

Chapter 3

Lasers principles

In this chapter and the two following ones we shall describe the principle of the operation of lasers, their common features and the properties of the light they emit. Our aim is not to provide an exhaustive catalogue of the types of laser available at the time of writing. Such an account would, in any case, soon be obsolete. Rather, we shall use concrete examples of existing systems to illustrate important features or general principles. We restrict ourselves in this chapter to a rather simplified approach of main features of lasers. Most of our considerations apply as well to continuous lasers (cw lasers) as to pulsed lasers. But in this chapter we will consider mostly the laser steady-state behaviour. The next chapter will deal with a more quantitative development of the laser equations and their steady-state solutions, while chapter 5 will be devoted to laser dynamics and noise, and pulsed laser operation. For more advanced or specialized topics, we refer the reader to more specialized handbooks (see the bibliography at the end of Chapter 5).

The physical principles accounting for laser operation can appear quite straightforward. It is interesting to note, however, how painstaking our progress in understanding the principles that have led to lasers has been.¹ It is usually considered that the prehistory of the laser commenced in 1917 when Einstein introduced the notion of *stimulated emission*. In fact, Einstein was led to the conclusion that such a phenomenon must occur from considerations of the thermodynamic equilibrium of the radiation field and a sample of atoms at a finite temperature T . He found that the *Planck formula* for the *black-body* energy distribution could only be derived if, in addition to *absorption* and *spontaneous emission*, stimulated emission was assumed to occur. It is also worth noting that Einstein does not mention the fact that stimulated emission can lead to coherent amplification of light. Much work, both experimental and theoretical, was prompted by the Einstein's

¹For more information on the development of our understanding of lasers see M. Bertolotti, *Masers and Lasers, An Historical Approach*, Hilger (1983), C.H. Townes in *Centennial Papers*, I.E.E.E. Journal of Quantum Electronics, **20**, 545-615 (1984) or J. Hecht, *Laser Pioneers*, Academic Press (1992).

prediction of stimulated emission. This came to fruition in 1928 with the demonstration of the phenomenon in a neon discharge by Ladenburg and Kopferman.² Curiously, interest in stimulated emission declined thereafter, as few applications were foreseen in view of the belief that was widely held at the time that no system could be prepared in a state sufficiently far from *thermal equilibrium* that significant optical gain could be obtained. Furthermore, few physicists were sufficiently familiar with electronics to be aware of the possibility of using *feedback* to maintain oscillation in a system exhibiting gain. This meant that the only conceivable applications seemed to demand a gain much larger than one, which was believed to be unobtainable.

It wasn't until 1954, when Townes, Gordon and Zeiger demonstrated the ammonium maser, that the situation changed. Townes' contribution was to realize that if the amplifying medium were placed in a *resonant cavity*, oscillation could occur, even if the single-pass gain was small, provided that the gain was sufficient to compensate for the (small) losses of the cavity.³

The generalization of the *maser* ("microwave amplification by stimulated emission of radiation") in the microwave domain to the *laser* ("light amplification by stimulated emission of radiation") in the optical domain, was not immediate and gave rise to a number of controversies. It wasn't until 1958 that Schawlow and Townes published their suggestion that, in the optical domain, *two mirrors facing each other* could perform the function of the resonant cavity. Their article prompted furious activity amongst experimentalists which resulted in the realization of the first laser by Maiman in 1960. This device used ruby as the active medium, and settled a dispute with other physicists who believed they had demonstrated that laser action was not possible in such a material. As we shall show later, the gain mechanism in ruby is in some respects unique. The advent of the ruby laser was closely followed by demonstrations of laser action in a helium-neon gas mixture, by Javan, and in carbon dioxide by Patel. Amongst the striking advances that have followed have been the demonstration of high power lasers such as those employing gain media of neodymium-doped glass or of doped fibres, and tunable systems employing dyes or solid-state materials such as titanium-doped sapphire. The semiconductor revolution has provided lasers remarkable for their miniaturization, efficiency and price, which has enabled laser technology to become very widely spread, even in mass produced consumer products (such as CDs, DVDs, bar-code readers, printers...). The

²These authors did not, in fact, observe the amplification of an incident wave, but the modification of the refractive-index of the gas in the vicinity of the frequency of a resonance as a function of the difference in population of the two levels involved in the transition, which constitutes an indirect proof of the existence of stimulated emission (see section 2.5).

³The maser was discovered independently by the Russian physicists Basov and Prokhorov.

development of high-power fibre lasers has revolutionized the domain of industrial applications of lasers, such as cutting, welding, marking...

It was shown in chapter 2 that it is possible, for a medium in which a population inversion has been created, to amplify an incident light wave. In section 3.1 of this chapter we show how a light amplifier can be turned into a light *generator* when the amplifying medium is placed in a cavity which reinjects the amplified light into the gain medium and we study the influence of the cavity on the emitted light. Based on the important concept of *laser mode*, we then describe in section 3.2 the steady-state behaviour of the laser, and, in section 3.3, the spectral properties of the light it emits. We show that, depending on the characteristics of the cavity and the spectral broadening mechanism of the gain curve, the laser can have either a well-defined frequency (single-mode operation) or can emit several frequencies simultaneously (multi-mode operation). In the latter case, we show how it is possible to force single frequency output and give a first estimate of the degree of spectral purity thus obtained. In the next sections of the chapter we present some general techniques for obtaining population inversion, in order to achieve optical gain. We consider general schemes for three- and four-level systems and give, as concrete examples of each, descriptions of laser systems having played an important role in the field.

Complement 3A provides an overview of the properties of Fabry-Perot cavities, which were an important tool in spectroscopy before they were employed to provide feedback in laser systems. Whilst in this chapter and in Complement 3A we rely upon a plane-wave description of the light field, in Complement 3B we consider the non-uniform transverse intensity distribution in a real laser beam. We consider in particular the case when this corresponds to the fundamental transverse mode of the laser cavity, which has a Gaussian intensity distribution. We discuss also the *spatial coherence* of the emitted laser light. Complement 3C gives elements of *energetic photometry* of light sources which are essential for a deep understanding of the difference between incoherent light and laser light.

3.1 Laser oscillation. Threshold

We give in this introductory section a first simple account of laser oscillation, based on energy considerations. We will take full account of the *coherence* property of stimulated emission in the following chapters.

Lasers, like electronic oscillators, rely on the application of *positive feedback* to an amplifier for their operation. For a laser, the amplifier is the inverted atomic or molecular medium, in which optical gain is obtained (see Section 2.5). In a single pass through the amplifying medium, the optical intensity is increased by a factor $G = \exp gL_A$, known as the single-pass amplification coefficient, where L_A is the length of the amplifying zone and g the

gain per unit length, assumed here uniform inside the amplifying medium and given, according to the simplified model of Chapter 2, by Equation (2.101). The necessary feedback is realized by the mirrors which define the *laser cavity*. The laser cavity of Figure 3.1 is formed by three mirrors which cause a light ray to circulate in a closed triangular path, such that the ray coincides with itself after one round-trip. We assume that two of these mirrors are totally reflecting and that the third, the output mirror, has transmission and reflection coefficients R and T respectively, with $R + T = 1$.

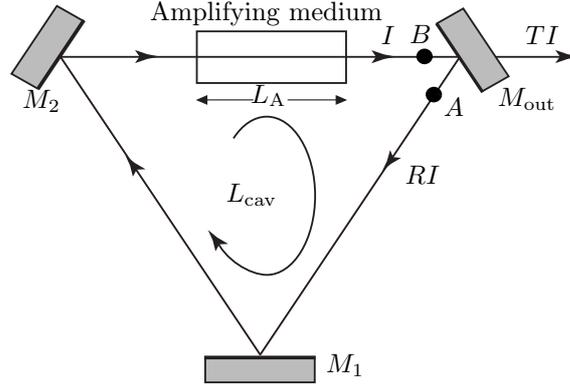


Figure 3.1: Ring-cavity laser. The mirror arrangement is such that the ray coincides with itself after a round-trip. The partially transmitting output mirror M_{out} reflects a part of the incident radiation back into the amplifier via the totally reflecting mirrors M_1 and M_2 . If the gain of the amplifying medium is sufficient to overcome the losses, a circulating light wave is established in the cavity.

Consider a light wave circulating in the cavity from a point A just after the output mirror (see Figure 3.1). If the intensity of the wave at this point is I_A , well below the saturation intensity I_{sat} of the amplifier (see Section 2.3.4) then its intensity at point B just before it strikes M_{out} on its return from one pass around the cavity is

$$I_B = G^{(0)} I_A, \quad (3.1)$$

where $G^{(0)}$ is the unsaturated amplification coefficient per pass in the amplifier:

$$G^{(0)} = \exp \left[g^{(0)} L_A \right], \quad (3.2)$$

$g^{(0)}$ being the unsaturated gain per unit length (Section 2.5.2). After reflection on M_{out} the intensity is reduced to $R G^{(0)} I_A$.

In order for oscillation to occur, the intensity of the beam after one cavity round-trip must exceed that with which it started out. For a weak input wave the *threshold condition* is given by

$$G^{(0)} R = G^{(0)} (1 - T) \geq 1. \quad (3.3)$$

In practice, it is necessary to consider, as well as the loss by transmission at the output mirror, other losses arising, for example, by absorption or light scattering, either within the cavity or on the mirror surfaces. We shall denote the sum of all these losses by an absorption coefficient A so that condition (3.3) becomes

$$G^{(0)}(1 - T)(1 - A) \geq 1 . \quad (3.4)$$

In the small loss and small gain limit $A, T, g^{(0)}L_A \ll 1$, often encountered in practice, this threshold condition becomes

$$G^{(0)} - 1 \simeq g^{(0)}L_A \geq T + A . \quad (3.5)$$

This condition simply expresses the fact that the unsaturated gain of the amplifier must be *greater than the total losses* in order for oscillation to commence. It is important to realize that *no incident beam* is necessary for this to occur; oscillation is initiated by noise, generally by photons spontaneously emitted by the gain region.

3.2 Steady-state intensity

3.2.1 Laser mode

The preceding section has allowed us to understand how oscillation starts from amplification of an initially spontaneously emitted photon. After a transient regime, the laser will eventually reach a steady-state regime. This means that if one looks at the cavity of Figure 3.1, all the characteristics of the field at point A , namely its intensity, phase, frequency, polarization, and transverse distribution, will no longer evolve with time. Since light propagates inside the cavity, steady-state means that the field has the same characteristics after one round-trip inside the cavity.

This identity of the laser field after one round-trip inside the cavity is the definition of a *laser mode*. It is the spatial distribution, amplitude, frequency, phase, and polarization of the field which is identically reproduced after each round-trip.

Regarding the spatial distribution of light in particular, the light beam which oscillates in the laser cavity cannot be treated as a light ray with vanishing transverse extension, as sketched in Figure 3.1. It also cannot be a traveling plane wave of infinite extension, as discussed in Section 2.2. When the transverse extension of the cavity mirrors is much larger than the light wavelength, however, we can use a toy model (the so-called “top-hat mode” model) where the beam is taken as a *cylindrically truncated traveling plane wave* of cross-sectional area S . Then all the notations of Section 2.2 apply, where z is the propagation direction following the intra-cavity beam of Figure 3.1. A more realistic description of laser beams is given in Complement

3B, but the top-hat model will allow us to understand the basic operating principles of the laser.

Let us remark that usually, there are several possible laser modes for a given configuration, that may or may not oscillate together (see section 3.3.2).

3.2.2 Steady-state intensity

a. Laser start-up

Once the threshold condition (Equation 3.4) is achieved, the gain per cavity round-trip exceeds the losses. This means that the net gain per round-trip is positive. The laser oscillation starts from an initial photon spontaneously emitted in the direction which is recycled by the cavity. At each round-trip, the light intensity is amplified by a factor $G(I)(1 - T)(1 - A)$. At the beginning, when I is small, the amplification coefficient $G(I)$ is not saturated and the product $G(I)(1 - T)(1 - A)$ is larger than 1, and I increases round-trip after round-trip.

The power circulating in the laser cavity cannot continue to increase for each round-trip up to infinity. In fact, the amplification coefficient $G(I)$ is usually a decreasing function of the intensity of the circulating wave because of saturation effects (see Section 2.3.4), as illustrated in Figure 3.2. One can see on this Figure that the gain decreases when I increases, till I reaches its steady-state value I_{ss} , which fulfills the following condition

$$G(I_{ss})(1 - T)(1 - A) = 1 . \quad (3.6)$$

The value I_{ss} is the value of the intensity for which *the gain exactly compensates the losses*.⁴

b. Steady-state intensity

When $I = I_{ss}$, the gain saturation is such that the laser intensity is left unchanged after one round-trip inside the cavity. Although the stability of this steady-state solution will be proved in Section 4.2, it is already easily visible from Figure 3.2. Indeed, if we imagine that I becomes slightly smaller than I_{ss} , Figure 3.2 shows that the saturated gain becomes larger than the losses and the intra-cavity intensity increases, while if at some point $I > I_{ss}$, saturation increases and the gain becomes smaller than the losses, leading to a decrease of I at each round-trip. This shows that above threshold, $I = I_{ss}$ is a *stable steady-state solution* for the laser evolution. There is usually only one such stable solution because $G(I)$ is a monotonic function of I .

