## CONTENTS

### Chapter 1. Basic principles

**A. Introduction**

**B. Development of the gas laser**

**C. Fundamental properties and definitions**

1. Output wavelength
2. Modes—spatial and temporal
3. Doppler width
4. Frequency separation between modes
5. Gain, loss, and useful output
6. “Hole burning”
7. Homogeneous and inhomogeneous broadening
8. Near-field and far-field optical patterns

**D. Physics of gas laser operation**

1. Stimulated emission versus spontaneous emission
2. Gain per unit length
3. The Bloch equations
4. Steady state equations of laser operation
5. Zeeman effect
6. Excitation mechanisms in gas lasers

### Chapter 2. Characteristics of practical gas lasers

**A.** Ranges of the operating characteristics of gas lasers listed according to type

1. Wavelength range in which useful transitions are found
2. Output power and gain in lasers of laboratory size 48
3. Efficiency in converting input electrical energy into laser light 50
4. Doppler width and coherence of laser output 51

B. Physical and engineering details of the various lasers 52
1. Helium-neon 52
2. Argon ion 60
3. CO₂ 65
4. Summary 67

Chapter 3. Modes and resonators 69
A. Spatial modes 69
1. Types of spatial mode. Importance of lowest-order mode 69
2. The equations of Boyd, Gordon, and Kogelnik 73
3. Practical resonator configurations and stability 74
4. Aligning a laser resonator 78
5. Conditions for operating in a single spatial mode 82

B. Temporal modes 85
1. Rules governing frequency spacing 85
2. Conditions for operating in a single temporal mode 86
3. The Lamb equations and the power dip 89
4. Schemes for stabilizing the frequency of a single-mode laser 93
5. Ring resonators 98

Chapter 4. The gas laser considered as a light source or local oscillator 101
A. Properties of an “ideal” Gaussian laser beam 102
1. Propagation as a plane or spherical wave 102
2. Effective brightness of source and focused spot 107

B. Effects of truncation and aberration on a focused Gaussian beam 113

C. Noise in the laser beam and in detectors 115
1. Theoretical noise limit 115
Contents

2. Extraneous noise sources—plasma and mode interaction noise 118
3. Detectors 121
D. Control of laser coherence by internal cavity perturbation
   1. Mode selection by interferometry 123
   2. AM phase locking 127
   3. FM mode coupling 130
   4. Application of internal perturbation methods to high-power single-frequency lasers 133

Chapter 5. Problems associated with laser applications 138
A. Long-distance metrology 138
B. Metrology by interferometry 141
   1. Reproducibility of present wavelength standards 141
   2. Problems in the coupling of interferometers to lasers 143
C. Data processing and holography 145
D. Optical testing 148
E. Scattering, including Raman and Brillouin scattering 149
F. High energy density 150
G. Safety problems associated with working with gas lasers
   1. Visible laser intensities of one to one hundred microwatts 152
   2. One to one hundred milliwatts 153
   3. One to ten watts, visible, near infrared, or CO₂ laser radiation 153
   4. One hundred watt beams, ion lasers, and CO₂ lasers 153

Bibliography 155

Appendix A. Selected list of important gas laser transitions 160
Appendix B. Exact wavelength of stabilized helium-neon laser 164
Author index 167
Subject index 169
CHAPTER 1

BASIC PRINCIPLES

A. INTRODUCTION

The gas laser is one type of device that is capable of generating coherent electromagnetic radiation at wavelengths shorter than those generally considered to be microwave frequencies. The other devices that have this property employ, as an active medium, crystalline solids, glasses, liquids, or semiconductors. The gas laser, however, has properties that are substantially different from those of the other types of lasers; thus it is desirable to treat them as a separate entity. Throughout this book, “laser” will refer solely to the gas laser.

We shall further restrict our discussion of gas lasers to only those that are excited or “driven” by electric discharge. The reason for this is that discharge-driven lasers are the most common gas lasers, and at present they are the only type available commercially. Experiments have been performed on gas lasers that are driven by optical pumping mechanisms similar to those employed in solid and liquid lasers; however, the experiments have not been pursued and there is little interest at the present time in exploring such lasers further. The reader who is interested in optical pumping of gas lasers can consult texts such as Smith and Sorokin [2] or the first comprehensive review paper of Bennett [8]. Another type of gas laser excitation that is of current research interest is excitation by chemical reaction, either photodissociation or rapid mixing of individual gaseous constituents. This topic, however, is not covered in the present book. In any event, the optical properties of such lasers are essentially identical to those of the discharge lasers, which are discussed in here.

Basically, a gas laser consists of a gas whose atoms or molecules are raised to higher energy states, and an “inversion” exists between a certain pair of energy levels in the gas. To understand the significance of
the inversion process, it is desirable to bear in mind the equilibrium conditions that normally exist among the energy levels of the gas that is maintained at a particular absolute temperature \( T \). A gas and temperature equilibrium, whether or not a discharge is maintained through it, has a population distribution among the various energy levels that is described as follows. Let the energy difference between any pair of levels be \( \Delta E \); then the ratio of the populations in the two levels is given by

\[
\frac{P(E + \Delta E)}{P(E)} = e^{-\Delta E/kT}.
\]

In all such cases, for positive values of the temperature \( T \), the upper energy level has less population than the lower energy level. This condition is required by the laws of thermodynamics. If incident radiation having an energy corresponding to the energy difference \( \Delta E \) between the levels is incident upon the gas, it will be absorbed, the absorption taking place by raising some of the atoms (or molecules)* in the lower state to the upper state. Spontaneous emission processes as well as thermal processes then remove atoms from the upper state back to the lower state and help maintain the temperature equilibrium of the gas.

However, under certain unusual conditions that may exist in an electric discharge, it may be possible to produce a situation in which, between a certain pair of energy levels, the upper level has a greater population than the lower level. This situation is referred to as an "inversion." When this happens, incident radiation, rather than being absorbed by the gas, will stimulate atoms in the upper state to release their energy faster than they would by spontaneous emission and actually contribute more energy to the incident beam than was there originally. This condition is analogous to the gain of an amplifier, and thus the gas discharge medium in a laser can be considered an amplifier of optical radiation. Further details of the physics involved in the laser emission from an inverted population are discussed under the topic of Physics of Gas Laser Operation (pp. 25-29).

As is well known, any oscillator can be turned into an amplifier by providing a suitable type of feedback. In a gas laser the feedback is provided by means of highly reflecting mirrors at either end of an optical path traversing the gas discharge medium. Thus energy that can be employed to start the oscillation begins somewhere in the optical path, traverses the gas, is amplified in the process, and returns eventually to

*The particles in the gas that interact with the laser radiation may be individual atoms, ions, or molecules, depending on the specific laser. In the remainder of the book, we frequently use "atoms" merely as shorthand to refer to whichever type of particle is appropriate.
Introduction

its original starting point; it is then amplified further, and so on, repeating the process, and this generates the amplification loop that is common to all oscillators. Another feature common to all steady-state oscillators is the fact that some sort of saturation condition must exist in order that the amplified wave does not built up to an infinite value. In a vacuum tube amplifier this saturation condition is provided by the characteristics of the vacuum tube (or other electronic device) that is employed in the circuit. In the gas laser, in fact, in all lasers, the saturation characteristic is provided in the details of the stimulated emission process, which determine in detail how the law of diminishing returns has to apply with respect to injection of additional amounts of energy into the oscillation as the energy content itself grows larger. Eventually, of course, a final limit has to be reached, determined by the rate at which the electric discharge pumps excited atoms into the upper state, which is responsible for the laser inversion itself. A problem in vacuum tube oscillators, which is apparently not present in lasers, may be that of starting the oscillation in the first place. In the laser, the spontaneous emission from the excited state in itself is always sufficient to provide an initial noise signal capable of starting the discharge. As is pointed out in Chapter 3, spontaneous emission noise is always present to some small extent even in a laser operating at high power, and it may be a contributing factor in the noise fluctuations of the laser output.

In summary, then, the gas laser may be thought of as a gaseous medium excited by electric discharge and containing within it a closed optical path in which optical energy can be contained for relatively long periods of time. Because the gain per unit length of most gas lasers is not unduly high, it is desirable in most cases to provide the optical resonator along the path of greatest length through the discharge. Thus a laser in its simplest form will look like Figure 1, which shows the bare essentials that are common to all gas lasers. These are (a) a long cylindrical tube containing the gaseous medium, (b) a means for exciting the discharge in the medium, and (c) a pair of mirrors facing each other,

Figure 1. Essentials of a gas laser.
Basic Principles

Figure 2. Brewster's angle window.

which constitutes the “resonator” for the laser energy. Often windows of suitable type are inserted between the mirrors and the gaseous medium, though this is not essential to the basic operation of the device. When windows are used, they are generally Brewster angle windows, illustrated in Figure 2. When the angle $\theta$ between the perpendicular of the windows and the laser axis is placed so that

$$\tan \theta = n,$$  \hspace{1cm} (2)

where $n$ is the index of refraction of the window, no reflection can take place at the surface of the window for the polarization shown in the figure. The other linear polarization suffers a high reflection and therefore gets attenuated too much to provide for laser oscillation, but this does not matter since oscillation in one polarization will effectively remove all of the available energy from the excited atoms provided that coupling exists in the atomic system within the upper and lower state for that particular polarization. In practice this is almost always the case, and the use of Brewster angle windows has the additional advantage that the output is plane-polarized, which is highly useful for many applications. The use of Brewster windows, in general, has been highly successful; Brewster windows when properly manufactured and oriented relative to the laser beam provide a negligible loss of laser energy, generally less than absorption and scattering losses in the mirrors themselves.

