficient to measure the time travel of the pulse and measuring the distance from the earth to the moon. With $\lambda = 694.3$ nm and $D \approx 1$ m, $\theta \approx 0.4$ µrad, which at the distance earth–moon of 300 000 km meant a 100 m radius circle. If one had replaced the ruby laser with a pulsed flash lamp, the size of the beam on moon would have been much bigger than the moon and there would be no chance for a single photon to make it to the moon.

The Equation 1.2 tells us also how small a waist can be achieved by focusing a large diameter beam. For instance, the military dream of focusing an infrared laser ($\lambda \approx 3$ µm) at 10 km distance to a spot size of only 1 cm imposes that the radius of the focusing lens or mirror be 1.8 m! The focal size is directly related to the intensity of the light that can be concentrated, as will be seen in the next section.

1.6.2 Intensity

Focusing light is akin to placing a jet nozzle on a water hose to squeeze a large flow into an intense jet. We all know how to start a brush fire with the lens of reading glasses (at least those of us who live in country where the sun shines). Large glasses may not be the fashion, but they are sure more efficient in starting a fire. The power of the sun radiation that is captured over 1 m^2 area is an *intensity* corresponding to $1\,\mathrm{kW/m^2}$ or $0.1\,\mathrm{W/cm^2}$. If we can make an image of the sun of 1 mm diameter with a lens, we reach and intensity of $100\,\mathrm{kW/cm^2}$, which is above the threshold to start a fire.

In fine applications like dealing with nanostructures and surgery (Chapter 3 and Section 4.4), it is not only the limited resource of energy that makes us use focusing elements, but also the fine control over the material. *How finely can we focus a light? Is there any special property specific to the lasers in that matter*?

The major advantage of laser light over other electronic and thermal light sources is its *coherence*. This is partly due to the difference in the nature of radiation as discussed in Section 1.2. In stimulated radiation all excited electrons follow the stimulated photon, the radiation happens at the same time and the same direction. *Temporal coherence* or coherence in time refers to radiation that is in phase in time. This is not enough to have a coherent source, just like it is not enough if chorus members all sing at the same time, each one at their own home. Somehow we need to have all these voices together to have coherent music. This is where size of the source plays its role. For a large source, even if all the points of the source emit photons at the same time, there will be a delay between the photons due to the extension of the source. Knowing that waves change sign completely in just half a wavelength distance, we prefer to have sources as small as possible, referred to as the *point source*. Having the same phase across the source size is coherence in space or *spatial coherence*.

A laser can produce a collimated beam that is coherent over its cross-section and that will focus on a minimum size spot (Figure 1.19a). The reverse is an ideal point source (Figure 1.19b), where all the emitted photons across the source are in phase. With a proper lens one can make a collimated beam. A different situation emerges with a filament from an incandescent lamp, as shown in Figure 1.19c. Every point of the filament emits photons of different phases. The rays issued from the filament are like spears put in a trash can, all pointing in different directions.

Given a collimated laser beam as in Figure 1.19a, Eq. (1.2) tells us the minimum spot size that we can achieve with a lens. If, for instance, we use a lens of diameter $D = 1\,\mathrm{cm}$ and focal distance $f = 10\,\mathrm{cm}$ to focus a laser beam, $\theta = R/f = 0.1$, and the spot diameter is $D = 0.6\lambda\theta \approx 6\,\mu\mathrm{m}$ for $1\,\mu\mathrm{m}$ wavelength radiation.

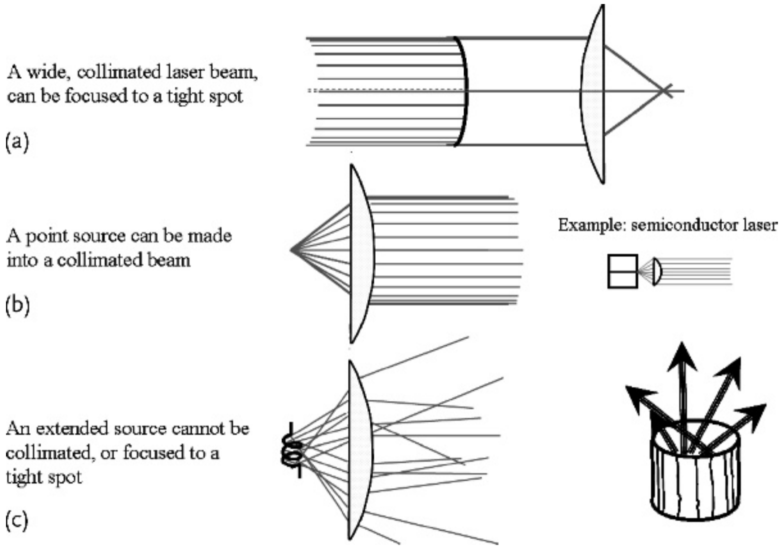


Figure 1.19 (a) The illustration of coherence at source is given by comparing a point source with an extended one like a flash lamp filament. If all rays come from a tiny (compared to the light wavelength) source, they all generate waves (photons) with the same oscillation phase. With ideal lenses and mirrors one can generate a parallel beam to infinity. (b) The same applies in reverse, if a beam is parallel it can be focused to one spot. (c) Since we do not live in an ideal word all these terms are relative, such that as the source get more extended the collimation is less. If the source emission is not spatially coherent, the radiation is like having spears out of a trash can that are all in different directions. An extended filament, even if emits light at the same time, will not generate a well collimated beam and will not be able to merge to a focal point. The smallest beam at the focus corresponds to imaging the filament. Please find a color version of this figure on the color plates.

1.7 How to Make the Shortest Laser Pulse

The methods of »Q-switching« and »cavity dumping« produce energetic laser pulses, but not the shortest possible. The shortest pulse that can be generated by the methods outlined in the previous sections correspond to the time travel in the length of the laser. Since this packet of energy propagates at the speed of light of $c = 299\,792\,458$ m/s or approximately 3×10^8 m/s, the pulse dumped from a 1 m long cavity will pass through an object in a time of $1/(3 \times 10^8) = 3 \times 10^{-9}$ s $=$ 3 ns. This is not a short time by the standard of »ultrashort pulses«, achieved two to three decades after the invention of the laser.

In Q switching and cavity dumping, the cavity is like a bucket or a dam that is filled with water and then drained partially or fully in a short time. An ultrashort pulse, in comparison, is like a high wave, compressing all the water of a dam, like a light bullet, rather than a filling material. In the language of athletics Q-switching and cavity dumping are analogous to walking through the laser cavity and a short pulse is like a »javelin long jump« that covers the whole length in a sudden high jump.

It may come as a surprise that the conformation of the laser producing extremely short pulses is as simple as the elementary device sketched in Figure 1.7b. Yet, each element of the cavity has to have its property accurately selected in order to develop the shortest pulse possible. Before going into the mechanism of pulse generation itself, let us first get into the dissection of a short pulse of light: how continuous waves can add up to create a short pulse.

1.7.1 The Nature of a Pulse

How do we construct a pulse in a mathematical language? We have seen that the nature of laser light is an electric field oscillating at a frequency characteristic of the color of the beam. A pulse is made of a combination of waves of different frequencies that add constructively at one point and destructively at other points. We will call $t = 0$ the instant at which the waves that constitute the pulse add at their crest. An example is illustrated in Figure 1.20, were five waves of increasing frequency are added on top of each other. The sum of these waves is shown with a thick red line. At the common crest ($t = 0$), the sum of the electric field of all these waves adds up to 5× the field of a single wave, which implies that the intensity is 25× that of a single wave. There is no violation of energy conservation, because after a few optical cycles, the fields add up to nearly zero.

Using our picture of radiation (Section 1.2) of girls on the slides, having many modes is like having different sets of slides that vary in height. The radiation (sliding) can be in phase if all the girls start at the same time, but eventually, since the slides are different, the landing time overlap is zero.

1.7.2 The Mode-Locked Laser and the Train of Pulses

In order to get shorter pulses, one has to manipulate the waves inside the laser resonator itself. Let us return to the analogy of the duckling paddling in a rectangular box of Figure 1.7c. It produces a series of waves that have different wavelengths, with the condition that all these wavelengths are a submultiple of the box length. All these wavelengths that fit into the duckling's pond are called modes of the pond (or laser cavity). If the duck were a real genius, he could perhaps paddle in such a way that all the waves were generated »in phase« as in Figure 1.20, creating a giant crest that would probably swallow him up. This is, however, the secret of ultrashort pulse generation in a laser: creating as many frequencies as possible, equally spaced, and adding them *in phase*.

This process of putting the modes in phase is called »mode-locking«. In the music world, this is as if many harmonics (10 000 to 1 000 000) of a note would be resonating together. The resulting pulse will propagate back and forth in the resonator, with a »splash« at the cavity end (the output mirror), creating a train of pulses at a regular interval, which is the cavity round trip time $t_{RT} = 2nL/c$, where L is the cavity length and c/n the speed of light in the medium of index of refraction n that constitutes the laser.

We have seen in Section 1.1.3.1 that there is a frequency of oscillation of the light field ν associated with a wavelength λ: $\nu = c/(n\lambda)$. The wavelengths $\lambda_1 = 2L/N, \lambda_2 = 2L/(N+1), \lambda_3 = 2L/(N+2), \ldots$ are spaced in frequency by $\Delta\nu = c/(2Ln)$. The longer the laser cavity L, the smaller the spacing between these wavelengths or »modes«

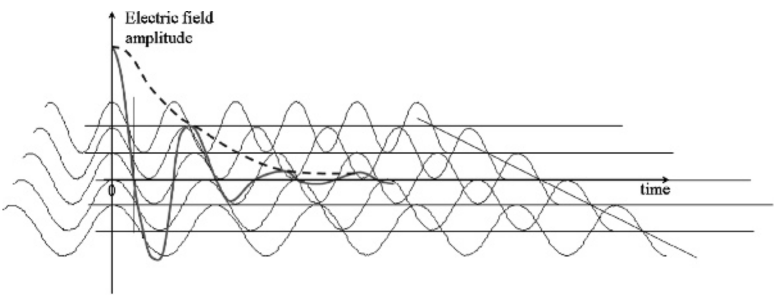


Figure 1.20 Five waves are added on top of each other such that at a given time all the crests coincide. Please find a color version of this figure on the color plates.

of the cavity, and generally the easier it will be to create short pulses. This remains true to the limit of an infinitely long cavity, as is the case when a short pulse is generated directly by propagation through an amplifying optical fiber. It also flows from the discussion of the preceding section that the shortest pulses require the broadest span in frequencies or the largest number of modes. For example, in a Ti:sapphire laser, the spacing between modes is typically 100 MHz, and the number of modes that constitute the pulse can be as high as one million.