⁴The transient evolution followed by the laser intensity, starting from a single spontaneously emitted photon, to reach the steady-state intensity I_{ss} , will be studied in Section 5.1.

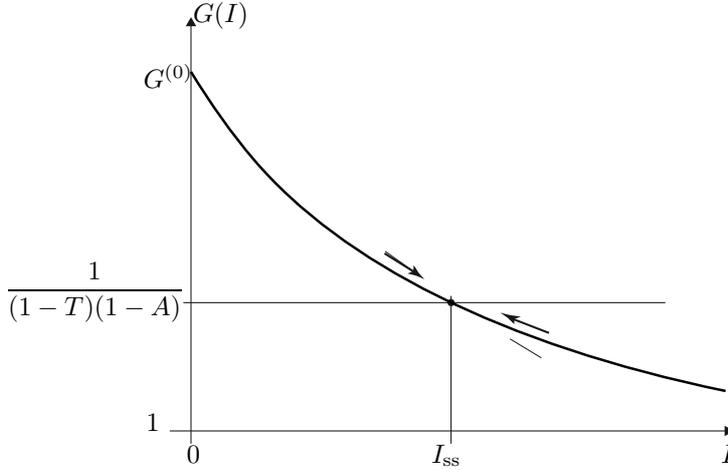


Figure 3.2: Amplification coefficient versus intensity circulating in the cavity. The steady state value I_{ss} is reached when the gain exactly compensates for the losses per round-trip inside the cavity.

We can give an explicit expression of I_{ss} in the case where we suppose that the gain and losses are small, i.e.

$$G(I) = \exp \left[g(I) L_A \right] \simeq 1 + g(I) L_A , \quad (3.7)$$

$$(1 - T)(1 - A) \simeq 1 - A - T . \quad (3.8)$$

Injecting equations (3.7) and (3.8) into equation (3.6), we obtain

$$g(I_{ss}) L_A = A + T , \quad (3.9)$$

which shows that, in steady-state regime, the round-trip gain exactly compensates the round-trip losses.

Using the simple model of Chapter 2 and assuming that the laser frequency is close to resonance ($\omega \simeq \omega_0$), equation (2.101) leads to the following saturation behaviour for the gain:

$$g(I) = \frac{g^{(0)}}{1 + \frac{I}{I_{\text{sat}}}} , \quad (3.10)$$

where $g^{(0)}$ is the unsaturated gain.

Injecting equation (3.10) into (3.9), we obtain the following expression of the laser intra-cavity intensity above threshold:

$$I_{ss} = I_{\text{sat}} \left(\frac{g^{(0)} L_A}{A + T} - 1 \right) = I_{\text{sat}} (r - 1) , \quad (3.11)$$

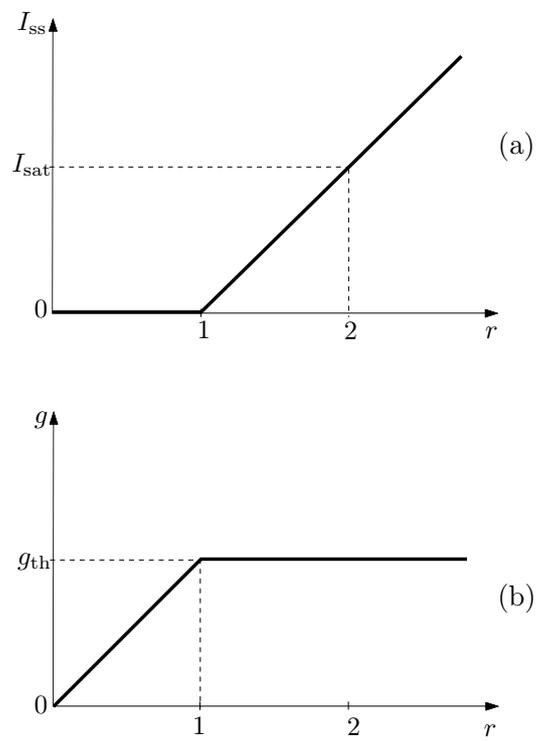


Figure 3.3: Evolution of (a) the laser intensity and (b) the laser gain versus excitation ratio r .

where we have introduced the excitation ratio r of the laser

$$r = \frac{g^{(0)} L_A}{A + T}, \quad (3.12)$$

i. e., the ratio of the unsaturated gain per round-trip to the losses per round-trip. Figure 3.3 reproduces the evolution of the laser intensity and the gain versus r . The threshold corresponds to $r = 1$, for which the unsaturated gain is equal to

$$g_{\text{th}} = \frac{A + T}{L_A}. \quad (3.13)$$

Below threshold ($r < 1$), the laser is off ($I_{\text{ss}} = 0$) and the gain is equal to the unsaturated gain. Above threshold ($r > 1$), the saturated gain $g(I_{\text{ss}})$ remains equal to the gain at threshold given by Equation (3.9). The fact that above threshold the gain is clamped at its threshold value is another consequence of the fact that, in steady-state regime, the gain exactly compensates the losses.

Many lasers employ linear as opposed to ring cavities (see Figure 3.4). If one considers such a cavity composed of a totally reflecting mirror, M_1 , and a partially transmitting mirror, M_{out} , two passes of the amplifying medium will occur for each reflection on M_{out} (one in the direction $M_s M_1$ and one on the return trip) so that condition (3.6) has to be replaced by

$$G^2(I_{\text{ss}})(1 - T)(1 - A) = 1, \quad (3.14)$$

where A corresponds to all the losses over one round-trip inside the cavity, except the output mirror transmission T . Moreover, in this case, the equivalent of the round-trip cavity length L_{cav} is given by twice the distance between the two mirrors:

$$L_{\text{cav}} = 2L_0. \quad (3.15)$$

Actually, the situation is more complicated, because the two waves traveling in opposite directions give rise to a standing wave inside the laser medium. The resulting spatial variation of the gain (due to the intensity variation) is responsible for subtle effects in such linear cavity lasers, as we will see in Section 3.3.3.

Comments

(i) The gain in the amplifying medium is due to the difference between the rates of stimulated emission and absorption along the laser transition. In expressions (3.5) and (3.7) (as in all the results of Sections 2.4 and 2.5), the decrease in intensity due to absorption *in the amplifying medium*, is responsible for saturation, and is included in $G(I)$, the net value of the amplification coefficient.

(ii) In the ring cavity of Figure 3.1 we have supposed that light circulates in the cavity in a given sense (here in the direction $M_{\text{out}} M_1 M_2$). In the general case, propagation in the

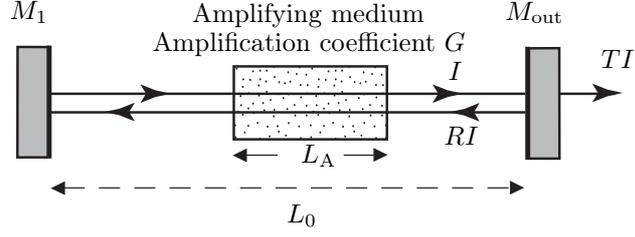


Figure 3.4: Linear cavity laser. In a round-trip, the beam passes twice in the amplifier. The round-trip length is $L_{\text{cav}} = 2L_0$.

other sense is equally possible, so that oscillation can take place in both senses. However, it is usual to insert elements in the cavity which induce different losses for the two propagation directions, and favors one sense, which alone oscillates.

(iii) It is easy to envisage ring laser cavities having more than three mirrors. In fact, it is usual to employ four.

3.3 Oscillation frequency

3.3.1 Longitudinal mode

In steady-state regime, as discussed in section 3.2.1, the intra-cavity field must be left unchanged after one round trip inside the cavity. If we approximate the field by a plane wave of frequency ω and wave number k , and assume for simplicity that the gain medium has an index that does not depend on the light intensity, this means that the round trip phase

$$\psi = k L_{\text{cav}} , \quad (3.16)$$

where L_{cav} is the optical length of the cavity, must be equal to an integer times 2π :

$$\psi = p 2\pi , \quad (3.17)$$

with p an integer. Consequently, the steady state laser can oscillate only at discrete values k_p of the wavenumber

$$k_p = p \frac{2\pi}{L_{\text{cav}}} , \quad (3.18)$$

corresponding to discrete values of the frequency ⁵

$$\omega_p = c k_p = p \Omega_{\text{cav}} , \quad (3.19)$$

⁵The quantity ω_p in Equation (3.19) is actually the angular frequency. However, we call it the frequency when there is no ambiguity.

where we have introduced the *free spectral range of the cavity*

$$\Omega_{\text{cav}} = 2\pi \frac{c}{L_{\text{cav}}} . \quad (3.20)$$

The corresponding wavelength λ_p satisfies

$$L_{\text{cav}} = p \frac{2\pi}{k_p} = p \lambda_p . \quad (3.21)$$

Condition (3.21) states that the optical length of the cavity should be an *integer multiple of the wavelength*. This multiple can be very large (of the order of 10^6) if the laser cavity length is of order 1 m whereas wavelengths are of the order of a micron. It can also be of the order of a few units in microcavity lasers for which L_{cav} is a few microns (see Figure 3.30).

Comment

It is important to realize the distinction between the absolute phase of the field, which can have any value, and the phase accumulated over one round-trip, which is subject to the constraint (3.17), or its time derivative, the frequency, which is subject to the constraint (3.19).

The unsaturated amplification coefficient, $G^{(0)} = \exp(g^{(0)} L_A)$, is a function of frequency, which usually has a bell-shaped form, centered on some frequency ω_M , as shown in Figure 3.5. (For the case of the simple model of Chapter 2, the explicit form of this curve can be deduced from Equations 2.84 and 2.102). If we assume that the maximum gain, which occurs when $\omega = \omega_M$, is larger than $1 + T + A$ then there will be a limited band of frequencies $[\omega', \omega'']$ for which the threshold condition (3.4) is satisfied (see Figure 3.5).

In fact, the laser does not emit light over the whole range of frequencies $[\omega', \omega'']$. This is because the frequencies at which the laser oscillates must satisfy the condition (3.19) (see the thin vertical lines in Figure 3.5). These discrete frequencies ω_p are associated with different *longitudinal modes* of the cavity corresponding to different values of the integer p . The frequency separation of these modes is equal to $\Omega_{\text{cav}}/2\pi = c/L_{\text{cav}}$ (which has a value typically of 5×10^8 Hz, for a cavity of length 60 cm, and which is much smaller than the optical frequency, of the order of a few 10^{14} Hz). The inverse of $\Omega_{\text{cav}}/2\pi$ is simply the time it takes for light to travel one cavity round-trip. This free spectral range is given either by Ω_{cav} (see Equation 3.20) or $\Omega_{\text{cav}}/2\pi$, depending on whether one is considering angular frequencies or simply frequencies, respectively.

The number of longitudinal modes for which (3.4) is satisfied can vary between 1 and 10^5 , depending on the nature of the gain medium. In the next section, we discuss the possibility of the coexistence of, or competition between, simultaneously oscillating modes.

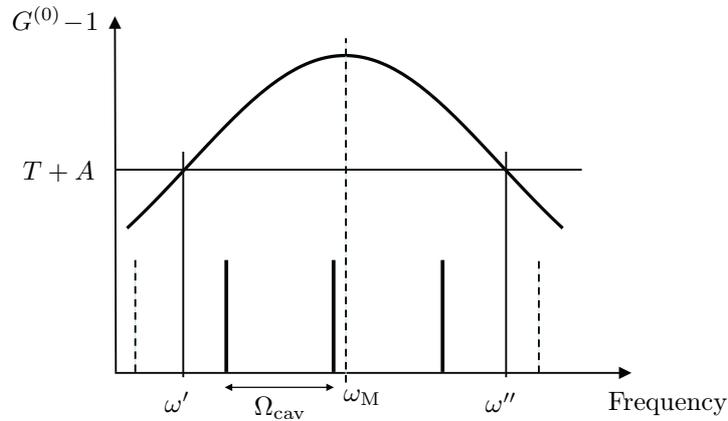


Figure 3.5: The thick line is the gain-frequency curve for a typical laser transition. The laser can only operate over the frequency range $[\omega', \omega'']$ for which the gain exceeds the losses. The vertical lines represent the frequencies of the longitudinal cavity modes falling within this frequency range.