Although there is no substitute for actual laboratory handling of a particular type of gas laser in order to understand its properties and characteristics, the following series of photographs may give some indication of the general nature of any given type of gas laser. Figure 3, which was taken with helium-neon lasers, shows the general size and structure of all neutral gas lasers. They are typically of the order of 1 m
Figure 3. (a) A laboratory-setup helium-neon laser showing, in detail, the plasma tube with Brewster angle windows and the resonator mirror mounts. (b) A large commercial laser with the cover removed. The rectangular structures under the plasma tube are magnets. (Photographs courtesy of Spectra-Physics, Inc.)
in length but may vary from 15 cm for very short single-wavelength lasers to 2 m or more for high power output. In the helium-neon laser the color of the discharge is somewhat more pale than that of a typical neon sign, owing to the presence of helium, which produces its spectrum lines in approximately the same intensity as those of the neon, which is the lasing material. The pressure at which neutral gas laser discharge tubes operate is somewhat lower than that of rare gas discharge tubes used for other purposes—approximately 1 to 2 torr for the laser versus 10 torr for other types of discharge tubes. The current densities employed are about the same—approximately 100 mA/cm$^2$ in dc excitation and an equivalent electron density in RF excitation—but all neutral gas lasers are subject to saturation effects if the current density is too high, and in some cases current densities are not as high as could be used merely for the purposes of exciting the emission spectrum. Both RF and dc excitation can be used with this type of laser, dc excitation usually being preferred because of its simplicity and relative lack of electrical interference with nearby laboratory equipment.

Figures 4 and 5 show two different types of noble gas ion lasers. These lasers are characterized by the fact that they must operate at very high current densities, often of the order of 300 to 400 amp/cm$^2$. For this reason, the plasma itself is at a very high temperature and water cooling or another special cooling means is required to maintain the integrity of the laser structure. The size of the laser structure is typically of the order of 1 m in length, although considerably shorter structures are possible, as shown in Figure 5. However, the requirements for cooling means and the attachments to the laser bore that provide the required current density—either electrodes for dc operation or coupling means for RF operation—also require a considerable breadth in the structure as well as weight. An additional factor contributing to the overall size and weight of the structure is the inclusion of a solenoid for generating an intense magnetic field along the axis of the laser tube, which generally provides a considerable increase in efficiency of the laser operation. Not shown in the photographs are the power supplies that must accompany each laser; since the required power input is often of the order of many kilowatts, this power supply in itself is quite massive and generally requires a special power source (three-phase power, for example) rather than the single-phase 110-V power which is available at most outlets. Thus, although the high-power ion laser does not represent a particularly large installation in terms of length, it does represent a considerable investment in power and in weight.

The method of introducing RF into an RF-driven ion laser is quite different from that of the simple capacitive electrode structures used in
Figure 4. (a) Plasma tube of a high-power argon ion laser excited by RF induction. The cooling-water jacket and magnetic field solenoid that generally accompany such a structure are not shown. (b) Commercial argon ion laser. The plasma tube in this laser is similar, though not identical, to that of Figure 4a. (Photographs courtesy of Spectra-Physics, Inc.)
the other types of lasers. Figure 6 shows schematic outlines of both dc- and RF-driven ion lasers, and it will be noted that the RF ion laser has to be coupled magnetically, as the secondary of a transformer, to the driving RF [9]. This type of coupling is permissible and efficient because the ion laser has such a high current density that the impedance per unit length of the discharge is low and therefore it constitutes a highly conducting current loop. An additional advantage of the RF-excited ion laser, which may possibly be useful in the future, is the fact that absence of any electrodes within the gas permits use of reactive materials as the laser medium. Figure 7 shows those elements that have been observed in CW ion laser action. Of these, chlorine, sulphur, and phosphorus have been known to generate high-power visible output in various parts of the visible spectrum under suitable conditions. These materials cannot possibly be used in a dc-excited ion laser because of their reactive chemical effects on electrodes.

The physical structure of the molecular laser is similar to that of the neutral gas laser, with two exceptions. First, since the wavelengths of
Figure 6. (a) Schematic diagram of dc-excited ion laser. (b) RF ring excitation of ion laser. (From [1].)
most molecular lasers are in the far infrared, well out of range of materials that are suitable for Brewster windows, it is almost always necessary to design such lasers with internal mirrors. Second, since the energy separations between transitions in the far infrared are of the same order

Figure 8. High-power CO₂ laser.
Introduction

of magnitude as thermal energy, it may be necessary to keep the gaseous medium cool in order to prevent upsetting the population inversion from ordinary thermal collisions. This is a particular problem in the CO₂ laser, where it has been shown that the efficiency of operation of the laser depends strongly upon the temperature of the walls of the discharge tube, cold walls being essential for efficient operation. Figure 8 shows a photograph of a high-power CO₂ laser in the laboratory. Note that long length is essential to high power, since it appears that power output is almost directly proportional to length and not to volume. Second, the walls of the discharge are water cooled; cooling with liquid nitrogen or some other refrigerated medium would be even better. The output of this particular laser was of the order of 50 W.

Pulsed lasers—particularly those in which a transient response from the gas is the only type of laser action expected—come in many shapes and sizes but generally have the same appearance as neutral gas lasers. There are two notable exceptions. A hollow cathode type of structure has been used to study ionized mercury, and this appears to have some promise for future applications inasmuch as it has the exceptional ability to operate with a very large laser bore diameter. Figure 9 illustrates this type of laser construction [10]. The other notable structure is one that has been used to obtain a high peak power pulsed output from nitrogen gas [11]. This is shown in Figure 10. It consists of many paral-

![Figure 9. Hollow cathode laser structure for use with mercury vapor. (From [10].)
lel electrodes placed so that the high electric field is across the axis of the tube. By this means, very high fields can be obtained with only "moderate" voltages applied to the electrodes (in at least one experiment, "moderate" was 20 kV). This type of structure probably produces the highest available peak ultraviolet power at the present time.

Other than the particular structures shown in Figures 9 and 10, pulsed lasers usually employ electrodes at either end of the tube similar to that of dc-excited neutral gas or molecular lasers. Excitation is then achieved by discharging a capacitor between the electrodes so that very high peak currents and voltages are available for periods of time of the order of 10 nsec to 1 \( \mu \)sec. RF excitation has been used on occasion, but since a certain number of cycles of the RF are required in order to define the nature of the input voltage, and since the RF is usually of the order of 10 MHz or less, the resultant pulses are relatively long compared to what is usually required for transient laser action, and the output is much more akin to that of a CW laser which is simply pulsed on for a short period of time.

**B. DEVELOPMENT OF THE GAS LASER**

The first concrete proposal for a gas laser is probably that of the historic paper of Schawlow and Townes [12], published in 1958. This paper, which bases its theory on the then-existing knowledge of stimu-

![Figure 10. Crossed-field laser structure to obtain very short duration high-voltage pulses. (Reproduced by permission from [11].)](image-url)
lated emission as related to microwave masers, proposed an optically pumped gas laser as a means of demonstrating the principle of laser action. The paper, considered to be the stimulus for all initial discoveries of lasers of all types, turned out to be surprisingly accurate in predicting the size, wavelength range, and output power for typical gas lasers having moderate input power excitation requirements, though it did not propose the use of a gas discharge for excitation. The proposal for employing a discharge to achieve population inversion and consequent laser action was made by, among others, A. Javan, then of the Bell Telephone Laboratories, about early 1960. Based on a detailed study of excitation processes and lifetimes in excited neon, he predicted possible laser action on the transition $2\Sigma_5 \rightarrow 2\Pi_9$ at $1.1177 \mu$. The first operating gas laser was, in fact, constructed by Dr. Javan and his associates, W. R. Bennett, Jr., and D. Herriott [13]. While this laser did oscillate on the predicted line, $1.1177 \mu$, it was found that much stronger oscillation was obtained on adjacent lines in the vicinity of 1.15 to 1.2 $\mu$. In fact, the $1.1177 \mu$ line has very rarely been seen from a helium-neon laser other than the first one; the author has observed it in some of his experiments, and undoubtedly it has been observed by others as well, but it is known to be a very difficult line to excite.