These million modes cover a frequency range of a hundred million MHz, or a wavelength range of more than 100 nm over which not only the gain, but the speed of light varies. Indeed, while the speed of light in vacuum c is a universal constant, the index of refraction n is a property of the matter and varies across the spectrum. This property of the medium is called *dispersion* and plays an essential role in the creation (and destruction) of ultrashort laser pulses.

If each wave that constitutes a pulse propagates at a different velocity $c/n(\lambda)$, what will be the speed of the wave packet or pulse (we will assume that there is still a pulse – we will show in the next section how it will be modified)? A pulse is the sum of groups of waves; the speed of the group of wavelengths (*group velocity*) differs from the speed at a particular wavelength. This can be visualized as a group of people walking on a street; there is a collective speed for the group. It can be shown that the pulse travels at a velocity v_g (the subscript »g« for »group«) that is generally slower than the velocity of the wave at the average wavelength.

Thus we see that a short pulse rattling back and forth in a *mode-locked* laser hits the end mirror at regular time intervals $t = 2L/v_g$, even though the wave velocity is not a constant across the range of the wavelength that is contained in the pulse. The output of the laser will consist in a train of pulses, equally spaced, like a pendulum ticking at a rate of $1/t_{RT}$. This is more than an analogy, as we will see that the mode-locked laser is used as an accurate time/frequency standard.

It remains now to be seen how the mode-locked laser can continuously produce a string of ultrashort *identical* pulses, despite the fact that the waves all have different velocities. This is the topic of the next section.

1.7.3 The Mechanism of Pulse Compression by Propagation

To understand how short pulses can be generated, we refer to Section 1.7.1 where it was shown that a pulse of light consists of a superposition of propagating waves that have a common crest propagating at a velocity called »group velocity«. The broader the range of frequencies involved, the shorter the pulse. There are two mechanisms in the production of ultrashort pulses in a laser: a mechanism – dubbed »nonlinear« because the speed of light depends on the pulse intensity, by which the range of frequencies is extended, and another one called »dispersion«, by which all the waves are synchronized to crest at the same point.

In the case of a laser, the waves that »fit in the box«, are equally spaced in frequency (that frequency spacing is the inverse of the round trip time of the laser cavity). In general, when such a pulse propagates back and forth between two mirrors, it will broaden in time. The reason is that there is a slight difference in velocity between the high frequency waves (short wavelength) and low frequency waves (long wavelength) that constitute the »wave packet«. This wavelength dependence of the velocity is called dispersion, and applies to water waves as well. How then can ultrashort pulses be generated, in presence of this ever present dispersion? The answer is in a nonlinear phenomenon, by which the wave velocity is linked to the intensity of the wave (Kerr effect). The speed of light is highest in vacuum: $c = 299\,792\,458\,\mathrm{m/s}$. As it propagates through matter, it is slowed down by the interaction with atoms and molecules, resulting in a velocity $v = c/n$ where n is the index of refraction. The electric field of the laser light (we will assume here that light is linearly polarized, with its electric field oscillating in a direction perpendicular to the direction of propagation) tends to orient the molecules, resulting in a higher interaction with the light field, and therefore a larger n and a slower wave velocity.

Thus, as the pulse intensity is ramped up, the wave decelerates, which implies that the optical cycles will stretch in time, or, in other words, the frequency decreases with time. The reverse occurs at the pulse tail. Near the peak of the pulse, the frequency sweeps from low to high: this is called »upchirp« (Figure 1.21a). This is the important first step in pulse compression: the sweep in frequency, such that the

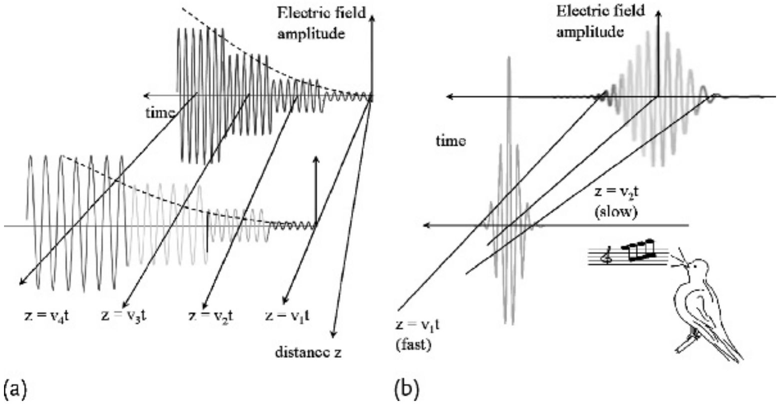


Figure 1.21 (a) Generation of new optical frequencies by propagation of a pulse in a medium where the wave velocity increases with intensity. The optical signal of increasing intensity is shown as having a stepwise increase on its leading edge (the time axis points to the left, for a better visualization of a pulse propagating to the right). The different intensities propagate at velocities v_1, v_2, v_3, and v_4 with $v_4 < v_3 < v_2 < v_1$. As a result, the pulse optical frequencies decrease with time along the pulse (or the wavelengths increase along the pulse) after some propagation distance. (b) Through the process shown in (a), one has generated an »upchirped« pulse, that is, a pulse of which the optical frequency increases with time. This is exactly what the little bird at the bottom of the picture is singing. This pulse is now sent through a medium with negative dispersion (i.e., a medium in which the higher frequencies (blue) propagate faster than the lower frequencies (red)). As a result, the tail of the pulse catches up with the pulse leading edge, resulting in pulse compression. Please find a color version of this figure on the color plates.

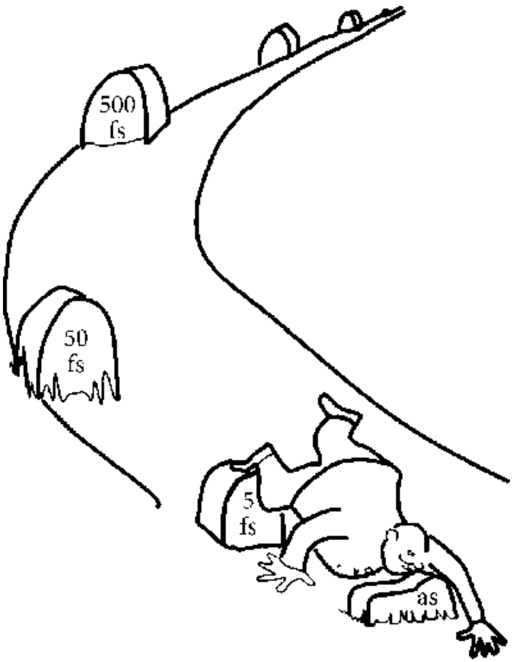
lower frequencies are pushed ahead of the pulse and the higher frequencies are pushed towards the trailing edge.

As we saw in Section 1.7.1, many frequencies are needed to make a pulse. Now the next question is how to stack those frequencies in phase to make a short pulse. The trick to continuously compress the pulse is to select a laser cavity configuration such that the lower frequencies propagate more slowly than the higher frequencies,[10] which implies that the leading edge moves towards the peak, and the trailing edge also recedes towards the peak. Hence, the pulse is compressed, as sketched in Figure 1.21b. It is the task of laser engineers to use optical components such as mirrors and prisms to modify the speed of light as a desired function of frequency.

10) This is called a cavity with negative dispersion.

1.7.4 The Real Thing: From Splashing Dyes to Crystal Lasers to Semiconductor Lasers

1.7.4.1 Milestones and Stumbling Blocks in Femtosecond Laser Development



From the nanosecond Q-switched and cavity dumped pulse, the path to ultrashort pulse generation has not been straight but full of stumbling blocks, the last one being the attosecond barrier. Historically, the field of ultrashort pulses has grown with dye lasers. It has been a remarkable circular evolution, as successive mechanisms were discovered, which led to new designs apparently repeating a much older one. The dye laser was chosen because of its high gain (up to a factor of 2 gain with continuous pumping, and a factor 3 with pulsed pumping, could be achieved through the 100 μm thick dye jet). Another advantage is that it is a visible laser, which, in addition to the beautiful palette of colors that it provides (Figure 1.22), implies rel-

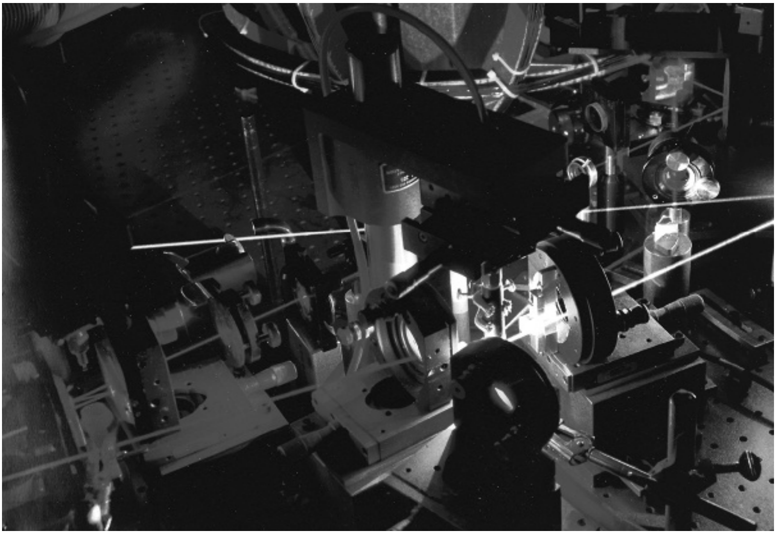


Figure 1.22 A ring dye laser. The dye jet is seen in the middle of the picture, pumped by the bright blue beam of an argon ion laser. The mirror in the foreground focuses the argon leaser beam at a tiny spot in the dye jet. This gain spot is at the focus of two curved mirrors, which direct a collimated beam in the other part of the cavity. The red intracavity beam of the dye laser can be seen on the picture. Please find a color version of this figure on the color plates.

ative ease of alignment because of the visible fluorescence[11] of the gain medium. The high gain made it possible to implement successive lossy mechanisms of pulse formation and compression inside the laser cavity. The first one that came to mind was saturable absorption, with an organic dye as saturable absorber, contained in a thin liquid jet similar to the jet contained the solution of the gain medium. The organic dye – Rhodamine 6G being the most popular gain medium – was dissolved in ethylene glycol and forced at high pressure through a thin (100 μm typically) stainless steel nozzle. The pump laser beam (argon ion laser of up to 20 W power) was focused into and nearly totally absorbed by that flowing sheet of liquid. It is because the focal volume was replenished at very high speed that such high power density could be injected into that gain medium. It should be noted that, in this configuration, the dye laser could handle higher power density than any other medium. Such properties explain why the dye

11) Radiation that is due to pumping with a different wavelength, usually a shorter pump wavelength.

laser has been one of the most popular sources for spectroscopy and short pulse generation for at least two decades. One can wonder why it seems to have vanished in the twenty-first century. Several factors contributed to its demise. First was the inconvenience of having to handle a noisy pump system and high pressure liquids. Not only did the laser produce a colorful display of colors, but most often the operator was covered from head to toe with red, orange, green colors, thanks to the occasional bursting of a high pressure hose. Second was the chemical hazard. While Rhodamine 6G was used as food coloring in the early 1960s, it was soon discovered to be carcinogenic, as most other organic dyes. A little known fact is that the viscous solvent, ethylene glycol, which is commonly used in car radiators, is listed as a hazardous liquid: when heated, its vapor can cause permanent brain damage.