3.3.2 Number of oscillating modes: homogeneous vs. inhomogeneous broadening

We showed in the preceding section that the oscillation condition (that the unsaturated gain should surpass the losses) could be satisfied for several longitudinal modes of the laser cavity (Figure 3.5). The number of modes on which the laser can oscillate is then of the order of the ratio of the width of the gain curve to the frequency interval c/L_{cav} between successive modes. For example, for the red line of the Helium-Neon laser at $0.633 \mu\text{m}$, the width of the gain curve is 1.2 GHz whilst for a typical round-trip cavity length of 0.6 m, the longitudinal mode spacing is 0.5 GHz. The laser therefore oscillates on two or three modes. It should be realized, however, that whilst the position of the gain curve is fixed in frequency, the comb of longitudinal cavity modes is displaced when the cavity length changes (for example as a result of thermal expansion of the mechanical structure holding the mirrors). Thus the laser, which might oscillate at a given instant on two modes might oscillate on three moments later.

This ability of the helium-neon laser to oscillate simultaneously on all the longitudinal modes lying within the range of its gain curve is by no means universal. For other lasers, such as the Nd:YAG laser or the Ti:Sapphire laser, fewer modes oscillate than one might expect. The difference in behaviour is directly related to the nature of the mechanism responsible for the broadening of the gain curve. In particular, whilst the gain curve of the helium-neon laser is *inhomogeneously broadened*, that of the Nd:YAG system is essentially *homogeneously broadened*.

A spectral line is homogeneously broadened if all the active centers

(atoms, ions, molecules) have the same spectral response. In contrast, a spectral line is said to be inhomogeneously broadened if the appearance of different frequencies in the response of the global gain medium is associated with active centers that have different spectral responses. For instance, Neon atoms with different velocities have different resonance frequencies, in the laboratory frame, because of the Doppler effect. Thus an inhomogeneously broadened curve is simply a juxtaposition of a number of homogeneous line-shapes centred on different frequencies that reflect the different responses of the various active centers (see Figure 3.6).

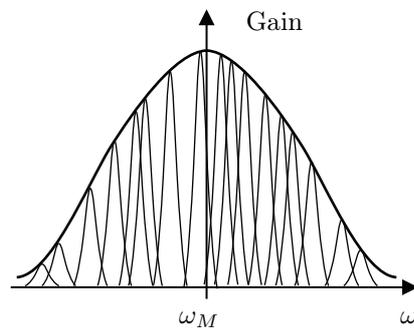


Figure 3.6: Inhomogeneously broadened gain profile. The total gain curve is the sum of individual gain curves corresponding to different atoms exhibiting different resonance frequencies.

The distinction between homogeneous and inhomogeneous broadening is crucial in determining the number of modes that oscillate. Recall that the steady state operation of a laser arises because gain saturation reduces the gain until it is precisely equal to the losses from the cavity (see Section 3.2.2). In the case of inhomogeneous broadening (Figure 3.6) the *gain saturation* (that is, the decrease in the gain as the intensity increases) only affects those atoms (or molecules) of which the resonant frequency is that of the light field (see Figure 3.7). The laser modes burn holes in the gain profile (so-called *spectral hole burning*) and the laser oscillation has no effect on the gain appearing at other frequencies. In this case oscillation can occur on all the cavity modes at the frequencies of which the gain is above the threshold value. This is the case with the Helium-Neon laser⁶, where inhomogeneous broadening is dominant because of the Doppler effect due to the motion of the atoms.

Consider now the case of a laser for which the gain curve reflects a homogeneous broadening of the laser transition (Figure 3.8). In this case the response of each atom is affected in the same manner by the gain saturation.

⁶At least when the gas pressure is low enough and the free spectral range large enough for the modes to interact with distinct velocity classes.

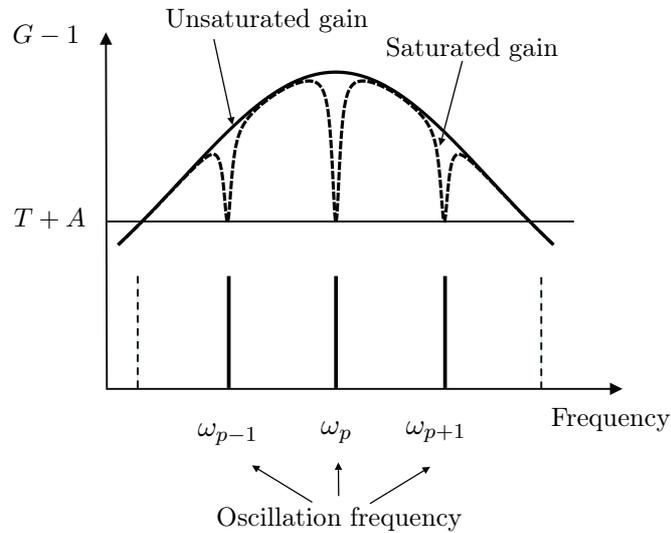


Figure 3.7: Effect of saturation on an inhomogeneously broadened gain curve: holes are “burnt” in the gain curve only at the laser oscillation frequencies. This allows several modes to oscillate simultaneously without competing against each other.

This leads therefore to the entire gain curve being lowered as the intensity in the laser cavity increases, so that, finally, oscillation occurs on a single longitudinal mode: the one lying closest in frequency to the maximum of the gain curve. Oscillation on this mode prevents laser emission on the other modes. This phenomenon is called “mode competition”.

Let us illustrate this phenomenon in the case where only two modes can oscillate, as sketched in Figure 3.9. These two modes have frequencies ω_p and ω_{p+1} . Figure 3.9 shows what happens when one progressively increases the unsaturated gain $G^{(0)}(\omega) - 1$. As long as it is smaller than the losses for all the cavity frequencies ω_p , the laser is below threshold and does not oscillate (Figure 3.9(a)). The laser reaches threshold when the unsaturated gain of the mode of higher gain is equal to the losses (Figure 3.9(b)). For stronger excitation ratios (Figure 3.9(c)), the gain is saturated by the oscillating mode and becomes smaller than the losses for all the other modes, which cannot oscillate although they would have been able to oscillate in the absence of saturation. The laser is thus always single-frequency and oscillates on the mode which has the stronger unsaturated gain.

In practice, it is unusual for the broadening of a laser gain curve to be purely homogeneous or inhomogeneous in nature. For the continuous lasers mentioned in Table 3.1, one of these types of broadening is dominant

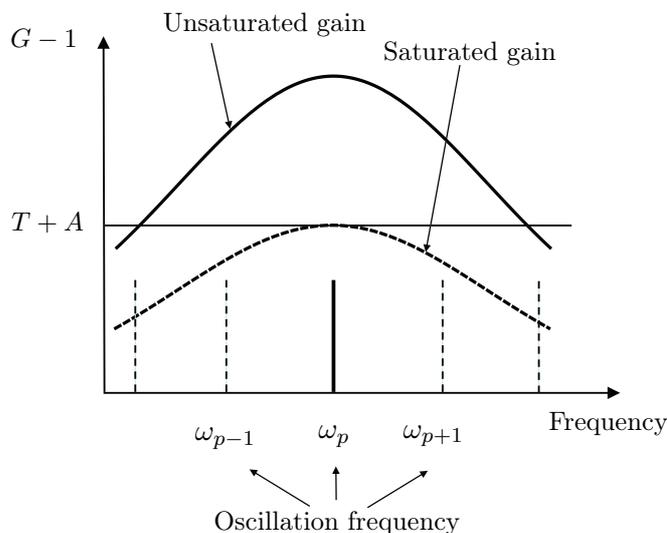


Figure 3.8: Effect of gain saturation in the case of a homogeneously broadened gain curve. The entire gain curve is depressed by the phenomenon of gain saturation, so that a single mode, that nearest in frequency to the maximum of the gain curve, can oscillate.

as indicated.⁷ However, very often the homogeneous and inhomogeneous widths can have comparable values, and the situation is not as clear cut as suggested by the simple reasoning sketched on Figures 3.7, 3.8, and 3.9.

Comment

Some systems can have a homogeneous or inhomogeneous broadening, depending on some parameter as the pressure in a gas laser. For example, the linewidth of the transition at $10.6\ \mu\text{m}$ in CO_2 gas at moderate pressure is extremely small (less than 1 MHz), but the thermal motion of the molecules of the gas leads to a Doppler width which is considerably larger (50 MHz at room temperature). This Doppler broadening is inhomogeneous, because their different velocities endow molecules with distinct resonance frequencies in the laboratory frame. In high pressure CO_2 laser systems, however, collisional broadening dominates the spectral width of the laser emission. This broadening is homogeneous since the average effect of collisions is the same for all molecules and their transition frequencies are affected in an identical manner.⁸

⁷Continuous laser emission is often referred to as cw laser emission, in opposition to pulsed emission. cw holds for “continuous-wave”.

⁸Homogeneous and inhomogeneous line broadening mechanisms are contrasted also in Complement 5B devoted to laser spectroscopy.

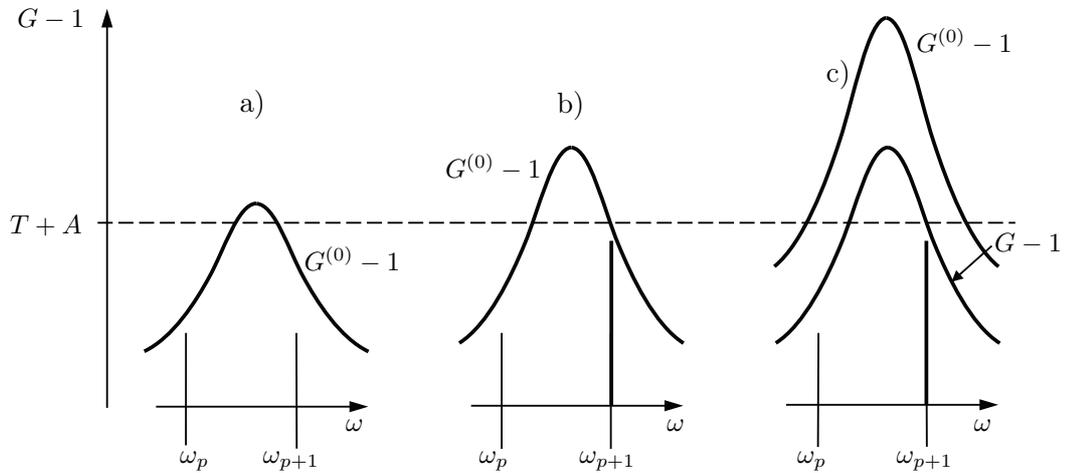


Figure 3.9: Mode competition in a laser based on a homogeneously broadened gain medium. Unsaturated gain $G^{(0)} - 1$ versus frequency. (a) The gain is larger than the losses at frequency ω_0 but smaller than the losses at the cavity mode frequencies: the laser cannot oscillate. (b) Mode $p + 1$ reaches threshold and starts to oscillate. (c) The two modes p and $p + 1$ are above threshold. However, once mode $p + 1$ oscillates, its saturated gain $G(\omega_p) - 1$ is equal to the losses and $G(\omega_{p+1}) - 1$ becomes smaller than the losses and mode p cannot oscillate.

3.3.3 Single frequency operation

For many applications (such as spectroscopy, holography, metrology⁹) it is desirable that the laser output should be as monochromatic as possible. It is therefore often necessary to ensure that a laser capable of oscillating on many modes simultaneously is constrained to oscillate on a single longitudinal mode (single-mode operation). Even in the case of dominant homogeneous broadening, it is usually necessary to take steps to enforce single-mode operation. A straightforward method consists of employing a laser cavity so short that the free spectral range $\Omega_{\text{cav}}/2\pi$ is greater than the width of the gain curve. Such a technique works for short helium-neon lasers ($\Omega_{\text{cav}}/2\pi = c/L_{\text{cav}} = 1$ GHz for $L_{\text{cav}} = 0.3$ m) and for some semiconductor lasers ($\Omega_{\text{cav}}/2\pi = 1000$ GHz). However, this method is not always applicable.

The simplest manner of ensuring that a laser operates on a single longitudinal mode is to introduce into the cavity a frequency-selective filter, of which the transmission profile is narrower than the gain curve and which ensures that the net gain coefficient (the product of the gain coefficient in the absence of the filter by the transmission of the filter) is above thresh-

⁹See Complement 5B.