This point is made in regard to the relation of theories of population inversion to the actually observed results. The history of the discovery of gas laser transitions has been one of extremely poor agreement between theories and subsequent experimental discoveries. The aforementioned example is one of the few in which a theory was actually able to predict the existence of a subsequently observed laser transition; another particularly successful one was the prediction by C.K.N. Patel [14] of the enhancement of CO$_2$ laser action at 10.6 $\mu$ through interaction with nitrogen. For the most part, however, the theoretical approach has not been the one that has resulted in the discovery of the many existing laser transitions. Most of these transitions were discovered by the simple expedient of preparing a laser structure according to the scheme of Figure 1, filling it with some gas at pressures of the order of 1 torr, employing mirrors that were highly reflective within some wavelength range, and “looking” to see if any laser light came out of the structure.

On the other hand, there is one area of laser research in which theory has been particularly helpful—in fact, essential—to the further development of the technique. This is the theory of optical resonator structure and modes. The original paper of Schawlow and Townes pointed out that the concept of modes as applied to a microwave cavity would not be appropriate in the optical region, and it suggested that the resonator
should perhaps consist of a Fabry-Perot type of structure, that is, a pair of flat mirrors enclosing the laser medium and directed toward each other. The exact nature of the modes of this structure was left open to question, and at least some investigators in the field felt that a “mode” of such a resonator might correspond to the “rings” or fringes that are observed when one looks at an incoherent light source through a Fabry-Perot interferometer. This partially misconceived notion persisted throughout the early work until the actual discovery of the gas laser, and the author can vouch for the fact that this misinterpretation of the true nature of an optical resonator mode provoked some lively discussions at scientific meetings around 1960.

It was not until after the discovery of the gas laser that the correct theory of the optical resonator began to emerge. Fox and Li [15] employed numerical integration on a computer to study energy storage in a plane-parallel resonator, and subsequently Boyd and Gordon [16] and Boyd and Kogelnik [17] did corresponding calculations for resonators employing curved mirrors. The results of this work form the basis for a good part of Chapter 3 of this book. Its importance in promoting the development of gas lasers after the initial discovery cannot be underestimated. Until this theoretical work was done, the gas laser was at best a marginally operating device whose oscillation depended upon almost impossibly precise tolerances in the end mirror adjustments. The theoretical studies on curved mirror resonators showed that resonators could be devised that were relatively insensitive to mirror adjustment and whose intrinsic losses could be much lower than those of a plane-parallel resonator, allowing observation of laser action in media with much lower gain than was heretofore thought possible. The plane-parallel resonator has since almost dropped out of existence for practical laser work, and all discoveries of new gas laser transitions have been performed with curved mirror resonators.

The next major experimental step was the discovery of visible laser output at 6328 Å by White and Rigden [18] in the spring of 1962. Besides providing a great stimulus for applications, the visible laser made possible a much greater understanding of the problems involved in optimizing laser structures and in demonstrating precisely what the difference was between a coherent and an incoherent beam of light. The work also stimulated additional research on other possible laser transitions and many more were found in neutral atom lasers whose structures were akin to that of the helium-neon laser, but all of these transitions were in the infrared. Such experiments continued for a period of about two years, during which considerable engineering development was done on the helium-neon lasers.
In 1963 the significance of employing pulsed operation to obtain laser transitions that could not be observed continuously was demonstrated by several investigators. Mathias and his co-workers experimented with molecular nitrogen and carbon monoxide, obtaining visible transitions from carbon monoxide [19]. However, the most important of these experiments was probably the discovery by Bell in the fall of 1963 of pulsed ion laser action in mercury vapor [20]. The discovery that laser action could be obtained from the ionic spectra of atoms suggested a whole new range of possibilities for laser operation in various materials, and a large number of investigators immediately shifted their efforts to the study of the ion lasers. An indication of the extent to which this occurred is the fact that the discovery of the argon ion laser, the most important type of ion laser, was made independently and almost simultaneously in at least four laboratories around the world [21]. Observation that argon and other noble gases could operate as ion lasers continuously as well as pulsed was made soon thereafter [22].

By using a combination of the successful techniques—weak glow discharges such as the helium-neon laser, pulsed discharges, and high current discharges such as are used in the ion lasers—investigators have come up with a total of approximately 1000 laser transitions in various gases. Although the rate at which new lines are being discovered has slowed down considerably in the last year or two, additional lines are still reported from time to time in the scientific journals.

The most recent important development in the laser field is probably the discovery that the CO₂ laser at 10.6 µ is capable of operation at undreamed-of power and efficiency. The present field of laser research appears to be aimed in two general directions: (a) efforts to raise the power and efficiency of existing types of laser, and (b) attempts to obtain a deeper knowledge of the basic physics of such lasers as the helium-neon lasers. Particularly to be noted in this connection is work on subtleties such as details of the line shape in lasers and the effects of weak magnetic fields on their operation. On the other hand, there is no reason not to expect additional important discoveries comparable to that of the CO₂ laser to occur in the near future.

The development of the gas laser up to the present time has suggested a broad classification of all lasers into four general types. This classification, which will be used freely in the text, and the general characteristics of lasers within each category, are as follows:

1. Neutral atom lasers. Lasers employing relatively weak discharges and having moderate power output and gain, the laser transition spectra being that of neutral atoms in the discharge.

2. Ion lasers. Lasers operating on transitions in ionized atoms in the
Basic Principles

16

These lasers generally operate at high current densities, high power inputs, and occasionally with relatively high power output.

3. Molecular lasers. Lasers somewhat similar to the neutral atom lasers but employing molecular spectra. They employ moderate current densities in the discharge but can occasionally be operated with large volumes and hence high input power. Many but not all lasers of this class emit in the far infrared.

4. Pulsed lasers. A heterogeneous group that overlaps into the other three classes but which includes many lasers that cannot be made to operate continuously for a number of reasons. Many lasers of this class have relatively high peak power output and gain, but because of the required pulsing conditions their average power is not particularly high.

The detailed characteristics of specific lasers in the first three classes are discussed in Chapter 2.

C. FUNDAMENTAL PROPERTIES AND DEFINITIONS

We define here terms that are used throughout the book to describe properties that are fundamental to the operation of any gas laser.

1. Output Wavelength

All gas lasers have the property, not shared to any great extent with the other types of laser, that the wavelengths of laser emission are precisely defined and occur only at certain very discrete parts of the spectrum; they are not “tunable” to any extent. Appendix A lists the wavelengths, in microns, of a number of representative gas laser transitions. All output wavelengths of all gas lasers occur at positions represented by spectrum lines in the emission spectra of the particular gas used in the laser. However, experience has shown that the spectrum lines that are observed in laser emission are very rarely the strongest lines seen in spontaneous emission; in fact, a number of laser lines represented the first experimental observation of the spectrum line involved. The reason for this lack of coincidence between laser transitions and strong spectrum lines is inherent in the physics of laser operation; laser operation demands the generation of a population inversion or “piling up” of an atomic population in the upper of the two energy levels that are involved in the laser transition. On the other hand, a strong spectrum line is strong precisely because there is a high probability that an atom that is deposited in the upper energy level will decay spontaneously at a rapid rate to the lower energy level.

A word should be said here regarding the notations used for the energy levels that give rise to laser transitions. A complete list of such
levels for atoms and ions is given in the work of C. E. Moore-Sitterly [23], now regarded by most investigators in the field as standard. We shall not have much occasion to use energy level notation in this book, but when it is used it will generally follow the notation of Moore. The helium-neon laser has, by convention, come to be described in a somewhat empirical notation known as the “Paschen” notation. This notation was developed before there was a complete understanding of the atomic structure of the neon atom, and it is used because it is simple to write and in any event possesses a 1 to 1 correspondence with the actual energy levels of the neon atom. For present purposes, it should be considered merely a code for designating the neon energy levels.

2. Modes—Spatial and Temporal

The end mirrors of a laser define a resonator, and as in all resonators there exist resonances or normal modes, field configurations that represent energy storage within the resonator and are self-sustaining within the resonator for relatively long periods of time. The modes of an optical resonator are similar to those of a microwave resonator in that they represent configurations of the electromagnetic field that are determined by boundary conditions. Here, however, much of the similarity ends. A microwave resonator typically is a box whose dimensions are defined by the coordinates \(x, y, \) and \(z\), and a mode may be defined by the symbol \(\text{TEM}_{mnq}\) where the numbers \(m, n,\) and \(q\) define the number of nodes in the \(x, y, z\) directions, respectively, of the standing wave within the resonator. A similar notation can be adopted for the optical resonator of a laser, with the \(z\) direction defined as the axis of the laser, that is, the line joining the two end mirrors. However, in a laser resonator mode \(\text{TEM}_{mnq}\) the number \(q\) defines the number of half-wavelengths of light between the two end mirrors and typically is a large number, of the order of \(10^6\). The numbers \(m\) and \(n\), on the other hand, are typically very small, on the order of 0, 1, or 2. The situation is illustrated graphically in Figure 11. Because of the essential asymmetry between the axes \(xy\) and the axis \(z\), and the large disparity in mode numbers, the properties of a mode configuration in the \(x,y\) directions are largely independent.