As mentioned above, the first continuously mode-locked laser had a saturable jet and a gain jet in the same cavity, as sketched in Figure 1.23a [13]. The pulse duration was 0.6 ps. Then a frantic competition started, with subsequent publications citing a pulse reduction by a small percentage, due to a new physical approach of configuration. Improved saturation absorption was achieved by what was later dubbed »colliding pulse mode-locking«; the implementation is sketched in Figure 1.23b [14]. The principle is that when two traveling waves »collide« head on, they create a standing wave where the crest and trough of each wave add up at one position. Water waves will double in amplitude. Light waves will double in electric field amplitude at the crests, but quadruple in intensity, which implies increased saturation. Two traveling colliding waves can be created by reflecting a wave against a wall – in the case of optics by reflecting a pulse on a mirror. The first »colliding pulse mode-locking« used a saturable absorber flowing against an end-cavity mirror [14] and resulted in the generation of 0.3 ps pulses. A prism in these first configurations was used as in a spectrometer to select the wavelength at which the saturable absorber can operate. As we have seen, the shorter the pulse, the larger the range of wavelengths that constitute the pulse, and therefore wavelength selection by a prism is contradictory to making a short pulse. The next increment in a shorter pulse was obtained by eliminating the bandwidth limitation of the prism and using a cavity mirror that uniformly reflects only the wavelengths over which the saturable absorber functions. This simplest cavity shown in Fig-

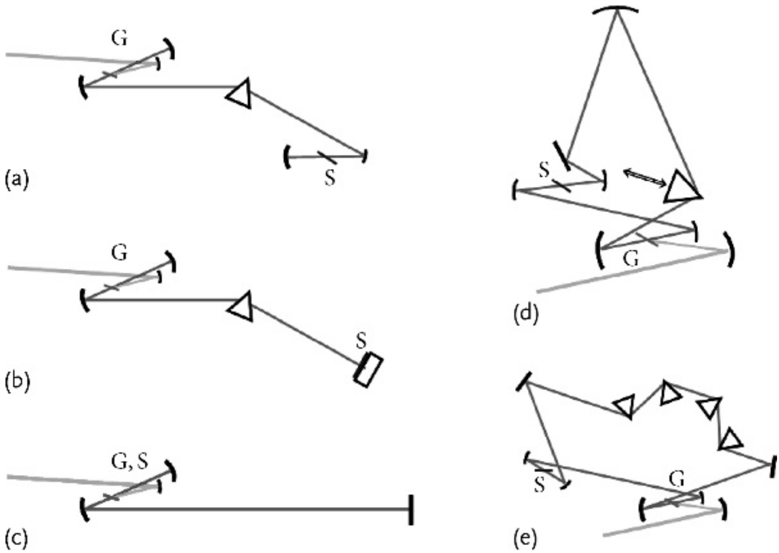


Figure 1.23 Successive configurations of mode-locked dye lasers. G is the gain jet – most commonly Rhodamine 6G – and S is the saturable absorber – most commonly diethyloxadicarbocyanine iodide. The configuration (a) of the early 1970s [13] gave way to (b) a »colliding pulse mode locked cavity« [14]. A simpler cavity in (c) used all dyes mixed in one jet, and mirrors with square spectral bandwidth [15, 16]. In (d), with the »colliding pulse« mode-locking implemented in a ring cavity, the prism is reinstated but at the center of curvature of a mirror [17]. In (e), a four prism sequence is introduced in the ring to produce negative dispersion [18]. Please find a color version of this figure on the color plates.

ure 1.23c [15, 19] produced pulses of 120 fs duration. It was realized that the thickness of the saturable absorber dye jet was setting a limitation on the length of the optical pulse. The next improvement that brought down the pulse width to 90 fs was a thinner saturable absorber dye jet, and the same colliding pulse mechanism as in Figure 1.23b but implemented by a cavity with a ring geometry [20]. The 90 fs shortened to 65 fs [21] after a group at Rochester reported 70 fs out of the configuration of Figure 1.23c pumped synchronously by a mode-locked argon laser [16]. As time seems to proceed backwards, the edge mirror was eliminated, the prism was reintroduced, but located at the center of curvature of a mirror, so that the wavelength selection it introduced was greatly reduced (Figure 1.23d). That configuration, which is identical to that of Figure 1.23a except for the ring geometry of Figure 1.23d, generated pulses of 55 fs duration [17].

The innovation was to compress the pulses inside the laser cavity by adjusting the amount of glass (translating the prism).

The mechanism is exactly that of »soliton compression«, mentioned earlier in Section 1.7.3. In this particular case, the dispersion of the prism is positive, compensating a downchirp caused by saturation of the dye (as the absorber dye saturates, its index of refraction decreases, which causes a downchirp). The soliton approach took a different twist, with the introduction of a four prism sequence to create negative dispersion (Figure 1.23d), and a positive *self-phase modulation*[12] in the solvent of a jet that, in this case, was no longer acrobatically thin. The balance of the Kerr effect and prism dispersion led to a pulse duration of 27 fs [18], which was the basis of the design of the solid state lasers that followed, such as the Ti:sapphire laser.

As the 1970s and 1980s were the decades of the dye laser, the 1990s saw the start of two decades of the ultrafast Ti:sapphire laser. Pulse durations of less than 12 fs in the early 1990s [22, 23] evolved to pulses of less than 5 fs (only two optical cycles) in 2001 [24]. This is the limit of what can be achieved directly out of a laser. The next era is in the attosecond range.

What made the Ti:sapphire laser more successful than the dye laser? The gain is lower, and the beam is barely visible. The output power of a laser is not determined by the gain but by the saturation intensity. A high gain implies a low saturation intensity, which explains that the mode-locked dye lasers had only an output power typically in the mW range. The Ti:sapphire laser has a thousand times lower fluorescence, much smaller gain, but a thousand times larger saturation intensity. Output powers of the order of 1 W can be routinely obtained from the Ti:sapphire laser. The wavelength in the near infrared is barely visible. The configuration is similar to that of the dye laser, as shown in Figure 1.24. The control of dispersion of this laser is made with a prism sequence, the first one of which is seen on the left of the picture. In subsequent lasers, the function of negative dispersion is achieved with special coatings of mirrors. The

12) Self-phase modulation is a change of phase or speed of light across the pulse due to the Kerr effect. This is a self-induced phenomena, because the medium response varies with the intensity of the pulse. A positive self-phase modulation results in a frequency sweep across the pulse, having lower frequencies at the front and higher frequencies at the tail of the pulse.

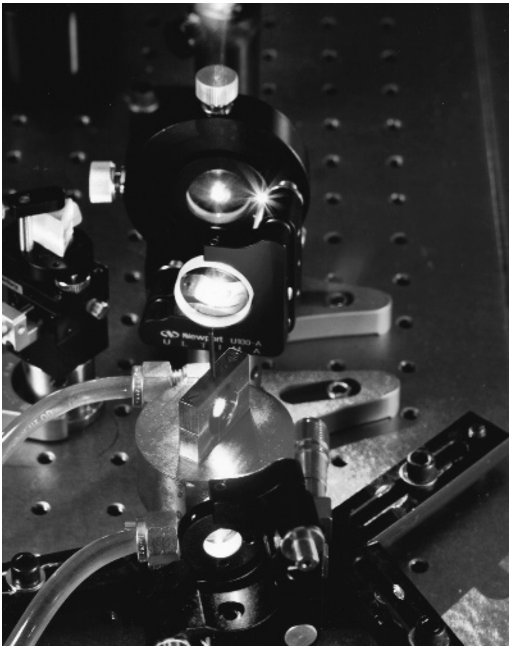


Figure 1.24 A Ti:sapphire laser. The Ti:sapphire crystal is pumped through one of the two curved mirrors projecting the gain spot into the cavity by a frequency doubled vanadate mirror. Please find a color version of this figure on the color plates.

mode-locked Ti:sapphire lasers used for ultrashort pulse generation have a pulse duration that can be as short as 8 fs. Since the speed of light is approximately 0.3 μm/fs, this represents a light bullet of only 2.4 μm, of three optical cycles at the wavelength of 0.8 μm.

It is not certain, however, that the Ti:sapphire laser has a bright future. Its biggest shortcoming is the cost of acquisition and operation. It requires a semiconductor pumped vanadate laser operating at 1064 nm, which is frequency doubled (see Section 1.11) to 532 nm, to pump the Ti:sapphire. This chain of lasers is not only expensive in acquisition, but also inefficient. The future most likely belongs to semiconductor lasers, with high wall-plug efficiency, which are slowly approaching pulse durations comparable to those of the Ti:sapphire laser.

1.7.5 Water Waves and Laser Pulses

1.7.5.1 From Water to Lasers: The Soliton

The equation that describes the compression mechanism of the previous section is famous in physics and in mathematics: it is the »nonlinear Schrödinger equation«. It has been used to describe the evolution of a pulse in a cavity, leading to a steady state solution which is a stable pulse of well defined shape, energy, and duration, propagating back and forth in that laser cavity. The nonlinear Schrödinger equation in space is also used to describe the spatial profile of a beam that collapses and creates its own waveguide in a nonlinear medium (which can be any medium, including air, as will be seen in a subsequent chapter). The soliton is one steady state solution of the nonlinear Schrödinger equation. The phenomenon was discovered long before the mathematical equation even existed, by the engineer and shipbuilder John Scott Russell (1808–1882) [25] who named it »the wave of translation«. In 1834 he was riding by the Grand Union Canal at Hermiston, Glasgow and observed that when a canal boat stopped, its bow wave continued onward as a well defined elevation of the water at constant speed. It was only in the 1960s that the name »soliton« was coined when the phenomenon was rediscovered by the American physicist Martin Kruskal. Details of the early history of the soliton can be found in a book by Robin K. Bullough [26]. The soliton took up so much importance in all areas of physics and laser optics that a recreation of the observation of John Russell was organized for the 150th anniversary of his observation. Figure 1.25 shows a picture of a re-creation of that event.

