INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE

H4.SMR/453-8

TRAINING COLLEGE ON PHYSICS AND CHARACTERIZATION OF LASERS AND OPTICAL FIBRES

(5 February - 2 March 1990)

BASIC LASER PHYSICS

THE LASER PRINCIPLE

O. Svelto

Centro di Elettronica Quantistica
e Strumentazione Elettronica
Istituto di Fisica del Politecnico
Milano, Italy

Basic Laser Physics

C. Sretto

The Laser Principle [First Lecture]

1

Introductory Concepts

1.1. SPONTANEOUS AND STIMULATED EMISSION, ABSORPTION

Quantum electronics can be defined as that branch of electronics where phenomena of a quantum nature play a fundamental role. This book will deal with a particular aspect of quantum electronics, namely, the physical principles of lasers and their behavior. Before going into a detailed discussion of the subject, it seems appropriate to devote a little space to an explanation, in a very simple way, of the ideas behind the laser.

A laser exploits three fundamental phenomena that occur when an electromagnetic (e.m.) wave interacts with a material, namely, the processes of spontaneous and stimulated emission and the process of absorption.

1.1.1. Spontaneous Emission (Fig. 1.1a)

Let us consider two energy levels, 1 and 2, of some given material, their energies being E_1 and E_2 ($E_1 < E_2$). As far as the following discussion is concerned, the two levels could be any two out of the infinite set of levels possessed by the material. It is convenient, however, to take level 1 to be the ground level. Let us now assume that an atom (or molecule) of the material is initially in level 2. Since $E_2 > E_1$, the atom will tend to decay to level 1. The corresponding energy difference $(E_2 - E_1)$ must therefore be released by the atom. When this energy is delivered in the form of an e.m. wave, the process will be called spontaneous (or radiative) emission. The frequency ν of the radiated wave is then given by the expression (due to Planck)

$$\nu = (E_2 - E_1)/h \tag{1.1}$$

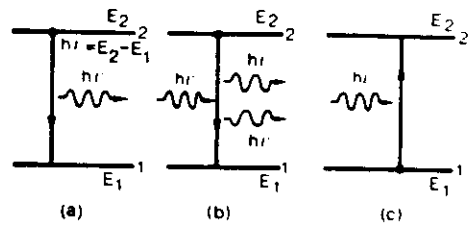

FIG. 1.1. Schematic illustration of the three processes: (a) spontaneous emission; (b) stimulated emission; (c) absorption.

where h is Planck's constant. Spontaneous emission is therefore characterized by the emission of a photon of energy $h\nu = E_2 - E_1$, when the atom decays from level 2 to level 1 (Fig. 1.1a). Note that radiative emission is just one of the two possible ways for the atom to decay. The decay can also occur in a nonradiative way. In this case the energy difference $E_2 - E_1$ is delivered in some form other than e.m. radiation (e.g., it may go into kinetic energy of the surrounding molecules).

The probability of spontaneous emission can be characterized in the following way: Let us suppose that, at time t, there are N_2 atoms (per unit volume) in level 2. The rate of decay of these atoms due to spontaneous emission, i.e., $(dN_2/dt)_{sp}$, will obviously be proportional to N_2. We can therefore write

$$\left(\frac{dN_2}{dt}\right)_{sp} = -AN_2 \tag{1.2}$$

The coefficient A is called the spontaneous emission probability or the Einstein A coefficient (an expression for A was in fact first obtained by Einstein from thermodynamic considerations). The quantity $\tau_{sp} = 1/A$ is called the spontaneous emission lifetime. The numerical value of A (and τ_{sp}) depends on the particular transition involved.

1.1.2. Stimulated Emission (Fig. 1.1b)

Let us again suppose that the atom is found initially in level 2 and that an e.m. wave of frequency ν given by equation (1.1) (i.e., equal to that of the spontaneously emitted wave) is incident on the material. Since this wave has the same frequency as the atomic frequency, there is a finite probability that this wave will force the atom to undergo the transition $2 \rightarrow 1$. In this case the energy difference $E_2 - E_1$ is delivered in the form of an e.m. wave that adds to the incident one. This is the phenomenon of stimulated emission. There is, however, a fundamental distinction between the spontaneous and stimulated emission processes. In the case of spontaneous emission, the atom emits an e.m. wave that has no definite phase relation with that emitted by another atom. Furthermore, the wave can be emitted in any direction. In the case of

stimulated emission, since the process is forced by the incident e.m. wave, the emission of any atom adds in phase to that of the incoming wave. This wave also determines the direction of the emitted wave.