![Figure 11](image_url)  
*Figure 11. A schematic “snapshot” of the E-vector in a laser resonator mode at a particular instant of time. The example is shown for mode number \(n = 1\). The mode number \(q\) is, in practice, much larger than is indicated here.*
of the configuration in the $z$ direction, and vice versa. Thus it is common to speak of the configuration in the $x$, $y$ directions as the spatial or transverse mode configuration, and, quite independently, to speak of the $z$ direction as the longitudinal or temporal mode configuration. The reason for the word "temporal" will become clear shortly. For simplicity, we shall immediately drop the use of the word "configuration" when speaking of the resonator modes.

3. Doppler Width

A property fundamental to all spectra emitted by gas discharges is the shifting in frequency or wavelength caused by the motions of the emitting atoms. Although thermal motion takes place in three dimensions, only motion parallel to the line of sight is important in causing Doppler shifts, transverse motion causing only an exceedingly small second-order effect. If the motion of the atoms in the discharge is completely Maxwellian—completely thermal in nature—then the intensity distribution caused by the Doppler effect is given by

\[ I(\nu + \Delta \nu) = I_0 \exp \left( -\frac{Mc^2 \Delta \nu^2}{2\nu^2 kT} \right) \]

where

\[ \delta = \frac{2kT\nu}{Mc^2} \]  

Here $\nu$ is the unshifted emission frequency, $\Delta \nu$ the amount of shift, $I_0$ the central intensity of the Doppler-broadened line, $M$ the atomic or molecular mass, $c$ the speed of light, and $kT$ the thermal energy at the temperature $T$. We shall define $\delta$ to be the Doppler width of the line, that is, the frequency spread in which the line intensity varies from peak to $1/e$ of its peak intensity. It is customary in most spectroscopic literature to define the Doppler width as the spread between the two half power points of the distribution. This quantity is $1.6651 \times \delta$, as defined above.

The Doppler width of the emission line that gives rise to laser action is of fundamental importance in determining the operating characteristics of the laser. First, the gain of a laser is determined by, among other factors, the number of atoms in the laser that are capable of interacting with radiation having a specified frequency. If the total number of excited atoms in the laser have their interaction frequencies spread out over a range, then clearly there are fewer atoms available to interact at any given frequency. Second, the Doppler width determines a range of frequencies over which laser operation can actually occur, within any given laser transition. As an approximate rule, the intensity of laser
action at some frequency $\nu$ is proportional to that given by (3). The primary factors that enter into this distribution and the width $\delta$ are the mass $M$ and the temperature $T$. This assumes, of course, that the distribution is in fact Maxwellian with a well-defined temperature $T$. One would expect in some lasers, particularly the ion lasers, that other effects such as acceleration by fields would modify the distribution to something other than Maxwellian, but in the cases that have been investigated it appears that this is not so [24]; rather, a well-defined temperature $T$ can be defined in all cases.

4. Frequency Separation between Modes

If a laser is operating in one spatial mode but in more than one temporal mode, then the different temporal modes must be at different frequencies. This is the reason for use of the word "temporal" in this connection. The frequencies at which temporal modes can occur are set by the boundary conditions at the end mirrors, which determine that laser oscillation can occur only when there are an integral number of half wavelengths between the mirrors. In a mode denoted by the notation $\text{TEM}_{mnq}$, the number $q$ denotes the nodes in the electric field between the end mirrors. The number of half wavelengths in the resonator is then either $q$ or $q + 1$, depending in detail on the spatial mode involved (a question that is immaterial at this point). If $q'$ is an integer that denotes either $q$ or $q + 1$, as appropriate, then the wavelength of the laser oscillation is given by

$$\lambda = \frac{2L}{q'}$$

and the frequency is therefore given by

$$\nu = \frac{cq'}{2L}.$$  

The frequency separation between modes corresponding to $q'$ and $q' \pm 1$ is thus given simply by

$$\Delta \nu = \frac{c}{2L},$$

where $c$ is the mean speed of light in the laser medium and $L$ the length of the resonator (in fact $L$ is an "effective" length taking into account possible wavelength changes because of refractive index effects). The quantity $c/2L$ is thus fundamental to many details of laser operation. It is also, incidentally, the inverse of the time required for a light signal to travel a complete round trip through the laser starting at any given point and bouncing between the two mirrors.

Although the quantity $c/2L$ can be calculated easily enough, it is im-
important to have a "feel" for the magnitude of the $c/2L$ frequency spacing in the following discussions. Table 1 gives the $c/2L$ frequency spacing for typical resonator lengths used in gas lasers.

When the laser is operating in several spatial modes as well as in several temporal modes, then the frequency allocations of the various modes become a much more complicated matter. This is discussed in Chapter 3.

5. Gain, Loss, and Useful Output

The laser, like any classical resonant oscillator, can have its operation characterized in terms of the gain of the medium, the internal losses, and the useful output.

The gain of the medium can best be considered by treating it as an untuned amplifier. Such an amplifier would consist of the laser medium alone in the discharge tube, without the end mirrors; thus one could insert an optical beam or "signal" at one end and receive an amplified signal from the other end. Consider a weak signal inserted into a laser amplifier at some frequency $\nu$ in the vicinity of a laser transition. If the gain of the laser is moderate, the increase in signal intensity as a function of $\nu$, as $\nu$ is varied, will simply follow the Doppler curve given by (3). [Strictly speaking, the differential gain per unit length obeys (3); the total gain throughout the length of the laser tube constitutes the integral of the gain over the length, which does not differ significantly from (3) unless the gain is very high, in which case it has a sharper central peak than (3) would indicate.] In any event, the maximum gain for a weak signal, inserted into a laser amplifier, occurs at the center of the Doppler curve. This amount of gain is generally known as the peak unsaturated gain, and it is this quantity we refer to whenever we speak of "gain," unless otherwise specified. If a laser is in oscillating condition, and one attempts to amplify an external signal through the laser medium simultaneously, then the gain for this signal is defined as the saturated gain, and this is generally a considerably smaller number than the unsaturated gain. The difference between unsaturated and saturated gain is important primarily in discussions regarding competition effects between modes in lasers that are operating simultaneously in several modes.

The losses in a laser are any media that remove energy from the laser resonator. Typical sources of loss are: transmission of energy through the end mirrors, scattering by imperfections in the mirror surface, scattering and reflections from imperfect Brewster windows, and diffraction losses within the resonator. The useful output of the laser is usually taken by transmission through one of the end mirrors. If both
end mirrors are transmitting to some extent, then both transmissions may be considered useful output, or one may be considered useful output and the other a source of internal loss, depending on the application of the laser.

**TABLE 1**

<table>
<thead>
<tr>
<th>Length $L$ (cm)</th>
<th>$c/2L$ (MHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>1500</td>
</tr>
<tr>
<td>30</td>
<td>500</td>
</tr>
<tr>
<td>60</td>
<td>250</td>
</tr>
<tr>
<td>100</td>
<td>150</td>
</tr>
<tr>
<td>180</td>
<td>83</td>
</tr>
</tbody>
</table>

The general rule for stable operation of all lasers is this: The peak unsaturated gain of the laser medium must be greater than the sum of the losses and the useful output. This is elaborated upon in the section on Steady State Equations of Laser Operation.

6. **“Hole Burning”**

This term, which will be used frequently in the following discussions, is intended to describe graphically the process whereby energy for the laser operation is extracted from only certain groups of atoms occupying various positions under the Doppler profile. When a laser is oscillating, it is drawing energy out of those atoms that are at resonance with whatever modes are being sustained within the laser, and in the process of doing so the amount of population inversion which gives rise to the laser energy is reduced: the population is “drained” of its energy. Another way of stating this is to say that the laser mode reduces the saturated gain in the vicinity of its oscillation frequency. Consider, for simplicity, a laser operating in only one temporal mode. It will interact with the atoms whose positions under the Doppler profile lie within a certain range of the oscillation frequency. (How wide this range is depends in detail upon the atomic interactions and is discussed later.) If this range is appreciably narrower than the Doppler width itself, then the saturated gain is depressed in the region immediately around the oscillation frequency but is unperturbed elsewhere in the Doppler profile. We have thus “burned a hole” in the Doppler profile. If several temporal modes are operating, then, of course, holes are burned in the vicinity of each oscillation frequency.
Figure 12. Hole burning by a single temporal mode. (a) Mode frequency $v$ centered within the Doppler profile. (b) Mode frequency offset from Doppler center.