	Gain linewidth	Typical free spectral range $\Omega_{\text{cav}}/2\pi$	Number of possible longitudinal modes	Dominating broadening mechanism
He - Ne	1.2 GHz	500 MHz	3	Inhomogeneous (Doppler)
CO ₂ (high pressure)	0.5 GHz	100 MHz	5	Homogeneous
Neodymium : YAG	120 GHz	300 MHz	400	Homogeneous
Dye (Rhodamine 6G)	25 THz	250 MHz	10^5	Homogeneous
Titanium doped Sapphire	100 THz	250 MHz	4×10^5	Homogeneous
Erbium: Glass	1 THz	10 MHz	10^5	Depends on glass and temperature
Semi-conductor	1 THz	100 GHz	10	Homogeneous

Table 3.1: Characteristics of some continuous lasers including the approximate number of modes that can possibly oscillate

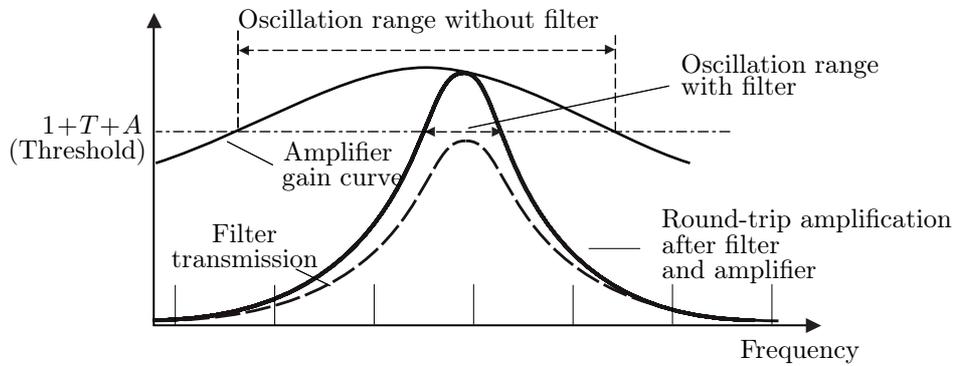


Figure 3.10: Laser frequency selection by an optical filter. The transmission curve of the filter (dashed line) is narrower than the gain curve of the amplifying medium and approaches unity at its peak. The round-trip gain of the laser at a given frequency is the product of the gain in the absence of the filter and the transmission of the filter at that frequency. Oscillation can occur when this product exceeds one. The thin vertical lines hold for the longitudinal mode frequencies.

old for a single cavity longitudinal mode only (Figure 3.10). In order to avoid introducing large losses, the filter must have its maximum transmis-

sion tuned exactly on the mode frequency. This is usually ensured by means of a servo-loop.

It is even often necessary, in order to achieve the required degree of frequency selectivity, to employ a filter composed of a number of separate elements with successively narrower transmission peaks (all of which must be accurately set to the same wavelength). For example, a single-mode Titanium-Sapphire laser includes a triple birefringent filter, which selects a frequency band 300 GHz wide, followed by a “thin” Fabry-Perot étalon (1 mm thick), which has a transmission curve 3 GHz wide and a “thick” étalon (5 mm thick), which selects a single cavity mode. Such a system necessitates the use of nested servo-locking mechanisms that can be difficult to set up.

It is generally more difficult to make a linear cavity laser work on a single longitudinal mode than one with a ring cavity. In the case of a linear cavity the intra-cavity light field is a standing wave, which comprises an alternating sequence of nodes and anti-nodes (see Figure 3.11). At the anti-nodes gain saturation occurs and diminishes the gain (so-called *spatial hole burning*). However, a given oscillating mode does not reduce the gain in the region of its nodes, where its intensity is low. This can lead to the oscillation of a second mode for which the anti-nodes are *spatially separated* from those of the first. In the case of the amplifying medium occupying the central region of a linear cavity, two adjacent modes ω_p and ω_{p+1} satisfy precisely this requirement (Figure 3.12), which explains the tendency of such systems to oscillate on several modes in the absence of frequency selective elements.

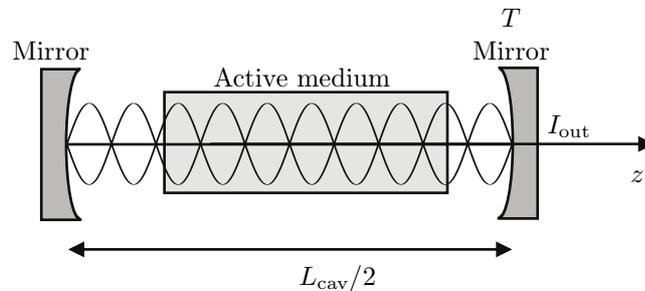


Figure 3.11: Laser based on a linear cavity sustaining standing waves.

The spatial gain modulation described above does not occur for a ring-cavity laser provided light is forced to circulate in a single direction; the intra-cavity light field is then a travelling wave, which has a spatially uniform intensity. This explains why a dye or a Titanium:Sapphire laser (which are essentially homogeneously broadened) tend to operate spontaneously single-mode with a ring cavity, but multi-mode with a linear cavity.

Once single-mode operation has been achieved, it is often desirable to be able to control the output frequency to a precision far exceeding the inter-

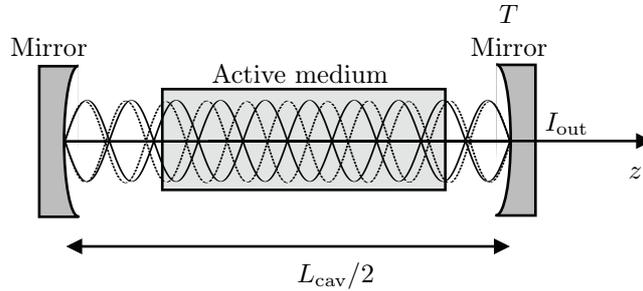


Figure 3.12: Standing wave structure of two successive longitudinal modes.

mode separation c/L_{cav} . The frequency of the p^{th} mode is given by (3.10). It is a function of the cavity length L_{cav} and a change in frequency equal to the inter-mode spacing is obtained for a change of L_{cav} equal to the optical wavelength:

$$\delta\omega = \Omega_{\text{cav}} = 2\pi \frac{c}{L_{\text{cav}}} \Rightarrow \delta L_{\text{cav}} = \frac{L_{\text{cav}}}{p} = \lambda_p . \quad (3.22)$$

Therefore, in order to set the laser frequency to a precise value, with an accuracy better than the cavity free-spectral range it is necessary to control the cavity length to better than one optical wavelength. This can be achieved by mounting one of the cavity mirrors on a piezoelectric transducer that allows the control of the cavity length and lock the laser frequency to an external reference, which is usually a narrow atomic or molecular resonance. In the best case, and with a lot of technological refinements, it has become possible to stabilize the frequency variation of lasers down to a value of the order of 1 Hz, which corresponds to a relative stability of the order of 10^{-15} ! For a laser with a 1-m cavity length, i. e., a free spectral range of 300 MHz, this corresponds to a control of the cavity length of 3×10^{-9} wavelength, i. e. of the order of 10^{-15} m, the size of an atomic nucleus. It may appear surprising to obtain such a precision with a mirror whose surface has a roughness of typically 10^{-10} m for the best polished surfaces. In fact, the precision of 10^{-15} refers to the average surface of the mirror on which the beam is reflected, with a typical diameter of 1 mm, on which the roughness is averaged out.

Comment

- (i) With semiconductor lasers the optical length of the cavity can be controlled by varying the temperature of the device or the electric current (on which the refractive-index depends).
- (ii) In some semiconductor lasers single-mode operation is ensured by the use of distributed feedback. The walls of the optical waveguide are modulated in the longitudinal direction

with a period of a multiple of the half-wavelength of desired operation: this Bragg-like structure reflects light only at this specific precise frequency, and the corresponding single longitudinal mode is oscillating.

3.4 Conditions for population inversion

In Sections 2.3 and 2.5 we described a first example of a system capable of amplification of an incident wave. It is a dilute medium composed of atoms with two excited states $|a\rangle$ and $|b\rangle$, of energies E_a and E_b respectively, in which the upper level¹⁰, $|b\rangle$, is more strongly populated than the lower level, $|a\rangle$, so that the competition between stimulated emission and absorption leads to net gain (Figure 3.13). Here we generalize the results of that discussion to more complicated, but more realistic systems.

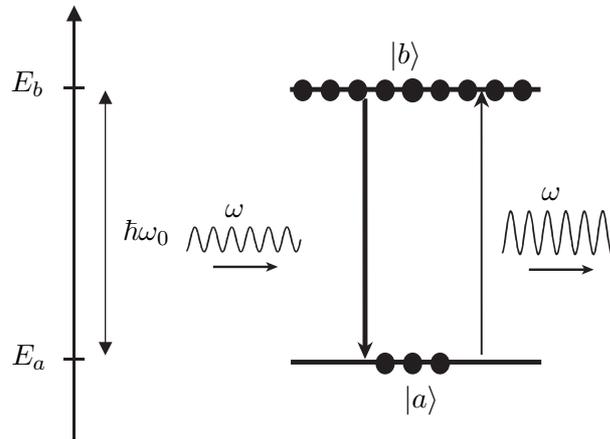


Figure 3.13: Atomic system providing laser gain. The upper level is more populated than the lower level of the transition, making stimulated emission dominate absorption.

3.4.1 Out of thermal equilibrium situation

For the systems that we shall consider in the following sections most of the results of Chapter 2 remain valid, notably the symmetric relationship between absorption and stimulated emission. If we consider a medium of thickness dz with a density of N_a/V atoms (or molecules) in the lower state and of N_b/V in the upper state¹¹, a wave of frequency ω propagating in the

¹⁰Here, we do not distinguish between a quantum state and an energy level, and treat the latter as if it were non degenerate.

¹¹Here V is the volume common to the amplifying medium and the laser cavity.

direction of positive z has its intensity changed according to (cf. Equation 2.101),

$$\frac{1}{I(z)} \frac{dI(z)}{dz} = g. \quad (3.23)$$

The gain per unit length g is proportional to the population difference (cf. Equation 2.125)

$$g = \frac{N_b - N_a}{V} \sigma_L(\omega). \quad (3.24)$$

This equation shows that the gain results from the competition between stimulated emission and absorption processes, the absorption being proportional to N_a and the stimulated emission to N_b . The laser cross-section, $\sigma_L(\omega)$ (see Section 2.6), is a positive number (with dimensions of an area). It generally has a resonant behaviour, around its maximum frequency ω_M . If we suppose that the light intensity is weak enough to allow us to neglect saturation, and if the medium is homogeneous, Equation (3.23) yields:

$$I(z) = I(0) \exp\left(g^{(0)} z\right). \quad (3.25)$$

If stimulated emission dominates over absorption, the gain is positive and the wave is amplified. This is obviously the case that is of relevance to the operation of lasers. It occurs in the case of population inversion

$$N_b > N_a. \quad (3.26)$$

Such a situation cannot happen when the system is in thermodynamic equilibrium, when the Boltzmann equation

$$\left(\frac{N_b}{N_a}\right)_{\text{th.equ.}} = \exp\left(-\frac{E_b - E_a}{k_B T}\right), \quad (3.27)$$

is satisfied. The realization of the required non-equilibrium state can only be achieved by exploiting the *kinetics*. Ideally, one would pump the population of the excited level $|b\rangle$ only and the lower level of the laser transition $|a\rangle$ would empty very quickly. In real systems, *relaxation processes* transfer population from the excited state $|b\rangle$ to the lower state $|a\rangle$, which suppresses population inversion. Progress in the development of new laser systems has essentially relied on the discovery of means for circumventing this problem in order to obtain population inversion. In the following we give some examples which illustrate how this has been achieved.

Comment

(i) Spontaneous emission from level b to level a is inevitable and has a deleterious effect on efforts to achieve population inversion. Spontaneous emission increases in importance as the frequency is increased, which explains why it is relatively easy to obtain laser oscillation in the infra-red but considerably more difficult in the ultra-violet. X-ray lasers

exist, but require extremely intense pumping processes.

(ii) The *optical parametric oscillator* is another kind of optical oscillator. The amplification process relies on the phenomenon of parametric mixing, which is an exchange of energy between a pump light wave and the emitted wave in a nonlinear medium.

3.4.2 Two levels are not enough

Let us show that for a closed two-level system (of which the lower and upper levels are denoted $|a\rangle$ and $|b\rangle$ respectively) it is not possible to create a steady population inversion by way of a pumping mechanism describable in terms of rate equations for atoms analogous to those of Section 2.1.4:

$$\frac{d}{dt}N_b = w(N_a - N_b) - \frac{N_b}{\tau_b}, \quad (3.28)$$

$$\frac{d}{dt}(N_a + N_b) = 0, \quad (3.29)$$

where w is the pumping rate from level $|a\rangle$ to level $|b\rangle$ and τ_b is the lifetime of the upper level. In the steady state we find, from (3.28)

$$\frac{N_b}{N_a} = \frac{w\tau_b}{1 + w\tau_b}, \quad (3.30)$$

which shows indeed that the population of the excited state is always less than that of the ground state.

Comments

(i) We saw in Chapter 1 (Section 1.4) that when a coherent, monochromatic wave interacts with a two-level system which is initially in the ground state, it is possible, after a period of time, to find a probability of 1 that the system is in the excited state (Rabi precession). A population inversion is therefore realized transiently. However, this inversion necessitates a coherent incident wave identical to that which would be created by laser action.