In this case, too, we can characterize the process by means of the equation

$$\left(\frac{dN_2}{dt}\right)_{st} = -W_{21}N_2 \tag{1.3}$$

where $(dN_2/dt)_{st}$ is the rate at which transitions $2 \rightarrow 1$ occur as a result of stimulated emission and W_{21} is called the stimulated transition probability. Just as in the case of the A coefficient defined by (1.2), the coefficient W_{21} also has the dimension of (time)$^{-1}$. Unlike A, however, W_{21} depends not only on the particular transition but also on the intensity of the incident e.m. wave. More precisely, for a plane e.m. wave, it will be shown that we can write

$$W_{21} = \sigma_{21}F \tag{1.4}$$

where F is the photon flux of the incident wave and σ_{21} is a quantity having the dimensions of area (it is called the stimulated-emission cross section) and depending only on the characteristics of the given transition.

1.1.3. *Absorption (Fig. 1.1c)*

Let us now assume that the atom is initially lying in level 1. If this is the ground level, the atom will remain in this level unless some external stimulus is applied to it. We shall assume, then, that an e.m. wave of frequency ν given again by (1.1) is incident on the material. In this case there is a finite probability that the atom will be raised to level 2. The energy difference $E_2 - E_1$ required by the atom to undergo the transition is obtained from the energy of the incident e.m. wave. This is the absorption process.

In a similar fashion to (1.3), we can define an absorption rate W_{12} by means of the equation

$$\frac{dN_1}{dt} = -W_{12}N_1 \tag{1.5}$$

where N_1 is the number of atoms (per unit volume) that, at the given time, are lying in level 1. Furthermore, just as in (1.4), we can write

$$W_{12} = \sigma_{12}F \tag{1.6}$$

where σ_{12} is some characteristic area (the absorption cross section), which depends only on the particular transition.

In the preceding sections, the fundamental principles of the processes of spontaneous emission, stimulated emission, and absorption have been described. In terms of photons, these processes can be described as follows (see Fig. 1.1): (1) In the spontaneous emission process, the atom decays from level 2 to 1 through the emission of a photon. (2) In the stimulated process, the incident photon stimulates the $2 \rightarrow 1$ transition and we then have two photons (the stimulating plus the stimulated one). (3) In the absorption process, the incident photon is simply absorbed to produce the $1 \rightarrow 2$ transition. Finally, it should be noted that $\sigma_{12} = \sigma_{21}$, as Einstein showed at the beginning of the century. This shows that the probabilities of stimulated emission and absorption are equal. From now on, therefore, we will write $\sigma_{12} = \sigma_{21} = \sigma$, and σ will be referred to as the transition cross section.* The number of atoms per unit volume in some given level will be called the *population* of that level.

1.2. THE LASER IDEA

Consider two arbitrary energy levels 1 and 2 of a given material and let N_1 and N_2 be their respective populations. If a plane wave with an intensity corresponding to a photon flux F is traveling along the z direction in the material, the elemental change of this flux due to both the stimulated emission and absorption processes in the shaded region of Fig. 1.2, according to equations (1.3)–(1.6), is given by

$$dF = \sigma F(N_2 - N_1)\, dz \tag{1.7}$$

Equation (1.7) shows that the material behaves as an amplifier (i.e., $dF/dz > 0$) if $N_2 > N_1$, while it behaves as an absorber if $N_2 < N_1$. Now, it is known that, in the case of thermal equilibrium, the energy-level populations are described by Boltzmann statistics. So, if N_1^e and N_2^e are the thermal equilibrium populations of the two levels, we have

$$\frac{N_2^e}{N_1^e} = \exp\left[-\frac{(E_2 - E_1)}{kT}\right] \tag{1.8}$$

where k is Boltzmann's constant and T the absolute temperature of the material. We then see that, for the case of thermal equilibrium, we have $N_2 < N_1$. According to (1.7), the material then acts as an absorber at frequency ν, and this is what happens under ordinary conditions. If, however, a nonequilibrium condition is achieved for which $N_2 > N_1$, then the material will act as an amplifier. In this case we will say that there exists a *population inversion* in

* The discussion here only applies to nondegenerate levels. For the case of degenerate levels the reader is referred to Section 2.8.