1.7.5.2 From Laser to Water: The »Freak Wave«

It is a water problem (the goal of John Russell was to design the best dimensioned canal for navigation) that led to the discovery of solitons, which brought a revolution in the field of lasers and fiber communications. Most recently, the same »nonlinear Schrödinger equation« that became an essential tool to understand the propagation of ultrashort optical pulses came under renewed interest to solve a puzzling problem for oceanographers. It was the biggest mystery of the seas. Large ships disappeared over centuries, without leaving any tracks. This led to the creation of myths of sea monsters and then that of »gi-

Figure 1.25 Recreation of the soliton in 1984 – soliton home page, Heriot Watt University (Reprinted by permission from Macmillan Publisher Ltd: *Nature*, **376** (1995), 373).

ant waves«. These stories were dismissed as seamen's tales, until two well documented events in 1995.

The first was an automated recording at the Norwegian Oil platform Draupner-E, which indicated in the New Year night a single wave of 26 m height. The second was on 11 September of the same year, when the luxury liner Queen Elizabeth 2 on its way to New York was hit by a giant wave (report of Ronald Warwick, Captain of the Queen Elizabeth 2, interviewed on the BBC Radio 2 on 14 November, 2002). All oceanographic models so far predicted a statistical distribution of waves not exceeding 7 m, while the freak waves recorded reached more than 30 m.

The wave equations used were similar to the linear wave propagation in optics, including a frequency dependent wave velocity. The breakthrough came when the same nonlinear term as in optics was introduced: an intensity (in this case height) dependent wave velocity. One can justify this behavior by arguing that the higher the wave, the less its velocity would be affected by friction with the deeper water. Unlike the case of lasers where there is a deterministic, controlled distribution of waves, the random distribution of water waves in a storm

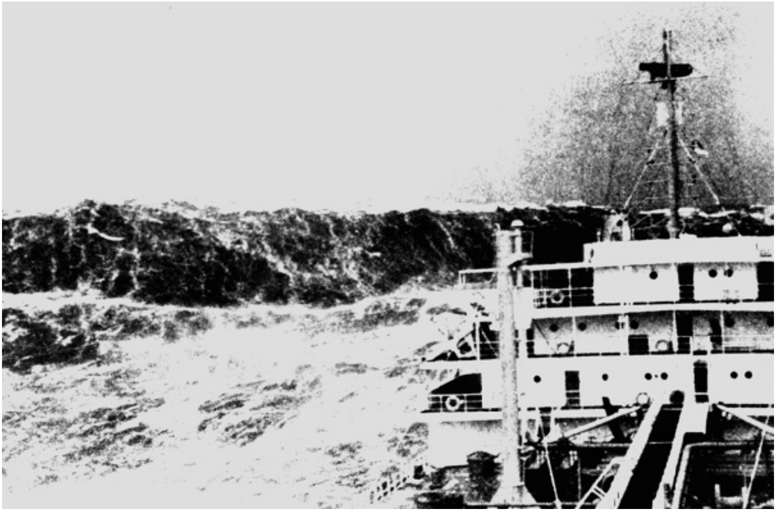


Figure 1.26 Huge waves are common near the 100-fathom curve on the Bay of Biscay. Published in the 1993 Fall issue of Mariner's Weather Log.

have only a very small, but still existing probability to be in the condition of »pulse compression« defined in Section 1.7.3, hence the freak occurrence of the »giant wave«.

There is a more scientific name than just »rogue« or »freak« for large events that have a nonzero probability of occurring. The *probability distribution* – that is, the probability that a wave of size h occurs, has a characteristic heavy tailed »L-shape«. In stark contrast to *Gaussian probability*, where the probability of being in the mean (middle) value is the most likely, events much larger than the mean occur with significant probability.

1.7.5.3 Back from Water to Lasers: Freak Colors

The »rare but significant« statistic [27] had stirred some excitement among mathematicians and physicists, who started to search for similar statistical events in other areas [28]. Once again, theoretical findings of water waves found their analogy in laser physics. This time, it is not »freak waves«, but »freak colors« as shown in Figure 1.27.

One of the key characteristics of laser light is the ability to form a nearly perfectly, uniform collimated beam (see Section 1.6.1). Driving your car at night with the best large laser headlight, you would be able to uniformly illuminate a wide road at 1000 km distance! This would

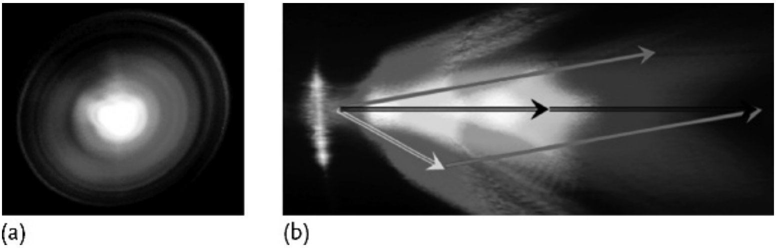


Figure 1.27 (a) The annular display seen on a door, 25 m away from the intense short pulse laser source. The beam has collapsed into a single filament, which has produced these concentric colorful cones. (b) A filament produced in glass. The color spectrum is displayed horizontally (from long wavelength to the left, to short wavelengths to the right), and the angle the individual rays make with the horizontal is the vertical axis. The intense beam propagates along an horizontal axis (shown in black). The directions of the rays is indicated for different colors. Please find a color version of this figure on the color plates.

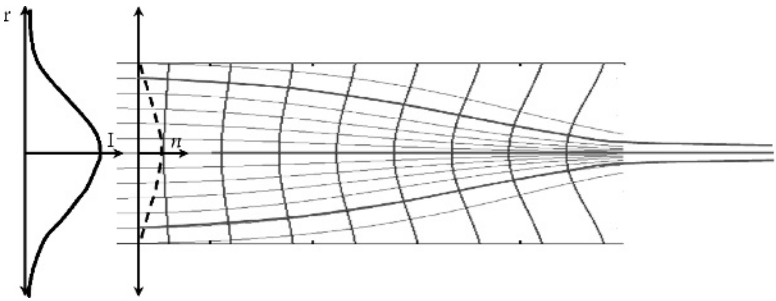


Figure 1.28 Formation of a »filament« or a self-guided intense beam in air. A laser beam is propagating from left to right, with an intensity versus radius sketched at the extreme left. Given a sufficiently high intensity, the index of refraction will increase proportionally to the intensity, as indicated by the dashed curve. This spatial dependence of the index of refraction results in a lensing effect. The wavefronts will curve, and the rays, perpendicular to the wavefronts, will bend towards the axis. A waveguide is formed by the beam itself, similar to the fiber sketched in Figure 1.12. Please find a color version of this figure on the color plates.

work only if our earth were flat, and if there were no distortion due to atmospheric turbulence. However, that ideal headlight would no longer give you good uniform illumination if you were to try to increase the optical power beyond reasonable limits. At powers exceeding 10 GW (10 GW or 10 000 MW), the hitherto uniform beam will collapse into tiny light needles or »filaments« (Chapter 7, Section 7.1). Of course this can only be observed for very short times: generating such high power, your car battery would run empty in less than a μs.

The physical origin of such collapse is the same intensity dependent wave velocity that was exploited in Section 1.7.3 to generate ultrashort pulses. However here, at these intensities, it is the index of refraction (Kerr effect) of air that becomes intensity dependent, creating a lensing effect where the intensity is the highest, as sketched in Figure 1.28. The beam is focused by that self-induced lens, until it reaches a very small diameter; it creates its own waveguide, guiding light in a similar manner as a fiber (Figure 1.12). The spatial shape (intensity versus radius) of these filaments is also governed by the »nonlinear Schrödinger equation«, another contribution from water science to the laser optics field. These filaments are »solitons in space«, as opposed to the »solitons in time« presented in the previous section. However, a dashing display of colors is associated with these filaments, as shown in Figure 1.27. Beauty, however, does not come with Gaussian statistics. It was recently suggested that the widest display of colors is related to the freak wave statistics. For instance, the appearance of the violet–blue in Figure 1.27 is characterized by the same »L-shaped« statistics as the freak oceanic waves [29].

1.8 Ultrashort Ultraintense Laser Pulses

When dealing with pulses one must consider the extension of pulse in time. One joule per second output of a continuous laser with spot size of 1 cm^2, gives the intensity of 1 W/cm^2. For a pulsed laser output of 1 J and 10 fs long and the same size, the intensity is 10^{14} or 100 million million Watt per square centimeter. Short pulses are used whenever extreme high intensity is needed, at the expense of the exposure time.

1.8.1 Definition of »Intense«

The perception of what is an »intense« laser pulse varies over an immense range. For the medical user, a beam that burns the skin can be qualified as intense, and this requires no more than 10 or 100 W/cm^2. A chemist will consider high the intensity that will break a chemical bond, which can be in the range of MW/cm^2 to GW/cm^2. The most power hungry are the physicists. At the lower end of the scale, the atom physicist will only get excited if the field of the laser[13] reaches a sufficient value in one oscillation to extract an orbiting electron from the atom (Figure 1.29a), and pull it back at the next half cycle of the field (Figure 1.29b). The disturbed electron comes back to the atom with a vengeance, producing a burst of x-rays of incredible short duration, in the attosecond range (Figure 1.29c). One attosecond, 1000 times shorter than the femtosecond, is in the same ratio to the second as the age of the universe is to the second (10^{18}). The intensities to reach electron extraction – return to the atom – are now in the terawatt (10^{12}) to petawatt (10^{15} W/cm^2) range. To move away from the abstraction of these numbers, let us try to get a feeling for what these intensities mean by considering the pressure of light. The photon has no mass, but it acts like a solid particle, having a momentum $h\nu/c$. When bouncing off a surface such as a mirror, it induces a recoil momentum $Mv = 2h\nu/c$, where v is the velocity given the mirror of mass M. A beam with a flux of N photons/(cm^2 s) has an intensity I (in Watts/cm^2) and exerts a pressure of $N2h\nu/c = 2I/c$ on the mirror. Sunlight has an intensity of 0.1W/cm^2. The pressure that sunlight exerts on the 1 cm^2 area of the reflecting »light mill« shown in Figure 1.29d is equivalent to one millionth of a 1 mg flea distributed over a cm^2 area! Definitely not sufficient to make a light mill spin.[14] At the intensity level of 10^{16} W/cm^2 used in attosecond pulse generation, the radiation pressure is 1000× higher than the pressure at the deepest point of the ocean (Figure 1.29e).

As impressive as these intensities may seem, there is another breed of physicists that looks upon them with disdain. At sufficiently high intensities, electrons can be accelerated during a half optical cycle to relativistic velocities, that is, velocities close to the speed of light. As one tries to accelerate an electron more by increasing the light field, the mass of the electron increases (to reach infinity at the speed of light). For the relativist plasma physicist, high intensity means above 10^{18} W/cm^2, for 1 μm radiation. That intensity corresponds to an op-

13) Remembering that laser light is an electric field that oscillates in time.

14) One may wonder what makes a light mill spin? Each of the four planes of the light mill has a black face and a mirror face. The radiation pressure acts on the reflecting surface. One may notice that the irradiated mill turns as if the light was pushing the *black* surface. This is because the black surface absorbs the photon and heats the surface, which partially vaporizes. It is the recoil from the molecules escaping the black surface that makes the light mill spin.