Figure 12 shows the process of hole burning for a single temporal mode, for the case where the mode frequency is at the center of the Doppler profile and where it is offset to one side. When it is offset, two holes are burned, one at the oscillation frequency and one at a position symmetrically located with respect to the center of the Doppler curve. The existence of the two symmetrically located holes has to do with the existence of a standing wave within the laser resonator. This standing wave can be considered to be made up of two traveling waves moving in opposite directions, thus there is symmetry within the laser with regard to motion in opposite directions along the axis. If there is a hole burned among atoms traveling at a velocity $v$ in one direction along the laser axis, it is because those atoms are at resonance, owing to a Doppler shift, with one of the traveling waves of the laser mode. However, because of the symmetry, this also means that atoms traveling at a velocity $-v$ will be at resonance with the other traveling wave. Hence there exist two symmetrically located holes. If the oscillating laser mode is at the center of the Doppler curve, then clearly only one hole can be burned instead of two, and this might be expected to cause a difference in output power under some conditions. The phenomenon does in fact exist
and is discussed in Chapter 3 under the topic of the power dip. Also, if a resonator is devised to use only traveling waves instead of standing waves, then clearly the symmetry with respect to Doppler line center no longer exists and only one hole is burned for any given oscillating mode. A short discussion of traveling wave resonators is given at the end of Chapter 3.

7. Homogeneous and Inhomogeneous Broadening

These terms relate to details of laser operation as discussed in the next section. However, since they are closely related to the question of hole burning, it is convenient to define the terms here. If a particular laser operates in such a way that the Doppler gain curve is uniformly "burned", that is, uniformly and thoroughly saturated, then it is said to be homogeneously broadened. If, on the other hand, it operates so that one or more discrete holes are burned leaving sections of unsaturated gain, then it is said to be inhomogeneously broadened. Homogeneous broadening can occur for one of two possible reasons: (a) the Doppler width is narrower than the theoretical width of the hole that can be burned (this occurs primarily in far infrared lasers), or (b) the laser is operating in many temporal modes spaced closely enough together so that no region of the gain curve is left unaffected by a temporal mode. Whether a laser is homogeneously or inhomogeneously broadened is therefore a question of design in some cases. Since inhomogeneous broadening represents the existence of untapped sources of laser power, as indicated by unsaturated portions of the gain curve, it represents less efficient operation of the laser than is the case with homogeneous broadening.

8. Near-Field and Far-Field Optical Patterns

The terms "near field" and "far field" occur frequently in the literature on lasers. We shall have occasion to use them in connection with discussions on laser spatial modes and on the propagation of the laser beam after it leaves the source. For the benefit of readers who have not previously encountered these terms, we give short definitions of them here.

When a beam of light is output from a laser or any other optical device, it has some sort of well-defined energy distribution as a function of position across the beam cross section. This is the near-field intensity distribution across the beam. The beam is propagated into space initially according to the laws of geometrical optics, but after a while diffraction effects tend to modify the intensity distribution in the beam. Eventually, a point may be reached where the beam has spread out to such a degree
that it covers a considerably larger area than would have been predicted by geometrical optics, and the intensity distribution in the beam is determined entirely by diffraction. This is the far-field distribution. More precise definitions in terms of transform properties are given in textbooks on physical optics; however, the preceding definitions will be adequate for our purposes.*

Our definitions are intended, in a sense, to apply to plane or approximately plane wave propagation, although they also apply to propagation in the form of spherical wave fronts. In particular, consider the near-field distribution to be in a converging spherical wave. According to geometrical optics, such a wave front should converge to an infinitely small point at the focus; however, we know that diffraction effects will always prevent the distribution at the focus from being infinitely small. It follows therefore that the distribution at the focal point of a converging wave is always the far-field distribution of the initial converging wave. A much more rigorous proof of this statement is given by Born and Wolf [25]. This fact is important in connection with the theory of laser resonators, as well as the treatment of propagation of laser beams given in Chapter 4 of this book.

There are two near-field intensity distributions whose far-field distributions will be of particular interest to us. One is the intensity distribution of a uniformly illuminated wave passed through a circular aperture of radius \( a \). The far-field pattern of this distribution is given by

\[
I(v) = I_0 \left( \frac{2 J_1(v)}{v} \right)^2,
\]

where \( J_1(v) \) is the Bessel function of order one and \( v \) is a length coordinate whose actual measurement depends on the wavelength of the light and the details of the optical system involved (it is defined more exactly in Chapter 4). The distribution of (8) is known as the Airy distribution and, when observed on a screen, is often known as the Airy disk. (The energy distribution in the Airy disk is shown in Figure 43.) Another near-field distribution that will be of interest to us is the Gaussian distribution given by

\[
I(r) = I_0 e^{-r^2/\alpha^2}.
\]

The far-field distribution corresponding to (9) is given by

\[
I(v) = \frac{I_0}{\pi} e^{-v^2}.
\]

*In textbooks on physical optics, the diffraction pattern that gives rise to the far-field intensity distribution is usually defined as the Fraunhofer diffraction pattern.
In other words, for the Gaussian distribution the far-field distribution has the same form as the near-field distribution. This fact makes the Gaussian distribution particularly important in discussions on the generation and propagation of laser modes.

D. PHYSICS OF GAS LASER OPERATION

The following discussion briefly outlines the topic, emphasizing specific points that are essential for the understanding of all gas lasers. For much more detailed discussions, the reader is referred to other texts that discuss laser physics in general [2-4] and to the earlier review articles on gas lasers such as those of Bennett [8,26], which discuss excitation and inversion mechanisms of gas lasers in much more detail that is possible here.

1. Stimulated Emission versus Spontaneous Emission

The basis for the physics of all maser and laser operation is the existence of stimulated emission from an upper energy level of a quantum mechanical system. Consider an atom having two energy levels separated by energy $E$, where $E = h\nu$. The atom is surrounded by an electromagnetic field which can be considered to be made up of waves traveling in various directions, and each of these waves in turn can be considered to be made up of a number, $n_\nu$, of quanta or photons, each of energy $h\nu$. If the atom is in the lower state, then the probability that the atom will absorb one of these quanta and arrive in the upper state is

$$p d\Omega = \frac{8\pi^3 e^2 \nu^3 d\Omega}{hc^3} |\mu|^2 n_\nu \cos^2 \theta.$$  \hspace{1cm} (11)

The probability that an atom in the upper state will emit a photon and drop to the lower state is

$$p d\Omega = \frac{8\pi^3 e^2 \nu^3 d\Omega}{hc^3} |\mu|^2 (n_\nu + 1) \cos^2 \theta.$$  \hspace{1cm} (12)

In these formulas $p$ is the transition probability per unit time, $d\Omega$ is an element of solid angle, $\nu$ is the optical frequency, $\mu$ is the matrix element for the transition, $\theta$ is an angle relating the polarization of the field to a "preferred" polarization of the atom, and $e$, $h$, and $c$ have their usual meanings as fundamental constants. These formulas are derived in most textbooks on quantum mechanics; the form used here was taken from Heitler [27]. For our present purposes, the term $n_\nu$ can be considered approximately equal to the number of photons existing in each mode of the laser resonator, if the active medium referred to is in the resonator. The terms involving $n_\nu$ in (11) and (12) thus represent the absorption
and stimulated emission probabilities, respectively, whereas the extra term—"1"—in (12) represents the spontaneous emission.

Because it does not depend on preexisting conditions in the electromagnetic field, in the spontaneous emission the phase of the output is completely unpredictable, only the rate at which the spontaneous emission occurs is predictable. Thus consider an ensemble of atoms capable of radiating spontaneously, such as those in the upper laser level of a gas laser. Although many atoms are emitting spontaneously at the same time, the fact that the phases are random means that mean-square averaging can be applied to the output of the ensemble taken as a whole; all phase relationships and interferences between the radiation of different atoms average out and the spontaneous radiation can be treated statistically merely as the sum of the radiation of the individual atoms. As a result it is isotropic in space and has the characteristics of "white noise" within the Doppler bandwidth of the emission. Spontaneous emission is always present in a laser therefore and represents a source of noise as well as a source of signal useful for starting the laser oscillation. However, only that part of the spontaneous emission which is within the solid angle subtended by the laser mode actually becomes part of the stored energy in the laser mode.

The stimulated absorption and emission, on the other hand, have the important property that they are coherent in phase with the electromagnetic field variations which produce them. The reason that this is so of course is buried in the fundamentals of quantum electrodynamics, but some insight can be obtained by going back to the original treatment of thermodynamic equilibrium in the radiation field by Einstein [28]. This theory, which preceded quantum mechanics by quite a few years, showed that the coefficients of stimulated absorption and emission have to be equal and opposite in nature, so that radiating matter can achieve thermodynamic equilibrium with the radiation field. Now, consider the stimulated absorption from the standpoint of a dipole or other multipole interacting with the radiation field. In order that energy be absorbed it is essential that there be a consistent phase relationship between the phase of the electromagnetic field and the phase of the oscillating dipole that characterizes the absorbing particle. If the absorbing particle is considered as an oscillating dipole, it must also produce its own radiation field, and the phase of the radiation field must be such as to subtract, by destructive interference, from the intensity of the incident field at large distances from the particle. If a consistent phase relationship did not exist, energy would not, on the average, be absorbed from the radiation field. The difference between absorption and stimulated emission, according to theories of radiation equilibrium, is only that the phase
then clearly we have gain and the quantity \( \alpha \) is the gain per unit length in the medium.