(ii) In the ammonium (NH_3) and hydrogen masers¹², the transition on which laser action occurs is related to a two-level system. In these devices a molecular beam is prepared in the excited state by a Stern-Gerlach type method allowing selection of the excited molecules only. This beam of totally inverted molecules, all in level $|b\rangle$, then traverses a resonant cavity which has a role analogous to that of the optical cavity of a laser. A significant fraction of the molecules is then transferred to the lower state a by stimulated emission. The molecules then leave the cavity. The molecules in state $|a\rangle$ are in this way physically removed from the cavity, thus ensuring that the population inversion is maintained. Such a system is, in fact, well described by the equations of Chapter 2, provided

¹²See C. Townes and A. Schawlow, *Microwave Spectroscopy*, § 15.10 and 17.7, Dover (1975). See also C. Townes' Nobel lecture, available at www.nobelprize.org

Γ_D is interpreted as the inverse of the mean time that the molecules spend in the cavity.

3.5 Three-level amplifiers

Since two levels are not sufficient to achieve population inversion, the simplest possibility is to look for systems exhibiting a third level. The first operational laser (the ruby laser) was precisely of this type. Decades later, a new three-level system, that of the erbium ion, Er^{3+} , embedded in silica, has found larger scale application as an amplifying medium in fibre-optic based telecommunications. We thus study here the distribution of populations in three-level systems, in order to ascertain the conditions in which laser amplification ending on the ground state can be obtained.

3.5.1 Three-level scheme: rate equations

Figure 3.14(a) illustrates the level scheme for a three-level system, which can give rise to laser emission on the transition linking the intermediate level $|b\rangle$ to the ground state $|a\rangle$. Pumping operates between level $|a\rangle$ and an excited state $|e\rangle$, which rapidly decays (with time-constant τ_e) to level $|b\rangle$. The spontaneous radiative de-excitation of level $|b\rangle$ to level $|a\rangle$ is, in the absence of stimulated emission, much slower: $\tau_e \ll \tau_b$.

The rate equations describing the pumping and relaxation processes in the three-level system of Figure 3.14(a) are the following:

$$\frac{d}{dt}N_e = w(N_a - N_e) - \frac{N_e}{\tau_e} \quad (3.31)$$

$$\frac{d}{dt}N_b = \frac{N_e}{\tau_e} - \frac{N_b}{\tau_b} \quad (3.32)$$

$$N_a + N_b + N_e = N, \quad (3.33)$$

where w is the pumping rate from level $|a\rangle$ to level $|e\rangle$. In the limit (realized in practice) in which $w\tau_e \ll 1$, the steady-state ratios of the populations (i. e. for which the time-derivatives in Equations 3.31 and 3.32 are zero) are:

$$\frac{N_b}{N_e} = \frac{\tau_b}{\tau_e} \gg 1, \quad (3.34)$$

$$\frac{N_e}{N_a} = w\tau_e \ll 1. \quad (3.35)$$

Neglecting $w\tau_e$ with respect to 1, and N_e with respect to $N_a + N_b$, one finds the population inversion

$$\frac{N_b - N_a}{N} = \frac{w\tau_b - 1}{w\tau_b + 1}. \quad (3.36)$$

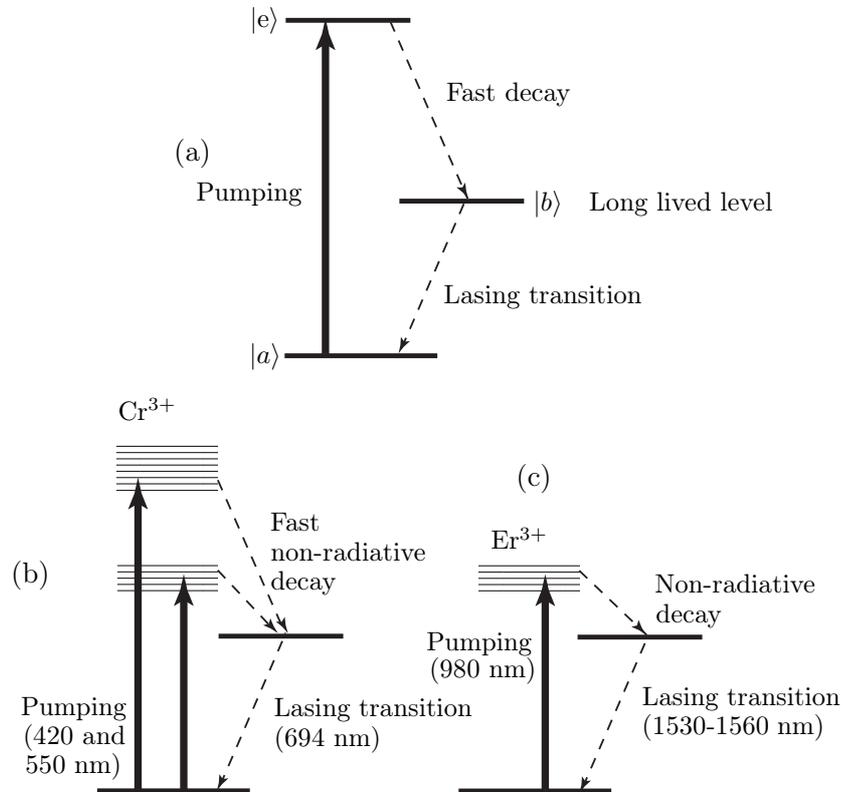


Figure 3.14: (a) Model energy level scheme for a three-level system, (b) its practical realization in the case of the Cr^{3+} ion in a ruby crystal and (c) in the case of the Er^{3+} ion in a glass (for example in a silica based optical fibre).

This result is illustrated in Figure 3.15, which plots the evolution of the population inversion $N_b - N_a$ versus pumping w . It shows that, for a three-level system, population inversion can only be achieved *if the pumping is sufficiently strong*:

$$w > \frac{1}{\tau_b}. \quad (3.37)$$

Equality in equation (3.37) corresponds to the transparency threshold of the medium. Below this value, the system absorbs, while above this value it exhibits gain.

It is clear from Figure 3.15 that a three-level system requires a strong pumping rate to provide enough gain for the laser to oscillate, while we will see below that this is not the case for a four-level system. Three-level systems play nevertheless an important role in some applications, as discussed below.

Comments

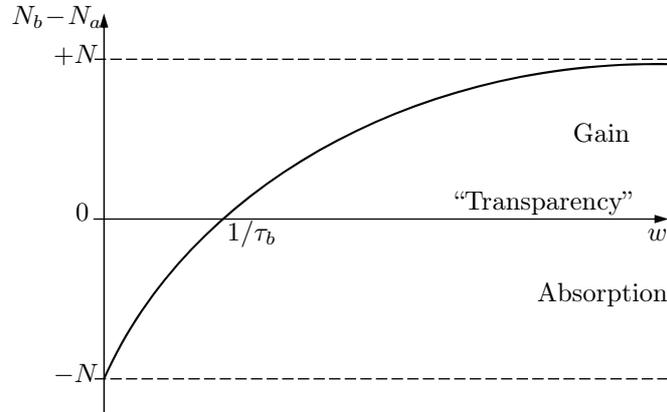


Figure 3.15: Evolution of the unsaturated population inversion $N_b - N_a$ versus pumping rate w for a three-level system. The transparency line corresponds to $N_b - N_a = 0$, where the system exhibits neither gain nor absorption.

(i) The condition (3.37) only ensures that the amplifying medium has a positive gain. If, in addition, the unsaturated gain fulfills the condition (3.5), laser operation starts, leading to a fast transfer from level $|b\rangle$ to level $|a\rangle$. The gain then immediately decreases, and one needs to take stimulated emission and absorption into account in the rate equations, as will be done in Section 4.1.

(ii) There exists another kind of three-level system where the lasing transition occurs between levels $|e\rangle$ and $|b\rangle$ of Figure 3.15(a) and level $|b\rangle$ decays very quickly to level $|a\rangle$. Such a scheme is close to the four-level system that will be described in Section 3.6.

3.5.2 Two landmark cases: ruby, erbium

a. The ruby laser

In the case of the ruby laser, the active element is the Cr^{3+} ion embedded in a crystal of aluminium oxide where it substitutes Al^{3+} ions in some sites. Figure 3.14(b) shows the energy levels of the Cr^{3+} ion in the crystal. Notice the existence of two absorption bands which can efficiently absorb the light emitted by a flash-lamp.

Figure 3.16 shows a sketch of the laser built by Maiman, the first to produce a coherent laser beam, in May 1960. The amplifying medium is a rod of ruby pumped by a flash-lamp in which it is enclosed. The end facets of the ruby rod are polished parallel and coated with a semi-reflecting silver coating to form a linear laser cavity. The laser emits pulses in the red at a wavelength of 694 nm.

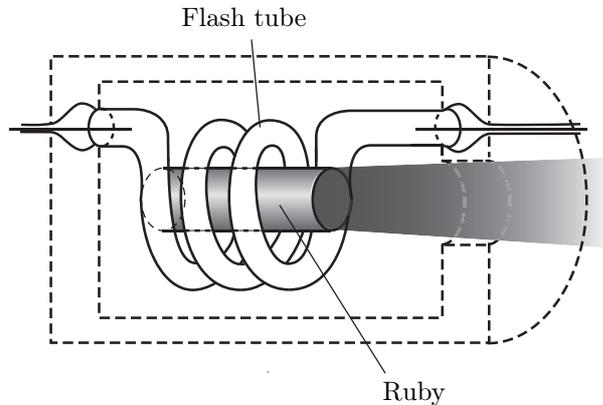


Figure 3.16: Schematic diagram of a ruby laser of the type built by Maiman. The cavity mirrors are deposited on the crystal faces.

b. The erbium Laser

The most used laser system that corresponds to the level scheme of Figure 3.14 is the erbium laser. This laser, which amplifies infrared light in the wavelength range of $1.52 - 1.56 \mu\text{m}$ has given a new lease of life to three-level lasers, at the end of the 1980's. Two particularly noteworthy characteristics of these lasers is that the emission wavelength corresponds to the minimum of the absorption in optical fibres and that it has a rather broad gain bandwidth. It is therefore ideally suited for applications in *telecommunications*. Erbium ions can be implanted into silica, yielding *erbium-doped fibres*. When excited by pump light, for example from a semi-conductor laser at $1.48 \mu\text{m}$ or $0.98 \mu\text{m}$, these fibres function as extremely efficient optical amplifiers, the so-called EDFA (Erbium-Doped Fibre Amplifier). When lengths of such amplifying fibre are inserted in long fibre optic cables they can regenerate the light pulses attenuated by the absorption losses in a rather broad frequency range, with negligible cross-talk between the different frequencies, enabling many communication channels of slightly different wavelengths to be carried and amplified simultaneously and independently (WDM: Wavelength Difference Multiplexing). Twenty years after the invention of the EDFA, communication rates over 10^{11} bit/s were achieved, and the EDFA market exceeded 1 billion €.

Laser oscillation is obtained if the erbium-doped fibre is placed in a suitable cavity. For example the mirrors can be integrated onto the end facets of the fibre itself. One obtains in such a way very convenient fibre lasers, which are totally reliable against mechanical vibrations, and turn out to be amazingly powerful, considering the small size of the fibre core (more than a kW cw in an almost TEM_{00} mode, see Complement 3B) and efficient (80% conversion efficiency from the pump light to the laser light). These power-

ful lasers have a wide range of industrial applications. Let us mention that other rare-earth ions can be used for the same purpose and provide other fibre lasers in other parts of the near infrared spectrum, as discussed below (Section 3.6.2).

3.6 Four-level amplifiers

We have just seen that the main drawback of three-level systems is that the lower level of the transition is also the atom ground state. This implies that the pumping needs to be strong enough to empty this level in order to reach population inversion. The four-level system permits to circumvent this issue.

3.6.1 Description of the four-level scheme

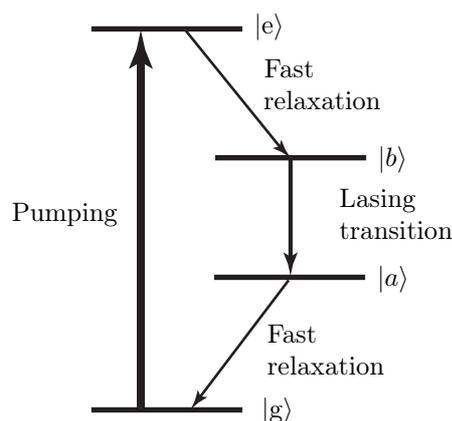


Figure 3.17: Level scheme for a four-level system. The laser operates on the transition $|b\rangle \rightarrow |a\rangle$. Levels $|e\rangle$ and $|a\rangle$ decay rapidly to levels $|b\rangle$ and $|g\rangle$ whilst a pumping mechanism populates $|e\rangle$ from the ground state $|g\rangle$.