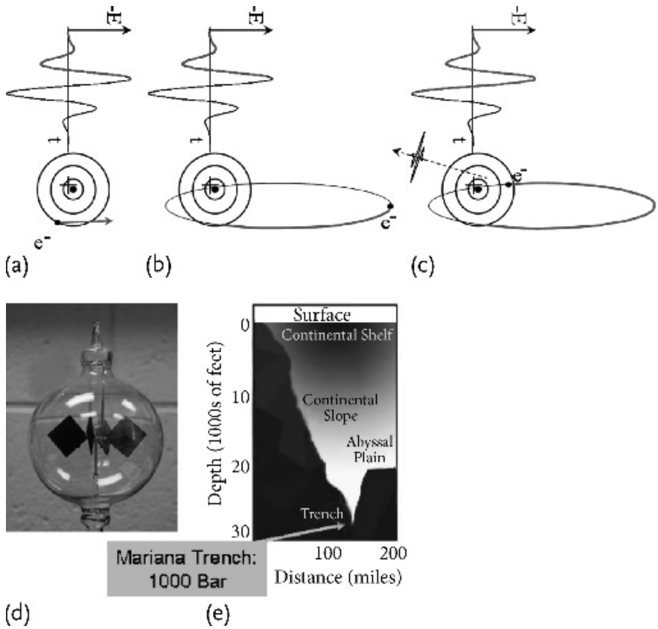


Figure 1.29 The electric field of a pulse, near the peak of a cycle, is sufficiently high to eject an electron from its orbit (a). The electron is accelerated away from the atom, then decelerated to a halt the next half cycle, as the light electric field reverses sign (b). During the next quarter cycle of the light pulse, the electron is accelerated back towards the atom. Numerous complex physical phenomena can occur during that recollision, one of them being the emission of an extremely short (typically 100 as) burst of x-ray radiation. (d) The pressure of sunlight or a He-Ne laser is not sufficient to make the »light mill« (or »Crookes radiometer«) spin. However, the radiation pressure at intensities used to create attosecond pulses is 10 000 times larger than that at the bottom of the ocean (e). Please find a color version of this figure on the color plates.

tical field strength that accelerates the electron to a speed at which the electron mass has increased by 50%.

There is a group of physicists even greedier for power: the particle physicist. For him, an intense laser field E is such that, for instance, an electron would be accelerated over a wavelength λ_c to an energy $eE\lambda_c \geq 2mc^2$ where m is the mass of the electron (or another particle). mc^2 is, according to Einstein, the energy equivalent to a particle of mass m at rest. As will be seen in Section 6.2, λ_c is the wavelength associated with the accelerated electron, or the »Compton wavelength«.

A back of the envelope calculation shows that the corresponding laser intensity is above 10^{23} W/cm^2. The concentration of energy is such that the creation of matter – a pair of electron–positrons – will emerge from vacuum. A plethora of other effects are expected at such intensity levels, which are totally beyond the scope of this overview.

One question to answer, however, is whether such intensities belong to science fiction or real research. Numerous countries have national facilities in the PW (10^{15} W) range. The peak power of the next generation laser that is contemplated is above the »exawatt« (or 10^{18} W) range. Such a laser is no longer at the scale of a national laboratory, but requires resources at the scale of a continent. The European Community has started such a project: the »Extreme Light Infrastructure« or ELI. Extreme is even an understatement when trying to describe the highest intensities contemplated in this project. To get another appreciation of the meaning of high intensity, let us embark on a thought exercise of what is required to keep a beam collimated, despite the diffraction effect mentioned in Section 1.6.1. In air, at atmospheric pressure, it has been established that a power of at least 3 GW peak power is required to create a self-induced of waveguide, as described in the previous section. The power required is inversely proportional to the density of air, since the intensity dependent increase of index or refraction is proportional to the number air molecules that create the nonlinear index. Could a filament be produced in the mesosphere (approximately 100 km altitude), for instance, to guide a reentry vehicle? The density of air there is 10^{16} molecules/cm^3, or 10 000 less than the 10^{20} molecules/cm^3 at sea level. The power required to create a light filament is then of the order of 10 000 GW or 10 TW. In deep space, the density of molecules drops down to $N = 10^5/\text{cm}^3$, and the power required to create a filament is then 10^{23} W. At this level of intensity, we are out of classical physics. Quantum field theory tells us of the existence of self-focusing »by vacuum polarization« and »birefringence of vacuum«; in short vacuum itself will cause the beam to self-focus at a power level of $\approx 10^{23}$–10^{24} W. For these powers, vacuum itself is a nonlinear medium, since it will absorb the light energy through creation of matter. There will also be nonlinear phase changes; all effects that are accessible in the range of power and intensities are contemplated by the ELI project.

1.8.2 How Ultraintense Pulses are Generated

We have described a laser source as an optical amplifier sandwiched between two mirrors. The shortest pulses are produced in a laser where a short pulse rattles back and forth in a laser cavity, hitting the end mirror at regular time intervals (Section 1.7). These mode-locked lasers are the source of the shortest optical pulses that can be generated (a couple of optical cycles), but the energy/pulse is generally small, at most a few nanojoules. Amplifiers are required to boost the energy of one or a few pulses from the train of pulses emitted by the laser. As in the case of the laser, the amplification process is limited by saturation of the gain medium. The pulse to be amplified has to be sent through several amplifier stages of increasing diameter, so that the intensity of the amplified pulse remains below the saturation intensity of the gain medium. The energy required to keep the media amplifying (maintain a population inversion) increases with the diameter of the gain medium (Figure 1.30). It is like threading a growing elephant through tubes, of which the diameter and strength have to increase with the growing elephant size. The problem can be solved by stretching the ultrashort pulse, so that its intensity is reduced by the stretching ratio and it can be amplified to considerably higher energies before its intensity reaches saturation or damage levels.

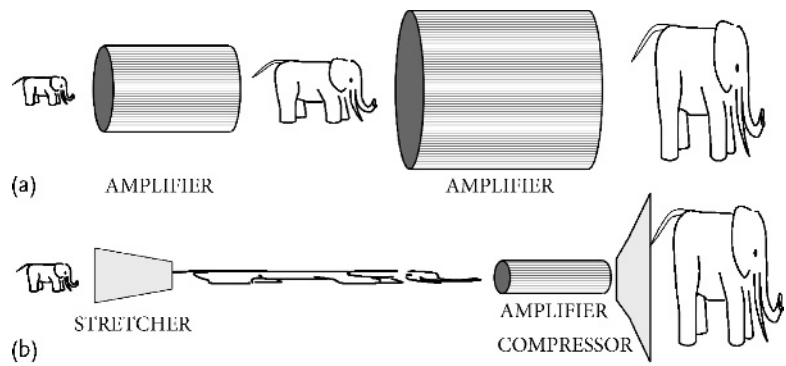


Figure 1.30 (a) Any amplifier stage has to be matched to the size of the amplified pulse. For a given amplifier diameter, the saturation limits the maximum pulse energy that can be extracted, as the diameter of the tube limits the size of the elephant that can pass through it. (b) Instead of increasing the diameter of the tube, why not stretch the elephant to make him pass through a narrower tube and give him the original proportion after amplification? Please find a color version of this figure on the color plates.

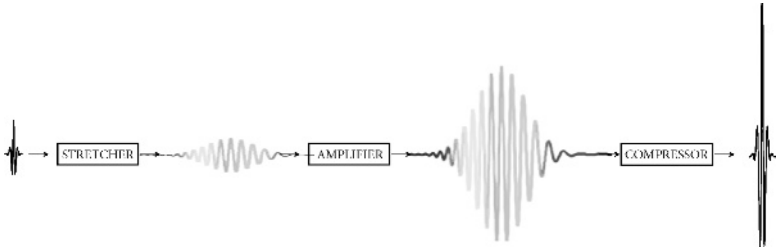


Figure 1.31 Chirped pulse amplification. An ultrashort pulse is sent through a stretcher, which can be constructed with discrete elements (gratings, prism) or be a piece of glass or a fiber. The different frequency components of the pulse are delayed with respect to each other, resulting in chirp and pulse stretching. The 10 000 to 100 000 times pulse stretching ratio allows considerable amplification before the damage limit to the components of the amplifier is reached. The action of the stretcher is reversed in the amplifier, resulting in the creation of an extremely intense ultrashort pulse. Please find a color version of this figure on the color plates.

The mechanism of stretching is dispersion, a mechanism that was discussed in Section 1.7.3, Figure 1.21. In a medium with dispersion (gratings, prisms, glass), different frequency components of the pulse spectrum propagate at different velocities, resulting in pulse stretching. This mechanism is linear and reversible as the sign of the dispersion is changed. »Chirped pulse amplification« (CPA) [30] refers to an amplifier system where an ultrashort pulse is stretched by four to five orders of magnitude, sent through several stages of amplification, before being compressed back to the original pulse width. The successive components, as sketched in Figure 1.31, are stretcher, (linear) amplifier, and compressor. In femtosecond high energy systems, the peak intensity reached are so high that the compressor has to be put in a vacuum chamber. Indeed, as will be seen in more detail in Section 7.1, the air itself is a nonlinear medium at these extremely high intensities.

1.9 Ultrashort Ultraprecise Laser Pulses

In the previous sections we introduced ultrafast events, which at first sight would have little relevance with the time scales associated with our daily life. By a bizarre twist of nature, it is the fastest event that ends up providing the most accurate time measurements for long duration events. Let us take, for instance, the rotation of the earth about its axis. Because of the huge mass involved, one would expect the length of the day to be a perfect clock for long term events, years, centuries, and millennia. The rotation of the earth is not constant: the tides cause a small friction that slows down the earth rotation. Paradoxically, motions of electrons in an atom, which are on the femtosecond scale, provide a more accurate clock than astronomical events. A quartz clock ticking at kHz has a considerably longer range accuracy than grandma's pendulum clock ticking at a second rate. Atomic clocks, however, ticking at petahertz (10^{15}), considerably outperform the quartz clock in long term accuracy. Atomic clocks are not only used as a time standard, but also as a distance standard. Here also, accuracy and stability are provided by microscopic standards rather than by large objects. The meter was initially defined at the time of the French revolution as being 1/100 of the decimal degree on the equator (or 140 000 of the circumference of the earth). This definition has been replaced by a smaller object: a block of quartz a meter long stored at constant pressure and temperature in Paris. However, even that object does not have the long range stability of the wavelength of the cesium clock. Presently, the length standard is defined through the time standard (the period of oscillation of the Cs clock, in the fs range) multiplied by the speed of light in vacuum (defined as a fixed number $c = 299\,792\,458\,\mathrm{m/s}$).