The calculation can be performed in terms of the Einstein \( A \) and \( B \) coefficients for thermodynamic equilibrium, as defined by Mitchell and Zemansky. Here \( A \) is the probability per unit time that an atom in the upper state will spontaneously emit a photon and drop to the lower state, and \( Bn_p \) is the probability that an atom is stimulated to absorb or emit a photon, as the case may be, into a mode of the electromagnetic field containing \( n_p \) photons. [There is thus a close relationship between the \( A \) and \( B \) coefficients and the quantities defined by (11) and (12). For simplicity, we assume that the statistical weight factors of the upper and lower states are equal. If they are not, a minor change in the formulation is required, as given by Mitchell and Zemansky.]

The relationship between the \( A \) and \( B \) coefficients is

\[
\frac{A}{B} = \frac{2\hbar v^3}{c^2}.
\]  

In order that the concept of a transition probability be valid, it is necessary to assume that the radiation field and absorption resonance are not infinitely sharp but cover a small frequency range \( \Delta \nu \). We suppose that there are \( \Delta N' \) atoms per cubic centimeter in the lower state and \( \Delta N \) in the upper state resonant over the frequency range \( \Delta \nu \). Mitchell and Zemansky then show that

\[
-(\alpha \Delta \nu) = \frac{\hbar \nu}{4 \pi} B (\Delta N' - \Delta N).
\]  

This could also be written as

\[
\alpha = \frac{\frac{\hbar \nu}{4 \pi} B (N' - N)}{\Delta \nu},
\]  

expressing the fact that, for given total populations \( N' \) and \( N \) in the upper and lower states, \( \alpha \) is inversely proportional to the frequency range, a fact which is frequently noted elsewhere in this book. Note also that \( \alpha \) is positive and there is net gain instead of absorption whenever \( N' \) is greater than \( N \), that is, whenever there is a population “inversion.”

Returning now to (14), we note that the \( A \) coefficient is the inverse of the lifetime of the upper state against spontaneous emission to this particular lower state. If we call this lifetime \( \tau \) and substitute in (14) and (16), we arrive at the following equation for the gain coefficient as a function of the lifetime \( \tau \):

\[
\alpha = \frac{\left( \frac{c^2}{8 \pi \nu^4} \right) (N' - N)}{\Delta \nu}.
\]
In this formula $N$ and $N'$ must have the units of total population per unit length in order that $\alpha$ have the correct units of inverse length. This formula is fundamental and applies to all lasers provided the appropriate quantity $\tau$ can be measured. It is sometimes convenient to express $\alpha$ in terms of the matrix element of the transition. If the matrix element is $\mu$, then the well known relationship between $\mu$ and $\tau$ is

$$\tau = \frac{3hc^3}{64\pi^4\nu^3\mu^2}$$

(18)

We have then

$$\alpha = \frac{8}{3} \left( \frac{\pi^2\nu^2\mu^2}{hc} \right) \frac{(N' - N)}{\Delta \nu}.$$  

(19)

In many lasers the total gain $\alpha x$ is small enough that the approximation

$$e^{\alpha x} \approx 1 + \alpha x$$

(20)

can be used. In these cases it has become convenient to speak of the gain in terms of a fractional increase in signal; thus one speaks of the gain as "percentage gain" where the percentage is $\alpha x \times 100$. Otherwise, a logarithmic unit such as decibel is more useful and is applicable to all situations. Moreover, if the gain is expressed in logarithmic units, then, if taken as a function of frequency under the Doppler profile, it is always proportional to the ordinate of the Doppler profile. For small gains, the quantity $(e^{\alpha x} - 1)$ is also approximately proportional to the ordinate of the profile, and this is the situation usually assumed in considering output power of a laser as a function of frequency deviation from the center of the Doppler curve. On the other hand, if the gain is high and one considers power output, significant deviations from the actual Doppler shape can be expected. Mitchell and Zemansky discuss the flattening effect of a large absorption coefficient upon the observed profile. Where there is gain instead of absorption, the effect is to narrow the profile. If the natural linewidth* is very narrow, then the half power point on the Doppler Gaussian becomes, in fact, the square-root point for the total gain, so that, if the gain is very high, the observed profile may be very narrow indeed. A limit is eventually set by the fact that the observed profile cannot be narrower than the natural linewidth. Use of this phenomenon has been made to measure natural linewidths of laser transitions [30].

*The natural linewidth is the range of frequencies over which atoms having identical velocities will be in resonance. To a good approximation, it is the sum of the $A$ coefficients of the upper and lower states.
3. The Bloch Equations

A description of the dynamical behavior of the atoms interacting with the laser radiation fields is perhaps obtained most easily by use of equations derived by Bloch [5] for use in nuclear magnetic resonance. It may seem surprising that a theory intended for use with nuclear magnetic resonance should have application in laser theory, but this is so due to a general unity in form imposed by quantum mechanics. It may be desirable to go briefly into the reasons for this. The original analysis of nuclear magnetic resonance concerned a system involving only two energy states. In a laser, the atoms involved in laser action have more than two energy levels, but as a rule only two of these levels—the upper and lower laser states—are actually involved in the dynamical behavior of the system to first approximation. Thus in both cases we have two-level systems subject to perturbation by resonant applied fields and to restoration of equilibrium conditions by relaxation processes. Now, a fundamental axiom of quantum mechanics is that dynamical variables of a system are associated with operators on the wave functions of this system. For a system consisting of only two energy levels, such a wave function might take the form

\[ \psi = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \]  

(21)

where the wave function is a two-element vector and \( a_1 \) and \( a_2 \) are the probability amplitudes for finding a particle in states 1 and 2, respectively. The operators on a two-element vector are four-element matrices, and it can readily be shown that there can be only four linearly independent \( 2 \times 2 \) matrices, including the identity matrix. (The Pauli spin operators are an example of a set of such matrices.) Thus there can be only four independent dynamical variables of a two-level quantum mechanical system, and if the variables are specified for one such system, the variables in any other system will be isomorphic to them in some sense. Since four such dynamical variables are defined by the Bloch equations, they can be taken over directly for use in a laser system. The remaining question is the extent to which the relaxation processes described by Bloch can be employed in describing lasers. These relaxation processes, describing simple exponential approaches to thermal equilibrium, are probably somewhat simpler than those that actually exist in lasers, but the approximation is good enough to make the equations useful for most purposes. For examples of the fruitfulness of this approach the reader is referred to Bloembergen [6] and to work on the transient response of laser systems to very sharp pulses, a few examples of which are given in the bibliography [31-33]. Many other examples also exist in the literature.
In the Bloch formulation, the identity matrix corresponds to the total available population; for present purposes it can be assumed to be constant and ignored. The remaining three dynamical variables are denoted by \( u \), \( v \), and \( M_z \), whose corresponding operators are as follows:

\[
\begin{align*}
\hat{u} & \approx \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cos \omega t + \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \sin \omega t, \\
\hat{v} & \approx -\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \sin \omega t + \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \cos \omega t, \\
\hat{M}_z & \approx \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.
\end{align*}
\]

Here, \( \omega = 2\pi \nu \) and \( \nu \) is a frequency which is generally chosen to be equal to the frequency of the "driving" field that generates the resonant behavior. Note that this is not necessarily the resonant frequency of the atoms themselves; like any other resonant system, an atomic resonance can be driven slightly off-frequency, though with reduced response.

The physical significance of the dynamical variables is as follows: \( M_z \) corresponds to the saturated population difference between the upper and lower laser levels; \( v \) corresponds to an oscillating dipole moment which is either plus or minus 90° out of phase with the driving electromagnetic field. It can be shown easily that this dynamical variable interacts directly with the energy content of the driving field, either doing work upon the field, as in a laser, or absorbing energy from the field in an absorption line. Whether it absorbs or emits energy depends, of course, upon the sign of \( v \). The remaining variable \( u \) is the in-phase component of oscillating dipole moment. This component, because of its phase relationship, cannot change the energy of the driving field but instead changes the phase velocity for propagating waves. It is responsible for anomalous dispersion in absorption lines and for similar effects in lasers. However, our primary concerns here are the variables \( v \) and \( M_z \).

It is also necessary to define the relaxation terms to be used with these variables. Rather than use the original Bloch notation of relaxation times \( T_1 \) and \( T_2 \), we shall employ a notation more in accord with spectroscopic practice and define relaxation rates \( \gamma_{ab} \) and \( \gamma'_{ab} \). \( \gamma_{ab} \) is the spontaneous emission or hard collision relaxation rate for the laser transition. It represents the average rate at which atoms interacting with the applied field are removed from the interaction, either by spontaneous emission processes or by hard collisions that remove the atoms to entirely different states. The quantity \( \gamma'_{ab} \) includes not only the relaxation rate \( \gamma_{ab} \) but also additional terms caused by soft collisions, which may cause the phase of the wave function of an interacting atom to wander but
not completely destroy its coherence. More detailed discussions of the meanings of these terms have been given by other authors [34], and it is also not obvious where one draws the line between “hard” and “soft” collisions, but we are concerned here only with the basic essentials of the theory.