The energy level scheme of a four-level laser is depicted in Figure 3.17. A *pumping* mechanism (optical pumping driven by the light emitted by a lamp or by an auxiliary laser, excitation by an electric discharge etc.) transfers atoms from their ground state $|g\rangle$ to an excited state $|e\rangle$. The *rapid relaxation* of this level then transfers its population to the upper laser level $|b\rangle$. The spontaneous radiative decay between $|b\rangle$ and $|a\rangle$ is assumed to be slow compared to the relaxation of level $|e\rangle$ to level $|b\rangle$. Finally, the atoms in the lower level of the laser transition $|a\rangle$ decay to the ground state by a rapid relaxation process. The characteristic de-excitation times of the excited levels are denoted by τ_e, τ_b and τ_a such that $\tau_a, \tau_e \ll \tau_b$. Level e is pumped at a constant rate w . We can write the rate equations describing the kinetics

of the system and deduce the population difference $N_b - N_a$ on which the amplification depends.

When the laser intensity is sufficiently weak that the effect of absorption and stimulated emission on the transition $|b\rangle \rightarrow |a\rangle$ can be neglected (i.e. in the unsaturated regime), the following rate equations can be written

$$\frac{d}{dt}N_e = w(N_g - N_e) - \frac{N_e}{\tau_e} \quad (3.38)$$

$$\frac{d}{dt}N_b = \frac{N_e}{\tau_e} - \frac{N_b}{\tau_b} \quad (3.39)$$

$$\frac{d}{dt}N_a = \frac{N_b}{\tau_b} - \frac{N_a}{\tau_a} \quad (3.40)$$

$$\frac{d}{dt}N_g = \frac{N_a}{\tau_a} - w(N_g - N_e) . \quad (3.41)$$

These equations automatically ensure the conservation of the total population:

$$\frac{dN}{dt} = 0 , \quad (3.42)$$

$$N_a + N_b + N_e + N_g = N . \quad (3.43)$$

Let us write the steady state equations ($dN_i/dt = 0$ with $i = g, a, b, e$), neglecting small terms associated with $\tau_a, \tau_e \ll \tau_b$, and assuming a weak pumping regime ($w\tau_e \ll 1$). We obtain

$$N_e \simeq w\tau_e N_g , \quad (3.44)$$

$$N_b \simeq w\tau_b N_g , \quad (3.45)$$

$$N_a \simeq w\tau_a N_g , \quad (3.46)$$

hence the relative population inversion

$$\frac{N_b - N_a}{N} \simeq \frac{w\tau_b}{1 + w\tau_b} . \quad (3.47)$$

The population inversion can be noticeable if level $|b\rangle$ has a long lifetime. Such a four-level system is ideal for laser operation, as we shall see in the following.

Comparison with that obtained for the three-level laser configuration (Equation 3.36 and Figure 3.15) is performed in Figure 3.18. It shows that contrary to the three-level system, the four-level system exhibits gain as soon as the pumping rate is different from zero ($w > 0$). This is due to the fact that in the four-level system, the lower level of the transition (level $|a\rangle$ in Figure 3.17) is not the atom ground state, and it thus unpopulated even in the absence of the pump.

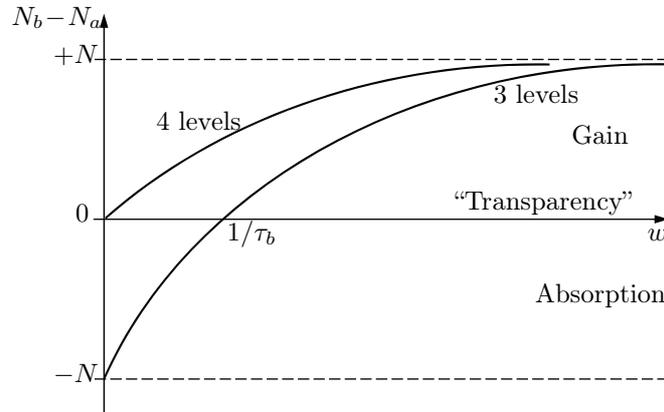


Figure 3.18: Evolution of the unsaturated population inversion $N_b - N_a$ versus pumping rate w for a three-level and a four-level system.

Comments

(i) If the pumping mechanism arises from the absorption of light, Equations (3.38–3.41) are the rate equations discussed in Chapter 2. When the pumping occurs by electronic collisions, as in an electric discharge, it is an incoherent process, leading also to rate equations.

(ii) In practice, laser media are almost never perfect 4-level systems like the one described in the simple model above. In particular, extra decay mechanisms hinder the creation of the population inversion, requiring an increase of the pumping rate.

3.6.2 Two emblematic examples: neodymium, helium-neon

a. The neodymium laser

Figure 3.19 shows the energy levels of a Nd^{3+} (neodymium) ion when embedded at low concentration in glass (neodymium-doped glass) or in a crystal named YAG (neodymium-doped yttrium aluminium garnet). Pumping is achieved using an exterior light source, usually either an arc lamp or a diode laser, which excites ions into excited bands from which they decay rapidly into the upper laser level, thanks to a *non-radiative* relaxation process due to the coupling with the vibrations of the material.

Laser emission occurs principally on the transition at $1.06 \mu\text{m}$ in the infrared. Other transitions are capable of laser action, but with much reduced efficiency. They are consequently seldom used. In the pulsed operating mode, the energy output can reach several Joules in a pulse of duration 1 ps to 1 ns, with repetition rates ranging from a few Hertz to the kHz domain. In

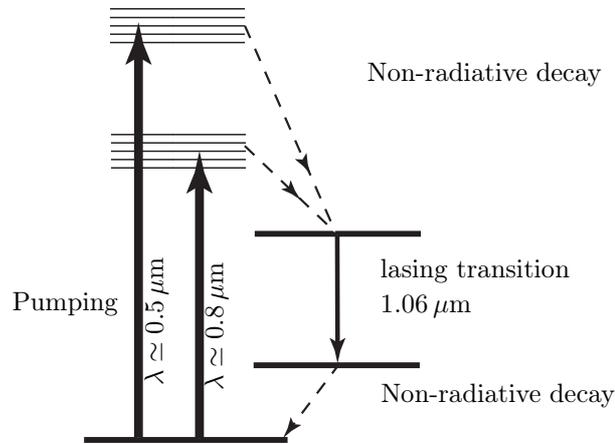


Figure 3.19: Nd:YAG laser medium. Energy level scheme for a neodymium ion embedded in a solid such as glass or crystal, showing discrete levels and bands. The transitions used for pumping have wavelengths in the regions of $0.5 \mu\text{m}$ and $0.8 \mu\text{m}$. Laser emission occurs at $1.06 \mu\text{m}$.

continuous-wave (cw) operation, powers exceeding 100 W can be obtained. For many applications a nonlinear crystal (see PHY432) is placed at the laser output, or even inside the laser cavity, in order to produce a coherent beam of the second harmonic of the laser output wavelength, at 532 nm (in the green), or of the third harmonic at 355 nm in the ultra-violet.

Neodymium lasers are highly efficient (the global efficiency ranging from 1% for a flash-lamp pumped system, to 50% for some laser diode pumped systems) and have a vast range of applications. Because of their high peak power in pulsed operation (see Figure 3.20a) they are used, for example, in experiments aimed at initiating thermonuclear fusion (see Complement 4C). These lasers are also used as cw sources (Figure 3.20b). In this case pumping by a semiconductor laser at $0.8 \mu\text{m}$ is possible. Frequency-doubled, they are frequently used as pump sources for tunable lasers (Section 3.6.4). These green pump lasers are widely used because of their small physical size and good efficiency.

b. The helium-neon laser

The helium-neon laser, invented in 1960, is still widely used. It employs a pumping mechanism of some subtlety (Figure 3.21). A continuous electric discharge in a mixture of helium and neon gases excites helium atoms to a number of *metastable* levels, which have a very long lifetime because they have no electric dipole coupling to the atomic ground state. In subsequent collisions of an excited helium atom with a neon atom in its ground state, the stored internal energy can be transferred, putting the neon atom into an

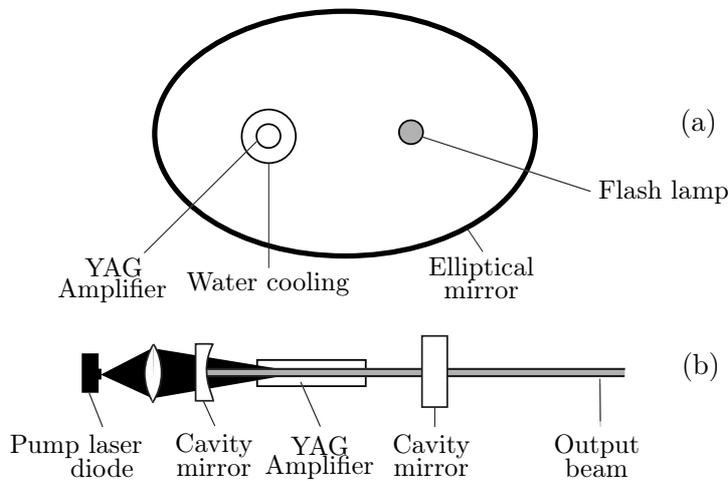


Figure 3.20: Nd:YAG laser source. a) Cross-section of the amplifying zone of a flash-lamp pumped neodymium YAG laser. The lamp and gain rod are situated at the foci of an ellipse. Such a laser can deliver an average power of 1 to 100 W in the infra-red. It consumes an electrical power in the range of 0.1 to 10 kW. b) Laser diode pumped Nd:YAG laser. Such a laser can have an efficiency exceeding 50%.

excited state of similar energy. Laser action can be obtained on transitions of which the excited levels are those pumped by the collisional energy transfer. Lasing can be obtained on lines at $3.39 \mu\text{m}$ and $1.15 \mu\text{m}$ as well as on the well-known, red transition at 633 nm .

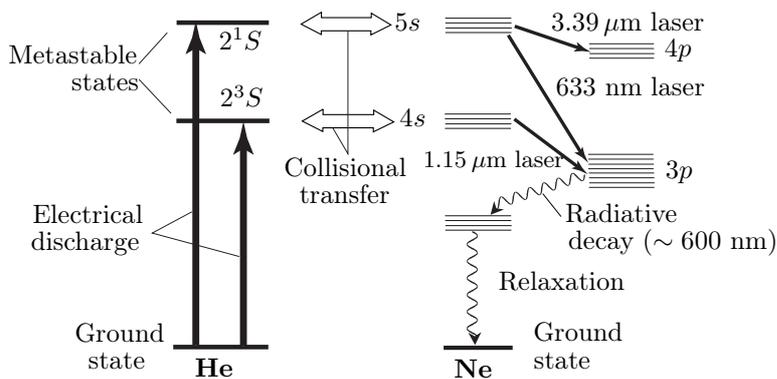


Figure 3.21: The energy levels of helium and neon involved in the operation of the helium-neon laser. The energy stored by the electronic excitation of metastable levels of helium is transferred to the neon atoms in collisions.

Figure 3.22 shows a much used configuration for the cavity of a helium-neon laser. The gas is enclosed in a tube with end windows inclined at *Brewster's angle* which cause no reflection loss for light passing through them

with its linear polarization vector in the plane of the figure. The orthogonal light polarization does experience a reflection loss and consequently has a higher threshold gain. The laser therefore oscillates only on the low loss polarization mode and the output is *linearly polarized*. The laser transition that oscillates (e. g., the 633 nm transition) is selected by the mirrors which reflect only a narrow band of wavelengths. The output power of such a laser is typically of the order of a few mW. Some versions differ from the configuration shown in that the mirrors are internal to the vacuum envelope so that the output is unpolarized.

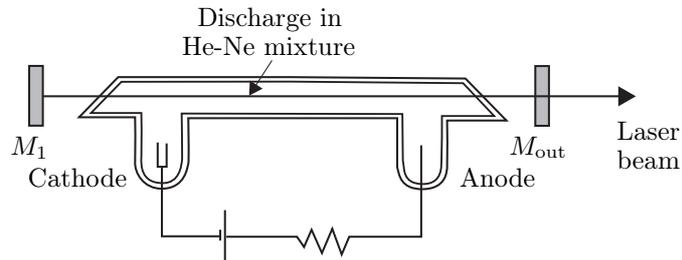


Figure 3.22: Scheme of a helium-neon laser producing a linearly polarized output beam.