While accuracy may come as a surprise, it should be obvious that the shortest clock period will provide the most sensitivity or resolution in time measurement, since the measurement is performed with finer tick marks. A seemingly impossible challenge is to count time, when the time interval between clock ticks is only a couple of fs, a million times faster than standard electronics can resolve. This is where the laser made a remarkable revolution, which was recognized by the Nobel Prize in Physics in 2005. It refers to a property of the mode-locked laser that has puzzled researchers for some time [31]. The sketch of Figure 1.32 will help explain the dilemma. We have seen that lasers used for generating ultrashort pulses by mode-locking have gain over a very broad range of wavelengths. Therefore, if they are not mode-locked (see Section 1.7.2), they can be operated over a broad range of wavelengths. As we tune that wavelength, we can record the round trip rate of that cavity on a frequency counter. Since the speed of light in the components of the cavity is a function of the wavelength, we

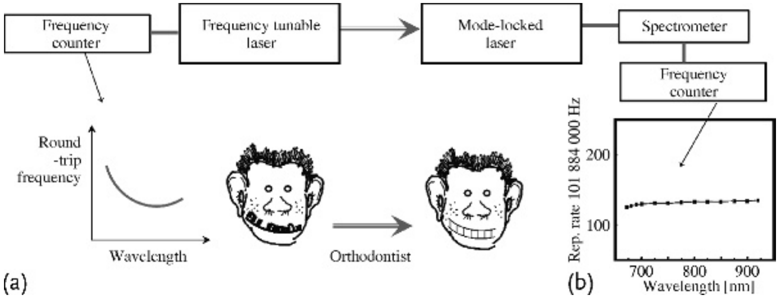


Figure 1.32 On the left: the output of a tunable laser (such as a dye laser or a Ti:sapphire laser) is sent to a detector, connected to a frequency counter. The latter records the repetition rate (round trip rate) of the laser as a function of wavelength. The repetition rate varies across the spectrum. The transformation of the laser from a continuously tunable into a mode-locked laser is equivalent to sending the laser to an orthodontist. As the output of the laser is sent to a spectrometer and thereafter to the frequency counter, the same number (repetition rate of 101 884 130 Hz) is observed across the spectrum.

will record a curve as shown in Figure 1.32a. What is measured is the separation between the optical frequencies, or *modes,* that can be produced by this laser. These optical frequencies form a *frequency comb*, of which the spacing between teeth is not constant. If, by a strike of magic, we make the same laser produce ultrashort pulses, we can perform a similar experiment, consisting in measuring the pulse rate after selecting a particular wavelength region. We find that the repetition rate is constant over the spectrum. The *frequency comb* produced by the laser in mode-locked operation has a constant spacing between teeth [32]. Mode-locking the laser is like sending the laser to the orthodontist, it corrects the tooth spacing. The mechanism by which the teeth spacing is »corrected« is rather complex [31] and is related to the soliton operation discussed in Section 1.7.3. Having a perfect comb in frequency has important consequences in metrology. It means first that this comb can be used to measure frequency differences, just as a metric ruler is used to measure distances. It also implies that, in the time domain, the pulse train that is issued from the mode-locked laser can be represented by a single frequency wave (labeled the »*carrier*«) that is sampled at regular intervals (which correspond to the round trip time of the laser cavity), as shown in the top part of Figure 1.33. In frequency, the mode-locked laser can be represented by a comb of equally spaced teeth, as shown in the bottom part of Figure 1.33.

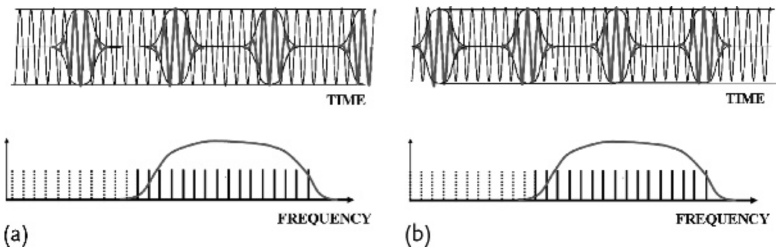


Figure 1.33 The mode-locked pulse train, in the time domain (top) is shown as a wave at a single carrier frequency, modulated by and envelope. The corresponding comb of optical frequencies is represented on the bottom. (a) A general case where the envelope spacing is not an integer number of optical cycles. (b) The envelope spacing is an integer number of light periods, and, in addition, the carrier to envelope phase is zero. Please find a color version of this figure on the color plates.

For each pulse, the phase of the carrier at the peak of the envelope is called the carrier to envelope phase (CEP). In Figure 1.33a, the spacing between *envelopes* is not an integer number of optical cycles of the carrier and, therefore, varies from pulse to pulse. The frequency at which this CEP varies (the ratio of CEP for two successive pulses to the pulse spacing) is called the carrier to frequency offset (CEO). The situation represented in Figure 1.33b is that where the envelope spacing is an integer number of light periods. Thus the time between the pulses is defined with the precision and accuracy of the light period (typically 2.5 fs) and is of the order of a nanosecond, and thus can be electronically counted. If the carrier frequency is linked to that of an atomic clock, one has created a time standard of femtosecond precision. In the frequency domain, one tooth of the comb has been defined with absolute accuracy. The spacing between teeth is also defined with the same accuracy since it is being forced to be an integer number of light periods (forcing the CEO to be zero). The frequency comb is then a frequency ruler of utmost accuracy and precision, which is particulary useful in astronomy.

1.10 Ultrashort Ultrasensitive Laser Pulses

The near perfection of the laser output itself makes it very sensitive to any perturbation. As an example, if we consider only the propagation of the frequency comb of zero CEO shown in Figure 1.33b, the velocity of the pulse envelope will generally be different from that of the carrier. As a result, the CEP will change; a change that can be exploited to monitor the properties of the medium traversed. Highest sensitivity can be achieved if this change in CEP is exploited inside the laser itself, as will be analyzed in more detail in Chapter 8. Other properties of ultrashort laser pulse that can be exploited to sense properties of the traversed media are the possibilities of high power pulses to project high intensities at distance through filaments (as discussed above in Section 8) and create other wavelengths through the nonlinear process to be described in Section 1.11 that follows. Methods to exploit filaments for remote sensing will be the topic of Chapter 7.

1.11 The Nonlinear Wizard: Juggling with Frequencies

1.11.1 Stacking up Photons

The nonlinear wizard is not a scary cross-sighted scientist, but instead a jewel: a small crystal. The generation of harmonics is common with electronic amplifiers. If a pure note is given by the amplifier, and we turn up the volume, the sound will be distorted. When a wave is no longer a pure sinusoidal wave, that means it contains higher frequencies – the harmonics. The first of these harmonics is the »second harmonic«, that is, an electromagnetic oscillation as twice the frequency of the light or half the wavelength.

For instance, the invisible, IR Nd:YAG laser or the vanadate laser at 1064 nm can be made to generate in a crystal a second harmonic at the green wavelength of 532 nm. This has become the most common source of green light used in green laser pointers, in powerful (up to 20 W) continuous lasers to pump Ti:sapphire lasers, or for rock concerts.

There are some stringent conditions to be met by the crystal to generate efficiently the second harmonic. The first condition is a lack of symmetry: the crystal structure should not have a center of symmetry. The second condition is called »phase matching«. The generating wave (called the »fundamental wave«) will create its second harmonic at any point of the crystal. The fundamental propagates at the velocity of the wave $v_{w1} = c/n_1$ where n_1 is the index of refraction of the crystal for the fundamental wavelength along the direction of propagation and polarization of the fundamental. Figure 1.34 illustrates three cases of second harmonic generation in a crystal with a »snapshot« of the waves passing through the crystal at a certain instant. The wave F is the fundamental wave, the same in all three cases. In all three cases, a second harmonic is generated at the entrance plane A, and propagates through the crystal at a velocity $v_{w2} = c/n_2$, where n_2 is the index of refraction of the crystal for the shorter second harmonic wavelength, along the direction of propagation and polarization of the second harmonic. Case (a) corresponds to the case where the crystal is phase matched, that is, $n_2 = n_1$, and the second harmonic and the fundamental propagate at the same velocity in the crystal. Consequently, the second harmonic generated in the plane B, SH_B is in phase with the second harmonic generated in A that has propagated to B. The two signals add in phase to create a larger second harmonic field SH. Case (b) shows the more general case when the crystal is not phase matched, $n_2 \neq n_1$ and the second harmonic SH_A generated at the entrance plane A of the crystal, having propagated to a plane B, may destructively interfere with the second harmonic SH_B created at the plane B. The resulting second harmonic field SH is zero. In the case of phase matching (a), the second harmonic generation can then be very efficient: up to 80% of the fundamental light can be converted into the second harmonic. However, one cannot always find in a crystal an orientation for which $n_2 = n_1$. In that case, there is another possibility to generate a second harmonic efficiently by using the property that the crystal has no center of symmetry. In (c), the second harmonic generated at the entrance plane A and propagated to plane B is SH_A as in case (b). However, by reversing the crystal orientation

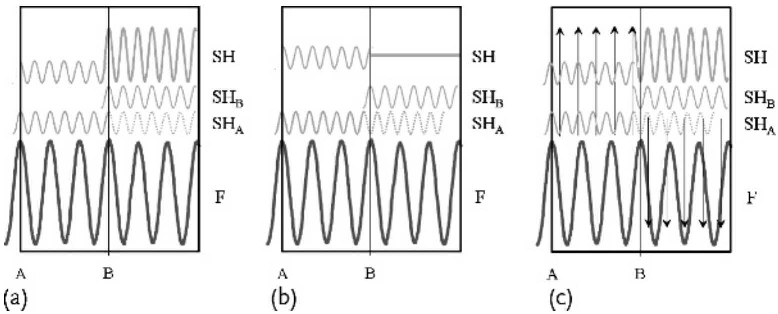


Figure 1.34 Simplified sketch to illustrate the second harmonic generation in a nonlinear crystal. F is the fundamental wave sent through the crystal. In (a) the crystal is phase matched: the second harmonic generated at the entrance plane *A* propagates to the plane *B* (SH_A), and adds constructively with the second harmonic produced by the fundamental at plane *B* (SH_B, to form a large second harmonic field SH). In (b), case of a nonphase matched crystal, the second harmonic SH_A is out of phase with the second harmonic produced by the fundamental at plane *B* (SH_B: the two fields cancel out to form a zero second harmonic SH). (c) is the case of a periodically poled crystal. With the optic axis oriented as shown by the arrows, the propagation and second harmonic generation from plane *A* to *B* are the same as in case (b). In case (c), the optic axis is reversed after plane *B*, and the second harmonic generated in *B* is reversed as compared to cases (a) and (b). Consequently, the two fields SH_A and SH_B add in phase to create a large second harmonic SH. Please find a color version of this figure on the color plates.

at plane *B*, the phase of the second harmonic generated at that plane, SH_B, is also reversed, and therefore adds in phase with SH_A. The crystal where the optical axis is flipped periodically is called a periodically poled crystal. This technique is called quasi-phase matching and is used most often with lithium niobate crystals, which have a very high nonlinearity and good transparency. The acronym PPLN is quite often used to designate these crystals. The nonlinearity is the highest along an optic axis of the crystal, the z-axis. Poling is achieved by applying thin electrodes (10–40 μm wide, depending on the length of the individual domain desired) on the faces perpendicular to that axis. The »poling« or orientation of the axis is made by applying a very high electric field at high temperature on these electrodes.