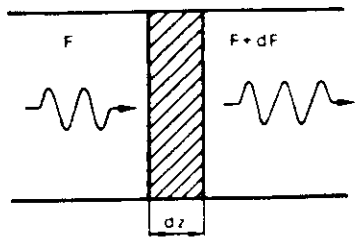

FIG. 1.2. Elemental change dF in the photon flux F for a plane e.m. wave in traveling a distance dz through the material.

the material, by which we mean that the population difference ($N_2 - N_1 > 0$) is opposite in sign to that which exists under ordinary conditions ($N_2^e - N_1^e < 0$). A material having a population inversion will be called an *active material.*

If the transition frequency $\nu = (E_2 - E_1)/h$ falls in the microwave region, this type of amplifier is called a *maser* amplifier. The word "maser" is an acronym for "*m*icrowave *a*mplification by *s*timulated *e*mission of *r*adiation." If the transition frequency ν falls in the optical region, the amplifier is called a *laser* amplifier. The word "laser" is again an acronym, obtained by substituting the letter l (*l*ight) for the letter m (*m*icrowave).

To make an oscillator from an amplifier, it is necessary to introduce a suitable positive feedback. In the microwave range this is done by placing the active material in a resonant cavity having a resonance at the frequency ν. In the case of a laser, the feedback is often obtained by placing the active material between two highly reflecting mirrors (e.g., plane-parallel mirrors, see Fig. 1.3). In this case, a plane e.m. wave traveling in a direction orthogonal to the mirrors will bounce back and forth between the two mirrors and be amplified on each passage through the active material. If one of the two mirrors is made partially transparent, a useful output beam can be extracted. It is important to realize that for both masers and lasers, a certain threshold condition must be fulfilled. In the laser case, for instance, the oscillation will start when the gain of the active material compensates the losses in the laser (e.g., the losses due to output coupling). According to (1.7), the gain per pass in the active material (i.e., the ratio between the output and input photon flux) is $\exp[\sigma(N_2 - N_1)l]$, where l is the length of the active material. If the only losses present in the cavity are those due to transmission losses, the threshold will be reached when $R_1 R_2 \exp[2\sigma(N_2 - N_1)l] = 1$, where R_1 and R_2 are the power reflectivities of the two mirrors. This equation shows that the threshold is reached when the population inversion reaches a critical value $(N_2 - N_1)_c$ known as

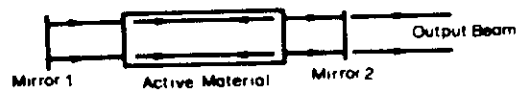

FIG. 1.3. Scheme of a laser.

the *critical inversion* and given by

$$(N_2 - N_1)_c = -\frac{\ln(R_1 R_2)}{2\sigma l} \tag{1.9}$$

Once the critical inversion is reached, oscillation will build up from the spontaneous emission. The photons that are spontaneously emitted along the cavity axis will, in fact, initiate the amplification process. This is the basis of a laser oscillator, or laser, as it is more simply called. Note that according to the meaning of the acronym "laser" as discussed above, the word should be reserved for lasers emitting visible radiation. The same word, is, however, now commonly applied to any device emitting stimulated radiation, whether in the far or near-infrared, ultraviolet, or even in the x-ray region. To be specific about the kind of radiation emitted one then usually talks about infrared, visible, ultraviolet or x-ray lasers, respectively. Note also that one usually refers to a solid-state, liquid, or gas laser depending upon the state of the active material.

1.3. PUMPING SCHEMES

We will now consider the problem of how a population inversion can be produced in a given material. At first sight, it might seem that it would be possible to achieve this through the interaction of the material with a sufficiently strong e.m. field at the frequency ν given by (1.1). Since, at thermal equilibrium, level 1 is more populated than level 2, absorption will in fact predominate over stimulated emission. The incoming wave would produce more transitions $1 \rightarrow 2$ than transitions $2 \rightarrow 1$ and we would hope in this way to end up with a population inversion. We see immediately, however, that such a system would not work (at least in the steady state). When in fact the condition is reached such that the populations are equal ($N_2 = N_1$), then the absorption and stimulated processes will compensate one another and, according to (1.7), the material will then be transparent. This situation is often referred to as two-level ***saturation.***

With the use of just two levels 1 and 2, it is impossible therefore to produce a population inversion. It is then natural to question whether this is possible by some suitable use of more than two levels out of the infinite set of levels of a given atomic system. As we shall see, the answer in this case is positive, and we will accordingly talk of a three- or four-level laser, depending upon the number of levels used (Fig. 1.4). In a three-level laser (Fig. 1.4a), the atoms are in some way raised from the ground level 1 to level 3. If the material is such that, after an atom has been raised to level 3, it decays rapidly to level 2, then in this way a population inversion can be obtained between levels 2 and 1. In a four-level laser (Fig. 1.4b), atoms are again raised from the ground level (for convenience we now call this level 0) to level 3. If the atom then