3.6.3 Fibre lasers

Fibre lasers, already mentioned in Section 3.5.2 about Erbium Doped Fiber Amplifiers, are now among the most efficient laser sources, able to emit powers of the order of 10 kW in continuous wave regimes. They consist in an optical fibre doped with active ions. Two mirrors located at both ends of the fibre close a cavity around the fibre. These mirrors can be either bulk mirrors, as in Figure 3.23, or deposited on the fibre ends, or Bragg mirrors, which consist in a longitudinal refractive index modulation directly engraved inside the fibre with a period equal to half the emission wavelength (in the material). When pumped by a laser (typically a semiconductor laser), such a doped fibre provides gain for the guided light.

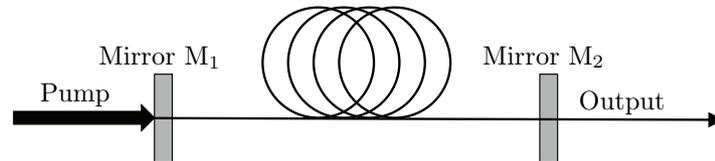


Figure 3.23: Fibre laser. Mirror M_1 transmits the pump wavelength and is fully reflecting for the laser wavelength. Mirror M_2 is the output coupler.

One of the major advantages of fibre lasers is that the pump light remains guided all along the doped fibre, which is typically a few meters long. This wave guiding permits to achieve a huge single-pass amplification at the laser wavelength, although the gain per unit length is moderate. Several fibre designs can be used to improve the energy transfer from the pump to the laser, such as the double-clad architecture illustrated in Figure 3.24. In such fibres, the pump light, which is usually emitted by a powerful laser exhibiting a poor spatial mode quality, is injected in the relatively large clad surrounding the small core that contains the active ions and in which the fibre laser light will propagate. Along propagation inside the double-clad structure, the pump light is progressively transferred from the clad to the core, where it excites the active ions and provides gain. Moreover, the large side surface permits efficient dissipation of the heat generated inside the fibre. Together with the laser radiation being guided in the core, this permits to maintain a very good beam quality in spite of large thermal dissipation. This is yet more true in the case of ytterbium-doped fibres, where the energy difference between the pump photons (around 950 nm) and the laser photons (around 1030-1050 nm), the so-called *quantum defect*, is small enough that the part of the pump energy that is not converted into laser light is relatively small (compared for example to the case of Nd:YAG, see Figure 3.19). Reduced heat production, efficient heat dissipation, and wave-guiding are crucial advantages, explaining why ytterbium-doped fibre lasers are among the most powerful and efficient lasers, and are ubiquitous in industrial applications such as welding, cutting, engraving, marking, etc.

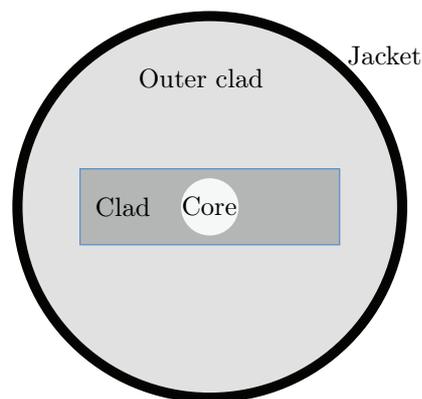


Figure 3.24: Transverse structure of a double-clad fibre used in high-power fibre laser. Laser light is confined in the core, which is surrounded by the clad that guides most of the pump light. The jacket is a coating that protects the fibre from the environment.

3.6.4 Tunable lasers

The helium-neon or Nd:YAG lasers described in the previous paragraphs, as many others, have a very narrow gain curve, which extends over a range of a few gigahertz or a few tens of gigahertz, respectively. It is therefore not possible to widely tune the output wavelength (visible light extends over a bandwidth of some 2×10^{14} Hz). It was to meet the need for light sources with a frequency adjustable over a broader spectral range that tunable lasers were developed. They are based on extended gain curves. This can arise, for example, if the lower level of the laser transition belongs to a *continuum* (or dense quasi-continuum) of levels. This configuration is encountered in dye molecules in a liquid solvent or for some metallic ions in a crystal, such as the titanium ion embedded in a sapphire crystal (see Figure 3.25). Because of fast relaxation in each of the two bands, the system is an effective four-level one.

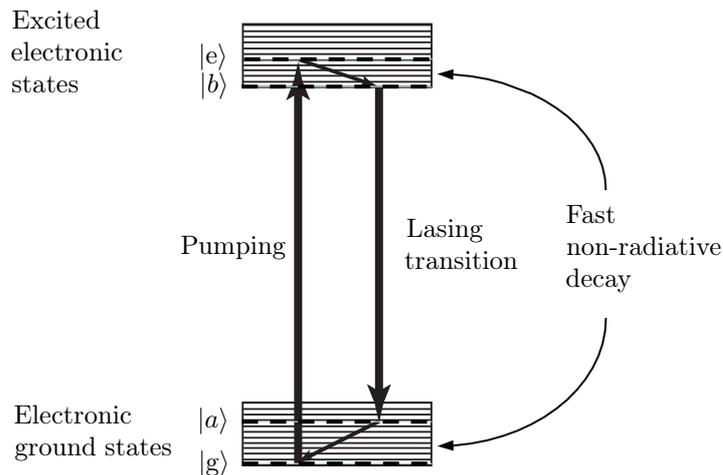


Figure 3.25: Energy level scheme showing the mechanism of amplification in a tunable laser. The energy levels of the active medium (dye laser, or titanium doped sapphire laser) comprise bands separated by electronic transitions. Non-radiative relaxation within a given band towards the lowest level of the band takes place very quickly (on a timescale of order 1 ps). The system is thus equivalent to a four-level system ($|e\rangle$, $|g\rangle$, $|a\rangle$, $|b\rangle$), as in Figure 3.17. The tuning range of the output is determined by the energy width of the lower band of the laser transition. The laser frequency is selected by intracavity filters.

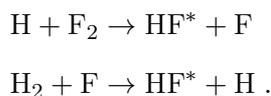
Selection of the laser frequency is achieved by the positioning in the laser cavity of *frequency-selective* elements, which induce losses for wavelengths outside the desired operating range (see Figure 3.10). A Titanium:Sapphire laser, for example, pumped by a frequency doubled YAG laser at $0.53 \mu\text{m}$,

is tunable over the range [700 nm, 1100 nm] (that is, over a range of 200 THz, 8 orders of magnitudes larger than than the few MHz linewidth of the emitted light). For a pump intensity of 10 W the useful output power is more than 1 W in single mode operation.

3.6.5 Molecular lasers

Population inversion between the vibrational levels of carbon dioxide molecules (CO₂ laser) is achieved in an electric discharge. These systems can be very powerful (10 kW or more) and they have industrial and military uses. The laser transition, which is at a wavelength of 10.6 μm in the far infra-red, involves two distinct molecular rotational/vibrational states (i.e. states involving different modes of oscillation for the relative motion of the atoms constituting the molecule).

There are many other molecular laser systems that emit in the infra-red. In some of these systems the active molecules of the gain medium are created directly in the excited state by a chemical reaction, which automatically ensures the existence of a population inversion. One such system is the hydrogen fluoride (HF) laser in which excited molecules HF* are created in the reactions



Such a laser works in a pulsed mode and can emit very high peak powers: typically 4 kJ in a pulse of duration 20 ns, giving a peak power of 200 GW. Note that such a laser can operate without a supply of electricity, since energy is converted from a chemical form; it is necessary only to have a supply of the reactant gases.

If the laser transition is between distinct *electronic* levels (Figure 3.26), the wavelength of the emission is often in the ultra-violet, and is widely used in photolithography and production of microelectronics chips. Commonly used molecular lasers are the *excimer* lasers, employing the ArF or KrF molecules, which emit in the ultra-violet, at 195 nm and 248 nm respectively. In these systems the lower level of the laser transition is unstable (see Figure 3.26) because there is no stable bound state of the two atoms (only the excited states are stable). The problem of the depopulation of the lower laser level *a* of the laser transition (Figure 3.17) is therefore elegantly solved since the ground state molecule very quickly dissociates, liberating its two constituent atoms.

3.6.6 Semiconductor lasers

Semiconductor lasers (also known as “diode lasers”) are the cheapest and most widespread lasers in use. They operate essentially in the red, infra-

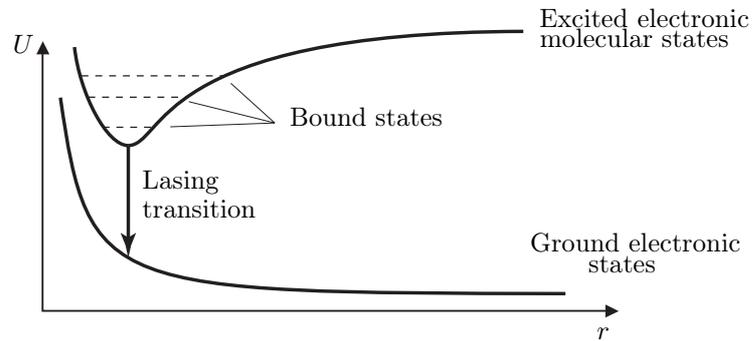


Figure 3.26: Energy levels scheme for an excimer molecular laser. The interaction energy of the two atoms is shown as a function of their separation. In the excited molecular state, the interaction potential has a well which supports bound vibrational-rotational states of stable molecules (e. g. ArF) produced by a discharge in a mixture containing a rare gas and a halogen (e. g. Ar and F₂). The upper laser level is one such bound state of ArF. The lower level of the laser transition is a dissociative state, which leads to the fission of the molecule into atomic Ar and F. Here again the system is an effective four-level one, with very efficient emptying of the lower level of the laser transition.

red, and blue spectral regions. Every CD, DVD or blu-ray player contains several ones (emitting at wavelengths respectively around 780 nm, 650 nm, or 405 nm). Fibre-optics based telecommunications rely heavily on devices operating at 1.3 μm and 1.5 μm (the spectral region in which optical fibres exhibit minimum dispersion and maximum transmission, see Section 5A.4 of Complement 5A).

The emission of light in a diode laser occurs in the junction region of a forward-biased semiconductor diode, formed of heavily doped materials (see Figure 3.27). Such a junction is formed by a contact between a *p*-type and a *n*-type regions of the same semiconductor. Figure 3.27(a) shows the evolution of the conduction band and valence band energies as a function of the direction *z*, which is perpendicular to the contact plane between the two regions, when no voltage is applied on the junction. These two bands are separated by the so-called gap energy E_g of the considered material. One can see that the junction area corresponds to a strong electric field that opposes the diffusion of electron and holes across the junction. Then the holes are confined in the *p*-doped region while the electrons are confined in the *n*-doped region, and there is no region where these two kinds of carriers can meet. To create a region where both electrons and holes are present, one needs to apply a positive voltage of the order E_g/e across the junction (so-called positive bias), leading to the situation of Figure 3.27(b). The *electron-hole recombination* can then occur at the junction and liberates energy in the

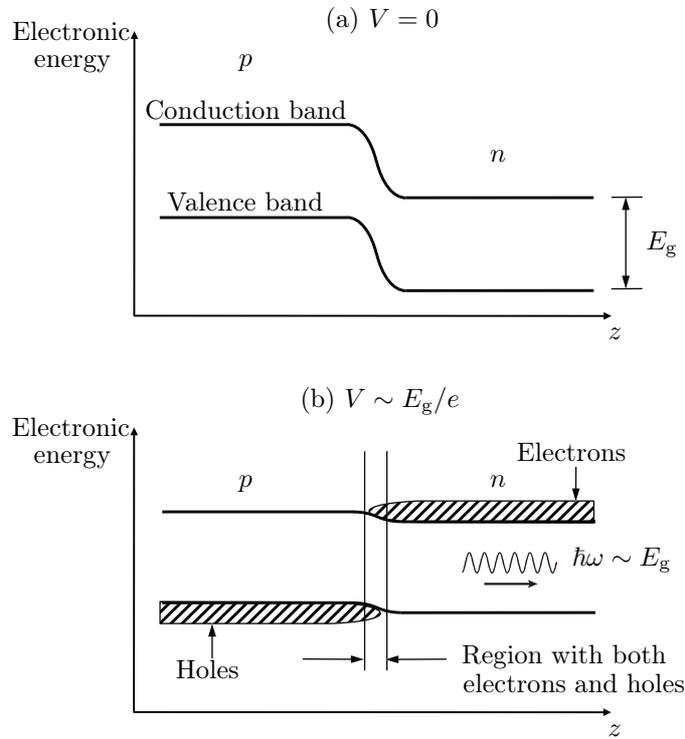


Figure 3.27: Energy-band diagram of a $p - n$ junction in a semiconductor laser when (a) no voltage is applied on the junction and (b) when a positive potential difference is applied between zones p and n (so-called forward bias). Light emission occurs in association with electron-hole recombination in the part of the junction where holes and electrons overlap. The emitted photons have an energy $\hbar\omega$ close to the gap energy E_g of the semiconductor.

form of a photon of energy $\hbar\omega \approx E_g$ where E_g is the energy gap separating the top of valence band from the lower edge of the conduction band (this band-gap is of the order of 1.42 eV for gallium-arsenide semiconductor alloy, GaAs, equivalent to a wavelength of 870 nm). In the regime in which only spontaneous emission occurs, one has a light-emitting diode (LED), which is the basis of many instrument displays and lightning devices. If, however, the injected current density is large enough, a regime can be reached in which stimulated emission is predominant: the system then exhibits a very large optical gain. Like in excimer lasers, the lower level a of the laser transition is almost instantaneously depopulated, here because of the electron-hole recombination.