Nonlinear optics can easily be described through the photon picture. Two photons of the fundamental radiation are converted into one photon of the second harmonic of twice the photon energy, under the influence of the crystal. The two generating photons do not need to be of the same frequency: if one photon has the energy $h\nu_1$ and the other the energy $h\nu_2$, they can combine in the crystal to form the *sum frequency* photon of energy $h\nu_3 = h\nu_1 + h\nu_2$. As almost everything is reversible with laser light, there is also a process of *difference frequency generation* where two light beams at optical frequencies ν_3 and ν_2 are sent into a crystal to generate a wave at frequency $\nu_1 = \nu_3 - \nu_2$. As in the case of the second harmonic generation, the process has to be phase matched to be efficient: all three waves have to propagate at the same velocity.

Two photons of equal frequency or wavelength can add up in second harmonic generation or »frequency doubling«, or they can subtract from each other resulting in zero frequency generation, or optical rectification. In that case, a field of zero frequency, which is a continuous field, is applied to the crystal. The reverse of that operation is the electro-optic effect, which was introduced in Section 1.5.3.3, by which the sum of a zero-frequency photon (continuous electric field) and an optical photon at frequency ν_1 is made, to result in an optical photon still at frequency ν_1 with slightly different phase and/or polarization.

1.11.2 Parametric Generation and Amplification

In the previous section a higher frequency photon was created by adding two lower energy ones. Two photons are lost from the fundamental beam, in order to create a single photon of higher energy. In the reverse process, if starting with a higher energy photon of energy $h\nu_3$, there is *gain* for radiation at the frequency $\nu_s < \nu_3$. This is called *parametric gain* and is fundamentally different from stimulated emission gain. In the presence of a beam of frequency ν_s, the »pump« photon of energy $h\nu_p$ gives birth to two photons, respectively, of energy $h\nu_s$ (the »signal photon«) and $h\nu_i$ (the »idler photon«), such that $h\nu_3 = h\nu_s + h\nu_i$ (Figure 1.35). As in the case of gain with population inversion, the photon $h\nu_s$ is undistinguishable from the other photons of the beam at frequency ν_s. Despite the similarity, parametric gain is fundamentally different from gain produced by a population inversion, as sketched in Figure 1.35. In this case, we consider a square pulse pump (in time), that is, the pump beam is turned on for a very short time, then turned off. The propagation of a weak pulse in a medium with population inversion is similar to a skier going downhill (Figure 1.35a). The signal pulse extracts more and more energy from

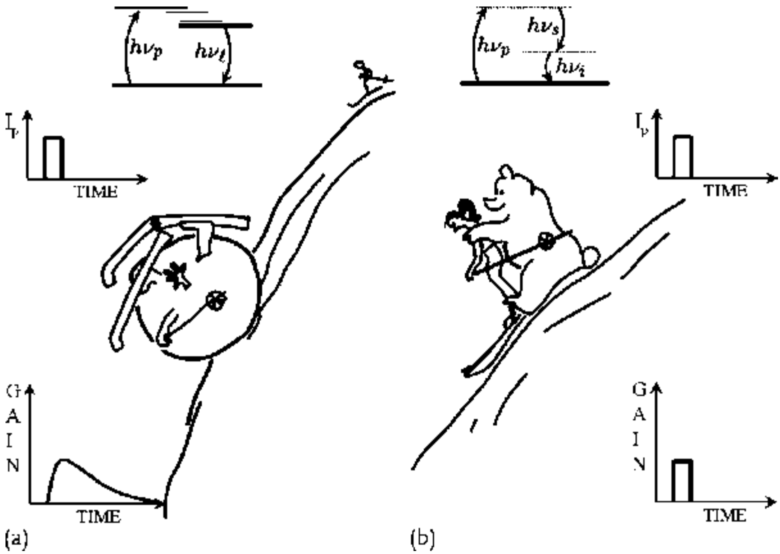


Figure 1.35 Comparison of gain through population inversion and parametric gain. (a) In the case of gain through population inversion, the signal pulse extracts more and more energy from the medium as it grows. (b) In the case of parametric gain, the signal pulse only gains energy as long as the pump is present.

the medium as it grows, because the gain has a lifetime of the order of microseconds in the case of crystal host lasers such as Ti:sapphire, Nd:YAG, Yb:YAG, etc., which is a very long time in the laser world. For a parametric gain however (Figure 1.35b), the signal pulse at ν_s only gains energy as long as the pump (the bear at ν_p) is present. With parametric gain, there is no background noise from fluorescence, as there is with a population inversion gain medium.

1.11.2.1 Optical Parametric Oscillators and Optical Parametric Generators

As we have seen in Section 1.3.3, a laser is an amplifier with the feedback of some mirrors, which can be arranged in a linear configuration or a ring cavity. An optical parametric oscillator is very similar to a laser with gain. The main difference is that the gain medium is a crystal, and that the pump radiation has to be provided by a laser. The main use of the optical parametric oscillator is to be a source of laser radiation in the infrared, since it provides photons of lower energy than the pump photons. At each round trip through the cavity of the beam at frequency ν_s, pump photons of energy $h\nu_p$ are replaced

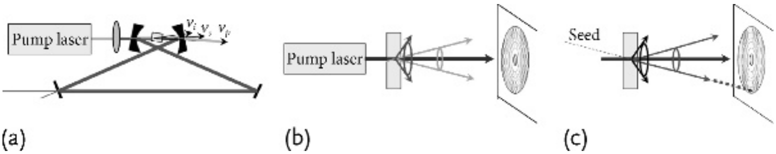


Figure 1.36 (a) An optical parametric oscillator in a ring cavity. (b) An optical parametric generator. (c) A seeded optical parametric generator. Please find a color version of this figure on the color plates.

by two photons of energy $h\nu_s$ and $h\nu_i$. The photon of energy $h\nu_s$ is called the signal photon and adds up coherently (that is, in phase) to the beam of wavelength $\lambda_s = c/\nu_s$. In general, the cavity will only be resonant for the signal wavelength, as sketched in Figure 1.36a. There would be too many constraints if both the idler and signal were resonant with the cavity. Indeed, the idler frequency would be simultaneously fixed by two conditions that, in general, may conflict: to be the difference of the pump and signal frequencies and at the same time a mode of the cavity.

We have seen in the case of a nitrogen laser (see Section 1.4.2.2) that a mirrorless laser can operate if the gain is sufficiently high: the fluorescence is amplified along the longest dimension of the gain volume. In the case of the pumped nonlinear crystal, there is no fluorescence to be amplified. Yet, with sufficient excitation, a crystal used for optical parametric oscillation may produce a rainbow of colors, as illustrated in Figure 1.36b. Complementary colors for which the sum of the optical frequencies equal that of the pump are generated at angles such that all three waves remain in phase throughout the crystal. This is called an optical parametric generator (OPG). Since there is no fluorescence to be amplified, what is the radiation that is being amplified? We have seen that vacuum itself is filled by a sea of photons (cf. Section 1.1.1). It is this vacuum photon or zero-point energy that seeds the optical parametric generator. In the typical arrangement of Figure 1.36, the pump radiation will be green laser light near 500 nm. The signal will be in a range of 700–1100 nm and the idler at the complementary wavelength, so that the sum of photon energies (idler and signal) equals the pump energy.

There is a remarkable application to the OPG: that of being a source of light that has its own intensity calibration. The intensity of the vacuum radiation is known from theory. The gain along a path of amplified radiation can be measured by sending seed radiation along that same path and measuring the ratio of amplified to initial intensity (Figure 1.36c). The amplified seed radiation is seen as a bright spot within the ring of optical generation of the same color in Figure 1.36c. Absolute calibration of a light source by such an arrangement has been demonstrated with high power pump lasers [33]. If miniaturized and implemented with efficient, compact lasers, it can become a valuable tool for spacecrafts and satellites in need of detector and source calibration that is not dependent on lamp standards that are susceptible to aging.

1.11.3 Too Many Photons to Swallow: Multiphoton Absorption

The phenomena described in the previous sections have a remarkable property: no energy is lost from the light waves. The material acts as a catalyst; exchange two or more photons for a higher energy photon. In harmonic generation, the crystal is a generous social worker, definitely not a capitalist: he does not keep any energy for himself. The sum of the product of the photon energies by their number is conserved in harmonic generation. We have seen in Sections 1.3.2 and 1.5.2 that an atom/molecule with energy levels separated by the photon energy will absorb the photon. The total energy of the radiation-matter system is conserved, but energy has been transferred from the beam to the matter. In another situation, a medium that is transparent to low intensity radiation may become absorbing for high intensity light, if there are some energy levels at twice the photon energy. The mechanism of multiphoton absorption is important for depositing energy inside a transparent medium. Indeed, when focusing an intense pulse inside a material, the intensity increases along the beam to reach a maximum at the focus. Likewise, the absorption will increase to a maximum at the focus. In the case of a two-photon absorber, the beam absorption, or the number of pairs of photons absorbed, is proportional to the *square* of the intensity of the laser radiation. For n-photon absorption, the attenuation is proportional to the nth power of the exciting beam intensity. Radiation may be emitted by fluorescence following multiphoton absorption. This radiation is quite different from the harmonic generation discussed in Section 1.11.1. Harmonic generation of intense radiation at the wavelength λ_0 is instantaneous and at a wavelength that is an exact submultiple of the fundamental $\lambda_n = \lambda_0/n$. Fluorescence following multiphoton absorption is not instantaneous [34], and the radiation that follows the excitation, at a wavelength $\lambda_F < \lambda_0/n$, lasts for a »fluorescence lifetime« in the range of ns to ms. Applications of multiphoton absorption are discussed in Chapter 5.

Ultrashort pulses can typically eject an electron from an atom, in a processus called multiphoton ionization. This phenomenon has an impact on filamentation (negative lensing due to the electron plasma – see Section 1.7.5.3), attosecond pulse generation (see Chapter 6), and plasma shutter.