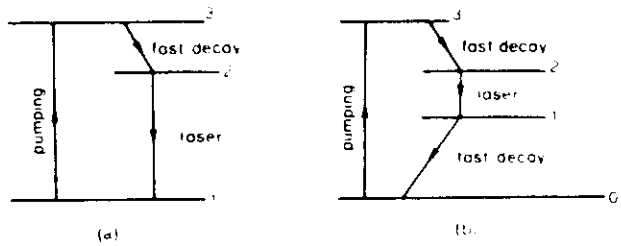

FIG. 1.4. (a) Three-level and (b) four-level laser schemes.

decays rapidly to level 2, a population inversion can again be obtained between levels 2 and 1. Once oscillation starts in such a four-level laser, however, the atoms will then be transferred to level 1 (because of stimulated emission). For continuous wave (cw) operation of a four-level laser it is necessary, therefore, that the transition $1 \rightarrow 0$ should also be very fast.

We have now seen how one can use three or four levels of a given material to produce population inversion. Whether a system will work in a three- or four-level scheme (or whether it will work at all!) depends on whether the various conditions given above are fulfilled. We could of course ask why one should bother with a four-level scheme when a three-level scheme already seems to offer a suitable way of producing a population inversion. The answer is that one can, in general, produce a population inversion much more easily in a four-level than in a three-level laser. To see this, we begin by noting that the energy differences between the various levels of Fig. 1.4 are usually much greater than kT. According to Boltzmann statistics [see, e.g., equation (1.8)] we can then say that essentially all atoms are initially (i.e., at equilibrium) in the ground level. If we now let N_t be the total number of atoms per unit volume of material, these will initially all be in level 1 for the three-level case. Let us now begin raising atoms from level 1 to level 3. They will then decay to level 2, and if this decay is sufficiently fast, level 3 will remain more or less empty. In this case, we first have to raise half of the total population N_t to level 2 in order to equalize the populations of levels 1 and 2. From this point on, any other atom that is raised will then contribute to population inversion. In a four-level laser, however, since level 1 is also initially empty, any atom that has been raised is immediately available for a population inversion. The above discussion shows that, whenever possible, we should look for a material that can operate as a four-level system rather than as a three-level system. The use of more than four levels is, of course, also possible.

The process by which atoms are raised from level 1 to level 3 (in a three-level scheme) or from 0 to 3 (in a four-level scheme) is known as *pumping*. There are several ways in which this process can be realized in practice, e.g., by some sort of lamp of sufficient intensity or by an electrical discharge in the active medium. We refer the reader to Chapter 3 for a more detailed discussion of the various pumping processes. We note here, however, that, if the upper

pump level is empty, the rate at which the upper laser level 2 becomes populated by the pumping, $(dN_2/dt)_p$, can in general be written as

$$\left(\frac{dN_2}{dt}\right)_p = W_p N_g \tag{1.10}$$

Here N_g is the population of the ground level (i.e., level 1 or 0 in Figs. 1.4a and 1.4b, respectively) and W_p is a coefficient that will be called the *pump rate*. To achieve the threshold condition, the pump rate must reach a threshold or critical value that we shall indicate by W_{cp}. Specific expressions for W_{cp} will be obtained in Chapter 5.

1.4. PROPERTIES OF LASER BEAMS

Laser radiation is characterized by an extremely high degree of (1) monochromaticity, (2) coherence, (3) directionality, and (4) brightness. To these properties a fifth can be added, viz., (5) short time duration. This refers to the capability for producing very short light pulses, a property that, although perhaps less fundamental, is nevertheless very important. We shall now consider these properties in some detail.

1.4.1. Monochromaticity

Briefly, we can say that this property is due to the following two circumstances: (1) Only an e.m. wave of frequency ν given by (1.1) can be amplified. (2) Since the two-mirror arrangement forms a resonant cavity, oscillation can occur only at the resonant frequencies of this cavity. The latter circumstance leads to the laser linewidth being often much narrower (by as much as six orders of magnitude!) than the usual linewidth of the transition $2 \rightarrow 1$ as observed in spontaneous emission.

1.4.2. Coherence

To first order, for any e.m. wave, one can introduce two concepts of coherence, namely, spatial and temporal coherence.

To define spatial coherence, let us consider two points P_1 and P_2 that, at time $t = 0$, lie on the same wave front of some given e.m. wave and let $E_1(t)$ and $E_2(t)$ be the corresponding electric fields at these points. By definition, the difference between the phases of the two fields at time $t = 0$ is zero. Now, if this difference remains zero at any time $t > 0$, we will say that there is a perfect coherence between the two points. If this occurs for any two points of the e.m. wave front, we will say that the wave has *perfect spatial coherence.* In practice, for any point P_1, the point P_2 must lie within some finite area