Within the framework of these definitions, then, the differential equations governing the dynamical variables $u, v,$ and $M_z$ are

\[
\begin{align*}
\frac{du}{dt} + \gamma_{ab} u + \Delta \omega v &= 0 \quad (25) \\
\frac{dv}{dt} + \gamma'_{ab} v - \Delta \omega u &= 0 \quad (26) \\
\frac{dM_z}{dt} + \gamma_{ab} M_z - \omega_1 v &= (N' - N) \gamma_{ab}. \quad (27)
\end{align*}
\]

Here, the following additional definitions are needed: $\Delta \omega = 2\pi (\nu_0 - \nu)$, where $\nu_0$ is the natural resonant frequency and $\nu$ is the driving frequency. $\omega_1$ is defined by

\[
\omega_1 = \frac{|\mu| E}{\hbar}
\]

where $\mu$ is the matrix element and $E$ the mean-square value of the driving field; this is the strength of the atomic interaction. $N' - N$ is, of course, the unsaturated or initial value of population difference as used in (16) to (19).

There are numerous transient (time-dependent) solutions of these equations that have interesting properties; these have been discussed by Bloch [5], Torrey [35], Hahn [36], and many others. We shall be more concerned with the steady state solutions. These, obtained by solving (25) to (27) with the differentials set equal to zero, are

\[
\begin{align*}
u &= -\omega_1 (N' - N) L (\Delta \omega) \\
M_z &= \left[ 1 + \left( \frac{\Delta \omega}{\gamma_{ab}} \right)^2 \right] (N' - N) L (\Delta \omega),
\end{align*}
\]

where $L (\Delta \omega)$ is a general resonance expression common to all three of these formulas and is given by

\[
L (\Delta \omega) = \frac{1}{1 + \left( \frac{\Delta \omega}{\gamma_{ab}} \right)^2 + \omega_1^2}.
\]
The quantity $M_z$, as we have pointed out, is directly proportional to the saturated population difference for a homogeneous population whose resonant frequency is $v_0$. The energy actually injected into the laser resonance mode by the laser medium is proportional to $v \left| E \times v \right|$; since we are generally concerned with extremely small variations in $v$ we shall consider $v$ to be constant. For further development of these equations, we wish to express energy injection in terms of the energy already present in the laser, which we define as $W$. $W$ is proportional to $E^2$. By substituting $W$ into (28), (30), and (32) and collecting all constants, we find this simple expression for the rate of energy injection:

$$\frac{dW}{dt} \propto v |E_v| \propto \frac{GW}{1 + \left( \frac{\Delta \omega}{\gamma'_{ab}} \right)^2 + SW},$$

where $G$ is the unsaturated gain and $S$ will be referred to as the saturation constant.

4. Steady State Equations of Laser Operation

We are now in a position to calculate the power output of a gas laser as a function of its gain, losses, and mirror transmission. However, it becomes necessary at this point to differentiate between homogeneously and inhomogeneously broadened lasers. The Bloch equations [(29) to (32)] can be used directly for the homogeneously broadened case, because they apply to ensembles of atoms whose resonances are all essentially at the same frequency. We shall derive the steady state conditions for homogeneous broadening first, going on to the inhomogeneous case later. It should be borne in mind that “steady state” applies to the output of almost all gas lasers including pulsed lasers, because the relaxation times involved in reaching dynamic equilibrium are generally shorter than the pulse length, even when submicrosecond length pulses are used.

The increase in energy in a laser resonator due to the active medium is proportional to the $v$ term, or to (33). With the driving field at frequency $v_0$, and reduced to simplest terms, the energy increase can be written as follows:

$$\frac{dW}{dt} = \frac{GW}{1 + SW}.$$  \hspace{1cm} (34)

The removal of energy from the resonator is due to the output mirror transmittance $T$ and undesirable losses $A$. Thus for the losses we have

$$\frac{dW}{dt} = -(A + T) W.$$  \hspace{1cm} (35)
At steady state, the energy increase equals the energy removal so that we have

\[(A + T) W = \frac{GW}{1 + SW}. \tag{36}\]

If the driving field is not at frequency \(\nu_0\), then, referring to (33), it will be seen that the effect on (34) is merely that of changing the proportionality constants \(G\) and \(S\), but the form of the equation itself is unchanged.

The output power \(P\) of the laser is then:

\[P = TW = \frac{T}{S} \left( \frac{G}{A + T} - 1 \right). \tag{37}\]

Negative values of \(P\) clearly have no meaning and thus the equation states the obvious—that steady oscillation is obtained only when the gain is greater than the loss. Note that the saturation constant \(S\) appears only as a proportionality term. From (32) it will be seen that \(S\) is determined by relaxation rates and by the matrix element of the transition; thus in a given laser its value is out of the control of the experimenter. Its numerical value can be determined by experiment if this seems desirable.

Equation 37 forms the basis for calculating the mirror transmittance that gives optimum output power for fixed gain and loss. This optimum transmittance is

\[T_{opt} = \sqrt{GA - A}. \tag{38}\]

This relationship is similar to but more complicated than the well known relationship in electric circuits that states that optimum efficiency is obtained when the load equals the internal resistance. It applies, strictly speaking, to the case of homogeneous broadening with moderate gain—of the order of not more than 40 or 50% per pass. For higher gain one must take into account the exponential growth of the oscillation signal, whereas (38) merely assumes a lumped active element emitting power into the resonator mode. Rigrod [37] has given expressions for the case of arbitrary gain, and has plotted the values of \(T_{opt}\) for various gains and losses. Even these formulas, however, do not take into account the effect of spontaneous emission. All of the aforementioned formulas—either (37) or Rigrod’s formulas—predict zero power output if the transmittance of one mirror is equal to unity (perfect transparency) because then, in principle, there is no longer any laser resonator. In fact, under conditions of sufficiently high gain it is possible for the spontaneous emission from one end of the laser to be amplified sufficiently to saturate.
Physics of Gas Laser Operation

the active medium at the other end of the laser (or at the same end if there is a mirror at the opposite end). This phenomenon has been given the name of superradiance*. It has been observed numerous times in lasers of high gain, particularly pulsed lasers where very high gains per pass can be achieved for short periods of time. Because of the absence of a structure that defines modes, however, a superradiant laser does not generate light having a very high degree of coherence.

The calculation of steady state conditions for an inhomogeneously broadened line differs from that for homogeneous broadening in that one must consider the existence of a driving field at fixed frequency $\nu$ and an ensemble of particles covering a range of values of $\Delta \omega$. At each of these values of $\Delta \omega$, there is a value of $M_z$ given by (31) that is determined by substituting for $\Delta \omega$ in that equation. Figure 13 shows $M_z$ as a function of $\Delta \omega$ for several intensities of the driving field. This is quite

---

*The term *superradiant state* was originally coined by R. H. Dicke [38] to describe certain coherent radiating states in the active medium. Its use for the purpose described here is thus a misnomer; however, it has become so common in the literature that we shall use it to mean amplified spontaneous emission.
obviously the hole-burning phenomenon described earlier. Note that for very strong driving fields the hole becomes quite wide, enabling the field to draw energy from atoms that are quite far removed from the resonant frequency.

For inhomogeneous broadening, the quantity \( v \) must be replaced by an integral of the contributions to \( v \) from atoms covering the entire range of resonant frequencies. Such a computation, applied to steady state laser conditions, was first derived by J. P. Gordon at the Bell Telephone Laboratories. Because the computation is in the form of unpublished notes not generally available, however, we shall derive it independently here. Clearly, the integral must take the form

\[
\nu_{inh} = \int_{-\infty}^{\infty} v(\nu_0 - \nu) F(\nu_0) \, d\nu_0, \tag{39}
\]

where \((\nu_0 - \nu) = \Delta \omega / 2\pi\) as before. The integral is thus of the general form

\[
\nu_{inh} = \int_{-\infty}^{\infty} \frac{v_0 F(\nu_0)}{(1 + SW) + \left[ (\nu_0 - \nu) / \gamma_{ab} \right]^2}, \tag{40}
\]

where \(v_0 \propto v(\Delta \omega = 0)\). The function \( F \) represents the distribution of atoms over the frequency range, for example, the Gaussian curve of the Doppler profile. If \( F \) is taken to be the Gaussian, the integral can be solved but not in terms of elementary functions [39]. A simpler distribution, which is nevertheless physically meaningful, is a Lorentzian distribution centered on the driving frequency \( \nu \), given by

\[
F(\nu_0) = \frac{1}{1 + \left[ (\nu_0 - \nu) / \delta' \right]^2}, \tag{41}
\]

where \( \delta' \) is the half-width at half-maximum of the distribution. Substitution of (41) into (40) gives an integral that can be solved analytically. The result is

\[
\nu_{inh} = \frac{\gamma_{ab} \nu_0 \pi}{1 + \frac{1}{\sqrt{1 + SW} \delta^2}} \left( \frac{1}{\sqrt{1 + SW} \delta} - 1 \right), \tag{42}
\]

where \( \delta = \delta' / \gamma_{ab} \). There are two limiting cases. For \( \delta \) very small, (42) reduces to the same form as the result for homogeneous broadening. This corresponds to a hole width larger than the distribution width. In the other limit — \( \delta \) very large — the result is
This is the result obtained by Gordon. We can now write it in the form

\[
dW = GW \frac{\gamma_{ab} v_0 \pi}{\sqrt{1 + SW}} \, dt \tag{44}
\]

and, combining this with (35), we obtain the following for the power output of an inhomogeneously broadened laser:

\[
P = \frac{T}{S} \left[ \left( \frac{G}{A+T} \right)^2 - 1 \right]. \tag{45}
\]