1.11.3.1 Plasma Shutter – How It Leads to a Better Understanding of Absorption

The plasma shutter deserves a few lines of introduction, as one of the oldest and simplest applications of multiphoton ionization. It is generally used in combination with a TEA laser (see Section 1.4.2.1). The MW peak power of that CO_2 laser is focused in air (or any other gas), creating a plasma of electrons by multiphoton ionization. These electrons are accelerated by the field of the laser, to an energy sufficient to ionize all the other surrounding molecules, a process called cascade ionization. The latter cascade ionization results in a total ionization of the focal volume, or creation of a »plasma« of electrons and ions as dense as the atmosphere. Within this volume, the air has become a perfect conductor and has all the properties of a metal, including the reflectivity of a metallic surface. The air that was initially transparent to the CO_2 laser pulse has suddenly become totally opaque and reflecting, and therefore the pulse emerging from the focal spot has its tail abruptly cut. By creating pulses with an abrupt (picosecond) cut off, the plasma shutter has been an early tool for subnanosecond pulse shaping [35]. The experiment performed with such a plasma shutter, sketched in Figure 1.37, leads to an understanding of absorption as an interference. In Figure 1.37a, a powerful beam of a CO_2 laser is sent through an absorbing cell of sufficient length, so that the complete nanosecond pulse is absorbed and the detector placed at the end of the tube does not observe any signal.

The phenomenon of absorption can be understood in the following way. The photons that are absorbed excite molecular dipoles, at the frequency of the light. We have seen that an oscillating dipole is a source of radiation (see Figure 1.1 at the beginning of this chapter), which

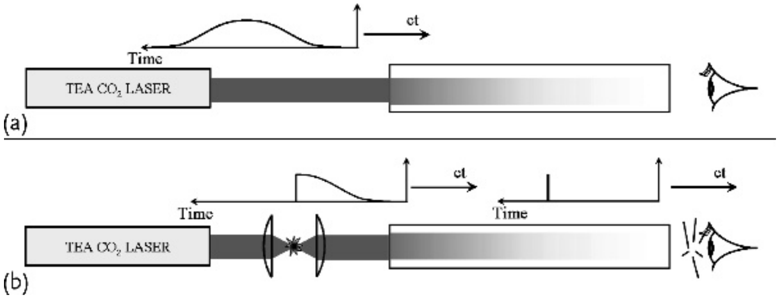


Figure 1.37 Absorption can be interpreted as being due to the incoming light exciting dipoles that radiate with opposite phase. In (a) A CO_2 laser pulse[15)] is sent through a resonant absorber (typically heated CO_2). The excited molecules emit a wave that destructively interferes with the incoming wave, resulting in a null signal on the detector. In (b), a plasma is created in air, which abruptly cuts the tail of the ns pulse. There is no more incoming wave to destructively interfere with the radiation from the excited molecules. An intense pulse is observed on the detector, which has a duration determined by the time that the excited molecules remain in phase. That time is related to the collision time between an excited molecule and any other atom or molecule in the cell. Please find a color version of this figure on the color plates.

is emitted at the same frequency as the exciting radiation, but exactly 180° out of phase. Waves that are out of phase tend to cancel each other. If a number of dipoles equal to the number of photons in the incoming pulse are reradiated out of phase, the incoming pulse is completely destructively interfered. The plasma shutter can provide a dramatic demonstration of this effect. In Figure 1.37b, the plasma abruptly cuts the tail of the strong CO_2 laser pulse, near the peak of its intensity. All the dipoles that were excited by the slowly rising pulse continue to reradiate, but, since there is no more incoming wave to cancel, the detector will be blinded by an intense pulse. This pulse will last as long as these dipoles have not lost their relative phase by collision. The time it takes for dipoles to dephase by collision is inversely proportional to the pressure and of the order of ns torr (depending on the particular molecules involved). For instance, at 20 torr pressure, the pulse created in this manner would have a duration of the order of 50 psr.

15) In the figure, the source is labeled as a TEA laser. In general the bandwidth of this laser is too large compared to that of the absorber, which is a low pressure heated CO_2 cell. Some »tricks« are needed in the actual experiment to make the source a narrow line, such as injecting it with another laser, or inserting a low pressure CO_2 laser in the same resonator.

References

1 Haisma, J. and Spierings, G.A.C.M. (2002) Contact bonding, historical review in a broader scope and comparative outlook. *Mater. Sci. Eng. Rep.*, **37**, 1–60.

2 Haisma, J. (1962) *Construction and properties of short stable gas laser*. PhD Thesis, Rijksuniversiteit te Utrecht, Utrecht, the Netherlands.

3 Strong, C.L. (1974) The amateur scientist. *Sci. Am.*, **230**, 122–127.

4 Snitzer, E. and Osterberg, H. (1961) Observed dielectric waveguide modes in the visible spectrum. *J. Opt. Soc. Am.*, **51**, 499–505.

5 Snitzer, E. (1961) Optical Maser action of Nd^{3+} in barium crown glass. *Phys. Rev. Lett.*, **7**, 444–448.

6 Heard, H.G. (1968) *Laser Parameter Measurements Handbook*, John Wiley & Sons, Inc., New York.

7 Knight, J.C., Birks, T.A., Russell, P.S.J., and Atkin, D.M. (1996) All-silica single-mode optical fiber with photonic crystal cladding. *Opt. Lett.*, **21**, 1547–1549.

8 Kumar, V.V.R.K., George, A.K., Reeves, W.H., Knight, J.C., Russell, P.S.J., Omenetto, F., and Taylor, A.J. (2002) Extruded soft glass photonic crystal fiber for ultrabroad supercontinuum generation. *Opt. Express*, **10**, 1520–1525.

9 Bhagwat, A.R. and Gaeta, A.L. (2008) Nonlinear optics in hollow-core photonic bandgap fibers. *Opt. Express*, **16**, 5036–5047.

10 Slepkov, A.D., Bhagwat, A.R., Venkataraman, V., Londero, P., and Gaeta, A.L. (2008) Generation of large alkali vapor densities inside bare hollow-core photonic bandgap fibers. *Opt. Express*, **16**, 18976–18983.

11 Trum, H.M.G.J. and Diels, J.C. (1970) Quiet turbine for high speed optical applications. *J. Phys. E: Sci. Instrum.*, **3**, 564.

12 McClung, F.J. and Hellwarth, R.W. (1962) Giant optical pulsations from ruby. *J. Appl. Phys.*, **33**, 828–829.

13 Ippen, E.P., Shank, C.V., and Dienes, A. (1972) Passive mode-locking of the CW dye laser. *Appl. Phys. Lett.*, **21**, 348–350.

14 Ruddock, I.S. and Bradley, D.J. (1976) Bandwidth-limited subpicosecond pulse generation in mode-locked CW dye lasers. *Appl. Phys. Lett.*, **29**, 296.

15 Diels, J.C. and Sallaba, H. (1980) Black magic with red dyes. *J. Opt. Soc. Am.*, **70**, 629.

16 Mourou, G.A. and II, T.S. (1982) Generation of pulses shorter than 70 fs with a synchronously pumped CW dye laser. *Opt. Comm.*, **41**, 47–49.

17 Dietel, W., Fontaine, J.J., and Diels, J.C. (1983) Intracavity pulse compression with glass: a new method of generating pulses shorter than 60 femtoseconds. *Opt. Lett.*, **8**, 4–6.

18 Valdmanis, J.A., Fork, R.L., and Gordon, J.P. (1985) Generation of optical pulses as short as 27 fs directly from a laser balancing self-phase modulation, group velocity dispersion, saturable absorption, and saturable gain. *Opt. Lett.*, **10**, 131–133.

19 Diels, J.C., Stryland, E.W.V., and Benedict, G. (1978) Generation and measurements of pulses of 0.2 psec duration. *Opt. Comm.*, **25**, 93.

20 Fork, R.L. and Shank, C.V. (1981) Generation of optical pulses shorter than 0.1 ps by colliding pulse mode-locking. *Appl. Phys. Lett.*, **38**, 671.

21 Shank, C.V. (1982), Progress in ultrashort optical pulse generation,

Paper WE3, presented at the XIIth Int. Quantum Electron. Conf.
22 Asaki, M.T., Huang, C.P., Garvey, D., Zhou, J., Kapteyn, H., and Murnane, M.M. (1993) Generation of 11 fs pulses from a self-mode-locked Ti:sapphire laser. *Opt. Lett.*, **18**, 977–979.
23 Stingl, A., Spielmann, C., Krausz, F., and Szipöcs, R. (1994) Generation of 11 fs pulses from a Ti:sapphire laser without the use of prisms. *Opt. Lett.*, **19**, 204–406.
24 Ell, R., Morgner, U., Kärtner, F.X., Fujimoto, J.G., Ippen, E.P., Scheuer, V., Angelow, G., Tschudi, T., Lederer, M.J., Boiko, A., and Luther-Davis, B. (2001) Generation of 5-fs pulses and octave-spanning spectra directly from a Ti:sapphire laser. *Opt. Lett.*, **26**, 373–375.
25 Russell, J.S. (1844) Report on waves, in *Report of the fourteenth meeting of the British Association for the Advancement of Science*, York, pp. 311–390; plates XLVII–LVII.
26 Bullough, R.K. (1988) The wave, »par excellence«, the solitary, progressive great wave of equilibrium of the fluid – an early history of the solitary wave, in *Solitons* (ed. M. Lakshmanan), Springer Series in Nonlinear Dynamics, pp. 150–281.
27 Coles, S. (2001) *An Introduction to Statistical Modeling of Extreme Values*, Springer-Verlag, London.
28 Solli, D.R., Ropers, C., Koonath, P., and Jalali, B. (2007) Optical rogue waves. *Nature*, **450**, 1054.
29 Kasparian, J., Béjot, P., Wolf, J.P., and Dudley, J.M. (2008) Optical rogue wave statistics in laser filamentation. *Opt. Express*, **17**, 12070–12075.
30 Maine, P., Strickland, D., Bado, P., Pessot, M., and Mourou, G. (1988) Generation of ultrahigh peak power pulses by chirped pulse amplification. *IEEE J. Quantum Electron.*, **QE-24**, 398–403.
31 Arissian, L. and Diels, J.C. (2009) Investigation of carrier to envelope phase and repetition rate – fingerprints of mode-locked laser cavities. *J. Phys. B: At. Mol. Opt. Phys.*, **42**, 183001.
32 Jones, R.J. and Diels, J.C. (2001) Stabilization of femtosecond lasers for optical frequency metrology and direct optical to radio frequency synthesis. *Phys. Rev. Lett.*, **86**, 3288–3291.
33 Migdall, A., Datla, R., Sergienko, A., Orszak, J.S., and Shih, Y.H. (1998) Measuring absolute infrared spectral radiance with correlated visible photons: technique verification and measurement uncertainty. *Appl. Opt.*, **37**, 3455–3463.
34 Rudolph, W. and Diels, J.C. (1986) Femtosecond time resolved fluorescence. *Picosecond Phenomena V* (ed. A.E. Siegman), Springer-Verlag, Berlin, pp. 71–74.
35 Yablonovitch, E. (1974) Self-phase modulation and short pulse generation from laser-breakdown plasmas. *Phys. Rev. A*, **10**, 1888–1895.