The first diode lasers were operational as early as 1962, but their use was for a long time restricted to specialized laboratories: it was necessary to maintain their temperature at that of liquid nitrogen in order to limit the

non-radiative relaxation rates. Moreover, a very high current was required for the threshold to be reached. In the 1970s, developments in semiconductor components fabricated from gallium arsenide and similar materials allowed considerable progress to be made. It became possible using multi-layer structures (termed *heterostructures*, see Figure 3.28) to confine the recombination to a layer of thickness $0.1 \mu\text{m}$ and width of a few μm . The threshold current was in this way greatly reduced.

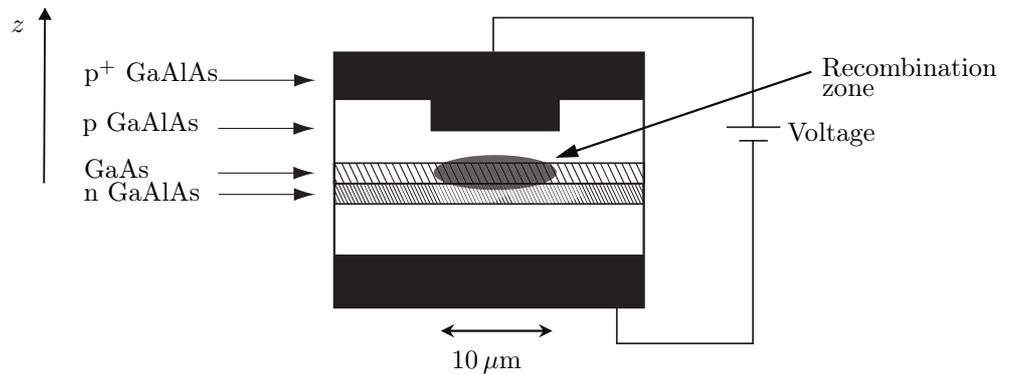


Figure 3.28: GaAlAs/GaAs heterostructure laser. Schematic diagram of the various layers in the active-region of a diode laser. The layers are deposited along direction z . Several layers each with different gaps are superimposed. Such a structure is used for laser emission at room temperature, with an injected electric current of much less than 1 A.

Stimulated emission occurs in the thin active layer, within which the light is guided, as in an optical fibre (the refractive-index is higher than in the surrounding material). Transverse confinement results in a waveguide of section $1 \times 10 \mu\text{m}^2$ (see Figures 3.28 and 3.29). Two parallel facets, cleaved orthogonal to the active channel, serve as the mirrors of a monolithic linear cavity of length typically $400 \mu\text{m}$ (see Figure 3.29). Since the 1990's, commercially available diode lasers operate at room temperature and provide coherent light at powers of several watts in the c.w. regime for an input current of the order of a few amperes. By using semi-conductors of different gaps, one is able to cover the whole near infra-red, red, and blue-violet spectrum with a set of diode lasers. As in other types of lasers, the shorter the wavelength the more difficult to achieve the regime where stimulated emission dominates, and it is only in the 1990's that blue and violet diode lasers were invented. Based on InGaN/AlGaIn junctions, they operate around 400 nm and are used in more advanced generations of optical storage discs and RGB video-projectors.¹³

¹³These technological breakthroughs are also at the origin of the development of blue

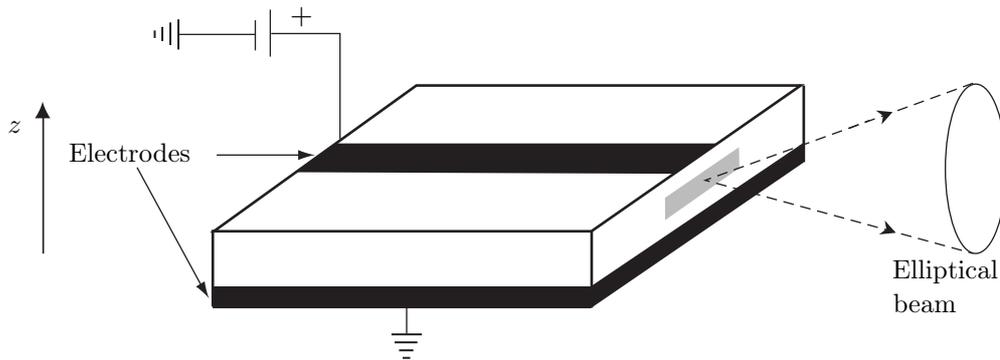


Figure 3.29: Light emission from a diode laser occurs from the end facets (grey shaded on the figure) of the recombination zone (of typical dimensions $1\ \mu\text{m} \times 10\ \mu\text{m} \times 400\ \mu\text{m}$). Because of diffraction, the emitted beam has a far field elliptical cross-section, with a vertical divergence of the order of 30° and a lateral divergence of a few degrees. This can be compensated by reshaping optics.

Because of diffraction, the emitted beam is elliptical and highly divergent (see Figure 3.29); collimation must be achieved with the aid of external optical components. These lasers are nevertheless widely used because of their ease of operation, their overall efficiency which exceeds 50%, and their compactness. Assembled by hundreds in tight packages, they constitute lasers arrays of high brightness and efficiency in the cw or quasi-cw regime (up to 100W cw, with a conversion efficiency from electrical power of about 75%).

VCSELs (Vertical Cavity Surface Emitting Lasers) are another kind of semi-conductor lasers (see Figure 3.30). The laser cavity is made of different layers of semiconductors, so that the output beam is perpendicular to the surface of the device. The mirrors of the laser cavity are made by layers of alternated indices, which constitute a thick optical Bragg grating which reflects the light at a wavelength determined by the thickness of the layers. The gain medium is a layer of semiconductor which is so thin that it constitutes a 1-dimensional quantum well (along the vertical direction) having discrete electronic levels between which the inversion takes place when the system is electrically or optically pumped. The VCSEL generally oscillates in the fundamental mode of the cavity, for which the intra-cavity standing wave has a single antinode overlapping with the quantum well. These lasers are interesting for their good optical quality and the possibility of integrating many of them on the same chip.

light-emitting diodes which have enabled bright and energy-saving white light sources. This invention led to the Nobel prize in physics awarded to Isamu Akasaki, Hiroshi Amano, and Shuji Nakamura in 2014.

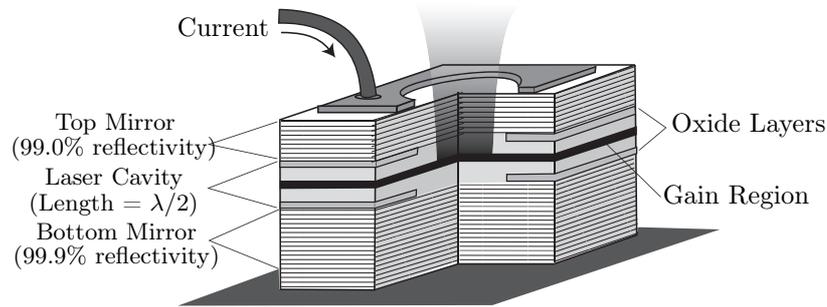


Figure 3.30: Scheme of a Vertical Cavity Surface Emitting Laser (VCSEL). The total length of the laser cavity is half a wavelength, and supports one antinode only, which overlaps with the gain region, a quantum well for electrons.

3.7 Conclusion: why laser light?

This chapter has allowed us to introduce, in a rather heuristic manner, the main features of lasers: the concept of mode, the notion of threshold, and some pumping configurations leading to population inversion. This is sufficient to get a basic understanding of what a laser source is, either cw or pulsed. The two following chapters will give a more formal derivation of the laser equations of evolution. They will be used to describe *quantitatively* many important features of lasers, including their dynamical behaviour and its fluctuations. This will also allow us to show that beyond its applications, the laser is also a fascinating object for the physicist. Indeed, it permits to illustrate and study many phenomena occurring in many different areas of physics.

But at the end of this chapter the reader may ask the question: why so many efforts to get laser light? At the beginning of the XXIst century, a frequency doubled YAG laser capable of emitting 15 Watts costed around 50.000 Euros. A 100 W incandescent lamp, on the other hand, costed a few Euros. That lasers are employed at all indicates, therefore, that the light they emit must have some rather remarkable properties. The principal of these is that the admittedly smaller energy of the laser can be concentrated in an extremely narrow region of space – or of directions, and of time – or of frequencies. This leads to the generation of energy densities that are enormously superior to those obtainable with ordinary sources.

Complement 3C gives some details about the fundamental reasons why laser light can be much more concentrated than light emitted by a usual, incoherent, light source. Let us mention here that the crux is the fact that for a beam issued from an incoherent source, and propagated through any kind of passive instrument, there exists a conserved quantity, the radiance. More precisely the radiance, which has units of Watts per unit area and solid

angle, either is conserved, in the ideal case of a perfect instrument without any loss, or it decreases. This property is as fundamental as the second principle of thermodynamics, and it has a close relationship to the fact that the entropy of an isolated system can only increase, or remain constant in the ideal case. As a consequence, one cannot obtain an irradiance, i.e., the power per unit surface, obtained after any kind of instrument, larger than the initial radiance multiplied by a factor of π . It means for instance, for the light issued from the sun and concentrated with a large magnifying glass, a maximum irradiance of a few 10^3 W/cm². In contrast, because of its spatial coherence, the whole power of a laser beam can be concentrated over a spot with a size of the order of the square of the wavelength, λ^2 . For a laser delivering 10 Watts at $\lambda \simeq 1\mu\text{m}$, one obtains an irradiance of 10^9 W/cm², 6 orders of magnitude more than for light from the sun.

The gain is yet more striking if one considers the fact that the light from the sun, or any kind of incandescent light, is spread over a bandwidth of more than 10^{14} Hz, while the light from a single-mode laser can easily be emitted in a linewidth of 10^6 Hz, or less. This means another factor of 10^8 in favor of laser light, when one considers the spectral irradiance, i.e., the power per unit surface and unit frequency, a quantity relevant for instance when one wants to excite an atomic transition.

Rather than concentration in space and frequency, laser light can be concentrated in conjugate variables, i.e., direction and time (see Complement 3C for more details). Concentration in direction is done by passing a laser beam through a telescope in the “wrong” direction, entering at the eye-piece and emerging at the large mirror. The beam, with a diameter of 1 m or more, has thus a divergence less than 10^{-6} rad, and ends as a disk of 300 m in diameter on the moon. This is used to measure the distance earth-moon by sending laser pulses onto retroreflectors planted on the moon during Apollo and Lunokhod missions.

Concentration in time is spectacular, since it is possible to emit pulses of duration less than one femtosecond, in the attosecond range. The magnitude of the instantaneous electric field then reaches values much higher than the Coulomb field keeping an atomic electron bound to the nucleus, giving access to the large-intensity regime of matter-light interaction where many new effects happen, such as high harmonic generation. Femtosecond pulses are also able to break chemical bonds, and thus allow one to act upon materials by means different from simple heating.

The answer to the question “What makes laser light different ?” is thus crystal clear: thanks to its coherence properties, laser light has the capability of being concentrated to a degree impossible with light from any other source. Its *temporal coherence* allows its energy to be concentrated either in time or in the frequency domain, whilst its *spatial coherence* allows it to be concentrated in space (i.e., focused onto a small area) or to be concentrated in direction (i.e., transformed into a wide, diffraction limited, highly

parallel beam). It is these possibilities that have enabled the laser to find such wide-ranging applications from those in fundamental research to those in technology designed for the mass market.

In fact, the coherence of a single-mode laser beam can be considered a consequence of a yet more fundamental property, the fact that *in a laser mode the number of photons is very large compared to 1*, in contrast to the case of a non-laser source, such as a discharge lamp, an incandescent lamp, or the surface of the sun. As explained in Complement 3C, the number of photons per mode, given by the Bose Einstein distribution at a temperature of 5800 K, does not exceed 10^{-2} in solar light, at the maximum of spectrum. But even for a modest Helium-Neon laser delivering a few milliwatts, the number of photons in a mode is huge, typically larger than 10^8 , and this remains true for each mode of a multi-mode laser. When we remember that all the photons in the same mode of the electromagnetic field are indistinguishable bosons, we have the reason why a single mode laser beam is fully coherent. Once again, we find that the basic principles of lasers are closely linked to fundamental physical laws.