The important fact about (45) is that under some conditions the power varies with the square of the gain rather than with of the first power, as in (37). This can be seen qualitatively in terms of the hole burning concept, since higher field intensities within the laser not only result in higher output power but also in a wider hole being burned so that more atoms are able to emit their energy into the laser mode. This relationship is very important in certain types of lasers that exhibit inhomogeneous broadening characteristics, particularly the CW ion lasers.

There is no simple expression for the value of \( T_{\text{opt}} \) in the inhomogeneously broadened case. Rigrod [37] has calculated \( T_{\text{opt}} \) for homogeneous and inhomogeneous broadening and for various gains and losses. The general form of these results is shown in Figure 14. Figure 14a shows the general nature of the power output curve as the transmittance \( T \) is varied from 0 to 100%. Figure 14b shows the power output at \( T_{\text{opt}} \) as a function of the losses \( A \) in the resonator for both the homogeneously and inhomogeneously broadened cases. The following important difference should be noted. Most lasers operate with relatively small losses, corresponding to operation close to the \( Y \) axis of Figure 14b. Under such conditions, if the losses are reduced still further, for example, by a factor of 10, the power output that can be obtained at the new value of optimum transmittance is changed hardly at all for homogeneous broadening. However, in the case of the inhomogeneously broadened line, the new value of \( T_{\text{opt}} \) is smaller, thus permitting less leakage of energy out of the resonator and building up a greater intensity in the standing wave. This in turn increases the interaction of the resonator mode with the atoms in the inhomogeneously broadened line and results in a considerable increase in the available power. If (24) is taken literally, this curve increases without limit as it approaches the
Y axis. This, of course, is not the case in practice, since a leveling off will occur at the point at which the “hole” becomes equal to the width of the Doppler curve. However, the dramatic increase in available output power from an inhomogeneously broadened laser as the losses are reduced has been observed experimentally and is of particular importance in small, single frequency lasers where the ratio of gain to fixed losses is not overly large.

5. Zeeman Effect

The Zeeman effect is a well known spectroscopic effect involving splitting of atomic spectrum lines when a magnetic field is applied. It enters into the present discussion only indirectly, because magnetic fields are sometimes applied to lasers for various reasons, with consequent effects on the gain and output linewidth. The following is intended as the briefest possible introduction; for more detail the reader may consult any textbook on atomic physics or spectroscopy.

The Zeeman effect in its simplest form is the splitting of a spectrum line into several components with well-defined polarizations, the splitting being generally proportional to the magnetic field. Figure 15 shows

![Figure 14. Typical curves of output power as a function of mirror transmittance and loss.](image)
the components and polarizations observed when the magnetic field is parallel and perpendicular, respectively, to the line of sight. It should be emphasized that Figure 15 shows only one simple case of Zeeman effect; in many specific examples each of the polarization components shown is further split into other components. The amount that the outer components are displaced from original line center can vary widely for different spectrum lines but is generally in the vicinity of 1.4 MHz/G of applied field. In the case of a field perpendicular to the line of sight the spectrum line is split into three linearly polarized components; with a longitudinal field there are two circularly polarized components whose senses of polarization are opposite to each other.

In a laser, application of a strong magnetic field will not, in general, split the line into distinct spectral components unless the splitting is appreciably greater than the Doppler width. If the splitting is only comparable to the Doppler width, an effective broadening of the line occurs, with consequent reduction of gain in any given frequency range and with the gain somewhat stronger in preferred polarizations than in others. Strong longitudinal magnetic fields are particularly important
for efficient operation in the CW ion lasers, and they can be expected to provide appreciable Zeeman splitting there. A signal of arbitrary polarization injected into one end of an ion laser can be considered decomposed into two circularly polarized components, one of which will be amplified preferentially over the other by the laser medium because it may fall within one or the other of the Zeeman-split components. If Brewster angle windows are present, however, they treat the amplified laser radiation as made up of two linearly polarized components, one of which is partially reflected out of the cavity. Thus, in establishing static equilibrium in a laser mode, the laser medium tends to amplify circularly or elliptically polarized waves, whereas the Brewster angle windows tend to preferentially attenuate certain linearly polarized components of the mode. The problem has been treated analytically by Sinclair [40], who shows that when steady state is reached the Brewster angle windows do not actually remove a great amount of energy from the mode unless the resonator losses are very high. This is the reason that Brewster angle windows can be used in commercial CW ion lasers.

In an internal mirror laser there are no polarization determining elements, and modes can exist with arbitrary polarizations. In this case application of even very weak magnetic fields may cause interesting effects such as slight differences in oscillation frequency between modes that have the same $\text{TEM}_{mnq}$ configuration but different polarizations. These effects are being studied by a number of investigators.

6. Excitation Mechanisms in Gas Lasers

An understanding of the excitation mechanisms that provide the population inversion in gas lasers is of great interest to physicists who are concerned about the basic processes in lasers. To a person who is using lasers, however, the question is only secondary, for two reasons. First, he can use only such lasers as are known to exist, and he has absolutely no control over the excitation processes that make them work. Second, theories of excitation processes are ex post facto in nature, and so far they have not been particularly useful for predicting new lasers. For these reasons, we give only a brief introduction to the problem of excitation in this book. The reader who is interested in further details should consult review articles in the literature that deal with the question explicitly. Among these are the following: for the neutral atom lasers, the review papers of Bennett [8, 26]; for the ion lasers, the second paper by Bennett [26] and a paper by E. I. Gordon et al. [41]; and for molecular lasers, the work of C. K. N. Patel [14].

An implicit requirement for any steady state inversion is that the lifetime of the lower laser state be shorter than that of the upper state.
Other than that, in both pulsed and CW lasers, it is necessary that the rate of population of the upper state be greater than that of the lower state. There are, in general, three different ways of populating the upper state.

1. Direct excitation by electron collision of atoms in some highly populated lower state, such as the ground state or an intermediate metastable state.

2. Multiple step processes in which an atom must undergo several discrete events before arriving in the desired energy state. These may consist of successive excitations by electron collisions, excitation to a higher energy level followed by cascading through spontaneous emission down to the upper laser state, or any combination of these.

3. Resonant processes whereby an atom excited to a level very close in energy to the desired level is able to raise a lasing atom to the desired upper state, giving up its own energy in the process.

Direct electron excitation is probably the most important populating method in most gas lasers employing moderate discharge currents; this includes neutral atom lasers except the helium-neon laser and most molecular lasers. The notable exception is the helium-neon visible laser, for which excitation of the lasing neon atoms is provided by resonant transfer of energy from excited helium atoms in one of the metastable states. The originally discovered helium-neon infrared laser transitions around 1.15 μ can apparently be populated either by direct collision or by resonant energy transfer from helium metastables. In this case, laser operation has been obtained in pure neon, indicating that direct collision in itself is sufficient to provide population inversion, but the process is apparently also greatly assisted by helium metastables, if they are present. Another situation of “mixed” processes is in the high-power CO₂ laser at 10.6 μ. Here it has been shown that a large population inversion is generated in CO₂ by resonant energy transfer from excited nitrogen atoms, but operation at somewhat lower power can also be obtained without nitrogen, or in a mixture of CO₂ and helium, indicating that other mechanisms must also be present to generate the population inversion. In this particular case, the other mechanism is probably a cascade process from above. Finally, in the CW ion lasers, there has been some debate as to whether excitation to the upper laser level is through a one-step or multiple-step process. There is at present considerable experimental evidence that the excitation process for these lasers depends on the length of time the discharge has been turned on. If the lasers are operated in short pulses, of the order of several microseconds in length, the excitation process is that of a single step from the
neutral atom ground state, but after a short period of time this mechanism becomes inadequate to maintain the inversion and multiple-step processes take over. Experimentally, this can be seen by a sharp pulse at the instant of turn-on of the discharge followed by a gap or greatly reduced power output, followed again by a slow rise to eventual CW operation in which multiple-step processes are apparently dominant [42]. Some additional details of these processes as they apply to the particular lasers mentioned are given in the next chapter.